Imperial College London

Mass Spectrometric Characterisation of Immunologically Important Glycosylation

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

Submitted by

Dongli Lu

Department of Life Sciences

Faculty of Natural Sciences

Imperial College London

South Kensington

London

SW7 2AZ

April 2019

Declaration of Originality

I hereby declare that the works presented in this thesis have not been submitted to any institution for acquiring a qualification at any level. All the research outcomes are originated from my own independent works at Imperial College London. Information from other sources is indicated and acknowledged.

Dongli Lu

1st April 2019

Declaration of Copyright

Dongli Lu

1st April 2019

Acknowledgement

"At long last, I would be able to say a few words of my ..."
—Edward VIII

It is now a serious moment in my life, it is a serious time to say my thanks to people who have supported me in the last four years to reach this milestone. I am very grateful that I could be guided into the field of Glycobiology by Professor Anne Dell and Dr. Stuart Haslam to discover the wonder of the invisible sweety world. I do thank Professor Dell and Dr. Haslam very much for their inspirational advice in my study and kindly supports during the worst period of my personal life. It is my honour to be here and to study. My experience in this laboratory would be the most unforgettable things to go along with me in my life as it is so precious.

I now would say thank you to all my previous and current colleagues for their selfless supports in my study: Dr. Gang Wu, Dr. Tiandi Yang, Dr. Qiushi Chen, Dr. Nan Jia, Dr. Laura Bouché, Dr. Grigorij Sutov, Dinah Rahman, Dr. Linda Ibeto and Dr. Anja Krueger. Specical thanks must give to Professor Howard Morris, Dr. Maria Panico, Dr. Aristotelis Antonopoulos and Dr. Paul Hitchen for their expertises in helping me to learn glycomics, proteomics and glycoproteomics. I also need to say thank you for all my collaborators: Professor Gabriela Dveksler, Professor Gary Clark, Professor Richard Pleass and Professor Alan Salama; without your supply of interesting materials, I would never reach here to complete this thesis.

I also need to say thank you to all my friends in London, China and around the world. You made me happy when the sky was grey, you would never know, dear, how much I love you all. But we would be together in heart forever. Special thanks must give to Dr. Yuhua Xia for his great supports and assistance in the year of thesis-writing.

At last but not the least, I must to thank my parents for their great supports during my study abroad. They gave me the greatest understandings in the difficult time when I was suffering. They might not be the best of best parents, but they are the only best in my life. I do appreicate for all the things they have done for me. I am indeed extremely lucky to be their child from the beginning of knowing the world. I owe them too much but I would do whatever I could do to never disappoint them.

Dongli Lu,
1st April, 2019

To
X and Jean-Jacques Rousseau

Publication

Blundell, P. A., **Lu, D**., Wilkinson, M., Dell, A., Haslam, S. & Pleass, R. J. 2019. Insertion of N-Terminal Hinge Glycosylation Enhances Interactions of the Fc Region of Human IgG1 Monomers with Glycan-Dependent Receptors and Blocks Hemagglutination by the Influenza Virus. **J Immunol**, 202, 1595-1611.

Abstract

Glycosylation is a vital modification for many macro-biomolecules to obtain structural heterogeneity and functional diversity. In the immune system, nearly all engaged molecules are glycosylated, and their glycosylation status directly correlates with their functions. Therefore, structural studies of glycosylation are important and fundamental to allow understanding of the basic mechanisms underlying the immune system. In this thesis, comprehensive mass spectrometric based glycomic and glycoproteomic strategies were applied to investigate the glycosylation of the following immunologically important glycoproteins and animal models: pregnancy specific glycoprotein 1 (PSG1), canine melanoma, fragment crystallisable (Fc) fusion protein and cryoglobulins.

Glycomic and glycoproteomic studies of native PSG1 showed that four of the seven N-glycosylation sites on this protein were occupied by glycans, which were dominated by mono- and bi-antennary complex structures with sialylation and partial core-fucosylation. N-glycosylation of recombinant N-terminal domain of PSG1 from Chinese Hamster Ovary (CHO) systems differs from the native materials for universal core-fucosylation and extensive poly-LacNAc elongation. N-glycosylation of the new Expi-CHO system was characterised as abundant truncated multi-antennary complex structures.

Analysis of the glycoprotein and glycolipid glycans expressed by stage III canine melanoma cells indicates that the glycosylation of this animal model is highly consistent with that of human melanoma with an exception of species-specific expression of Gal-α-Gal epitopes. These results support the idea that canine melanoma should be an excellent animal model for the pre-clinical studies of melanoma.

Glycomic studies of the Fc fusion protein with engineered additional N-glycosylation sites indicated an improvement of sialylation and branching on N-glycans in a CHO cell system. Identical fusion proteins expressed in a HEK293 cell system exhibited more diverse epitopes and considerable antennary truncation.

Finally, analysis of the N-glycosylation of cryoglobulin indicated that glycosylation might not play a vital role in the formation of the cryoprecipitate of mixed-type cryoglobulins.

List of Abbreviations

ADCC	Antibody dependent cell mediated cytotoxicity
BCR	B cell receptor
C1GalT1	Core 1GalNAc beta 1,3 galactosyltransferase 1
CA125	Cancer antigen 125
CerGlcT	Ceramide glucosyltransferase
CHO	Chinese hamster ovary
CID	Collision indued-disassociation
CLTA	Cytotoxic T lymphocyte antigen
CRD	Carbohydrate recognition domain
CTB	Cytotrophoblast cell
CTL	Cytotoxic T lymphocytes
CTLD	C-type lectin like domain
Da	Dalton, mass unit
DC	Dendritic cell
DC	Direct current
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
dNK	Decidual natural killer cell
ECM	Extracellular matrix
EDEM	ER Degradation-enhancing α-mannosidase-like protein
EI	Electron impact ionisation
ER	Endoplasmic reticulum

ERAD	ER associated degradation
ESI	Electrospray ionisation
ESL-1	E-selectin ligand-1
ETD	Electron transfer disassociation
Fc	Fragment crystallisable region
FcgammaR	Fc gamma receptor
Fuc-T	Fucosyltransferase
GC	Gas-phase chromatography
GlcNAc-T	GlcNAc-transferase
GSL	Glycosphingolipid
hCG	Human chorionic gonadotropin
HEK	Human embryonic kidney
HEV	High endothelial venules
Hex	Hexose
HexNAc	N-acetylhexosamine
HILIC	Hydrophilic interaction chromatography
HLA	Human leukocyte antigen
Hu-FEDs	Human fetoembryonic defence system
ICAM	Intracellular adhesion molecule
Ig	Immunoglobulin
IL	Interleukin
ITAM	Immunoreceptor tyrosine-based activating motif
ITIM	Immunoreceptor tyrosine-based inhibitory motif
LacNAc	N-acetyllactosamine
LC	Liquid-phase chromatography
m/z	Mass to charge ratio
MAG	Myelin-associated glycoprotein
MALDI	Matrix-assisted laser desorption ionisation
MHC	Major histocompatility complex
MS	Mass spectrometry

MS/MS	Tandem mass spectrometry
NK	Natural killer cell
OST	Oligosaccharyltransferase
PAMP	Pathogen-associated moelcular pattern
PNGase F	Peptide N-glycoside F
poly-LacNAc	Poly-N-acetyallactosamine
ppGalNAc-T	Polypeptide GalNAc-transferase
PPR	Pattern-recognition receptor
PSG1	Pregnancy specific glycoprotein 1
PSGL-1	P-selectin glycoprotein ligand-1
Q-TOF	Quadrupole time-of-flight mass analyser
QMS	Quadrupole mass analyser
RF	Rheumatoid factor (Chpater 6)
RF	Radio frequency (Chapter 1)
SHP	Src homology region 2 domain-containing phosphatase
ST6GalNAc-I	α-GalNAc alpha2,6 sialyltransferase
STB	Syncytiotrophoblast cell
T-antigen	Thomsen-Friedenreich antigen
T_C	T cytotoxic cell
TCR	T cell receptor
TGF	Transforming growth factor
T_H	T helper cells
TIM-3	T-cell immunoglobulin mucin 3
TOF	Time-of-flight
TOF/TOF	Tandem time-of-flight analyser
TSLP	Thymic stromal lymphopoietin
UGGT1	UDP-Glc:glycoprotein glucosyltransferase 1
VCAM	Vascular cell adhesion molecule
VEGF	Vascular endothelial growth factor
XIC	Extracted-ion chromatogram

Contents

Index of Figures

Index of Tables

Chapter I. Introduction

Chapter 1 Introduction

1.1 Glycobiology

Glycobiology is a field in life sciences which studies the nature of carbohydrates in respect of their structure, biosynthesis, function and evolution in all living systems (Cummings and Pierce, 2014). The term carbohydrates was coined in the late 19th century since their compositions fit an empirical formula of $C_n(H_2O)_n$ with possession of a carbonyl group as either an aldehyde or ketone (Seeberger, 2015). Emil Fischer firstly synthesised a structure of glucose, a common monosaccharide, and established the first classification and stereochemical projection of carbohydrates. However, such studies of carbohydrates were not immediately linked to their biological roles.

A century later, the term glycobiology was introduced in 1988 to emphasise the return of carbohydrates, renamed as glycans, from studies of chemistry to biology (Blow, 2009). It unified the modern understandings of cell and molecular biology with principles of biochemistry and carbohydrate chemistry (Varki and Kornfeld, 2015). This step forward was associated with the great developments of biological and analytical technologies, including mass spectrometry, glycan microarrays and gene sequencing. These technologies enable more comprehensive investigations of glycans and their receptors to elucidate their biological engagements. Due to the dramatic developments of glycobiology in the last few decades, glycans have been recognised as one of the four founding macromolecules of living systems along with nucleic acids, proteins and lipids (Hart and Copeland, 2010).

Glycans are widely distributed in nature via covalently linking to the other macromolecules as saccharide-chain to form glycoconjugates. Glycoconjugates in cells can be linked with nucleic acids or cytoplasmic proteins and therefore be involved in cytoplasmic signalling and gene expression. Glycoconjugates on cell surface in the form of glycoproteins and lipids can form a dense array of glycans that cover the surface, the so-called glycocalyx (Varki, 2007). The divergent functions of glycoconjugates are not only from the linking macromolecules, but also from the glycans themselves. Glycan sequences are not predictable from a gene template since only 1% of the genome of most species are related to glycan processing (Coutinho et al., 2003). The stereochemistry of each monosaccharide can further contribute to the complexity of glycan structures. Thus, the heterogeneity of glycoconjugates generates enormous inherent biological complexities.

In this thesis, I focus on projects characterising protein derived N-, O-glycans and glycolipid derived glycans from different living systems to understand their structural significance correlating to their functionalities in relevant events.

1.2 Monosaccharides

The simplest glycans are defined as monosaccharides which cannot be further hydrolysed to a smaller unit. The composition of all simple monosaccharides can be represented by the classic empirical formula, $C_n(H_2O)_n$, and the n commonly ranges from 3-9 (Seeberger, 2015). Monosaccharides are generally named by the number of carbons that they possess, such as pentose for five carbons and hexose for six carbons. Pentose and hexose are also the most common monosaccharides that can be found in mammalian systems.

Monosaccharides are a chain of hydroxymethylene units with a hydroxymethyl terminal at one end and a(n) aldehyde/ketone terminal at the other end (Seeberger, 2015). Due to the presence of a hydroxyl group, the carbon of hydroxymethylene is conferred with chirality. The chiral carbon can have two configurations, so that each monosaccharide can have 2^k possible stereoisomers, k corresponds to the number of chiral carbons. For instance, a six carbon hexose has four internal hydroxylmethylene, so it can have $16(2^4)$ different isomeric forms (Seeberger, 2015).

Following the nomenclature of organic chemistry, carbon atoms on monosaccharides are numbered from the aldehyde carbon as C-1 and ketone carbon as C-2 (Seeberger, 2015). There are two overall configurations, D and L, on each monosaccharide determined by a chiral carbon (configurational carbon) which is the furtherest one to the aldehyde/ketone group. Using the Fisher projection, the two overall configurations can be represented by the positioning of the hydroxyl group on the right (D-) or the left (L-) to the configurational carbon. Epimers are termed as a pair of monosaccharides which are only different from the configuration at one chiral carbon. The interchange of configurations of two epimers is named epimerisation. For instance, epimerisation of D-Gal at C-4 results in D-Glc, and the C-2 epimer of D-Glc is the D-Man (Figure 1.1).

Monosaccharides can be organised into two forms as acyclic and cyclic dependent on the reaction between the aldehyde/ketone group and one of the hydroxyl groups to form a cyclic hemiacetal or hemiketal. Generally, there are five or six-membered hemiacetals formed on monosaccharides named furanose and pyranose, respectively. The cyclic form of monosaccharide can be adopted into the Harworth Projection. Moreover, the formation of this cyclic hemiacetal introduces two more configurations onto monosaccharides as the aldehyde/ketone carbon is amended to a chiral one, named anomeric carbon. The two anomeric configurations are determined by the chiral consistency between anomeric and configurational carbons as α and β anomers. Simple monosaccharides can be further modified to form divergent monosaccharide structures. For example, N-acetylhexosamine is derived from hexose by a replacement of its C-2 hydroxyl group to an acetylated amino group. The N-acetyl-hexosamine can be further modified to larger structures named sialic acid by reacting with phospho-enolpyruvate. There are enormous potential variants of monosaccharide structures, due to stereo-
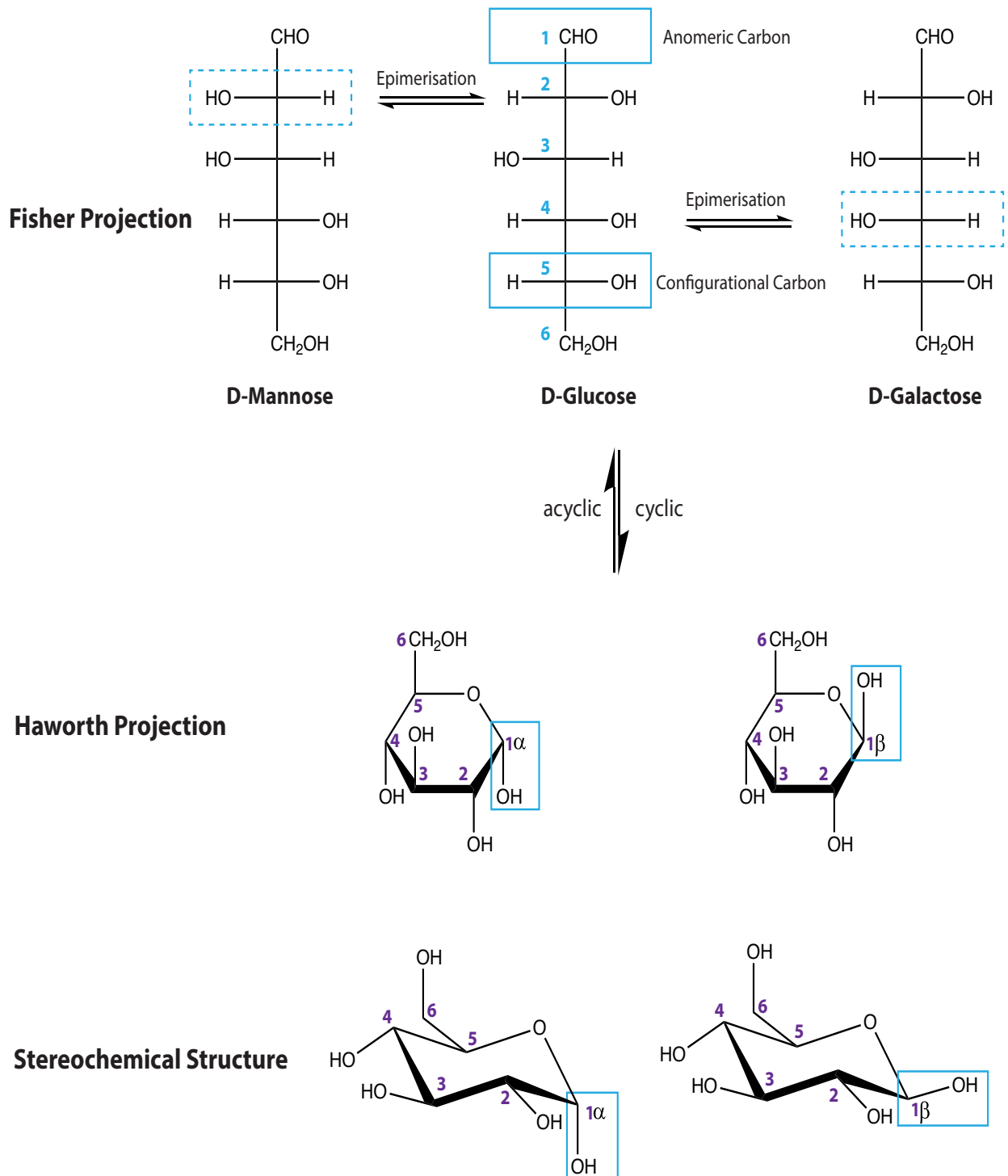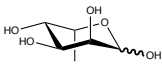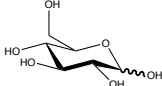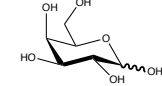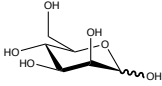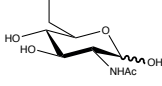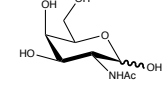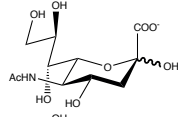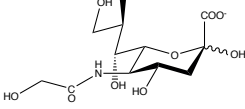
Figure 1.1 Structural presentation of Monosaccharides.
D-mannose, glucose and galactose are represented in Fisher Projection, and their epimerisation centres are highlighted with dashed blue square. Anomeric and configurational carbons are figured out with blue squares on acyclic D-glucose in the Fisher Projection, and the cyclic D-glucoses are presented in Haworth Projection and actual stereochemical structures with their anomeric carbons for α and β configurations highlighted with blue squares.

chemistry and modifications, whereas only eight monosaccharides are listed in this thesis which are commonly observed in mammalian systems (Table 1.1).

Glycosidic bonding is the main driving-force for monosaccharides to covalently connect with another monosaccharide and biomolecules. The bonding principally joints the hemiacetal group on the ano-meric carbon of a monosaccharide to a hydroxyl group of another molecule via formation of an acetal glycosidic bond. The glycosidic bond is the fundamental linkage of all glycans between connected monosaccharide residues (Seeberger, 2015). An example of glycosidic bonding is shown in Figure 1.2 as the formation of a disaccharide, Lactose. The C-1 anomeric carbon of a Gal and the C-4 carbon of a Glc are jointed together via a glycosidic bond to form the Lactose. The linkage of this glycosidic bond is defined by the anomeric configuration of the Gal and the carbon of the hydroxyl group involved on the Glc as β1,4. Thereby, the disaccharide can also be more accurately described as Galβ1,4Glc to reflect its composition and glycosidic linkage. The presenting orientation of the lactose sequence from left to right is defined by the reducing capability of the ends of the saccharides. As the Glc retains its hemiacetal group which still maintains its reducing power in reaction, this Glc is named the reducing-

Table 1.1 Common Mammalian Monosaccharides

Monosaccharide	Abbreviation	Symbol	Structure
L-Fucose	Fuc		
D-Glucose	Glc		
D-Galactose	Gal		
D-Mannose	Man		
D-N-Acetyl-Glucosamine	GlcNAc		
D-N-Acetyl-Galactosamine	GalNAc		
N-acetylneuraminic acid	NeuAc		
N-Glycolylneuraminic acid	NeuGc		

end of the lactose. On the other side, the hemiacetal group of Gal is engaged in the glycosidic bond and subsequently it has lost its reducing power, so that this Gal is named the non-reducing end. As glycans are connected by glycosidic bonds, there is a maximum of one reducing end on a glycan sequence and at least one non-reducing end. For simplicity a symbolic representation is applied in the presentation of glycan structures normally without presenting the glycosidic linkage of each bond.

The formation of glycans via glycosidic bonds is not energetically spontaneous. The formation of glycosidic bonds between monosaccharides requires free energy and enzymatic catalysation. Transformation of monosaccharide substrates to nucleotide sugars stores required energy on the conjugates (Table 1.2). Reactions for the formation of glycosidic bonds are catalysed by a group of enzymes named glycosyltransferases between targeted saccharide sequence and nucleotide sugars utilising the stored energy. As the example of Lactose formation (Figure 1.2), the Gal substrate is converted into UDP-Gal as a nucleotide sugar donor, and an enzyme, β-4 galactosyltransferase, specifically catalyses the assembly of the UDP-Gal donor onto the C-4 carbon of Glc acceptor. As specified in the name of the enzyme, glycosyltransferases generally have specificity for their donors and catalysing linkages for the products. There are exceptions for the linkage specificity, such as Fuc-T III, which can catalyse the addition of Fuc on to acceptors with either α 1,3 or 1,4 linkage.

Reversibly, monosaccharides can also be cleaved off from glycans, specifically or non-specifically, by a group of enzymes named glycosidases. For example, there is a pair of sialidases, sialidase A and S, which are able to remove sialic acids from glycans sequences. Sialidase S can specifically hydrolyse sialic acids with α2,3 linkage from glycan sequences. But sialidase A can unselectively release sialic acids with all known linkages as α2-3, 6, 8 and 9 from glycan sequences.

Due to the specificity and non-specificity of glycosyltransferases and glycosidases, the proliferation of glycan sequences does not strictly follow a template-based approach, and this could allow the structural diversity of glycan sequences to reach a higher level than the other template-based bioproducts as proteins, DNAs and RNAs.

Table 1.2 Nucleotide Sugars for Common Mammalian Monosaccharide

Monosaccharide	Nucleotide Sugar
Glc	UDP-Glc
Man	GDP-Man
Gal	UDP-Gal
Fuc	GDP-Fuc
GlcNAc	UDP-GlcNAc
GalNAc	UDP-GalNAc
NeuAc	CMP-NeuAc
NeuGc	CMP-NeuGc
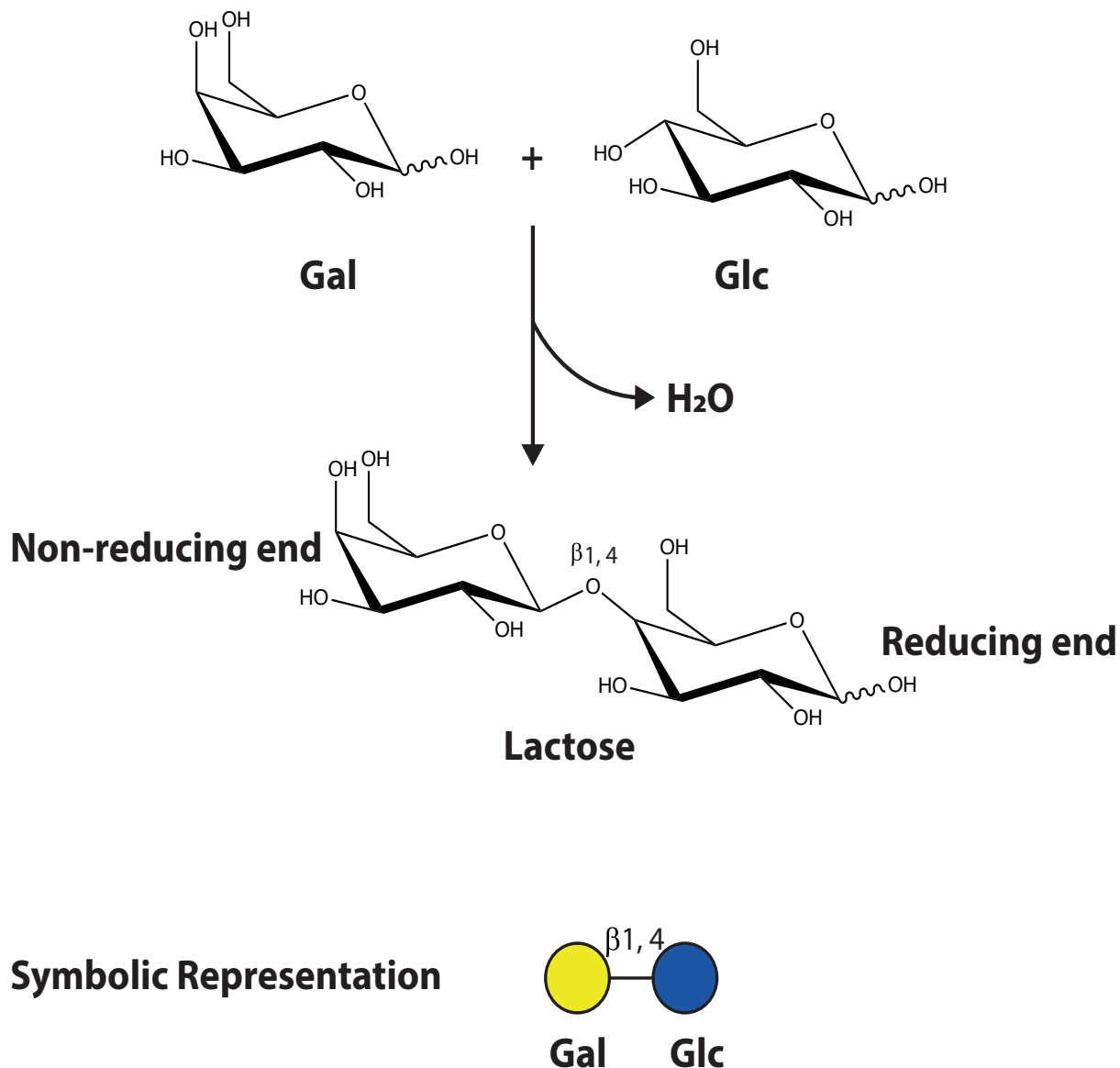
Figure 1.2 Formation of Lactose.
The formation of lactose requires a reaction between the hemi-acetal group of anomeric carbon of Galactose and C-4 hydroxyl group of Glucose. The monosaccharide substrates and di-saccharide lactose product are presented in stereochemical structures; the linkage of the formed glycosidic bond is indicated. Symbolic representation for the di-saccharide is shown, but normally the linkage information of glycosidic bond is not presented in the simplified presentation scheme.

1.3 Protein glycosylation

Protein glycosylation is a biosynthetic process in which glycan sequences are conjugated with protein sequences via glycosidic bonding between the reducing end monosaccharides and desirable amino acid side chains. Even though approximately 1% of genes encoded contribute to glycosylation related components, over 50% of proteins are estimated to be naturally glycosylated (Apweiler et al., 1999, Coutinho et al., 2003). There are two major types of glycosylation on proteins which are group-named by the elements of amino acid side chain that the glycan sequences attached onto, N- and O-(linked) glycosylation.

1.3.1 N-glycosylation

N-glycosylation is a process assembling a glycan sequence onto the amide nitrogen of selected Asn residues on a protein sequence via a β-N-glycosidic bond. N-glycosylation has a specificity for Asn residues with certain amino acid sequons. The required sequon must be in a tri-amino acid form as Asn-X-Ser/Thr in which the X amino acid residue cannot be proline. N-glycosylation is initiated at ER as the synthesis of an oligosaccharide precursor on a lipid dolichol. The entire precursor is then transferred onto a desired sequence and modified in the ER-Golgi complex to become a mature structure. Glycoproteins are finally expressed on the outer surface of the plasma membrane or secreted to the extracellular environment (Figure 1.3).

1.3.1.1 Precursor synthesis

The ER assembled N-glycan precursor consists of three types of monosaccharides, GlcNAc, Man and Glc forming a fourteen-residue sequence ($Glc_3Man_9GlcNAc_2$). The biosynthesis of this precursor (Figure 1.4) is carried out in two stages on the ER membrane from its cytoplasmic side to lumenal (Burda and Aebi, 1999).

The precursor commences to assemble by addition of a GlcNAc residue onto a dolichol via a pyrophosphate on the cytoplasmic side of the ER membrane. Another GlcNAc and five further mannose
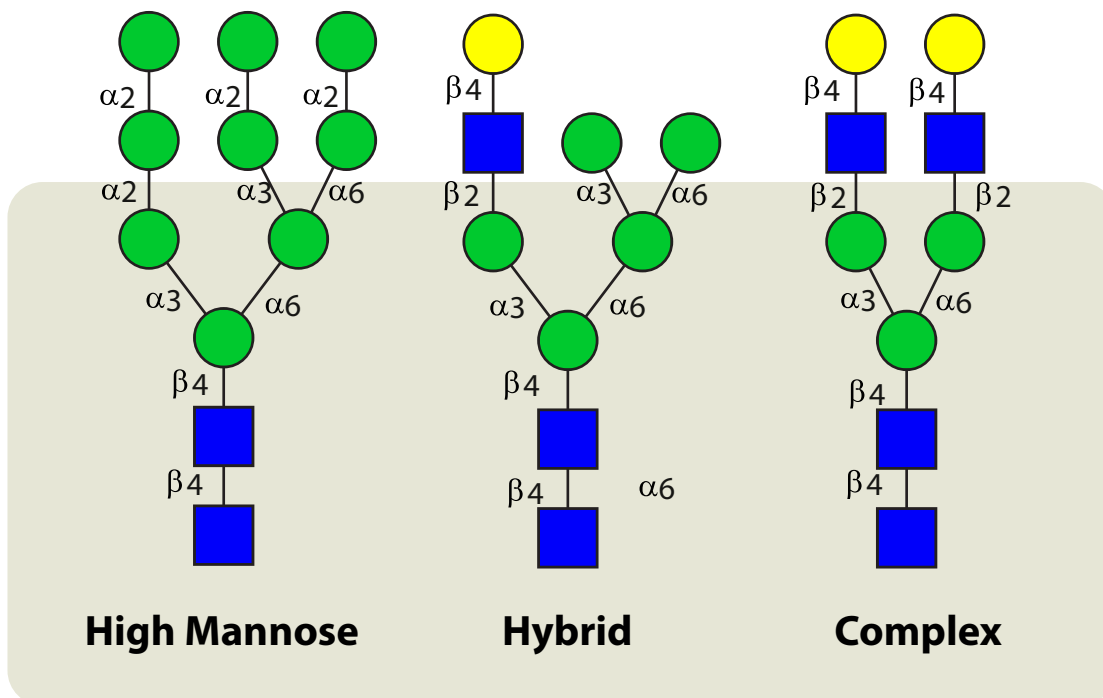
Figure 1.3 Common types of N-glycans.
There are three types of N-glycans which are widely observed on glycoproteins synthesised in mammalian systems named high mannose, hybrid and complex type. They all share a common core that is coloured by a square in grey.

residues are sequentially added onto the glycan sequence of the saccharide-dolichol complex to be a seven-monosaccharide sequence. The seven-monosaccharide sequence is then flipped inside the ER lumen.

As the sequence moves inside the ER lumen, a further four mannose residues are successively added from the supply of GDP-Man-Dolichol flipping from the cytoplasmic environment. The final three Glc residues from cytoplasmic UDP-Glc-Dolichol are then assembled onto the glycan sequence to complete the biosynthesis process.

Once the synthesis is completed in the ER lumen, there is an en bloc transfer of the fourteen-monosaccharide precursor onto the Asn site in a conserved tri-amino acid sequon of a translating and folding protein sequence. The en bloc transfer is carried out by an enzyme named oligosaccharidetransferase
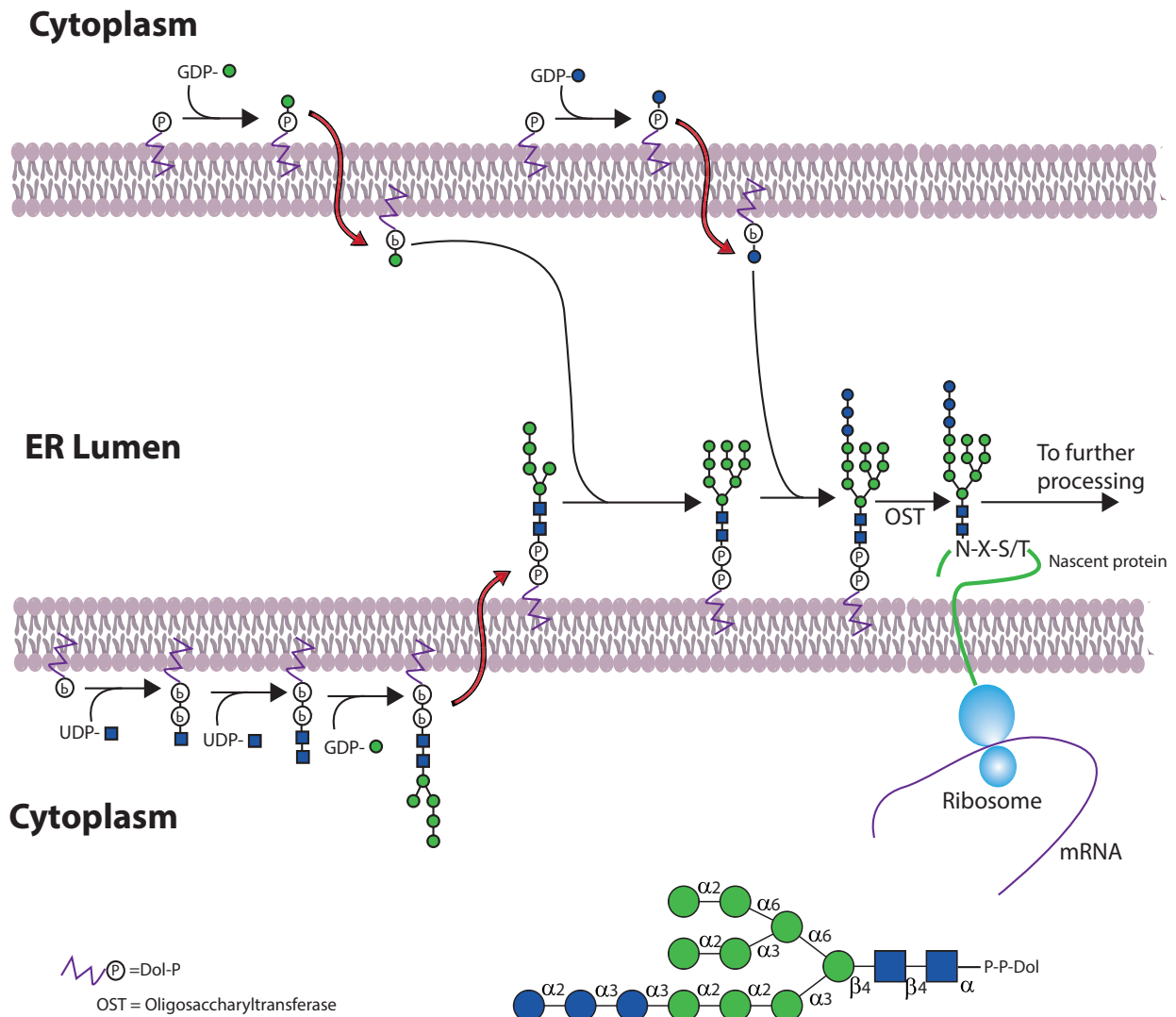
Figure 1.4 Biosynthesis of N-glycan precursors and en bloc transfer.
The initial $Man_5GlcNAc_2$ sequence is sequentially assembled onto a dolichol on the cytosolic side of ER and flips into the lumen. Man and Glc substrates are anchored onto Dolichol-P at the cytosolic side of ER and flips into the lumen to supply the substrates for precursor elongation to a 14-monosaccharide sequence as $Glc_3Man_9GlcNAc_2$. The 14-monosaccharide precursor is then transferred onto the N-X-S/T sequons of nascent proteins.

(OST) in the ER. This enzyme contains eight subunits as a hetero-oligomeric complex (Kelleher and Gilmore, 2006) and the catalytic domain of this enzyme is in one of these subunits named STT3 (Zufferey et al., 1995, Kelleher et al., 2003). There are two isoforms of this subunit that are found in mammalian systems, STT3A and STT3B. These two isoforms are differentiated by their catalysing specificities and activities (Kelleher et al., 2003). The STT3A isoform is the primary enzyme that transfers the N-glycan precursor onto nascent protein sequence with stringent selection of oligosaccharide precursors. STT3B mainly complements to glycosylate the remaining N-glycosylation sites after STT3A glycan-transfer (Ruiz-Canada et al., 2009). The other subunits may also be involved in this process as facilitating the presentation of protein substrates to the catalytic domain of this enzyme (Wilson et al., 2008).

1.3.1.2 Initial processing in the ER

The trimming of the glycan precursor on glycoprotein sequences is not only salient for the maturation of glycans, but also functionally contributes to the folding quality control of glycoproteins in the ER (Figure 1.5).

Terminal α1,2 Glc is immediately removed from the N-glycan precursor on a glycoprotein by ER glucosidase I after the en bloc transfer to prevent re-recruitment of the glycan sequence by OST. Glycoproteins with the remaining $Glc_2Man_9GlcNAc_2$ glycan sequence commence to enter protein folding process. Once misfolding occurred at this stage, the thirteen-monosaccharide glycan on the mis-processed protein sequence can be preferably bound by malectin (Schallus et al., 2008) to prevent its further processing and induce ER-associated degradation (ERAD) (Chen et al., 2011, Galli et al., 2011)

Once glycoproteins pass this control, ER Glucosidase II subsequently cleaves at least one Glc from their glycan sequence to continue the folding process. The folding process can be further assisted by a calnexin/calreticulin associated complex with accessory ERp57, a thioldisulfide oxidoreductase, and cyclophilin B, a peptidyl prolyl. With a Glc remains on the glycan sequence of the glycoprotein, the calnexin/calreticulin can associate with the glycoprotein via its affinity with this glycan sequence to rectify protein folding (Frickel et al., 2002, Kozlov et al., 2010). The remaining Glc is removed by ER Glucosidase II after this assistance and the glycoprotein enters the final folding check. The glycoprotein subsequently binds to UDP-Glc: glycoprotein glucosyltransferase 1 (UGGT1) (Labriola et al., 1995). During the association, the folding quality can be screened and a Glc is added onto the glycan sequence of misfolded glycoprotein to re-enter the calnexin/calreticulin folding mechanism. The misfolded glycoproteins can be subject to repeat the cyclic quality check until protein folding is corrected (Sousa and Parodi, 1995).

Before the departure from the ER, the terminal α1, 2Man on the mid-arm of $Man_9GlcNAc_2$ glycan

sequence is removed from most glycoproteins by α Mannosidase I (Stanley et al., 2015). The correctly folded glycoproteins subsequently pass for further modification in the Golgi complex. But still misfolded glycoproteins are recognised by ER-degradation-enhancing α-mannosidase-like proteins (EDEM) and guided to ERAD degradation (Olivari and Molinari, 2007).

1.3.1.3 Further processing in Golgi complex

Figure 1.6 shows the processing of N-glycan sequence in the Golgi once passed the ER compartments. Trimming for the α1,2 Man residues on the glycan sequence is continued in the cis-Golgi apparatus. Three Golgi mannosidases as IA, IB and IC are responsible for the trimming, and it is not inevitable to
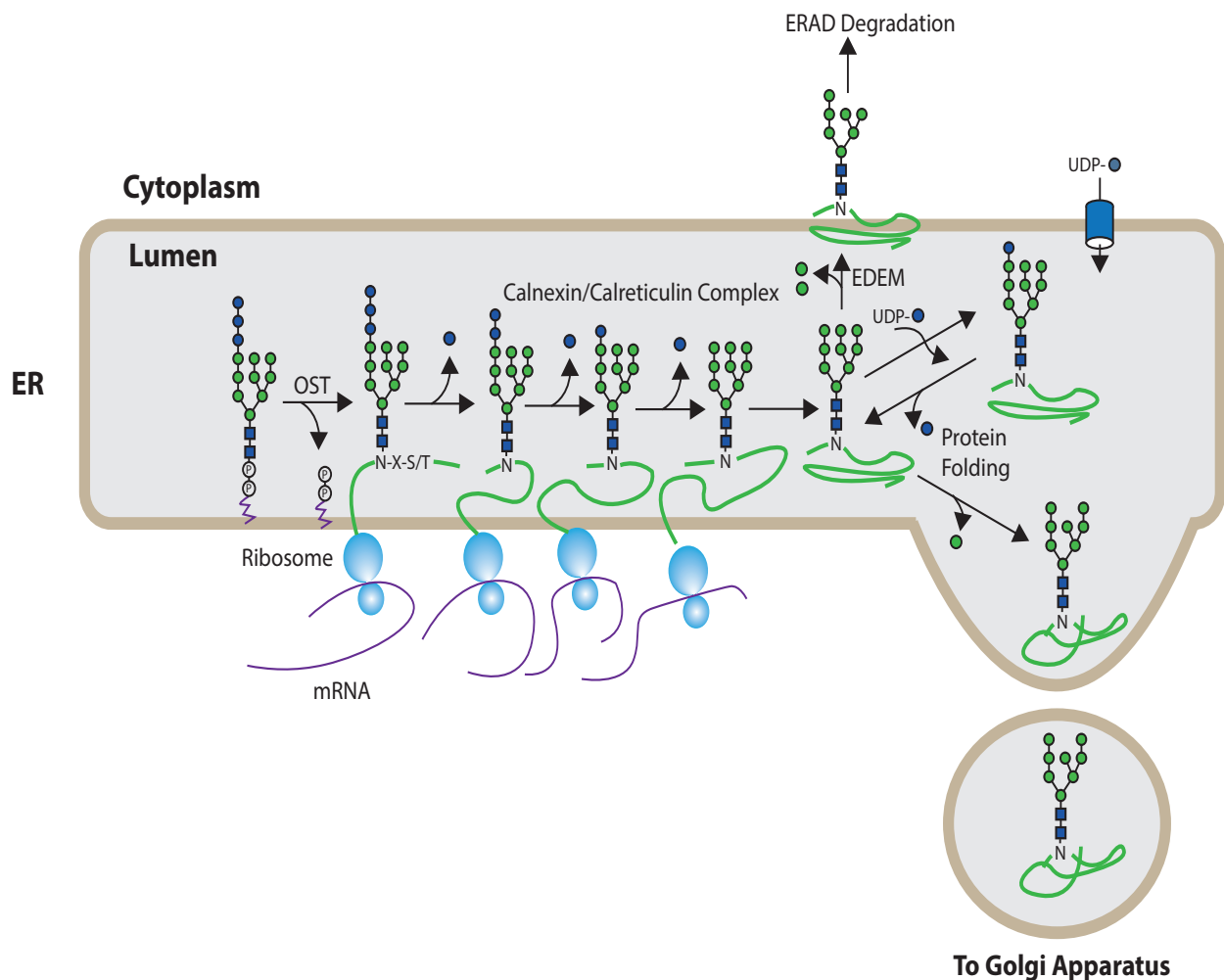
Figure 1.5 Glycan trimming and glycoprotein folding in the ER lumen.
The 14-monosaccharide precursor is en bloc transferred onto a nascent protein sequence, and ER Glucosidase I immediately cleaves terminal Glc to lead glycoproteins into the folding process. ER Glucosidase II cleaves the second Glc from the glycan sequence during a process of initial quality check by malectin. Calnexin/Calretinculin Complex then engages in the process to assist protein folding, and the last Glc is cleaved after this assistance. UGGT1 then associates with the proteins to check their folding, and a Glc residue would be added onto the glycan sequence if the protein was still mis-folded to re-enter the last process for correcting their folding. However, proteins unable to be correctly folded after this repeatable process are then led to ERAD degradation. The correctly folded proteins are then transported to the Golgi apparatus for further modifications on their glycan sequences. Adapted from (Stanley et al., 2015).
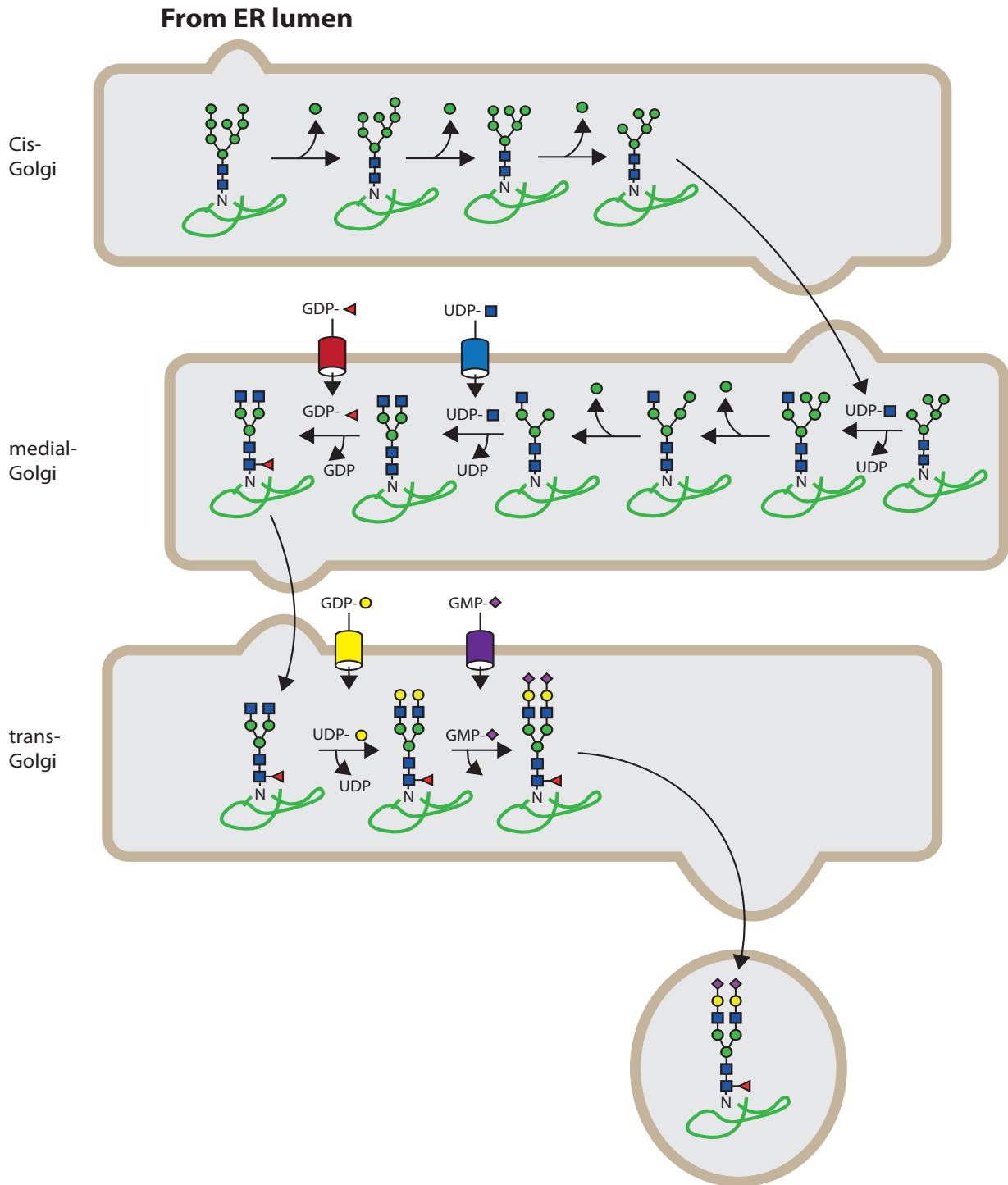
From ER lumen

Figure 1.6 N-glycan processing in the Golgi apparatus.
N-glycans are firstly transported from ER lumen into the Cis-Golgi, and Mannosidase IA, IB and IC can cleave mannose residues to form a glycan sequence with 5-9 Man. The glycan sequences without further processing are named high mannose type. The glycan sequence with 5 Man can be further processed in the medial-Golgi to initiate the first GlcNAc antenna by GlcNAc-T I. The glycan sequence with remaining Man on the 6-arm of trimannosyl core is named hybrid type. The 6-arm Man can be further trimmed by Mannosidase II, and subsequently more GlcNAc antennae can be initiated on the trimannosyl core under the catalysation of GlcNAc T family transferases (I-VI). In the trans-Golgi, the formed antennae can be further extended by various antennary modifications such as galactosylation and sialylation. The well-decorated glycoproteins are then secreted from the Golgi network to be presented to the extracellular environment. Adapted from (Stanley et al., 2015).

result in complete cleavage. Thereby, 5-9 Man residues can remain on the glycan sequence after cleavage, and these structures are also named high-mannose type once their glycoproteins pass the Golgi complex without further processing. In the medial Golgi apparatus, GlcNAc-transferase I (GlcNAc-T I) can assemble a GlcNAc onto the trimmed 3-arm of trimannosyl core on $Man_5GlcNAc_2$ structures. It initiates the assembly of hybrid and complex type structures.

Hybrid structures possess both GlcNAc and Man residues on their trimannosyl core. The GlcNAc as an antenna can be extended by other monosaccharides, and there should retain at least one Man residue on the 6-arm of the trimannosyl core on a hybrid structure. In addition, there could be an early modification of the conserved core structure, which adds an α1,6 fucose onto the reducing end GlcNAc by α1,6 fucosyltransferase (Fuc-T VIII) which is active after the addition of the first GlcNAc antenna by GlcNAc-T I (Stanley et al., 2015). Following initiation of the first GlcNAc antenna on the 3-arm, two Man residues on the 6-arm can also be cleaved by Mannosidase II to form complex type structures. Other antennae can be initiated on complex glycans by a series of GlcNAc-Ts which have their exclusive specificity to assemble GlcNAc residues onto the trimannosyl core with specific linkages (Figure 1.7).

Mammalian systems can assemble no more than four extendable antennae onto their complex structures, whereas several vertebrates can assemble more antennae onto their complex structures (Yamashita et al., 1982, Taguchi et al., 1995). Moreover, a β1,4 GlcNAc can be added onto the central
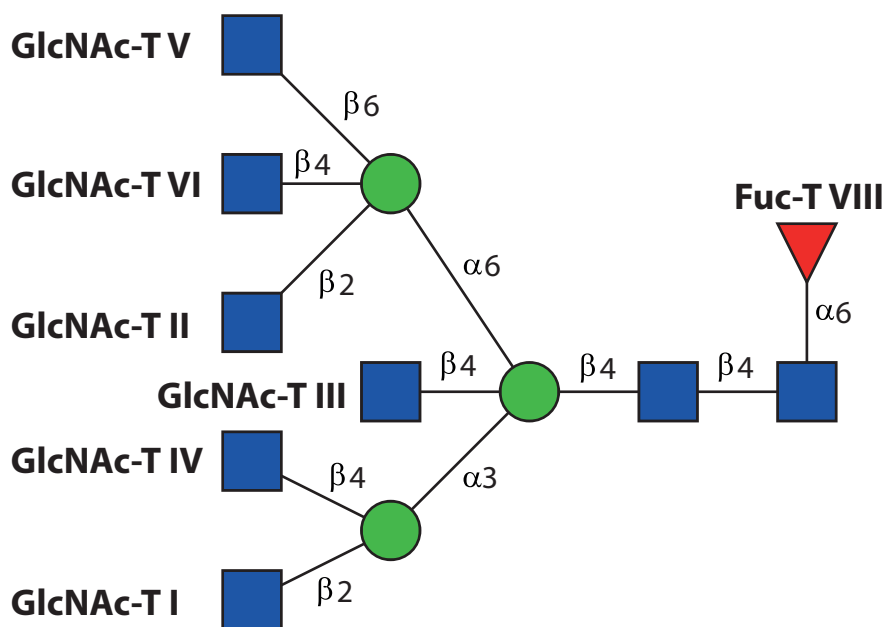
Figure 1.7 Branching and core-fucosylation of N-glycans.
The GlcNAc antenna is initiated by GlcNAc-T I on the 3-arm of trimannosyl core. Fuc-T VIII becomes active for adding a fucose onto reducing end GlcNAc named core-fucosylation after the GlcNAc addition by GlcNAc-T I. GlcNAc-T II, IV, V and VI and sequentially add four GlcNAc antennae onto the two arms of trimannosyl core. GlcNAc-T III may also be active to sequentially assemble an unextendible GlcNAc antenna at the centre of tri-mannosyl core for Hybrid and Complex type structures.
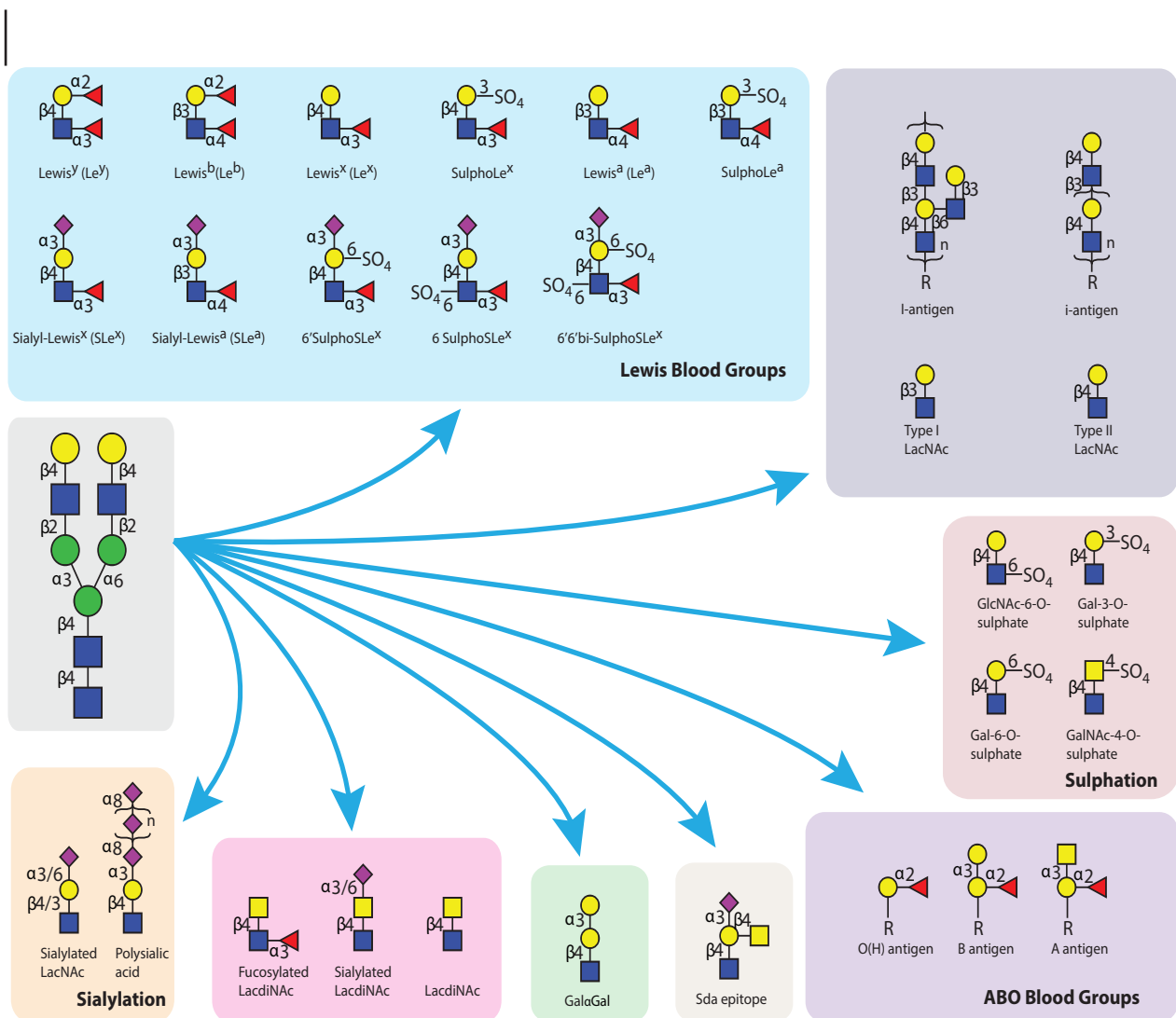
Figure 1.8 Major antennary modifications on N-glycans expressed in mammalian system.
The presented modifications mostly occur on the non-reducing ends of N-glycan antennae as terminal moieties. Sialylation generally terminates the potential for further elongation of antennae with the exception of polysialic acid, which can extend the antenna by a repeated assembly of sialic acids. ABO blood groups and GalαGal are also un-extendible moieties. LacNAc can be either a terminal or internal moiety on a N-glycan depending on status of their self-repeated elongation to be branched "I" or linear "i"-type poly-LacNAc sequence on an antenna. Sulphation can be found on Lewis group antigens, LacNAc and LacdiNAc moieties.

Man of the trimannosyl core of hybrid and complex structures by GlcNAc-T III. This GlcNAc is named bisecting GlcNAc and it is unable to be extended by another monosaccharide due to steric constraints. Hybrid and complex structures can be further extended and modified by various monosaccharides in the trans Golgi. The structural diversity and functional epitopes of N-glycans are largely contributed by the antennary modifications occurred in the trans Golgi (Figure 1.8).

1.3.2 O-glycosylation

In mammalian systems, O-glycosylation generally refers to the most common O-GalNAcosylation. Other O-linked glycosylations have also been identified such as O-Fucosylation, Mannosylation, Glucosylation and GlcNAcosylation in living systems (Wopereis et al., 2006, Gebauer et al., 2008).
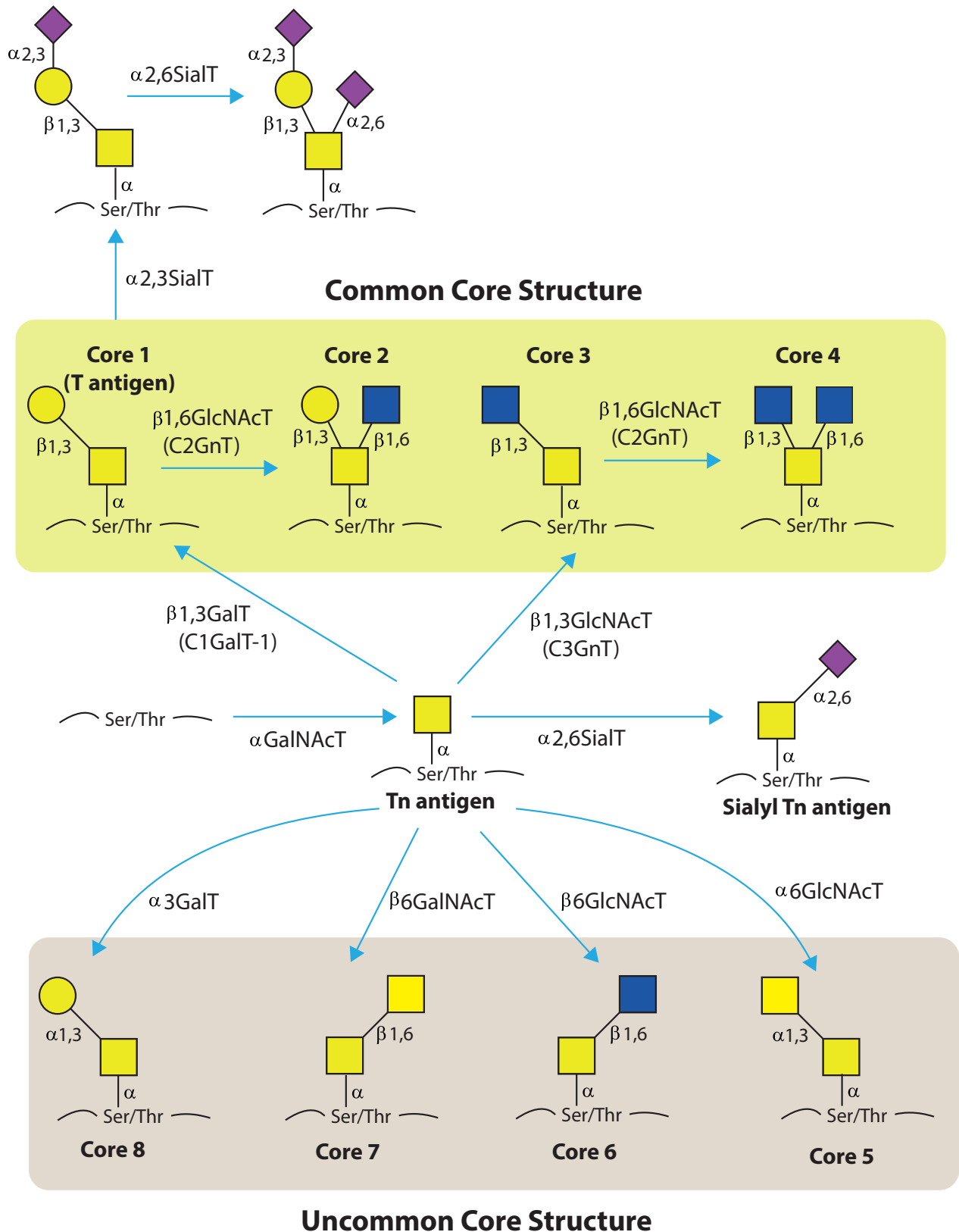
Figure 1.9 Core structures of O-glycans.
There are eight core structures that differ by the linkage and composition of their branching monosaccharides after the GalNAc residue. GalNAc and its sialylated form can be directly expressed onto glycoproteins in certain physiological conditions in a mammalian system. Sialylation on core 1 structure is commonly observed with a maximal of two sialic acid residues assembled on both monosaccharides of this di-saccharide sequence.

O-glycosylation assembles a glycan sequence that is initiated by adding a name referred monosaccharide residue onto the oxygen of the hydroxyl group of Ser or Thr residues via an α-O-glycosidic linkage. The initial monosaccharide for GalNAc O-glycosylation is assembled in the Golgi complex on targeted residues without requirement for a consensus sequence. Extending monosaccharides are sequentially assembled onto the glycan sequence without a pre-synthesised precursor.

For the O-GalNAcylation, the first GalNAc is assembled onto folded proteins via the catalysis by a group of glycosyltransferases named polypeptide GalNAc transferases (ppGalNAcT) to initiate the biosynthesis (Ten Hagen et al., 2003). Humans possess 20 ppGalNAcTs (Bennett et al., 2012) and they target Ser/Thr substrates with differentiated O-glycosylation status. For instance, pre-exisiting O-GalNAc on the sequence is required for some ppGalNAcT to be active to catalyse O-glycosylation on remaining unoccupied sites (Ten Hagen et al., 2003, Lira-Navarrete et al., 2015). This reflects that O-glycosylation should be a highly regulated procedure with a hierarchical processing sequence.

The initial GalNAc can be extended by Gal and/or HexNAc residues on its C-3 and C-6 hydroxyl groups to form eight di- or tri-saccharide core structures, named core 1-8 (Figure 1.9). Core 1 and 2 are commonly found in mammalian systems but core 3-8 have relatively restricted expressions to be tissue-specific and/or protein-specific (Brockhausen, 1999, Chai et al., 1992, Wopereis et al., 2006). Extension of these core structures are identical to that of N-glycans as they pass through the remaining Golgi compartments concurrently.

1.4 Glycosphingolipid

The major glycolipids found in mammalian systems are glycosphingolipids (GSL). The GSL is a conjugate of a glycan sequence with a lipid moiety, named ceramide (Figure 1.10), via a β-glycosidic bond on the C-1 hydroxyl group of the ceramide. The ceramide constitutes an amino alcohol base, sphin-
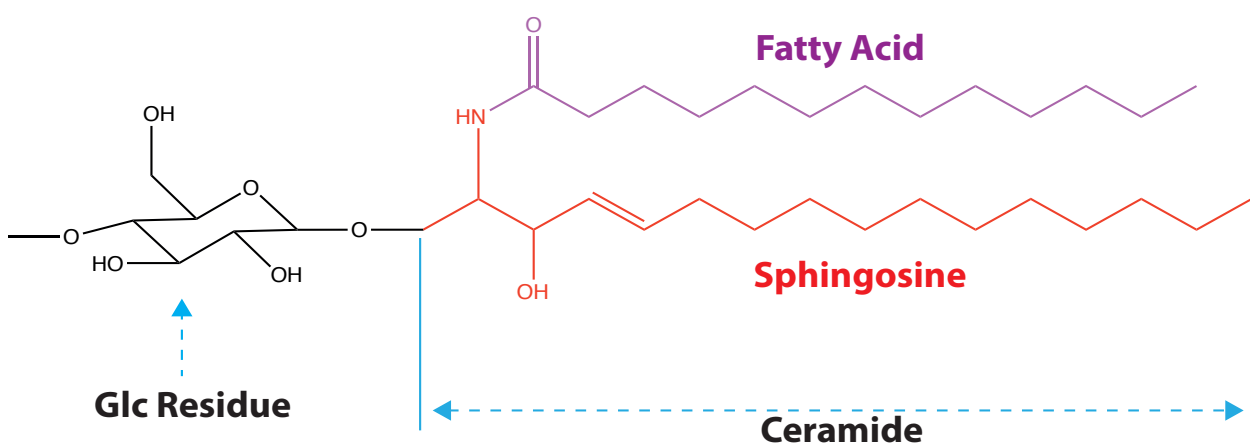
Figure 1.10 The GSL conjugate of a glucose and ceramide.
The sphingosine is coloured in red and the fatty acid is highlighted in purple. The conjugating glycosidic bond is in a β configuration.

gosine, joined with a fatty acid via amide linkage at C-2. The ceramide structures can vary by the length of the fatty acid (C14-C30), saturation and α-hydroxylation status (Schnaar and Kinoshita, 2015, Karlsson, 1970, Degroote et al., 2004). The glycosphingolipids are further sub-grouped by the first monosaccharide attached to the ceramide base such as galacto-(GalCer) and gluco-sphingolipids (GlcCer). Although the glycan sequence and ceramide base have their structural diversity respectively, the nomenclature of glycolipid structures is generally based on the glycans.

The ceramide is synthesised on the cytosolic surface of the ER to initiate the biosynthesis of GSLs (Degroote et al., 2004, Hirschberg et al., 1993). GalCer is synthesised in the ER by the addition of an initial Gal onto the ceramide that is catalysed by ceramide galactosyltransferase (CerGalT) after it flips into the lumen of the ER. The GalCer can be further modified in the Golgi as GM4 or sulfatide by sialylation or sulphation, respectively.

The initial Glc residue of GlcCer is assembled onto ceramide by ceramide glucosyltransferase (CerGlcT) at the cytosolic side of the Golgi after the transportation of ceramide from the ER. GlcCer is further matured in the lumen of the Golgi. In mammalian system, there are four common neutral core structures of GlcCer, Ganglio-, Globo-, Lacto- and neoLacto- series (Figure 1.11, Chester, 1998). The distribution and extension of these core structures exhibited a tissue-specific and development-related manner (Iwamori et al., 1984).

Sialylated GlcCer is collectively termed as ganglioside regardless of the class of core structures (Figure 1.12). The term, ganglioside, refers to the original discovery of sialylated GlcCer from ganglion cells in the brain (Kolter, 2012). In addition, the widely used Svennerholm nomenclature (Svennerholm, 1963) simplifies the representation of sialylated GlcCer structures to primarily two capital letters plus a number. The first letter "G" refers to ganglioside, the second letter indicates the level of sialylation as "a-, mono-, di- and tri-", and the number reflects the migration order of gangliosides in the original examination on a thin-layer chromatography.
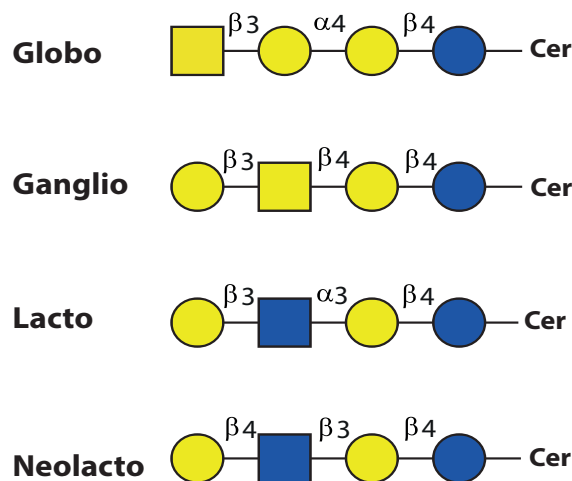
Figure 1.11 Common mammalian core structures of GlcCer.
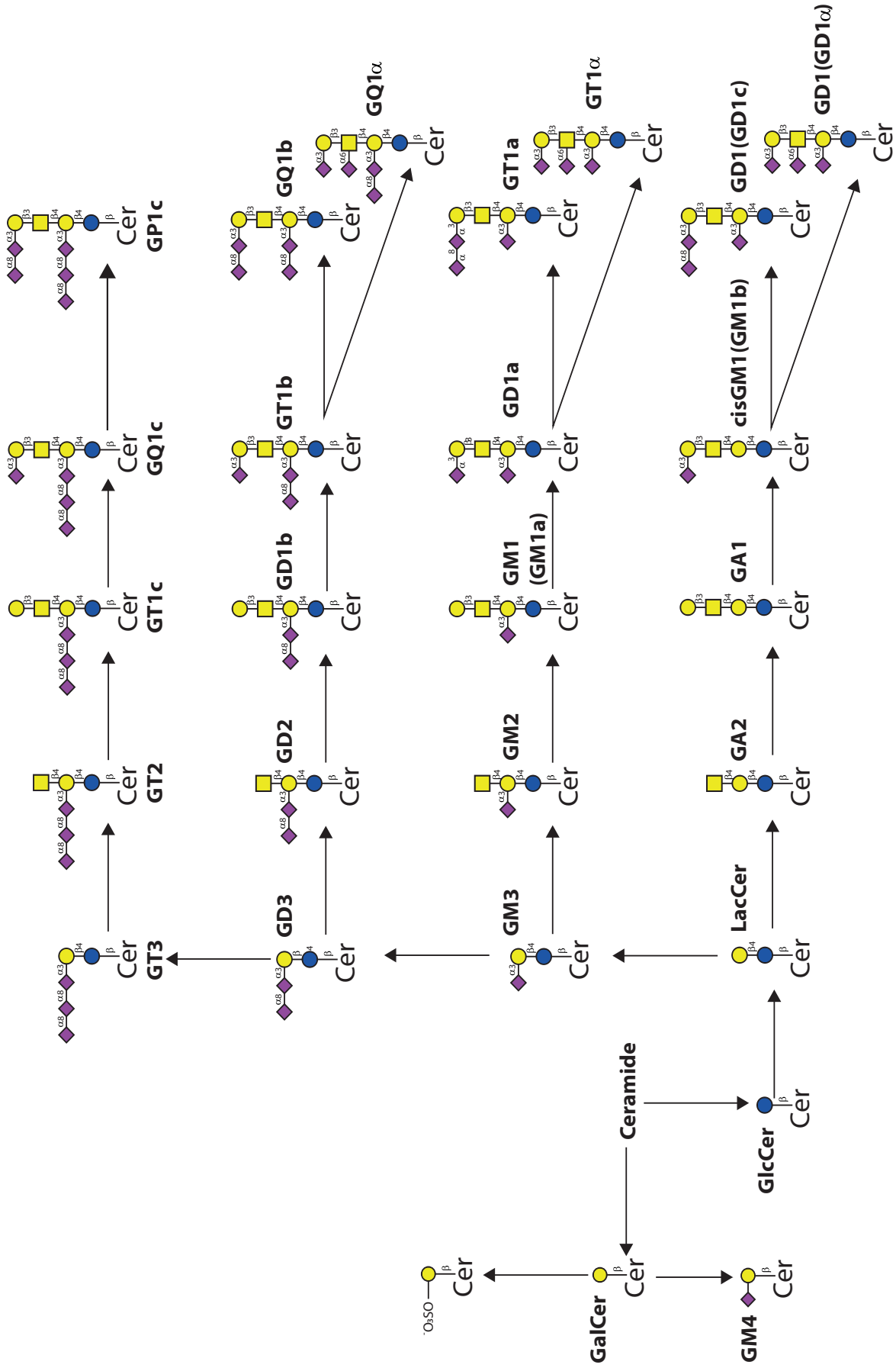
Figure 1.12 Biosynthetic pathways of GalCer and Brain Gangliosides.
Adapted from (Schnaar and Kinoshita, 2015)

1.5 Mass spectrometry

Mass spectrometry is a technology which records the spectra of mass to charge ratio (m/z) of ionised molecules.

There are three fundamental modules of a mass spectrometer: 1) an ion source, 2) a mass analyser and 3) a detector. The ion source is a module to facilitate the ionisation of molecules to form charged ions in gas-phase. The charged ions traverse a mass analyser which separate charged ions based on their mass to charge ratios. The ions separated by their m/z are recorded by the detector to form a spectrum presenting their signals with relative abundance.

1.5.1 Ion source

The initial step of a mass spectrometric experiment is to ionise the sample molecules to generate cations or anions. The ion source is the module of a mass spectrometer where the ionisation occurs. Currently, there are various techniques being applied for molecule ionisation in mass spectrometry. In this Section, EI (Electron Impact Ionisation), ESI (Electrospray Ionisation) and MALDI (Matrix Assisted Laser Desorption Ionisation) are focused on as these are the technologies which were applied in the studies documented in this thesis.

1.5.1.1 Electron Impact Ionisation

In 1918, A.J Dempster introduced the first EI ion source (Dempster, 1918). It is also one of the oldest ion sources which is still applicable in research for small molecules as it was originally designed.

In the EI source (Figure 1.13), electrons are generated by a heated filament and subsequently accelerated to 70 eV. The accelerated electrons are guided to collide with gas-phase analytes in a chamber. The analytes are then ionised. An electron is ejected to generate singly positively charged radical ions. Due to the high energy that the accelerated electrons transfer to the analytes, fragment ions of analytes are normally generated as well as, or instead of, intact molecular ions. Therefore, the actual detected ions in the mass spectrometer from this ion source are generally fragment ions of the analytes, and this technique is also referred as a "hard" ionisation. Even though it is normally unable to produce intact molecular ions, its fragmentation power is still recruited in analysing unknown compounds and defining structures of small molecules. As EI can only ionise thermally stable and volatile small molecules, this ion source is commonly coupled with gas-chromatography and a quadrupole mass analyser. The incorporated system can be applied for derivatised small molecules such as monosaccharides to determine their identities and glycosidic linkages by the information from the chromatographic retention and fragmentation by mass spectrometry.
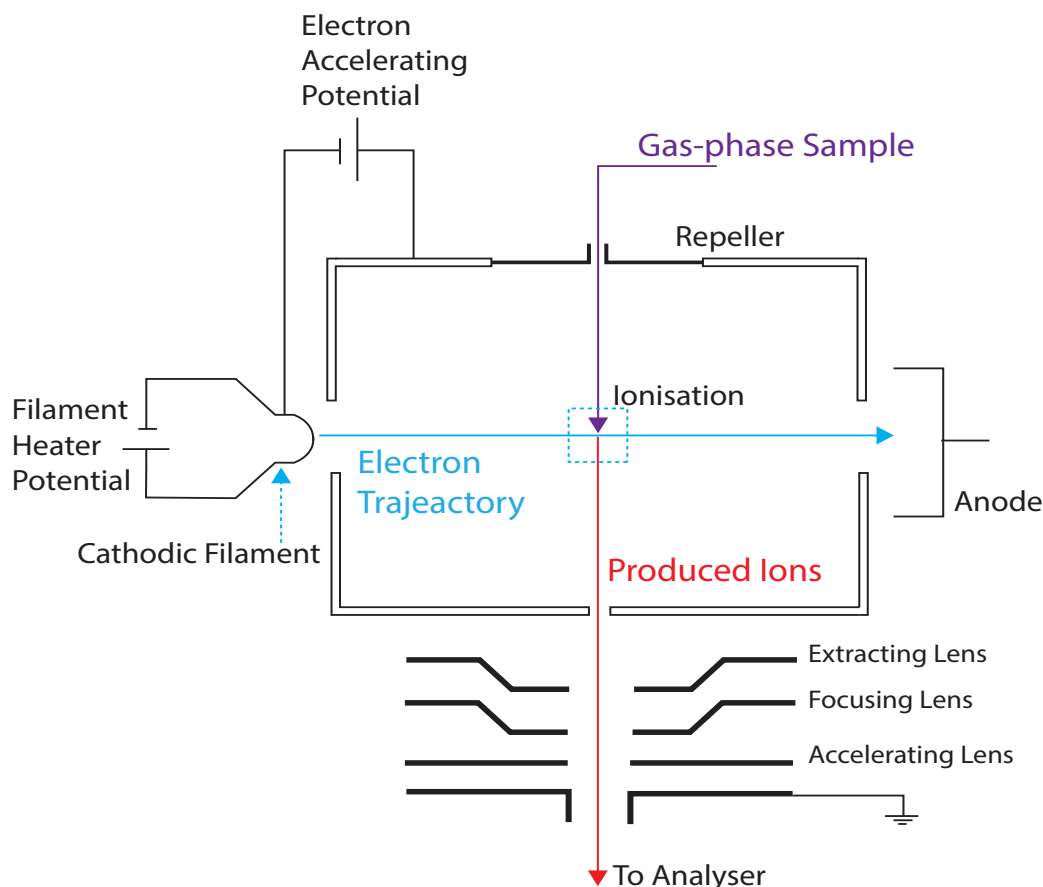
Figure 1.13 Schematic diagram of Electron ionisation.
Adapted from (Hoffmann and Stroobant, 2007)

1.5.1.2 Electrospray ionisation

In the 1980s, ESI was optimised by John Fenn (Yamashita and Fenn, 1984) to conduct analysis for large bio-molecules. It is a soft ionisation technique that produces mostly intact molecular ions which can be multiply charged. Due to its advantages in producing multiply charged ions, molecule ions can be detected with a lowered m/z and therefore the mass range of the mass spectrometer is increased. ESI is a widely applied ion source in analytical science and industry, especially in the studies of proteomics and glycoproteomics.

The formation of charged molecular ions is completed in the process of transforming analytes from liquid to gas phase and it commences at a charged spray-tip (Figure 1.14). The solution of analytes is loaded into a capillary tip which is charged with a high voltage. Due to the high voltage and thin capillary tip, the liquid at the capillary tip is polarised to form a cone-shape, named Taylor cone. Under the force of the field, the charged droplets continuously emerge like a jet from the Taylor cone when the electrostatic repulsion at the surface of the cone exceeds liquid tension. The droplets are guided by counter-electrodes toward the mass analyser (Figure 1.14b). The size of droplets is decreased by the evaporation of dissolving solvent under a drying gas during the journey. The shrinkage of droplet size leads to increased charge density and electrostatic repulsion. Once the repulsion exceeds the droplet
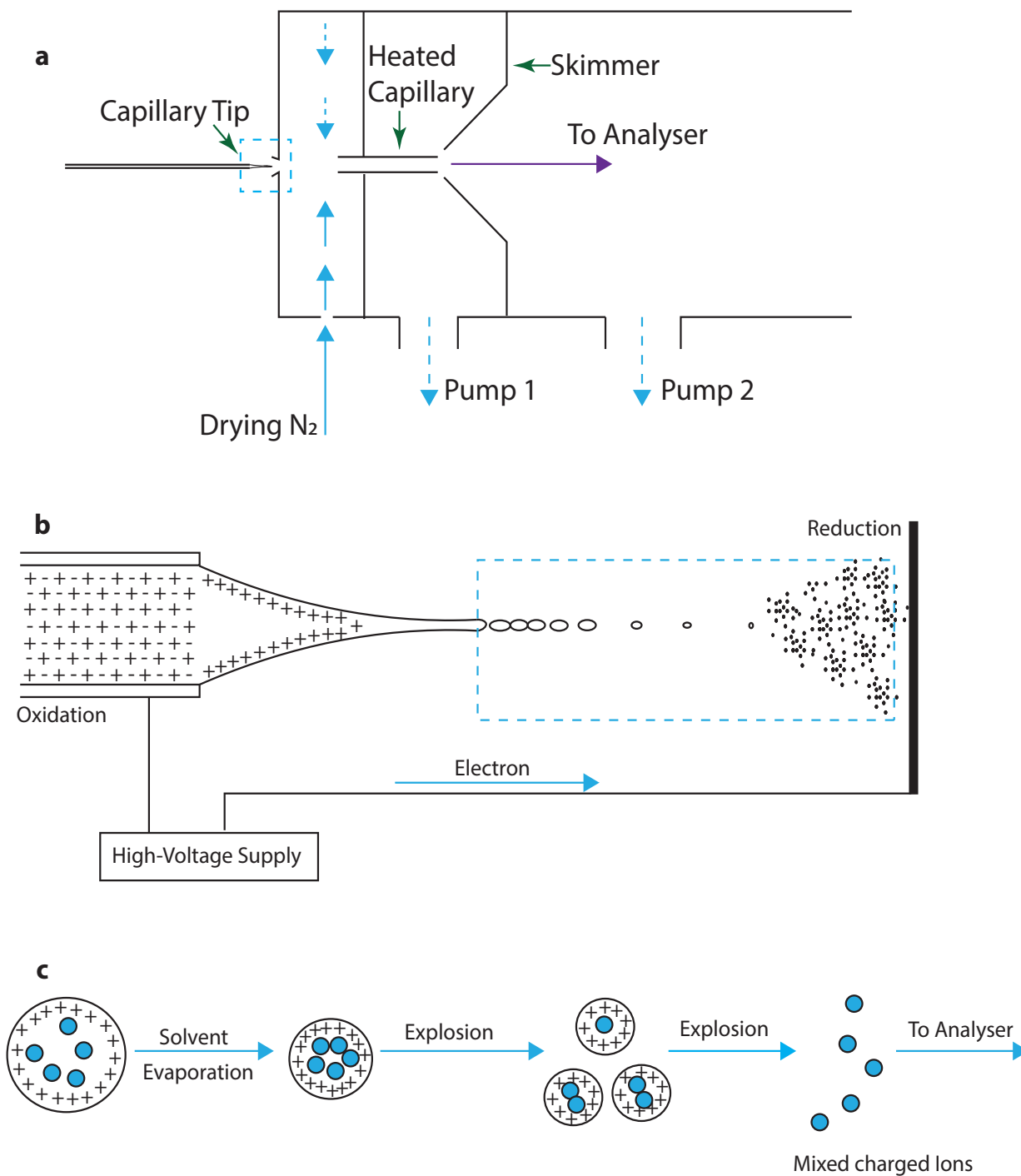
Figure 1.14 Schematic diagram and mechanism of electron spray ionisation (ESI).
Schematic diagram of a typical ESI ion source (a) and a zoomed-in area is highlighted with dashed blue square. The zoomed-in area of dash blue square (b) illustrates the spraying procedure of the ion source and a further area highlighted again with dashed blue square for ionisation process. The ionisation process (c) associates with droplet shrinkage and explosion. Adapted from (Cole, 2010).

surface tension it leads to an explosion of a droplet forming smaller droplets. This drying process constantly occurs until solvent-free molecular ions are finally formed (Figure 1.14c).

During the ionisation process, multiply charged molecular ions are produced in the ESI ion source by protonation/deprotonation e.g. $[M+nH]^{n+}$ or $[M-nH]^{n-}$ in either positive or negative mode of the mass

spectrometer. The multi-charge state of a molecular ion can facilitate the detection of large molecules, but it can also complicate the process of data interpretation.

ESI is normally connected with a liquid chromatography system to successively fractionate and ionise molecules. Mixed analytes can be sorted by the chosen liquid chromatography (LC) in that its stationary phase has differentiated association with the injected molecules based on their physicochemical property prior to reaching the mass spectrometer. The ESI ion source then constantly ionises fractionated molecules from the LC system to provide a set of simplified molecules for the mass analyser at each associated time point. ESI generally does not produce fragmented ions for structural definition, an extra MS/MS module can be incorporated into the mass spectrometer to fragment ionised molecules from ESI. Thus, this incorporated system is named online-LC-MS/MS.

1.5.1.3 Matrix-assisted laser desorption ionisation

MALDI is another mainstream soft ionisation technique which is widely used in studies of proteomics and glycomics. Koichi Tanaka (Tanaka et al., 1988) invented this technique in the 1980s while another group also developed this technique independently (Karas et al., 1985, Karas and Hillenkamp, 1988).

The general concept of MALDI ionisation is that analytes can be ionised by a laser induced irradiation with the assistance of a matrix (Figure 1.15). For the process of MALDI ionisation, analytes are firstly well-mixed with a matrix solution and spotted onto a metal plate. The spotted mixture is dried and spontaneously co-crystallised. The crystallised mixture is then hit by laser pulses. The matrix has a strong and accumulative absorption of the laser energy leading to a desorption of the co-crystallised mixture while transferring energies to embedded analytes. The analytes are then excited and sublimated into the gas phase. During this process, even a large molecule that may not be volatile and thermally stable can still be ionised as $[M+Na]^+$ and $[M+H]^+$ in positive mode without fragmentation. However, the exact mechanisms of the energy transformation and ionisation are still largely unknown. There is a widely accepted proposal that the formation of ions in the MALDI source could be related to a proton transfer and chemical processes in the excited mixture (Knochenmuss, 2006).

1.5.2 Mass analyser

After analytes are ionised a mass analyser is needed to separate ions based on their m/z ratio before the ions reach the detector to be recorded. Four different types of mass analysers are introduced in the following Section due to their engagements in the documented studies in this thesis.

1.5.2.1 Quadrupole and tri-quadrupole

In 1953, Paul and his colleague Steinwegen firstly developed the Quadrupole mass analyser (QMS)
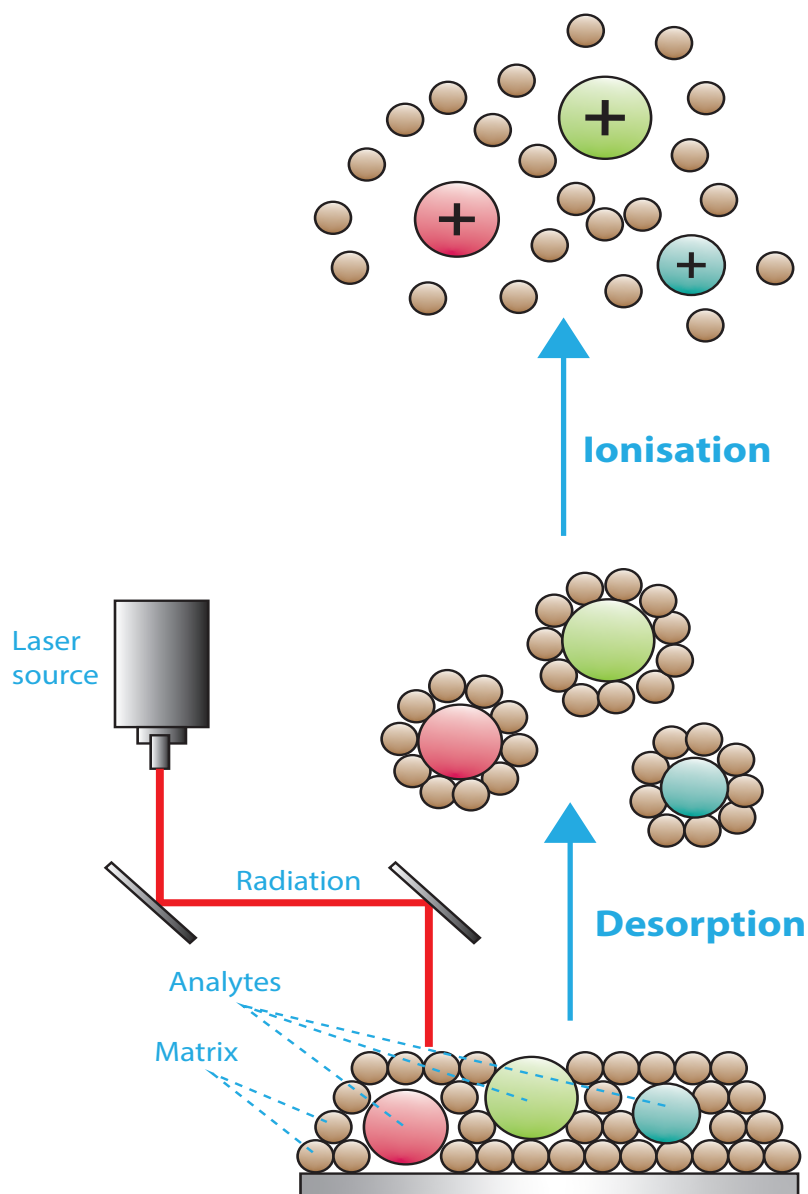
Figure 1.15 Schematic diagram of desorption and ionisation of MALDI.

utilising the electric fields of direct current (DC) and radio frequency (RF) voltages (Paul, 1960).

QMS consists of four metal rods in parallel, which are spatially organised to form a hyperbolic cross-Section (Figure 1.16). One pair of the diagonal rods is applied with a DC plus RF voltage while a mirror voltage applied on the other pair of rods. The dynamic quadrupole electric fields can only allow ions with certain m/z ratio to pass the rod fields under a certain voltage when they have a stable travel trajectory in the fields. The m/z of passed ions can be determined by the associated voltage after calibration. The other ions with unstable trajectories under a certain voltage would be filtered by the fields by colliding with the rods. Moreover, QMS can serve either as a mass filter to scan a selected ion under a certain voltage or a mass guider to allow a range of ions to pass through with oscillating voltages. In addition, QMS can be adapted to a collisional cell as it can guide ion collision inside this module with an inert gas.
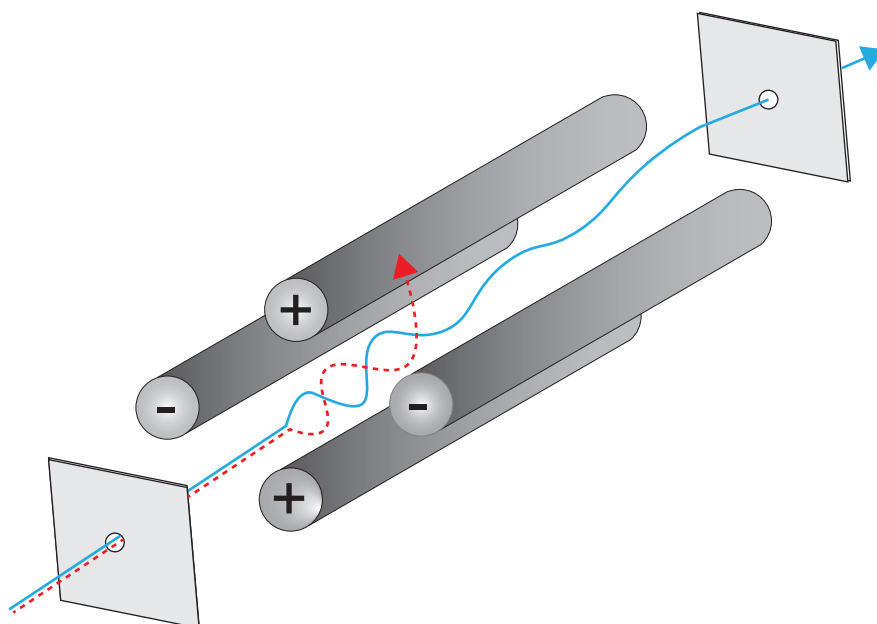
Figure 1.16 The basic geometry and components of Quadrupole mass analyser.

The parameters of QMS is not outstanding in terms of mass range and resolution, its compatibility, robustness and cost efficiency still allows it to be adopted widely in various mass spectrometry-based studies. GC-MS used in the study described in this thesis has a single QMS analyser coupled with the gaseous ion source for m/z determination.

More advanced QMS-associated geometry of mass spectrometry can utilise the capacities of QMS as a mass filter, mass guider and collisional cell in a single instrument.An example of combining different roles that a QMS can carry out is in a tandem-linked triple-quadrupole mass analyser as tandem mass spectrometry (MS/MS) for fragmentation experiments. The geometry for this instrument is obvious as three single QMS analysers tandemly connected shown in Figure 1.17. The first QMS can serve as a mass filter to select parental ions. The second QMS then induces the fragmentation of selected ions by collisions with an inert gas. The third QMS is used to determine the m/z for fragment ions from the previous QMS. Thereby, this instrument is also named as QqQ in which the lower case refers to the collisional cell. This instrument allows selection and fragmentation of a molecular ion with targeted m/z ratio. The signal-to-noise ratio of the generated spectrum can be improved, and another layer of structural information can also be elucidated. Another application of this tri-quadrupole geometry is on a Q-TOF type instrument, which is described in **Section 1.5.2.4**.
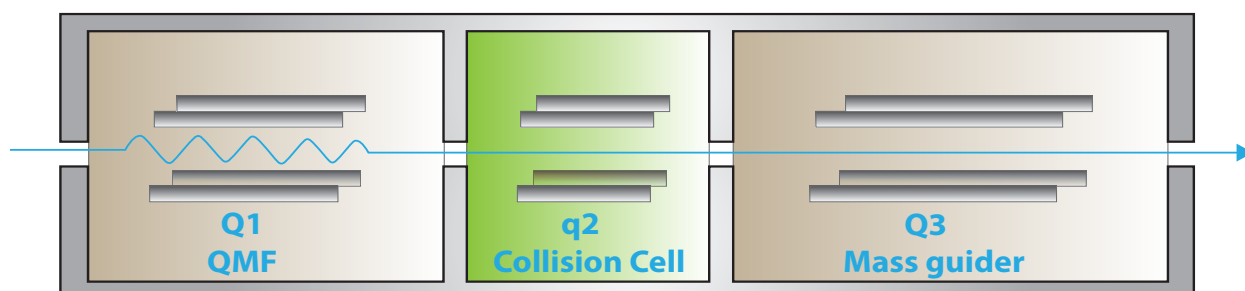
Figure 1.17 The geometry of a QqQ tandem Quadrupole mass analyser.

1.5.2.2 Time-of-flight

In 1946, the prototype of a time-of-flight (TOF) mass analyser was designed by Stephens to determine the m/z ratio of ionised molecules by examining their flight times in mass analyser after acceleration in an electric field (Stephens, 1946). In the same accelerating field, ions are accelerated to different velocities which correlate to their m/z ratio. They are then transferred into a vacuumed flight-tube and their flight-times to traverse this tube are recorded to determine their m/z. The m/z ratio of each ion negatively correlates to its initial velocity, so that the lower m/z ions have a shorter flight time in the tube than that of a higher m/z ion.

There are two advantages of a TOF analyser in practice. First, there is theoretically no upper-limit for m/z range in a TOF analyser, but practically, the detected highest m/z for a singly charged ion reached up to 330 KDa in a TOF analyser (Moniatte et al., 1996). Second, dissimilar to the described Quadrupole mass analyser in the above Section, there is no selection of ions in the mass analyser to transfer to the detectors. The transmission rate of ions is subsequently much higher in TOF mass analyser. Thus, a higher sensitivity is achievable to detect the signals for molecular ions with minor abundance.

In practice, the original linear TOF mass analyser is disadvantaged by low resolution. The low resolution is mainly due to two reasons: 1) the ion source may supply heterogeneously distributed ion batches to the analyser causing differentiated spatial entry location and even different entry time-point; 2) the initial kinetic energy may not be homogenous on the same ions, so that subsequently the ion velocity may be varied in the mass analyser for the same m/z ratio. The combined effects of these two obstacles generally lead to a low resolution by broadening peak width.

The first solution to improve TOF resolution was introduced by Wiley and McLaren in 1955, as delayed extraction (Wiley and McLaren, 1955). The general concept of this solution is that the energetically differentiated ions are delayed in transmission into the mass analyser by allowing more energy to be transferred into the low energy ions in the acceleration field (Figure 1.18). The molecular ions after ionisation enter a field-free area to allow free movements in a transient delay. The differentiated energy level of molecular ions allows them to distribute into the area differently as the higher energy ions move farther into the area than the lower energy ions. A pulsed voltage is then applied to accelerate all ions to enter mass analyser. The lower energy ions need to travel for a relatively longer distance in the acceleration field, and more energies are simultaneously transferred to them. Thereby, the varied velocities are relatively unified for the molecular ions with the same m/z.

In 1973, the reflectron electric field, as a second solution, was introduced by Mamyrin (Mamyrin et al., 1973) to improve the resolution of TOF mass analyser (Figure 1.19). The general concept of this approach is to extend the travel time of higher energy ions for a specific m/z in the flight-tube to match
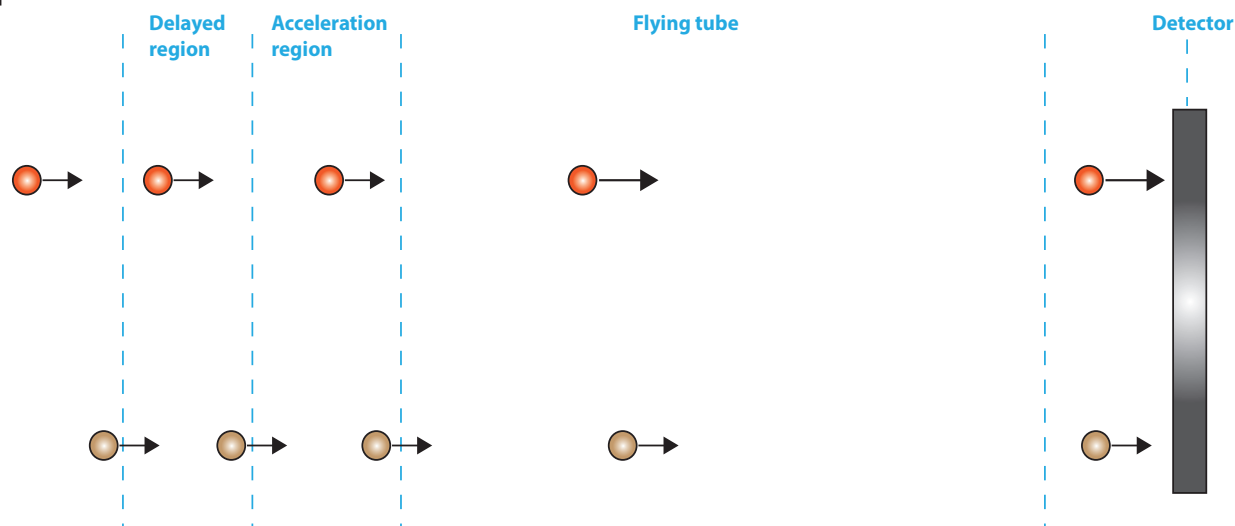
Figure 1.18 The schematic diagram of delayed extraction in TOF mass analyser

with their lower energy ions by a reversed electric field. The reflectron field is installed in the flight tube near the linear TOF detector as a series of electrodes. The reversed electric field can initially decelerate entered ions and then re-accelerate these ions in the opposite direction. The ions with higher kinetic energy move a longer distance into the electric field and re-accelerate to result in a delayed arrival at the detector which matches with that of the lower energy ions that travel a shorter distance in this reflecting field. Consequently, the resolution of the mass spectrum is improved as peak width is sharpened for signals. Due to the higher demand for TOF resolution, multi-reflectron incorporated TOF analysers were developed in the last decade.

1.5.2.3 TOF/TOF

TOF/TOF is a tandem mass spectrometric instrument which is developed from a basic TOF analyser (Figure 1.20). The TOF/TOF instrument would allow a fragmentation MS/MS experiment to be carried out directly in a TOF type analyser. The major geometric alterations in a TOF/TOF instrument are installations of a mass analyser as an ion selector and a collision-induced disassociation (CID) collisional cell with accelerating and decelerating electric fields in the original TOF flight tube. The collisional cells would allow to induce dissociation of glycosidic bonds of glycans by collision with filled inert gas, such as Argon. For a typical TOF/TOF experiment, the targeted parental ions for fragmentation are firstly selected by the ion selector and they are subsequently decelerated to enter the collisional cell which is filled by the inert gas. The parental ions are fragmented by collisions inside the cell. Generated fragment ions are then re-accelerated in the accelerating field to resume the flight in the second analyser. The fragment ions are eventually reflected after the reflectron field and detected by the detector. The studies carried out in this thesis mainly utilised a MADLI-TOF/TOF instrument named Applied Biosystems 4800 mass spectrometry for glycomics studies in both glycome screening and fragmentation based structural definition.
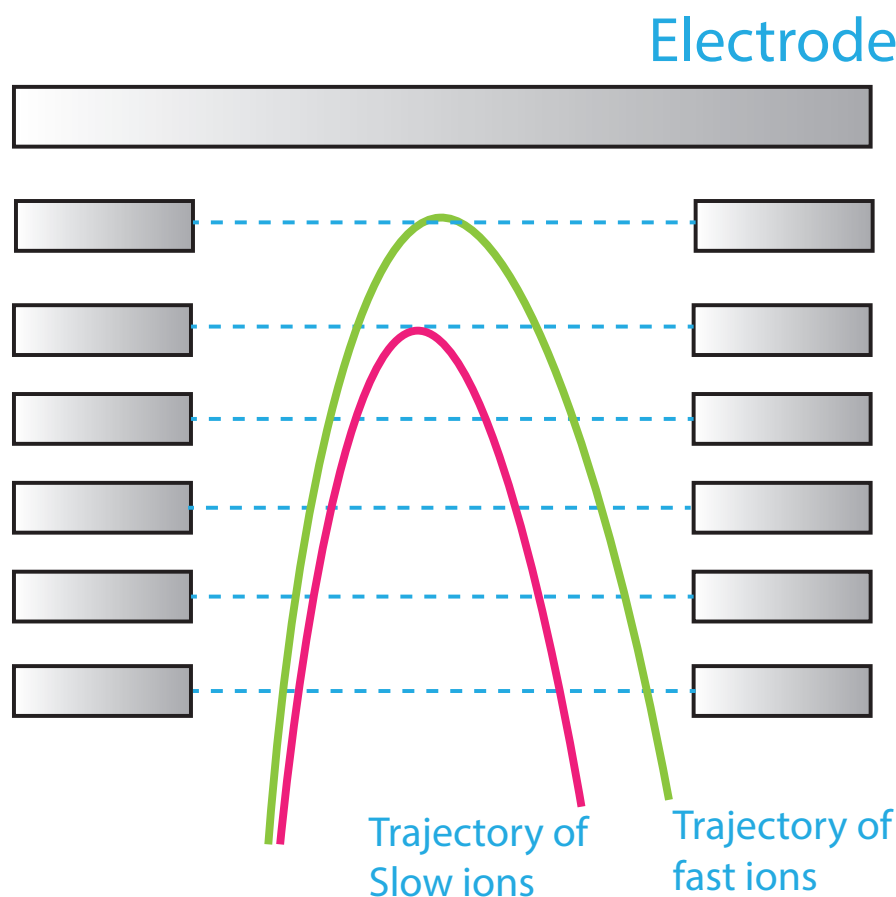
Figure 1.19 The schematic diagram of Reflectron Field of a TOF mass analyser.

1.5.2.4 Quadrupole Time-of-flight

In the 1990s, Prof. Howard Morris from Imperial College was the first to develop a hybrid type mass analyser, the Q-TOF, which is coupled with an ESI ion source (Morris et al., 1996). As the named suggests it was a hybrid analyser made up of a combination of several quadrupoles (QMS) and an orthogonal TOF mass analysers (Figure 1.21). The geometry of this instrument enables MS and MS/MS experiments to be firstly performed in a single instrument with an ESI ion source. There are two QMS type incorporated into the instrument to fulfil different roles. In the MS mode, these two QMS serve as ion guide and molecular ions from ESI source are gudied into the TOF analyser for m/z definition of intact molecular ions. In the MS/MS mode, the first QMS analyser serves as a ion selector to specifically select partental ion for the subsequent fragmenation analysis. Once selected parental ion passes through the QMS, it would be immediately introduced into the second QMS which is filled with inert gas to serve as a CID cell. The fragmentation occurs in this cell and fragment ions are then introduced into orthogonal TOF mass analyser. The orthogonal TOF mass analyser finally analyses the intact or fragment ions from QMSs and impacts of the initial velocities and spatial locations of ions on m/z determination are minimised in this analyser. High sensitivity, high mass range and high spectral resolution are achievable in this hybrid mass analyser. Two Q-TOF type mass spectrometry systems have been used in this thesis. The Waters Synapt G2-Si High definition MS and the Applied
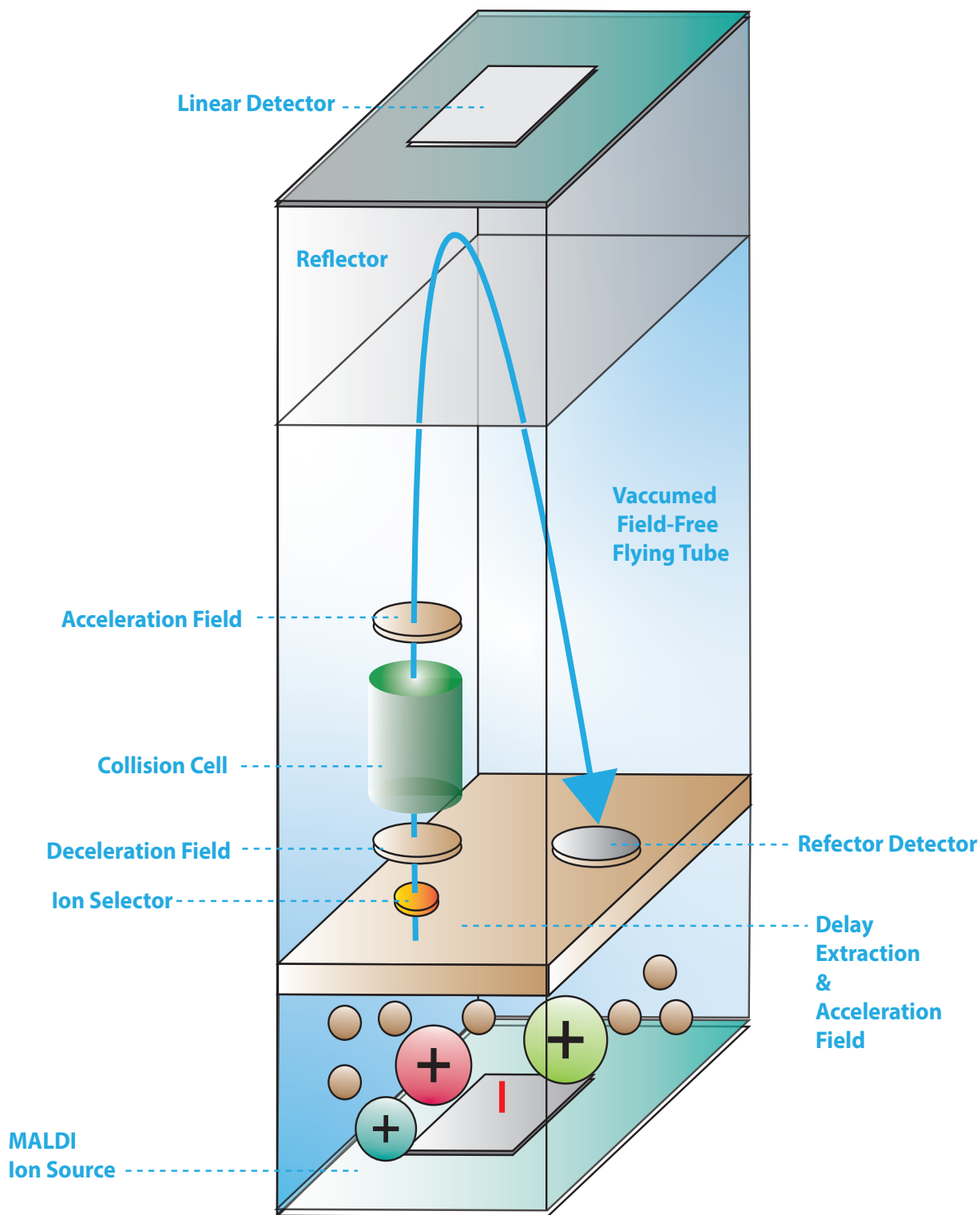
Figure 1.20 The geometry of a MALDI-TOF/TOF mass spectrometer.
The operation of the collision cell and relevant components is optional, so that the instrument can carry out MS and MS/MS experiments employing the same flying route.
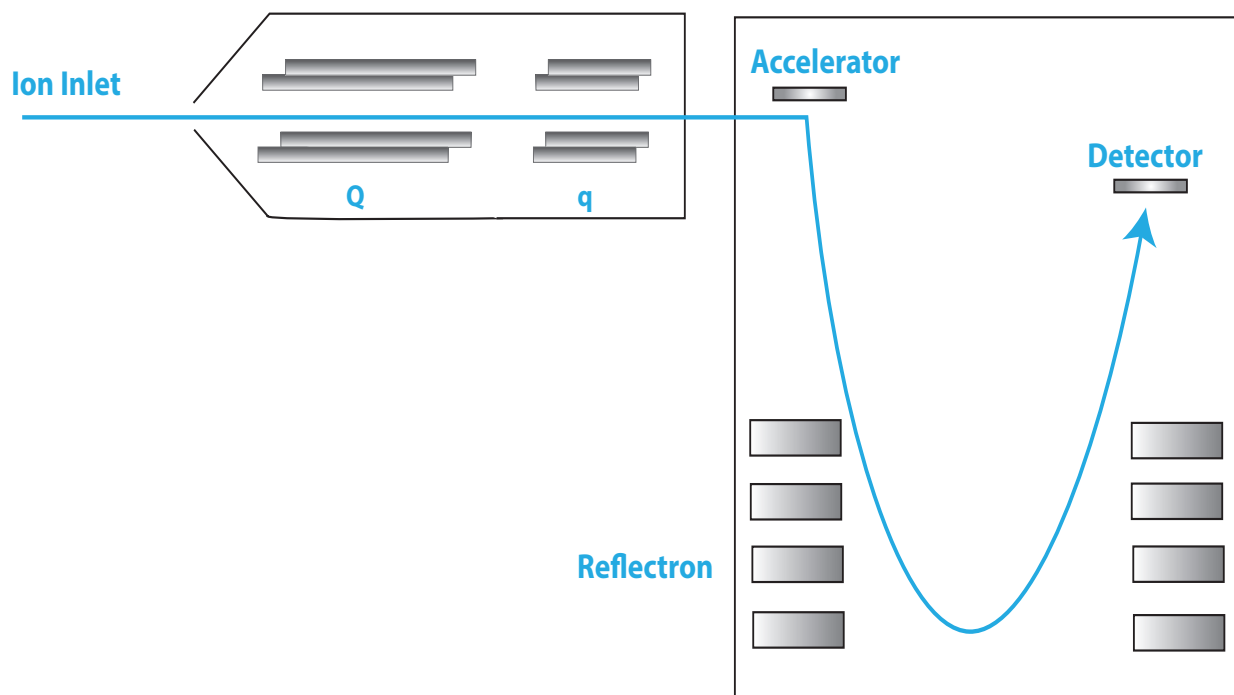
Figure 1.21 The principal geometry of a Q-TOF mass analyser.

Biosystems QSTAR mass spectrometer. The former was used for the study described in **Chapter 6** to obtain high quality mass spectrometric data for protein identification by proteomic. The latter instrument was applied in the study described in **Chapter 3** for glycoproteomic research.

1.5.3 Derivatisation

Glycans are natively highly hydrophilic which is not favoured for efficient ionisation in mass spectrometry. Derivatisation is introduced to increase the hydrophobicity of glycans and subsequently improve the efficiency of ionisation. There are several derivatisation strategies such as permethylation and peracetylation available for globally derivatising hydroxyl and amide groups on native glycans. Permethylation is a primary choice for derivatisation in glycomics studies. Permethylation can replace all the protons of hydroxyl and amide groups on glycans by methyl groups. Increasing the hydrophobicity by permethylation is not only beneficial for the ionisation process and sensitivity, but also can stabilise labile glycosdic bonds on glycan structures, such as sialic acid residues. Native sialic acids are labile in ion-source to cause loss of signals, whereas permethylation can improve their detection. Moreover, increased hydrophobicity of glycans by permethylation can also simplify the purification of glycans from salt and hydrophilic contaminants in practice.

For detailed structural definition, permethylation is also helpful in MS/MS fragmentation analysis and GC-MS linkage definition. In MS/MS, fragmentation of permethylated glycans generates consistent fragmentation patterns which can match with proposed fragmentation pathways (Dell et al., 1983). The cleavage of a glycosidic bond on glycans is distinguishable after permethylation as an ad-

ditonal loss of a methyl group on the reducing-end side fragments. In GC-MS, permethylation is a prerequisite for defining glycosidic linkages as it can pre-label hydroxyl groups which are not engaged in the formation of glycosidic bonds.

1.5.4 Fragmentation of glycans

In a low energy CID fragmentation approach, there are two common types of fragmentation observed on permethylated glycans as A type cleavage and β-elimination. The A-type cleavage preferably occurs on a monosaccharide with amide group as HexNAc (GalNAc/GlcNAc). The cleavage generates oxonium ions which can subsequently undergo a β-elimination of the group on C-3. Ions produced by
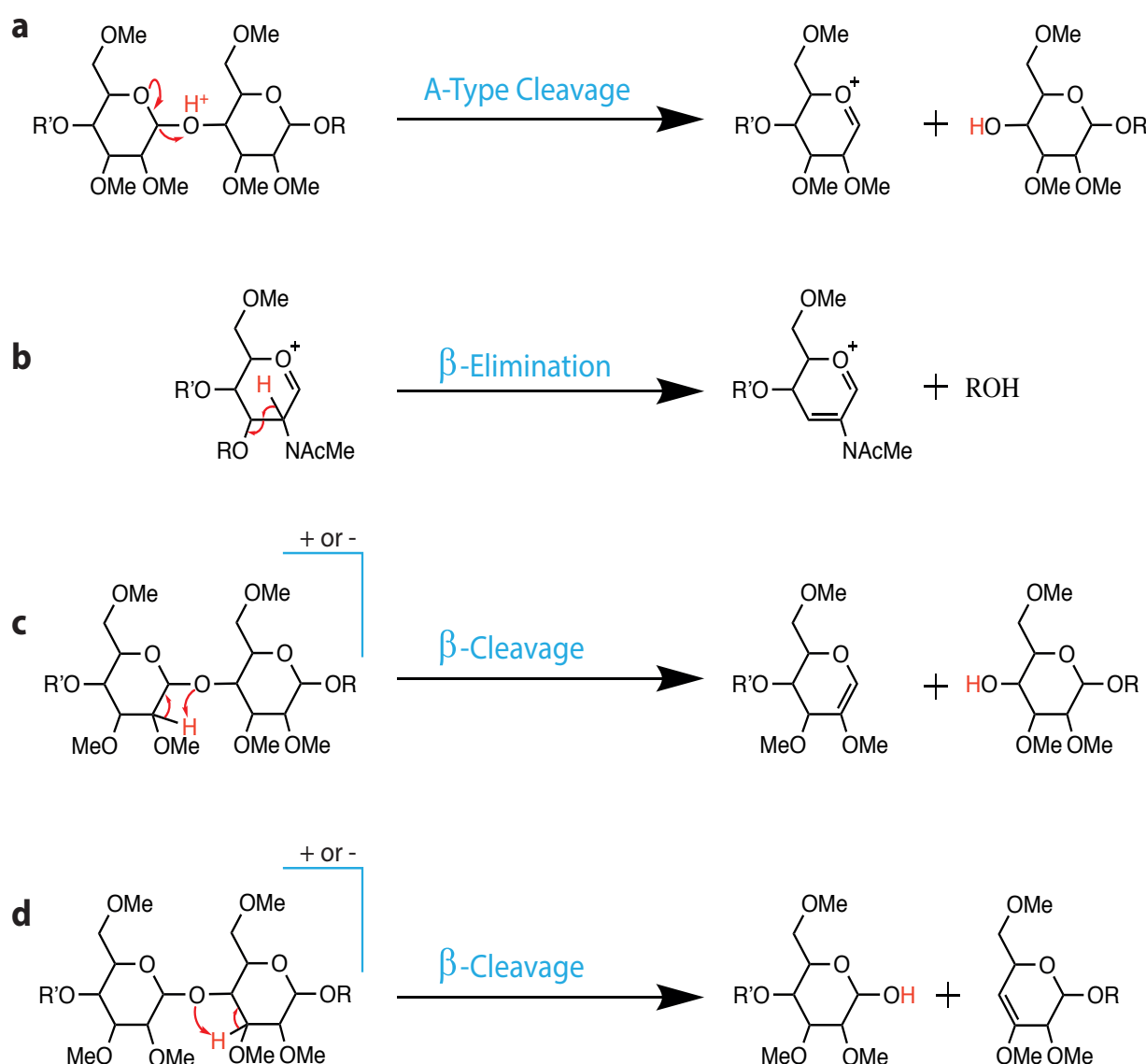
Figure 1.22 Fragmentation pathways for permethylated glycans.
Oxonium ions formed on the anomeric carbon side of glycosidic bonds after occurrence of A type cleavage (a). Once an amide group present at the C-2 of an oxonium ion, further β-elimination would occur for C-3 to remove its group (b). β cleavage can occur on both side of a glycosidic bond as (c) for non-reducing end side and (d) for reducing end side, and the fragmentation associates with a proton transfer as indicated by red arrows and labels.

this elimination are indicative for the occupation of C-3 on HexNAc. A β-cleavage can occur on either reducing or non-reducing end of a glycosidic bond. The cleavage of glycosidic bonds by this pathway is accompanied by a proton transfer (Figure 1.22).

In addition, breakage of the monosaccharide ring structure is also achievable under higher energy. In the cross-ring cleavage, sequential electron movements lead to eliminations on the ring structure forming smaller fragment ions with multiple bonds. Linkage associated structural information can be obtained from analysing these smaller fragments to reinforce the structural definition.

To consistently describe the fragmentation of glycans, Domon and Costello introduced an alphabetic nomenclature for naming fragment ions of glycans which is used in this thesis (Domon and E Costello, 1988). The fragmentation ions from the reducing end terminal are named X, Y and Z, whereas the ions from non-reducing end terminals are designated to A, B and C. The subscript number of fragment ions indicates the site of the cleaved bonds from their corresponding ends (Figure 1.23).

1.5.5 Peptide fragmentation

For describing peptide fragmentation, a similar nomenclature was introduced by Fohlman and Roepstorff (Roepstorff and Fohlman, 1984). This nomenclature is also used in this thesis for the analysis of peptide fragmentation in the glycoproteomic research in **Chapter 3**. This nomenclature defines the fragmentation ions based on the three different types of bond-cleavage. The three type of bonds that can be cleaved on a peptide sequence as $C\alpha$-C, C-N and N-$C\alpha$. The fragment ions from the N-terminal of peptides are named a, b and c, whereas the fragment ions from the C-terminal of peptides are
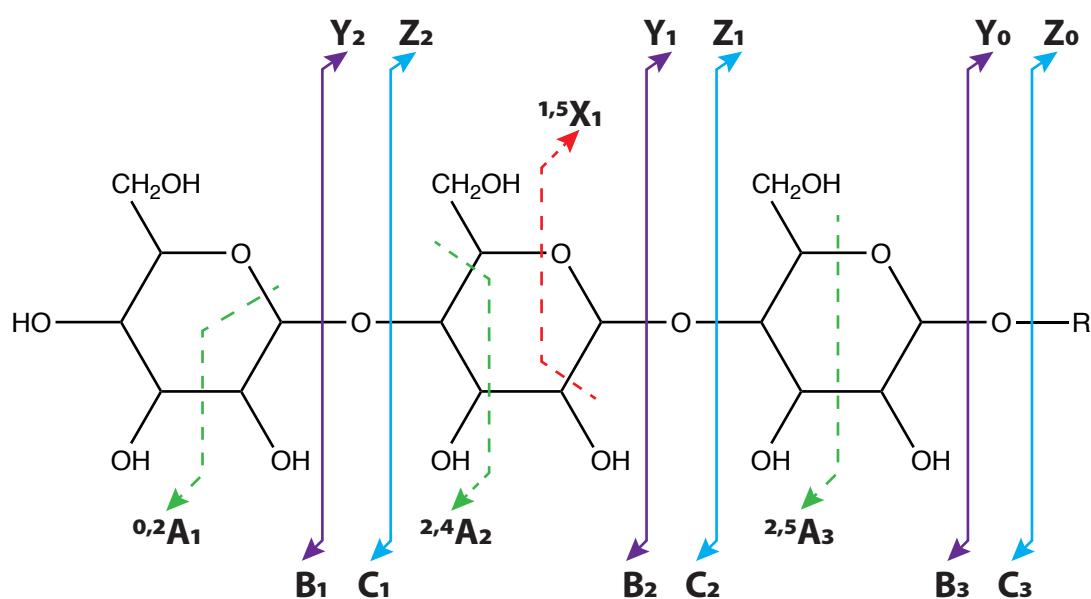
Figure 1.23 Systematic nomenclature of glycan fragmentations.
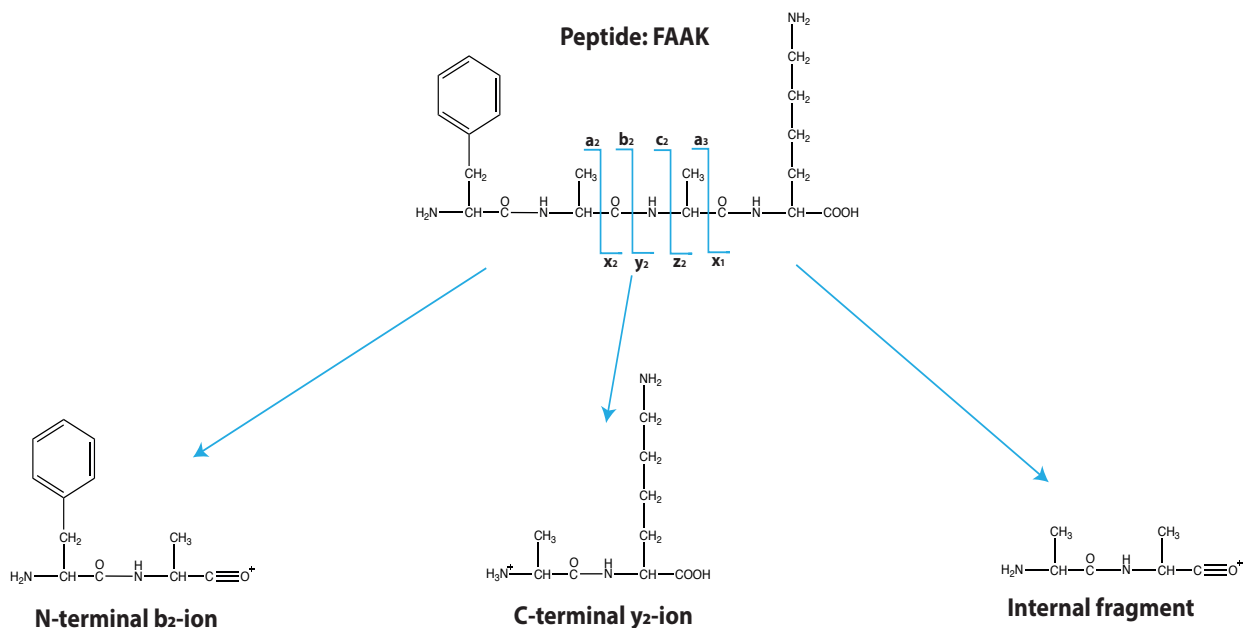
Figure 1.24 Systematic fragmentation nomenclature of peptides and BY & internal fragmentation.

named x, y and z. The subscript number of these fragment ions represents the site of bond cleavage on the fragment ion from their retained terminals. The most common fragmentation for peptides is observed on C-N bonds and the fragmentation ions for this cleavage are b and y ions. Internal fragmentation is also observable on peptide fragments and it is generally resulted from cleavages of two C-N bonds to lose both N- and C-terminals (Figure 1.24).

1.5.6 Linkage analysis

Fragmentation analysis for glycans by mass spectrometry can provide invaluable information for structural definition, but linkage specificity of glycosidic bonds is still largely unobtainable via the most common fragmentation techniques. Alternatively, GC-MS linkage analysis is introduced to overcome the limitation by collectively defining the linkages and compositions of monosaccharides which exist in a targeted glycan pool.

The experiments for GC-MS is typically a derivatisation for glycans after permethylation. A complete permethylation of glycans is a prerequisite as it labels all unoccupied functional groups on a glycan with methyl groups. The permethylated glycans are then hydrolysed in acidic solution to cleave all pre-existing glycosidic bonds. Monosaccharides are liberated from this procedure and then converted to alditols by reduction at their reducing ends. Acetylation is applied to reoccupy all freed hydroxyl groups from the previous processes. The thermally stable and volatile derivative of monosaccharides named partially methylated alditol acetates are injected into the GC-MS instrument. Due to the differentiated volatility of monosaccharide derivatives in the heated column, each of them is eluted into the mass analyser at a different time point. A comparison of fragmentation patterns and elution time of derivatives to pre-characterised monosaccharide standards can indicate the identity and linkage

specificity for each monosaccharide derivatives. Subsequently, all the linkage information can be mapped back to define interested linkages of an epitope or structural feature based on the knowledge of biosynthetic pathways. In addition, the relative abundance of each monosaccharide with a specific linkage can also be elucidated by normalising the intensity of their spectral signals.

However, there is still limitation for GC-MS linkage analysis. Sialic acid residues are labile in the harsh acidic hydrolysis, and no derivative can be generated for this group of monosaccharides. Therefore, the linkage information on these monosaccharides can only be partially and indirectly illustrated by the linkages of their potentially connecting monosaccharides via this approach. Some linkage information is completely lost such as that of the α2,8 linked poly-Sialic acids. Therefore, alternative approaches are required to study the linkage of sialic acid residues such as enzymatic digestion of sialic acids with specific linkage preference.

1.6 Immunity

Immunity is a self-defence system that is developed in vertebrates to prevent undesirable symbiosis and invasion with pathogens. Non-specific and specific immune systems are established in vertebrates to counter all kinds of potential pathogenic activities, named innate and adaptive immunity. Innate immunity is non-specific and provides universal but unselective protection, whereas adaptive immunity, is selective and provides specific pathogen-directed protection.

1.6.1 Innate immunity

Innate immunity is a universal protection and the first line of immune defence. It consists of four defensive systems that are organised in a hierarchical structure morphologically and functionally.

The anatomic barrier is the first established barrier for example at the surface of skin and mucous membranes. It is the first defence to prevent penetration of exogenous pathogens by mechanical and biochemical approaches. Skin cannot only prevent the entry of pathogens physically but it also produces sebum from the sebaceous glands of its dermis to inhibit the growth of microbes. Mucous membranes line the surface of four organ systems which require to be lubricative, including respiratory, alimentary, urogenital tracts and conjunctivae. The secretions of mucous membranes can assist to clear penetrating pathogens with antiviral and antibacterial substances. The inhabitation of microbes onto these organ systems is also manipulated by these secreted materials. Pathogens can also be propelled by the movement of a protrusion on the mucous membranes named cilia.

As the second barrier, physiological barrier has three main factors that contribute to the immune defence, as temperature, pH and soluble factors. The former two factors can constitute unfavourable environments for the pathogens to survive, whereas the soluble factors (lysozyme, complement, and interferon) act in a more active approach against pathogens. For instance, lysozyme from mucous

membrane can actively hydrolyse bacterial peptidoglycans. Virus infected cells can produce interferon to alert nearby cells by inducing antiviral activities. Moreover, the complement proteins can be activated by the immune systems to cause direct damage or trigger phagocytosis to engulf pathogenic cells.

The phagocytic barrier is an immune cell-engaged approach; three cell types, macrophages, neutrophils and monocytes are specialised to conduct phagocytosis in which process that antigens, including microbes, are engulfed and digested.

The inflammatory response is the last barrier that innate immunity recruited to counter tissue damages caused by pathogens. The inflammatory response is collectively referred to a series of events which are induced after tissue damages. Once damage occurs, blood is efficiently carried away after widening vessels surrounding the damage site which leads to an engorgement in the area. The observable physiological effects of this event is the redness of tissue and higher local temperature. Permeability is then increased for the capillaries around the area. Phagocytes and body fluids subsequently penetrate to the damaged tissue leading to swelling. There are multiple steps for phagocytes to migrate from capillary vessels to the damaged sites. Phagocytes are initially captured onto the surface of vessels via cell surface receptors. The phagocytes then pass through the endothelial barrier into the tissue via a process named extravasation. The phagocytes eventually arrive at the affected sites through chemotaxis and then commence phagocytosis.

The innate response is not pathogen-specific, some molecules in this system such as pattern-recognition receptor (PRRs) and secretory complement proteins are still able to recognise pathogen-associated molecular patterns (PAMPs) on pathogens and induce immune responses. As PAMPs are a wide range of molecules from exogenous resources, some glycan binding lectins are also engaged as PRRs to recognise glycan-based PAMPs.

1.6.2 Adaptive Immunity

Adaptive immunity is highly discriminative to antigens from "non-self" molecules and pathogens. It is capable of discerning minor differences between presented antigens as post-translational modifications including glycosylation . The system can then direct a highly specific immune response to eliminate the discerned "non-self" target. The highly specific recognition exhibited by adaptive immunity is conferred by its vast production of largely diversified recognition molecules to discern various features of an antigen. Besides, immunologic memory is another vital feature of adaptive immunity. Specifics of an immune response to antigens are generated and stored in memory cells. Once the system is challenged by the same antigen again, the memory cells would lead a more aggressive and immediate response from the adaptive immune system to the descriminated targets.

Fundamentally, adaptive immunity needs two types of cells to fulfil its functionality. The first type is antigen presenting cells including macrophages, B cells and dendritic cells(DC), which can engulf pathogens by endocytosis and process the molecules from pathogens to antigens. The second type is lymphocytes including T cells, natural killer(NK) cells and B cells, which can receive the antigens and/or signals from the antigen presenting cells to mediate pathogen-specific immune responses. B cells are a dual-member of both groups indicating its dual functions. In addition, NK cells play dual roles in both the innate and adaptive immune systems. The major effective members of the adaptive immune system are lymphocytes which are sourced from bone marrow and produced via a haematopoiesis process. They relocate into different lymphoid organs from bone marrow via circulation in lymphatic and vascular systems at different developmental stages.

The maturation of B cells is completed in the bone marrow before their circulation and relocation. A matured B cell expresses a B-cell receptor (BCR), which is a membrane-bound antibody molecule with an exclusive antigenic specificity obtained via gene rearrangements during maturation. Naïve B cells are the cells which have not encountered a designated antigen. Once its BCR specifically binds to the antigen and induces a primary response, the naïve B cells rapidly proliferate to a large population and further differentiates into memory B cells and effector B cells. Effector B cells then rapidly scale up the immune response by abundantly producing secretory antibodies to target the encountered antigens. Memory B cells record the features of encountered antigens and retain a long life-span to standby for further pathogenic encounters. They can trigger stronger responses with a more immediate effect once the designated antigens are presented again.

Different from the maturation process of B cells, T cells are matured in the thymus after their relocation from bone marrow. Each T cell expresses a specific antigen-binding molecule after maturation in the thymus. The antigen-binding molecule is named the T cell receptor (TCR), and its binding specificity is obtained by random rearrangements of related genes. Instead of a direct interaction with antigen, TCRs are required to bind with a membrane-bound antigen presenting complex from antigen-presenting cells. The antigen presenting complex is named major histocompatility complex (MHC). It is an assembly of membrane-bound glycoproteins in which specific molecules are trapped and processed from endogenous and exogenous sources to present to T cells. Two types of MHC molecules are expressed on cells as MHC I and MHC II. Nearly all nucleated cells have expression of MHC I molecules on their membranes. MHC I is capable of presenting endogenous antigens from viral-infected, cancerous and host cells. Besides, MHC II molecules are mainly expressed on antigen-presenting cells and responsible for presenting exogenous antigens obtained from phagocytosis of microbes and pathogens.

T cells can be further grouped into two major subtypes as T helper (T_H) cells and T cytotoxic (T_C) cells based on divergent functionality, MHC binding specificity and expression of membrane-bound

proteins. T_H cells can associate with MHC II complexes to recognise their presenting exogenous antigens, and a membrane bound glycoprotein called CD4 is also exclusively expressed on this subtype of T cells. Once the primary response of T_H cells induced by MHC molecules, T_H cells differentiate into different subtypes of effector T_H cells, such as T_H1 and T_H2, with rapid proliferation. Each subtype of effector T_H cells produces varied sets of cytokines to reinforce other cell-mediated innate and adaptive immune responses. For instance, the cytokine signals from T_H1, such as IL-2 and IFN-γ, can induce macrophages and T_C cells to target intracellular viral infections. T_H2 cells produce cytokine signals as IL-4 and IL5 for B cells to induce their antibody production for enhancing immune responses against extracellular pathogens. Besides, T_C cells express a membrane bound glycoprotein CD8 and their TCRs can exclusively bind with MHC I molecules to proliferate and become effector cells named cytotoxic T lymphocytes (CTL). Once the designated antigens on MHC I are presented to CTLs, they can be induced to perform cytotoxic activities targeting nucleated cells presenting antigens with pathogenic origins. Memory T cells also exist in T cell population, and they function similarly to B memory cell to record the structural feature of encountered antigens and induce highly specific and strengthened immune response in the subsequent antigen-exposure.

1.6.3 Antibodies

Antibodies are a group of glycoproteins, also named immunoglobulin (Ig), which can specifically bind to a feature of antigen to induce recognition by the immune system and effect subsequent immune responses. B cells express secretory antibodies into blood circulation, and membrane-bound antibodies on the cell surfaces are the major part of BCRs. The basic antibody is a Y-shape protein which consists of two Ig light chains and two Ig heavy chains. Each light chain is associates with a heavy chain by a disulfide bond. Two heavy chains are connected by disulfide bonds to form the final Y shape. Based on the sequence and domain arrangements of the antibody heavy chains, antibodies can be classified into five classes: IgG, IgA, IgD, IgE and IgM. There are one variable and few constant domains on an Ig heavy chain; for instance, three constant domains are on IgG, IgD and IgA as well as four constant domains on IgM and IgE. A linker sequence named hinge region is on the heavy chain of IgD, IgG and IgA between their CH1 and CH2 domains, but not on IgM or IgE. The light chain has one variable domain and one constant domain on its sequence. The variable domain has the most prominent diversity on its sequence among all known antibody domains, whereas constant domains do not exhibit diversity in the same class. There are two types of light chains and their definitions are based on the sequence of their constant domains as κ and λ. Corresponding to the nomenclature of antibodies, heavy chains are also defined by the sequence of their constant regions as α, γ, δ, ε and μ. In addition, only one type of light chain, either κ or λ, can be assembled onto a single antibody.

Antibodies can not only present as a monomeric structure, but also can be further polymerised to form a larger structure (Figure 1.25). IgM and IgA both can be polymerised forming monomers and polymers, while other Ig types can only form monomers. IgM is the major component which consti-

tutes the BCR once in a membrane-bound and monomeric form. IgM can also be secreted into serum with a pentameric-structure in assistance of a joint chain (IgJ). In the primary response of naïve B cells, pentameric IgM is firstly produced into serum to entrap the antigens, other types of Ig heavy chains are subsequently produced on antibodies via class switching process. In the class switching, the variable domain is preserved in the rapid production of antibodies to maintain antigenic specificity, but the constant region is changed from Ig μ chains to other heavy constant chains. Besides, IgAs in serum are primarily secreted as monomers. Once IgAs are secreted to external environments, they can polymerise to form di-, tri- and even tetramers. The formation of secretory IgA dimers requires a IgJ chain and a secretory component to connect the constant region of the heavy chains on two IgA monomers. In this formation, secretory IgA can entrap pathogens to prevent their contact with mucosal membrane via antigenic association.

Antibodies specifically recognise antigens via binding which occurs on the variable domains. In mucus material, the secretory IgA entraps pathogens via an antigenic binding which normally does not result in subsequent immunological responses. However, more commonly, antigen-antibody binding results in varied responses as a consequence of simultaneous interactions of the constant regions with cellular receptors of the immune systems. This is exemplified by the interactions and functionalities of IgG. IgG is the most abundant serum antibody which maintains a monomeric form in the circulation. There is a group of receptors (FcγR) on immune cells which can specifically recognise a region of Fcγ constant chains to trigger divergent cellular responses. The phagocytosis of macrophages and neutrophils can be enhanced once their FcγRs bind to an IgG constant chain. Cytotoxic NK cells can be guided and triggered to perform cytotoxicity activities via binding of their FcγRs with IgG constant regions to initiate antibody-dependent cell-mediated cytotoxicity (ADCC). The constant region of IgG
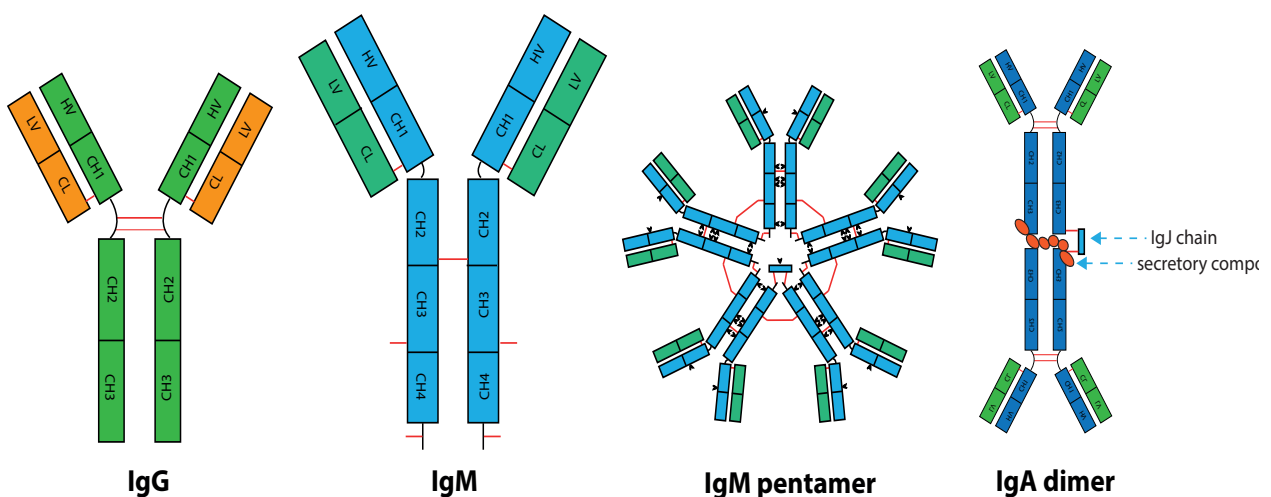
Figure 1.25 Schematic diagram of antibodies.
Monomeric IgG &IgM, IgM pentamer and IgA dimer. The IgM pentamer is formed via disulfide bonds between monomeric IgM in assistance with a J chain. The IgA dimer is linked by a J chain and a secretory component.

can also be bound by complement proteins to induce the complement cascade.

1.6.4 Glycan-binding proteins in the immune system

In the mammalian immune system glycosylation is common on immune cells and molecules, and it is of importance in assisting the immune system to fulfil its functionalities. Glycans have salient roles in immune activities to regulate migration and recognition of immune cells as well as their maturation and responses (Haslam et al., 2008). Glycan binding proteins, also named lectins, in mammalian immune system help to fulfil these roles via their interactions with glycans. Lectins are characterised by having a conserved carbohydrate-recognition domain (CRD) which binds with glycans (Drickamer and Taylor, 1993, Weis and Drickamer, 1996). Lectins in animal systems are further classified into five sub-families based on the sequential arrangements of their CRD domains and ligand binding specificities (Drickamer 1988). Three sub-families are widely observed in the immune system, C-type lectins, Siglecs and galectins. They are engaged in various events to regulate immune activities (van Kooyk and Rabinovich, 2008).

1.6.4.1 Galectin

Galectin is a lectin family which is characteristic for its β-galatoside binding preference and distinctive CRDs (Barondes et al., 1994). Fifteen galectin members are identified in mammalian systems, named galectin-1 to -15. Based on their sequence arrangements, these fifteen mammalian galectins are categorised into three types: proto-, chimera- and tandem repeat (Figure 1.26A). Nine of the galectin members are proto-type with a single CRD on their sequence as galectin-1, -2, -5, -7, -10, -11, -13, -14 and -15. They are normally dimerised to be homodimers with bivalent glycan binding affinity. Besides, Galectin-3 is the only chimera-type galectin which has a single CRD in its sequence. Galectin-3 can form homo-oligomers via N-terminal interactions to obtain multivalent glycan binding affinity. Tandem repeat type galectins (-4, -6, -8, -9 and -12) have two CRDs on their sequence to confer natural bivalent glycan binding affinity. The β-galactoside of the LacNAc disaccharide is a universally preferred ligand for all galectins. The binding activities of galectins to this preferred ligand can be manipulated by modifications on the non-reducing end Gal residues on a LacNAc unit, such as fucosylation and sialylation. Moreover, different galectin members still have differentiated binding affinity to LacNAc moities on different glycan structures. The ligation of galectins to glycans can cross-link LacNAc bearing glycoconjugates, mostly glycoproteins. This cross-linking can occur on cell membranes and the extracellular matrix (ECM) to regulate cell-cell, cell-matrix and membrane protein integrations (Figure 1.26B, Perillo et al., 1998). These interactions are vital for several immunological events to manipulate cell survive, growth and adhesion (Rabinovich et al., 2007, Rabinovich and Toscano, 2009). Examples of the biological functions of galectins in the modulation, development and functionalities of immune cells is outlined below.

Expression of Galectin-1 in thymic epithelial cells facilitates its engagement in T cell selection during the T cell maturation process (Baum et al., 1995). Galectin-1 has been shown to be able to selectively induce apoptosis of immature Thymocytes (Perillo et al., 1997). The differentiation of glycosylation on thymocytes was correlated to their maturation status. α2,6 linked sialic acid was exclusively found on matured medullary thymocytes (Baum et al., 1996). The α2, 6 linked sialic acid on T cells was able to inhibit galectin-1 binding and subsequent apoptosis (Amano et al., 2003). This is evident that a glycan-based selection occurs during T-cell selection. Moreover, expression of C2GnT was lower in matured medullary thymocytes than cortical thymocytes (Baum et al., 1995). This indicates a higher level of core-2 O-glycans on immature thymocytes and this branched structure would favour galectin binding as more potential ligands are available. It is possible that self-reactive T cells can exit from the thymus without being eliminated in the maturation process, but Galectin-1 can still have negative impacts on their survivals. This is because the apoptosis of T cells can still be triggered by galectin-1 (Perillo et al., 1995).

The binding ligand of Galectin-1 is primarily a galactoside, the interactions of Galectin-1 to T-cells are restricted to membrane glycoproteins. Galectin-1 was found to have affinity to CD43 on T cells in an early study (Baum et al., 1995). The binding is dependent on the Core-2 O-glycans of CD43 as antibodies for this type of glycans on CD43 could block the interaction. CD43 was then characterised to have heavy O-glycosylation on its sequence and serve as the major binding-ligand for galectin-1 on T cells (Hernandez et al., 2006). It is because the quantitative binding of galectin-1 to T-cells was reduced to around 50% without CD43 expression and apoptosis of T-cells was also induced. Dissimilar to the previously suggested Core-2 specificity for the binding between Galectin 1 and CD43, Core-1 O-glycan only CD43 was shown to have an equivalent effect as Core-2 O-glycan only CD43 to enable its binding with Galectin-1. As a previous study indicated that Core-1 sequences have a lower binding affinity to Galectin-1 than the ligand sequence on Core 2 O-glycans (Leffler and Barondes, 1986), it is very likely that Core-2 glycans are the preferred ligands for galectin-1, but quantitatively they are less expressed on CD43. A reinforcement of this suggestion was found by analysis of CD45 which is another major binding ligand for galectin-1 on T cells.

CD45 was first implicated as a potential ligand for Galectin-1 when T lyphoblastoid and Thymic epithelial cells binding of Galectin-1 was shown to be inhibited by an anti-CD45 antibody which had a specificity against its glycans (Baum et al., 1995). CD45 has isoforms on cells due to differential O-glycosylation. Lower weight isoform with relatively fewer O-glycosites required the presence of Core-2 O-glycans to be interactive with Galectin-1. Such a glycan is not required for its high weight isoform (Earl et al., 2010). Besides, the abundant N-glycans on CD45 were also shown to be engaged in the binding of Galectin-1. However, N-glycans with terminal α 2,6 sialic acid could cause a significant decline of binding affinity. CD7 was also identified as a binding ligand for galectin-1 and the absence of this membrane glycoprotein conferred resistance of Galectin-1 induced apoptosis on malignant
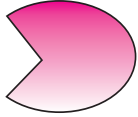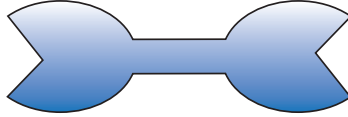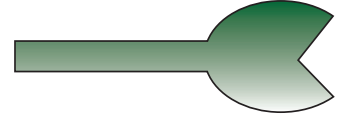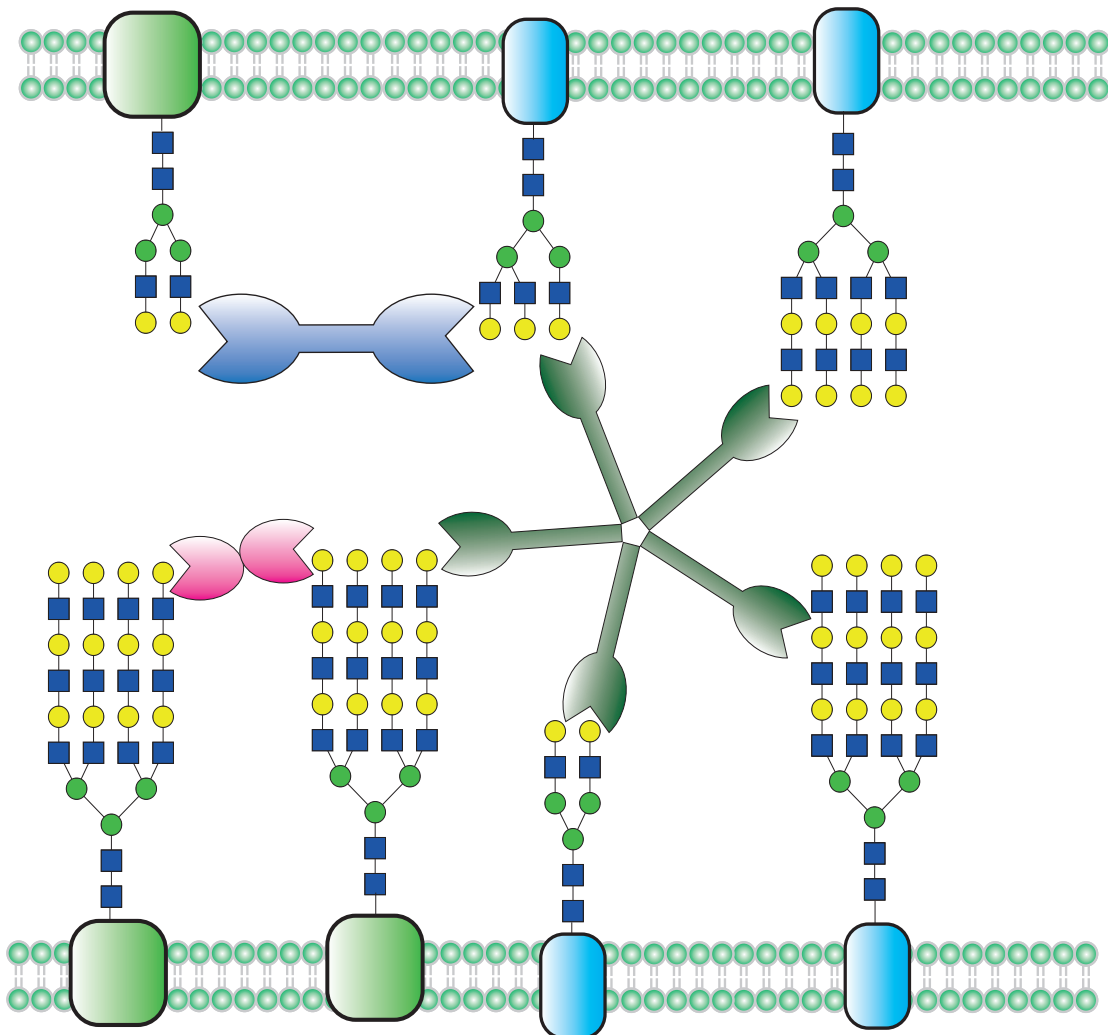
Figure 1.26 Schematic diagram of galectins and interaction.
Three types of galectins monomers are shown in (a), trans- and cis-cellular cross-linkings of membrane proteins via the ligation of galectins and β-galacotside (b).

T cells (Roberts et al., 2003). Sialylation on Core-1 O-glycans of CD7 was also correlated to binding resistance of galectin-1 and subsequent induction of apoptosis (Roberts et al., 2003). CD2 and CD3 are also expressed on T-cells, which have a glycan-dependent affinity with galectin-1, but the binding of CD3 to galectin-1 does not directly result in cell death.

The galectin-3 induced apoptosis was firstly described on multiple types of lymphocyte cells across human and mice, such as human T leukemia cell lines and activated mouse T cells (Fukumori et al., 2003). Glycosylation of CD7, CD29, CD45 and CD71 was shown to be engaged in galectin-3 induced apoptosis, because the exhibited apoptosis could be inhibited by the application of exogenous lactose. Both Galectin 1 and 3 can induce apoptosis signal for T-cells, but they might have different specificity and mechanisms to cause this consequence. In the maturation process in the thymus, Galectin-1 and -3 are both expressed to induce apoptosis for immature thymocytes by negative selection. Only thymocytes expressed either CD4 or CD8 as single positive cells would eventually be matured to be functional T-cells. However, expressing both or neither cell membrane proteins can be found on immature thymocytes and they are named double-positive or double-negative thymocytes. Galectin-1 could universally induce apoptotic signals for both types of thymocytes, whereas the apoptosis of double-negative thymocytes was preferably induced by Galectin-3 (Stillman et al., 2006). Other galectins are also engaged in controlling T-cell death, including galectin-2, -4 and -9. Galectin-2 and -4 are both expressed in the gastrointestinal tract, whereas they induced glycan-dependent T-cell apoptosis triggers different downstream mechanisms. Galectin-2 was shown to bind with CD29 and induce caspase-dependent pathway for T-cell apoptosis (Sturm et al., 2004), but Galectin-4 was shown to bind CD3 and induce calpain-dependent pathways (Paclik et al., 2008). Galectin-9 can interact with TIM3 via its glycans to induce apoptosis on T_H 1 cells (Zhu et al., 2005).

Except for apoptosis induction, galectin-3 was also shown to be a negative regulator for T cell activation. The clustering of the T-cell receptors is required to initiate T-cell responses, whereas galectin-3 cross-linking lattices on the T-cell surface can restrict the mobility of the T-cell receptors to limit the cluster to achieve activation threshold (Demetriou et al., 2001). GlcNAc-T V is a glycosyltransferase involved in the branching of N-glycans to form a β1,6 antenna. A deficiency of this enzyme in mice caused autoimmunity and increased T-cell receptor clustering. A similar effect was also observed on a control group once exogenous lactose was applied to compete for galectin-3 association (Demetriou et al., 2001). Galectin-3 is also engaged in retaining membrane expression of cytotoxic T lymphocyte antigen-4 (CTLA4) on the T cell surface (Lau et al., 2007). It is a negative regulator on T-cell membranes and arrests cell growth and proliferation at the late stage of T cell signalling (Rabinovich and Toscano, 2009)

Table 1.3 Binding specificity and immunological effects of some galectins

Galectins	Binding Specificity	Immunological Effect
Galectin-1	Terminal Gal enhanced binding with branched structures binding blockage with α2,6 sialylation of terminal Gal	Apoptosis of macrophages, thymocytes, T- and B cells Inhibition of T cell growth Cell adhesion Enhancement of phagocytosis by neutrophils Induction of oxidative burst of neutrophils B cell differentiation Mediation of cytokine secretion
Galectin-2	Terminal Gal higher affinity with α1,2 Fucosylation of terminal Gal blockage with α2,3/2,6 occupation of terminal Gal	T cell apoptosis macrophage secretion of lymphotoxin α
Galectin-3	Terminal and Internal Gal Poly-LacNAc moiety and/or branched glycans terminal binding blockage with α2,6 sialylation of terminal Gal	Mediation of T cell growth T cell apoptosis Induction of oxidative burst of neutrophils Enhancement of phagocytosis degranulation of mast cell Mediation of adhesion B cell differentiation Mediation of cytokine secretion
Galectin-4	High affinity for LacNAc with α1,3 Galactosylation and α1,2 Fucosylation	T cell apoptosis Induction of T cell cytokine production
Galectin-8	High affinity for LacNAc with α1,3 Galactosylation and α1,2 Fucosylation	neutrophils adhesion thymocytes cell apoptosis proliferation of T lymphocytes
Galectin-9	Branched and poly-LacNAc elongated glycans	apoptosis of T- and B-cell cytokine secretion adhesion and apoptosis of eosiinophils

Data from (Liu, 2005, Rabinovich and Toscano, 2009, Rabinovich et al., 2007, Tribulatti et al., 2007, Tribulatti et al., 2009)

1.6.4.2 Siglec

Siglecs are a family of lectins which have a binding specificity for sialic acids on glycans with an abundant expression on immune cells. Siglecs stand for sialic acid, immunoglobulin superfamily lectins (Crocker et al., 1998). The first four members, siglec 1-4, are also referred to as sialoadhesin, CD22, CD33 and MAG. There are fourteen members of this family of protein found in human and five found in mouse.

All members of this family are transmembrane proteins. At the extracellular side, they have an Ig-like V-set domain on their N-terminal with multi-numbers of Ig-like C2-set domains underneath. The V-set domain is engaged in sialic acid binding and the C2-set domains are mainly to extend the N-terminal domain into the extracellular environment. There are different types of arrangements for the transmembrane and cytoplasmic sequence of Siglecs (Pillai et al., 2012). The first type has at least one immunoreceptor tyrosine-based inhibitory motif (ITIM) in the cytoplasmic as found on Siglec-2, -3, -5, -6, -7, -8, -9, -10 and -11 in human; and Siglec-2, -E, -F and G in mouse. The second type has a positively charged residue in the transmembrane domain of the Siglec, which has affinity to a membrane protein, DAP12, that has an immunoreceptor tyrosine-based activating motif (ITAM). The members of this type include Siglec-14, -15 and –16 in human and Siglec-3, -15 and –H in mouse. The third type has neither ITIM nor ITAM, or a positive residue in the transmembrane domain, such as Siglec-1 in mouse and human. In addition, Siglec-4 does also not have ITIM or ITAM motifs on its cytoplasmic tail, but instead, it has a site on its sequence for Fyn phosphorylation. Due to its low expression in the immune system, Siglec-4 can be generally regarded as similar to Siglec-1 in functioning in the immune system, its Fyn signalling consequence is not further discussed (Figure 1.27).

Due to the differentiated domain arrangements, the biological engagement of siglec with each type of cytoplasmic arrangements is correspondingly different. For the first type, phosphorylation by Src family kinase for tyrosine on ITIM can be induced by Siglec-Sialic acid ligation. The phosphorylated Siglec cytoplasmic tail can further interact with SH2-domain-containing effectors, SHP-1 and SHP-2. These two effectors are protein tyrosine phosphatases which can further down-regulate the signal transductions of the other receptors (Ravetch and Lanier, 2000). For the second type, Siglec and DAP12 can be associated via interactions of their charged residues in transmembrane domains to indirectly induce phosphorylation for tyrosine on the ITAM motifs on DAP12. The phosphorylated complex can serve as a docking site for inducing subsequent phosphorylation by Syk family tyrosine kinases. Phosphorylation by this kinase family can either positively or negatively regulate downstream signalling of immune cells (Hamerman and Lanier, 2006). The third type, Siglec-1 or –4, could possibly serve as adhesion molecules due to the fact that they do not have relevant downstream signalling potential.

As indicated above, most Siglec family members can be engaged in cellular signal transductions, and
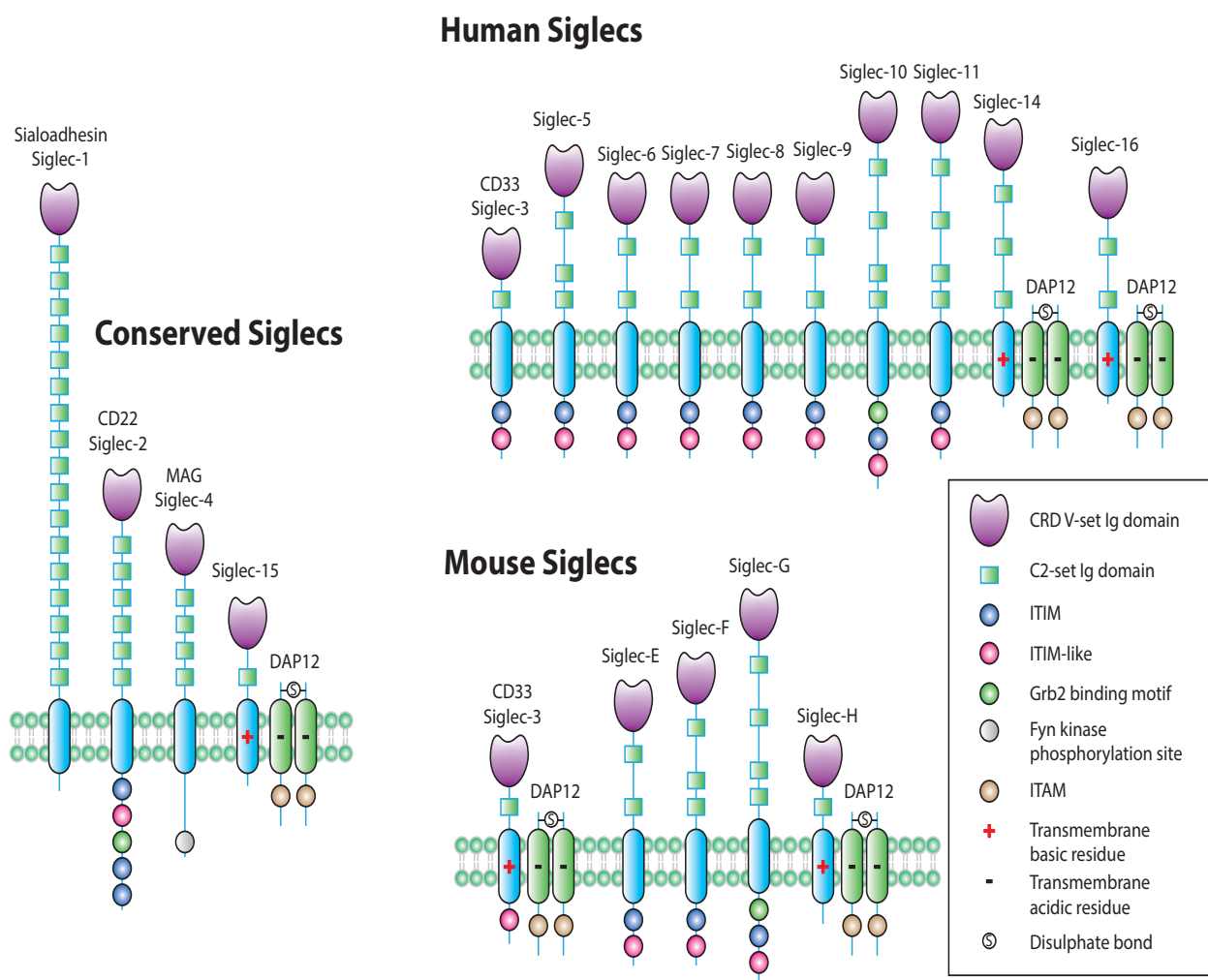
Figure 1.27 Schematic diagram of human and murine siglecs.
There are four conserved siglecs in mammalian systems. They differ from each other by the number of C2 set Ig domain, transmembrane charges and sequence arrangement of their cytoplasmic tails. Siglec with ITIM on their cytoplasmic tails can induce downstream signalling via phosphorylation, while Siglec with short cytoplasmic tail and basic transmembrane residue can attract a DAP12 dimer to trigger signalling via ITAM phosphorylation. Siglecs with neither a basic transmembrane residue nor an ITIM motif would generally be unable to trigger cellular signalling. Adapted from (Varki et al., 2015).

their involvements may generally lead to signal attenuation in the immune system. This signal attenuation could have impacts on unwanted aggressive activities, such as autoimmunity. Their immunosuppressive effects are implicated in studies of mainly B- and T- cells.

CD22 is mainly expressed on B cells, it has ITIM motifs on its cytoplasmic tail and a binding preference for sialic acids with α2,6 linkage (Sgroi et al., 1993). CD22 has one V-set Ig-like domain followed with six C-set Ig-like domains on its extracellular part and it has four ITIM signalling motifs inside the cytoplasm. Early studies investigated the B cell activities of CD22 deficient mice, they observed hyper-responsive activities of B cells, lowered triggering threshold by receptor cross-linking and increased B cell induced cell apoptosis (O'Keefe et al., 1996, Nitschke et al., 1997). Calcium influx, which is a downstream signalling event after B cell activation, was consistently elevated in studies using dif-

ferent CD22 knockout mice (O'Keefe et al., 1996, Nitschke et al., 1997). This could indicate a role for CD22 as a negative regulator for B cell activation. A further study then confirmed that the B cells of CD22 deficient mice could also produce autoantibodies against a range of self-biomolecules, including double-stranded DNA (O'Keefe et al., 1999). This is indicative for a salient role of CD22 in immune tolerance. A deficiency of Lyn, which is a Src family kinase for ITIM phosphorylation, or deletion of SHP-1 in B cells could both cause systematic autoimmunity in mice (Hibbs et al., 1995, Nishizumi et al., 1995). Modifying glycan binding sites of the CD22 V-set domain and applying synthetic sialioside both can lead to elevation of Calcium influx and corresponding down-regulation of tyrosine phosphorylation on CD22 (Jin et al., 2002, Kelm et al., 2002).

Acetylation of the hydroxyl group on C-9 of sialic acid residues (9-O-acetylation) can naturally occur in living system. Its inhibitory effect was found on CD22 as it can prevent the ligation between CD22 and sialic acid residues (Sjoberg et al., 1994). This chemical modification can be reversibly removed by sialic acid acetylesterase, and defects of this enzyme in mice can lead to production of autoantibodies and BCR activation (Takematsu et al., 1999, Cariappa et al., 2009).

Due to the limited expression of Siglecs, such as Siglec-7 and Siglec-9, on subsets of T cells (Nicoll et al., 1999, Zhang et al., 2000), their effects on T cell signalling were artificially assessed by gene transfection into Jurkat cells (Ikehara et al., 2004). Phosphorylation of Siglec-7 and –9 was increased by employing SHP-1, while a decreased phosphorylation of ZAP70, which is a downstream signal after TCR stimulation, was observed. Correspondingly, transcriptional activity of nuclear factor was affected by the presence of either Siglec in activated T cells (Ikehara et al., 2004). Mutations on an arginine residue on the glycan binding site of Siglec-7 and –9 can improve the reduced transcriptional activity. Thereby, the optimal activity of nuclear factor and T cell regulation could be achieved by the presence of both Siglec proteins with ligand bindings.

1.6.4.3 C-type lectin

C type lectins were originally discovered as a group of calcium dependent glycan binding proteins which could have divergent binding specificities and multiple C-type CRDs. Their interactions with targeted glycan moieties require the presence of a calcium cation stabilising their binding pocket to form a preferable binding site (Weis and Drickamer, 1996). The calcium cation is also engaged in lectin-glycan interactions as forming coordination bonds with glycans. The C-type CRD domain is identified as a sub-member of a superior domain group, named C-type lectin-like domains (CTLD). The definition of C-type lectin is also expanded to be a protein with at least one CTLD on its sequence (Dodd and Drickamer, 2001, Weis et al., 1998, Drickamer and Dodd, 1999). The expanded definition of C-type lectins is based on similarity of sequence rather than functions, and subsequently at least 1000 theoretical members of this lectin family have been identified via computational approach (Zelensky and Gready, 2005). In the identified members, there is a population of the members which

are unable to interact with glycans because they lack the glycan binding components in their CTLD. Instead, they are interactive with other biopolymers including lipids and proteins as well as inorganic molecules (Sancho and Reis e Sousa, 2012). There are two types of sequence motifs in the functional CRD including three amino acid residues which confer divergent specificity for their binding glycosides. The mannose/fucose specificity occurs in C-type CRDs with an E-P-N motif presented in their binding site, whereas the Q-P-D motif confers specificity for galactose (Drickamer, 1992). C-type lectins can be expressed in secretory or membrane-bound forms in mammalian systems, and they have a broad binding capacity to endo- and exogenous ligands to manipulate multiple biological functions. In immune systems, secretary C-type lectin can mainly interact with exogenous ligands to induce innate immune cell functioning as phagocytosis, inflammation and antigen presentation. Simultaneously, membrane-bound C-type lectins can modulate immune responses once they interact with endogenous ligands. In this section, the endogenous ligands that induced immune responses are mainly discussed for their correlation to the studies in this thesis.

Lymphocytes trafficking in the vascular system is a well-studied example for the engagements of glycan-lectin interactions of C-type lectins. There are three members of this C-type lectin family responsible for the trafficking, P-, L- and E-selectin. The capital letter for each of the selectin refers to the cell type of original discovery, platelets, leukocytes, and endothelial cells (Cummings and McEver, 2015). They share a similar sequence arrangement with a C-type CRD domain, an epidermal growth factor-like domain, several repeats of sushi domains, a transmembrane domain and a cytoplasmic tail from the sequence N-terminal to C-terminal, successively (McEver et al., 1995). The lymphocyte trafficking is directly initiated and manipulated by selectin-ligand interactions between vascular endothelial cells and platelets/leukocytes. For example, Leukocytes in vessels can be tethered by endothelial cells via selectin-ligand interactions and subsequent rapid-association/disassociation of selectin bonding allows Leukocytes to roll on the vessel wall endothelial cells. The contacts during the rolling allows the chemokine receptors on leukocytes to capture chemokine ligands on the contacting endothelial cells. The ligation of these two inter-membrane molecules induces the expression of integrins on the cell membrane of leukocytes. The newly synthesised integrins favourably bind to endothelial membrane protein ligands such as vascular cell adhesion molecule 1 (VCAM-1). The enhanced inter-membrane association reduces the rolling velocity of leukocytes and eventually leads to their exit from the vascular system.

L-selectin which is expressed on leukocytes whilst of P- and E-selectins are expressed on endothelial cells, not constitutive but inducible by inflammatory mediators (Lowe, 2002, Collins et al., 1995, Eppihimer et al., 1997). The differently expressed selectins play distinct roles in manipulating lymphocytes trafficking. For instance, P- and E-selectins expressed on endothelial cells manipulate the trafficking of lymphocytes from lymph nodes to damaged sites by interacting with ligands on leukocytes. L-selectins on lymphocytes mediate trafficking back into lymph nodes via ligation with glycoconjugate

ligands on endothelial cells, such as those found on high endothelial venules (HEV).

The glycan engagements of selectin-ligand interaction were described in early studies that indicated that destroying peptide could not inhibit interactions but applying exogenous saccharides could reduce the binding affinity between HEV and lymphocytes (Stoolman and Rosen, 1983, Stamper and Woodruff, 1977, Stoolman et al., 1984). The requirement of sialic acid residues for the binding was then elucidated in vitro and in vivo by applying sialidases, and the selectin-ligand interaction was disabled after such treatments (Rosen et al., 1985, Rosen et al., 1989). A LacNAc based tetra-saccharide epitope, Sialyl-Lewis x, with mono-sialylation and fucosylation was then characterised as a universal ligand for these three selectins (Berg et al., 1991, Polley et al., 1991). Sulfated glycan ligands, such as Sulfatides and sulphated Lewis antigens, were then found also to have affinity to L- and P-selectins (Aruffo et al., 1991, Green et al., 1992)

The major glycoconjugate ligands expressed on endothelial cells and lymphocytes have been characterised, and P-selectin glycoprotein ligand-1 (PSGL-1) was identified as a common ligand for all three selectins. Besides, each selectin has their own specific ligands expressed on lymphocytes or endothelial cells (Cummings and McEver, 2015).

E-selectin has three specific ligands identified as human L-selectin, CD44 and E-selectin ligand-1 (ESL-1). ESL-1 was firstly isolated from mouse neutrophils as a ligand for E-selectin by affinity isolation against a recombinant antibody-like E-selectin (Levinovitz et al., 1993). A further study treated ESL-1 with N-glycosidase F and then its E-selectin binding was abolished indicating a substantial engagement of N-glycans in this selectin-ligand binding (Lenter et al., 1994). Sialic acid was also required for the binding which is under a calcium dependent manner. L-selectin purified from human neutrophils was found having affinity to recombinant Ig-like E-selectin. Sialidase treatment but not endoglycosidase F treatment of purified L-selectin can abolish its binding with the recombinant E-selectin (Zöllner et al., 1997). This outcome indicates the importance of ligand sialylation for E-selectin binding and that N-glycosylation might not be essential for L- and E-selectin interactions. In addition, CD44 isolated from specific cells such as peripheral blood polymorphonuclear neutrophils were able to bind with E-selectin. However, CD44 from a bone marrow stromal cell line could not bind with E-selectin. This highlights the importance of post-translational modification for the binding affinity of CD44 to E-selectin. Sialylated and fucosylated N-glycans were also suggested in this study to play a vital role in binding mediation (Katayama et al., 2005).

PSGL-1 is the major ligand for P-selectin, and CD24 also has affinity to P-selectin. PSGL-1 is characterised as a mucin protein which is homo-dimerised by a disulfate bond with a sequence consisting of an N-terminal selectin-binding domain, several decametric repeats, a transmembrane domain and a short cytoplasmic tail (Cummings and McEver, 2015, Sako et al., 1993). In an early study, the engage-

ment of N-glycans for its P-selectin binding was excluded as peptide N-glycoside F (PNGase F) was unable to affect the binding, but instead removing sialic acid inhibited its binding against P-selectin (Moore et al., 1992). Even sialyl-Lewis x antigen was identified on its O-glycans with a substantial abundance, a site specific O-glycosylation on T57 at the N-terminal region of human PSGL-1 was shown to play a vital role for its binding to P-selelctin (Norgard et al., 1993, Liu et al., 1998). The heavy O-glycosylation on the decametric region of PSGL-1 might function to prevent coiling of the sequence to maintain the functionality of the N-terminal binding domain (Sperandio et al., 2009). High binding affinity between PSGL-1 and selectins requires sulphation on the tyrosine residues around the N-terminal of PSGL-1 (Wilkins et al., 1995). Besides, CD24 was also shown to specifically bind with P-selectin via an O-glycan dependent approach as endoglycosidase F treatment could not compromise the binding activities on neutrophil CD24 coated beads (Aigner et al., 1997).

L-selectin has multiple ligands such as membrane mucins which are widely distributed on HEV to mediate L-selectin-dependent lymphocyte homing (Cummings and McEver, 2015), including Glyc-CAM-1, CD34 and MAdCAM-1. GlyCAM-1 is a dominant L-selectin ligand in HEV and it was originally found as a sulphated glycoprotein (Imai et al., 1991). A following study indicated the importance of O-glycan sulphation for its L-selectin binding (Imai et al., 1993). Subsequently, three sulphated Sialyl-Lewis antigens were identified on its core-2 O-glycans as shown in Figure 1.28 (Hemmerich et al., 1994, Hemmerich et al., 1995, Hemmerich and Rosen, 1994). However, C2GnT I deficient mice, which are unable to synthesise core-2 O-glycans, did not exhibit declined lymphocyte homing but limited neutrophil recruitment to respond to inflammation (Ellies et al., 1998). Another study then indicated a compensation in C2GnT I deficient mice which express novel extended Core-1 O-glycans with sialylation and sulphation on HEV mucin ligands (Yeh et al., 2001). CD34, MAdCAM-1, and podocalyxin were then sequentially verified as L-selectin ligands, and they are identical to GlyCAM-1 as possessing O-glycan ligands of L-selectin with sulphation and sialylation (Berg et al., 1993, Imai et al., 1991, Baumheter et al., 1993).

DC-SIGN is a multi-functional C-type lectin which can facilitate the functionality of DC as it is a
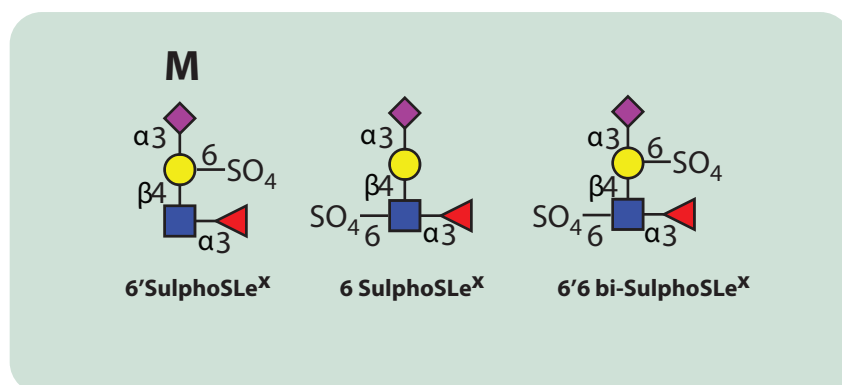
Figure 1.28 Three high affinity sulphated Sialyl-Lewis antigen ligands for L-selectins.
The major ligand is highlighted by a capital "M" above the designated epitope.

"DC-specific C-type lectin". DC-SIGN was firstly described for its reactivity to high mannose glycans on gp120, an HIV envelope glycoprotein (Curtis et al., 1992). The binding activity of DC-SIGN was then discovered in that it was capable of binding intracellular adhesion molecule 3 (ICAM-3) presented on resting T-cells (Geijtenbeek et al., 2000). Similar to selectins facilitating lymphocyte trafficking, DC-SIGN was shown to facilitate the trafficking of DCs to inflammation sites for capturing antigens and then secondary lymphoid organs for T-cell antigen-presentation. In the tethering and rolling of DCs on endothelial cells, DC-SIGN serves as the lectin to manipulate the processes via ligation with Lewis antigens, which are expressed on an endothelial membrane glycoprotein, named intracellular adhesion molecule 2 (ICAM-2) (Garcia-Vallejo et al., 2008). Once DC cells traffic back to lymph nodes, the antigen-presentation is commenced with a DC-SIGN mediated adhesion between DCs and T-cells. The Lewis antigens on ICAM-3 (Bogoevska et al., 2007) were indicated having high binding affinity for DC-SIGN to initiate the transient adhesion. This mediated close contact is beneficial for the formation of immune synapses and activation of resting T-cells (Steinman, 2000).

Antigen presentation by DC is vital for inducing T-cell directed antigen-specific immune responses. However, immature DCs back from inflammation sites are unable to induce the primary response of naïve T cells and their subsequent differentiation to effector T-cells. Instead, regulatory T cells can be induced, via the signals from immature DCs. Regulatory T cells can induce negative regulation of T-cell responses (Jonuleit et al., 2000). Neutrophils may have a role in the maturation of DC prior to its antigen presentation, and DC-SIGN plays a role in this process (van Gisbergen et al., 2005). DC could interact with neutrophils at the site of inflammation, and neutrophils expressed an integrin, named Mac-1, on which Lewis antigens were expressed. Immature DC could cluster with activated neutrophils via the ligation between Mac-1 and DC-SIGN. The activated neutrophils could produce TNF-α to induce DC maturation after this ligation (van Gisbergen et al., 2005) . Moreover, DC-SIGN was suggested to be unable to interact with Mac-1 expressed on the other leukocytes. This might reflect a differentiated glycosylation for this integrin in different cell lineages (van Gisbergen et al., 2005).

1.6.5 Antibody glycosylation

Antibodies produced by B cells are naturally glycosylated on the constant region of immunoglobulin heavy chains. The presence of glycans on the constant region enhances the structural heterogeneity of the antibody molecules and more importantly regulate their immunological activities. The regulation of antibody activities is exemplified by the N-glycosylation on IgG. IgG has a conserved N-glycosylation site on its constant region at N297 and deglycosylation of this site could abolish all effector functions of IgG via interactions with multiple receptors (Anthony et al., 2012). The levels of core-fucosylation, galactosylation and sialylation on IgG N-glycans have been well recognised for associating with its effector functions. The galactosylation of IgG from patient with autoimmune disease, such as rheumatoid arthritis, has been shown to be lower than healthy counterparts to exhibit truncated N-

glycans with GlcNAc terminals (Parekh et al., 1985). The level of galactosylation could increase on IgG in patients with rheumatoid arthritis during pregnancy or medical treatment with correspondingly reduced disease activity (van de Geijn et al., 2009, Pasek et al., 2006). Besides, the sialylation level on galactosylated N-glycan antennae has been associated with immunomodulatory effects and disease condition. A low level of sialylation was observed in various diseases such as rheumatoid arthritis and disease relapse could occur in patient with reduced sialylation (Kemna et al., 2017, Parekh et al., 1985). A high level of IgG sialylation can favourably induce interactions with immunomodulatory receptors (Anthony et al., 2011), especially α2,6-linked sialic acid residues, while impairing interactions with pro-inflammatory Fc receptors and complement cascade (Quast et al., 2015, Scallon et al., 2007). The presence of core-fucosylation on the reducing-end GlcNAc residue can reduce the ADCC activities via reducing their binding affinity to relevant receptors (Ferrara et al., 2011, Quast et al., 2015). Additionally, extra N-glycosylation sites could be found in the variable region of heavy and light chains on nearly 20% of serum IgG due to hypermutation in this region (Mimura et al., 2007, Zhu et al., 2002). The potential N-glycosylation in variable regions could assemble high mannose or complex glycans with heavy sialylation (Bondt et al., 2014) and the level of sialylation was associated with the half-life of the IgG molecules. The N-glycosylation of the variable region has impacts on antigen binding affinity (Coloma et al., 1999, Leibiger et al., 1999). However, the alteration of variable region glycosylation was not associated with the severity of arthritis of patients during pregnancy (Bondt et al., 2016).

1.6.6 Hu-FEDS hypothesis

During pregnancy, the maternal immune system is manipulated to accommodate the foetus as a genetic hemi-allograft, but the exact mechanism for this regulation is still largely unknown. In 1950s, Nobel laureate Sir Peter Medawar firstly proposed three hypotheses to explain the exceptional immune tolerance: 1) anatomical separation; 2) immaturity of foetal antigens or 3) inertness of maternal immune responses (Medawar, 1953). However, subsequent studies were not in support to all of these three hypotheses. First, the extra-villous cytotrophoblasts of foetal placental origin have direct contact with maternal uterus to remodel local vascular system to establish supplies of blood and nutrients from the maternal system for the foetus (Burton and Jauniaux, 2004). Second, injection of foetal materials was able to induce immune responses in animal models with mature immune systems (Billingham et al., 1956). Third, maternal system could induce robust immune responses to exogenous infections during pregnancy (Head and Billingham, 1986). More theories were then proposed to explain the possible mechanism of this immunomodulation, and the importance of glycosylation in the local environment during pregnancy is highlighted in one of these theories, the Hu-FEDS hypothesis.

The Hu-FEDS hypothesis stands for human fetoembryonic defence system hypothesis, which was proposed by Gary F. Clark in 1996 (Clark et al., 1996). In this hypothesis, Clark firstly suggested that glycan sequences on glycoconjugates in the reproductive system can serve as key molecules to ma-

nipulate immune recognition and subsequent immune responses (Clark et al., 1996). The hypothesis was initiated on the observation that the binding of human egg-sperm depends on the recognition of a glycan sequence, Sialyl-Lewis X. Their immune recognition on human eggs was also suggested to rely on this glycan sequence expressed on zona pellucida glycoproteins (Clark et al., 1995, Patankar et al., 1993).

There is accumulating evidence to support the immune-suppressive effects of glycan sequences from the studies of glycoconjugates ranging from glycoproteins to glycolipids in the reproductive system (Pang et al., 2016). The potential association of glycan sequences and the immunosuppressive effects of glycoconjugates in the human reproductive systems are exemplified below.

Glycodelin is a glycoprotein which is mainly expressed in the endometrium (Julkunen et al., 1986), but can also be detected in the male genital tract and immune cells of megakaryocytic lineage (Julkunen et al., 1984, Morrow et al., 1994). In female the expression of this protein commences from the luteal phase to the term of pregnancy (Dalton et al., 1995, Julkunen et al., 1985). Its expression peaks in the first trimester (7-11wk) of pregnancy to become a major protein observed in the endometrium (Julkunen et al., 1991). This protein is mainly secreted into amniotic fluids and the presence of this protein is dramatically declined after the first trimester of pregnancy (Julkunen et al., 1991, Julkunen et al., 1985). The immunosuppressive effect of glycodelin was firstly described in a study of decidual tissue extracts (Bolton et al., 1987). The tissue extracts were found to be able to suppress the proliferation of mixed lymphocytes and glycodelin purified from the extracts had a more potent effect. This effect can be abolished by applying antibodies against glycodelin and thereby this protein was suggested to be the major immunosuppressive factor in the decidual environment (Bolton et al., 1987). The glycodelin secreted in amniotic fluids, also named glycodelin A, was found to be able to inhibit the expression of IL-1, -2 and IL-2 receptor (soluble) in monocytes (Pockley and Bolton, 1989, Pockley and Bolton, 1990). Moreover, the cytotoxic activity of NK cells can also be inhibited by glycodelin A in a concentration dependent manner (Okamoto et al., 1991). Moreover, glycodelin A was also shown to downregulate expression of IgM and MHC II of B cells as well as inhibit the E-selectin mediated adhesion of neutrophils to vascular endothelium (Yaniv et al., 2003, Jeschke et al., 2003). All the evidence above indicates the importance of this glycoprotein in mediating the immune responses during the development of the local vascular system in the first trimester. Glycodelin A is the major glycoform, glycodelin S (seminal plasma), F (follicular fluid), C (cumulus matrix) and M (megakaryocyte) are also identified as glycoforms of glycodelin from different sources. Glycomic studies of the glycofoms of glycodelin, excluding glycodelin M, have shown the association of glycosylation on this glycoprotein to its immunosuppressive effects (Koistinen et al., 1996, Dell et al., 1995, Lee et al., 2009). Glycodelin A, F and C from the female reproductive system shared several structural features of their N-glycans in terms of expressing multi-antennary complex structures with type II LacNAc and Lac-diNAc antennae as well as Sda epitope (Lee et al., 2009). Antennary sialylation and fucosylation levels

were varied in between these glycoforms as glycodelin A was heavily sialylated whereas the other two glycoforms were less sialylated. Sialylation on glycodelin C was only found on Sda epitopes (Lee et al., 2009). The immunosuppressive effect of these glycodelins was also assessed on immune cells such as Jurkat cell lines. Inhibition of proliferation and IL-2 production as well as induction of cell death were observed on these cell lines with treatment of glycodelin A and F, but not glycodelin S and C (Lee et al., 2009). The glycosylation of glycodelin S had been characterised and only two N-glycosylation sites on glycodelin S are specifically occupied by high mannose glycans on site Asn-28 while the other site Asn-63 is occupied by bi-antennary complex N-glycans with LewisX/Y terminals (Morris et al., 1996). Additionally, glycodelin A from patients with gestational diabetes mellitus at the term of pregnancy exhibited higher presence of Sda epitopes and a decrease of α 2,6 sialylation (Lee et al., 2011). The described immunosuppressive effect correspondingly decreased in diabetic patients and could further be completely abolished after de-sialylation. The evidence above indicates that the antennary arrangements of their N-glycans, especially sialylation on their LacNAc or LacdiNAc antennae, could directly correlate to their immunosuppression effects (Lee et al., 2011).

CA125, cancer antigen 125, is an endometrial mucin, which was named for its initial detection by monoclonal antibodies for ovarian carcinoma cells (Bast Jr et al., 1981). This mucin has 24,000 amino acid residues and can accommodate a large volume of N- and O-glycans (Kui Wong et al., 2003, O'Brien et al., 2002) which can contribute equivalent molecular weight to its protein backbone. Its expression pattern in endometrium is coincident to that of glycodelin A during pregnancy (Dalton et al., 1995). N-glycosylation of CA125 characterised from OVCAR-3 cell lines is substantial as having high mannose and multi-antennary bisecting complex structures with antennary sialylation and/or fucosylation. O-glycosylation for this protein is dominated by core-1 and -2 structures (Kui Wong et al., 2003). The cytotoxic activities of NK cells could be suppressed by 50-70% via incubation with CA125 for at least 72h at a concentration between 10,000-100,000 U/ml (Patankar et al., 2005). The exposure of NK cells to CA125 can also lead to reduced expression of membrane protein CD16 and could consequently shift NK cells to a less potent subset as decidual NK cells (Patankar et al., 2005). Interestingly, the suppressive effect of CA125 for NK cells could be completely abolished by de-sialylation of CA125 as this mucin protein recruited siglec-9 to bind NK cells and attenuate their functional signalling, simultaneously (Belisle et al., 2010).

Sialylated glycolipids as gangliosides are dominated in the placenta in terms of abundance and immunosuppressive effect. GM3, GM1, GD3 and GD1a are abundant gangliosides in placenta (Dyatlovitskaya et al., 1990), and they could effectively suppress the lymphoblastic transformation and NK activities. Siglec-5, -7 and -8 can bind with glycan ligand on GD3, and the cytotoxic activity of NK cells can be mediated by the ligation between GD3 and siglec-7 (Rapoport et al., 2003, Nicoll et al., 1999).

1.6.7 Cancer glycosylation

The altered glycosylation of cancerous cells has been described since the 1950s (Ladenson et al., 1949, Nisselbaum and Bernfeld, 1956). Following developments of highly specific monoclonal antibody technology allowed structural characterisation of the cancer associated glycan epitopes which are defined as generally foetal-like oncofoetal antigens (Hakomori, 1983, Holmes et al., 1987). The glycan epitopes characterised by monoclonal antibodies are widely accepted to be named as tumour associated carbohydrate antigens (Hakomori, 1984).

The inherent heterogeneity of original cells determines that the alterations of glycosylation on cancerous cells could be on a wide range of structural features following a cell-specific, protein-specific and even site-specific manner. Two principal mechanisms, imcomplete and neo-syntheses, have been proposed for cell glycosylation during the transformation from original to cancerous cells based on the integrated effects of involved intrinsic factors (Kannagi and Hakomori, 1983). The incomplete synthesis can occur at the early stage of cancer transformation as the normal glycosylation machinery is disrupted to produce incomplete structures such as truncated Tn-antigen (Fu et al., 2016). The neo-synthesis refers to that of cancerous cells which decorate their glycans with featured moieties such as Sialyl-Lewis antigens and poly-LacNAc at advanced stages to favour their growth and malignancy (Walz et al., 1990, Kannagi et al., 2008, Sweeney et al., 2018).

As the biosynthesis of glycans is not based on templates, intrinsic factors play a role in directing the proposed mechanism to alter glycosylation in the transformation. There are three major factors that have impacts on the glycosylation during cancer transformation: 1) expression of glycosyltransferases; 2) availability of monosaccharide substrates; 3) localisation of catalysing components in the ER-Golgi pathway (Pinho and Reis, 2015). These factors can singly or synergistically drive the glycosylation machinery to produce the desired structural features during cancer developments. Thereby, there are five structural features widely observed on cancer glycans, fucosylation, sialylation, truncation, branching and elongation.

Sialylation is globally upregulated on N-, O- and GSL-glycans in cancers with mostly α2,3 or 2,6 linkages because of enhanced expression of relevant sialylatransferases (Kim and Varki, 1997). Terminal LacNAc moieties are generally decorated by a sialic acid residue on its Gal with either linkage (α 2,3 and 2,6) to form sialyl-LacNAc epitopes. Overexpression of Sialylatransferases for the assembly of both linkages (ST3Gal-I &ST6Gal-I) is correlated to the development and poor prognosis of various cancers (Shen et al., 2017, Wu et al., 2016, Lu and Gu, 2015). Sialyl-LacNAc epitopes can further serve as an acceptor for a Fuc residue to assemble on its GlcNAc to form either a sialyl-Lewis x or a antigens. Overexpression of sialyl-Lewis x antigens can result in a poor survival rate (Baldus et al., 1998), and this antigen and its sulphated forms can interact with selectins to facilitate the migration and tissue

invasion of cancer cells to distal locations (Nakamori et al., 1993). In addition, the increased sialylation on GSL-glycans can result in an overexpression of gangliosides; and poly-sialic acid with α 2,8 linkage can also be observed on cancer glycans (Todeschini et al., 2007, Falconer et al., 2012).

Both core and branch types of fucosylation can be found on glycans from cancerous cells. The branch fucosylation is catalysed by a group of fucosyltransferases (Fuc-T III-XI, excluding Fuc-T VIII) to assemble Lewis a/x and b/y antigens on the N-, O- and GSL- glycans of cancer cells (Miyoshi et al., 2008). Fuc-T III has dual linkage specificities to assemble either α1,3 or 1,4-linked fucose residue onto glycans. Core-fucosylation by Fuc-T VIII is normally overexpressed in cancers to add an α1,6-linked fucose residue onto the reducing-end GlcNAc on N-glycans. Overexpressing core-fucose on the glycans of epidermal growth factor receptor on cancer cells can result in enhanced signalling of this growth factor via elevating levels of dimerization and phosphorylation to facilitate the growth and malignancy of tumour cells (Takahashi et al., 2009).

The N-glycans on cancers can be generally well-branched to form multi-antennary structures. Hyperactivity of GlcNAc-T V can result in formation of a β1,6 antenna and increased formation of this antenna has been associated with enhanced metastasis (Dennis et al., 1987). Overexpression of this enzyme has been associated with cancer progression, while downregulating the expression of this enzyme could lead to suppression of the progress and metastasis (Demetriou et al., 1995). Elongation of cancer N-glycans by β1,3 GnTs to form linear poly-LacNAc is also dominant in several cancers and this structural moiety can facilitate the interaction of cancer cells to galectins promoting metastasis and tumour survivals (Rye et al., 1998, Jiang et al., 2018). The elongated poly-LacNAc chains could also facilitate to project away their terminal epitopes into the extracellular environments and to possibly enhance trans-cellular interactions via lectins (Pang et al., 2016).

Different from the neo-synthesis described above, incomplete synthesis in cancer cells is exemplified by the formation of truncated O-glycans. The most dominant truncated structures are T and Tn antigens with their sialylated forms. The formation of truncated structures is normally caused by the competition of glycosyltransferases for the same catalysing acceptors. For instance, the competition between C2GnT1 and ST3Gal-I for core 1 O-glycans associated with the progression of breast cancer could determine the O-glycan structures to be either Core-2 or sialylated T-antigens (Dalziel et al., 2001). Overexpressing of GalNAc α2,6 sialylatransferase (ST6GalNAc-I) in cancers with dysregulated enrichment in cis- and medial-golgi could result in a competition with core 1 GalNAc β1,3 galacosyltransferase 1 (C1GalT1) for the C-3 position to form the truncated sialyl-Tn antigen (Gill et al., 2010, Marcos et al., 2004).

Cancer cells share, to some extent, glycan moieties with those of the reproductive systems as oncofoetal antigens. These special antigens might confer immunological advantages for cancer cells as suggested in the Hu-FED's hypothesis. For instance, Lewis y antigen expressed as an oncofoetal antigen

in cancer cells could have immunomodulatory effect, as the presence of this antigen on lipopolysaccharide can induce IL-10 production from DC to favour immune tolerance via DC-SIGN signalling (Gringhuis et al., 2009). Sialyl-Lewis x is another example as it is not only an adhesion ligand for selectins to favour metastasis, but also a ligand for siglec-9 to possibly induce tolerance on innate immune cells (Avril et al., 2004). Sialyl-Lewis x antigens could also inhibit proliferation of T cells after antigen presentation and induction (Steinman, 2000).

1.7 Aim of this thesis

In this thesis I aim to characterise the glycosylation of several immunologically important species or proteins via comprehensive mass spectrometric strategies.

In each chapter, I specifically aim to:

1.	Characterise the glycosylation of pregnancy specific glycoprotein 1 from natural and recombinant sources via glycomics and glycoproteomics analyses (Chapter 3)

2.	Characterise the glycosylation of canine melanoma cells from a stage III patient by comprehensive glycomic strategies (Chapter 4)

3.	Characterise the glycosylation of fusion Fc proteins from different expression systems by glycomic analysis (Chapter 5)

4.	Characterise the glycosylation of murine mixed cryoglobulins by glycomic analysis to understand the pathogenetic mechanism of relevant diseases (Chapter 6).

Chapter II. Methods and Materials

Chapter 2 Methods and Materials

2.1 Biological materials

All biological samples were provided by collaborators for the projects included in the following chapters. The methodology of sample preparations prior to our experiments were kindly provided by them for the following sub-Sections.

2.1.1 Pregnancy specific glycoprotein 1

2.1.1.1 Native pregnancy specific glycoprotein 1

Pregnancy specific glycoprotein 1 samples were provided by Professor Gabriela Dveksler of the Department of Pathology, Uniformed Service University, Maryland, USA. The native proteins were harvest from healthy pregnant donors. Briefly, serum of healthy pregnant women was collected and pooled under strict regulation. Then the pooled serum material was passed through a self-assembled affinity column fixed with commercially available PSG1 antibodies.

2.1.1.2 Recombinant pregnancy specific glycoprotein 1 N-terminal domains with Fc-tag

PSG1 N-domains were transfected into CHO-K1 (transiently and stably) and Epxi-CHO (transiently) with a vector containing a Fc-tag, which introduced the Fc-tag at the C-terminal of PSG1 N-domain sequence. All post-transfection cell cultures were collected and purified on a Protein A column using an AKTA system.

2.1.2 Canine melanoma cells

Canine melanoma cells were kindly provided by Professor Gary Clark of the department of Obstetrics, Gynecology and Women's Health, School of Medicine, University of Missouri, Columbia, USA. The canine melanoma cells were collected from a dog which had been diagnosed with stage 3c melanoma cancer. Briefly, biopsies of the melanoma were surgically removed, and melanoma cells were then manually isolated from the biopsies with strict regulation and professional guidelines. The isolated melanoma cells were cultured and expanded in the laboratory to reach an optimal cell population. The expanded cell population was harvested from the cell culture and then lyophilised. A portion of the cells were provided for our glycomics analysis. The remainder were used for the vaccination regimen which is being piloted at the University of Missouri (Clark, 2017, Dell et al., 2018). Imaging revealed that this dog had developed a total of thirty metastatic tumours after two rounds of vaccination. However, one month after vaccination ended, imaging indicated that all tumours had been destroyed and

the dog remained free of metastatic disease for 20 months before dying due to an unrelated trauma.

2.1.3 Immunoglobulin G

2.1.3.1 Cryo-precipitate from transgenic TSLP mice

Cryo-precipitate and cryoglobulin samples were provided by Professor Alan Salama of the Centre of Nephrology, Royal Free Hospital, University College London, London, UK. A thymic stromal lymphopoietin(TSLP) overexpression model mouse was employed for production. Mice were sacrificed to collect sufficient amounts of murine serum; the collected murine sera were incubated in a 4 °C cold room for at least 24 hours to form a cryo-precipitate. The precipitates were then collected.

2.1.3.2 Cryoglobulin G from transgenic TSLP mice

The same TSLP overexpression model mouse was employed, and mice were sacrificed for serum collection. Collected murine serum was purified on a Protein G column following the manufacturer's instructions. A wild-type control of immunoglobulins, obtained via the same procedure, was also kindly provided.

2.1.3.3 Recombinant Fc fusion protein

Engineered IgG1-Fc fusion protein was kindly provided by Professor Richard Pleass of Liverpool School of Tropical Medicine, Liverpool, UK. The detailed preparation was described in (Blundell et al., 2019). Briefly, the sequence of Fc region was rearranged by addition of a tail-piece from the IgM C-terminus. Introduction or deletion of N-glycosylation sites were selectively conducted on the hinge region and/or CH2 domain and/or tail-piece. These mutants were further grouped by the presence of cysteine residues on the CH2 domain and/or engineered tail-piece via mutagenesis. All mutants were stably expressed in CHO and HEK293T cell systems. The supernatant of cell cultures was harvested from both cell lines after centrifugation. The supernatant was then purified by Protein G column.

2.2 Materials

2.2.1 Chemicals, reagents, enzymes and consumables

Alfa Aesar (Lanchashire, UK): Acetic anhydride and methyl iodide
Applied Biosystems (Warrington, UK): 4700 Mass Standards kit
BOC (Guildford, UK): Argon, helium and nitrogen compressed gases
Calbiochem (CA, USA): Endo-β-galactosidase (B. fragilis)
Fisher Scientific (Loughborough, UK): Slide-A-Lyzer dialysis cassettes (3.5KDa, MWCO)
Fluka (Poole, UK): DOWEX 50W-X8 (H) 50-100 mesh resin and potassium hydroxide
Promega Corporation (Hampshire, UK): Sequencing grade modified trypsin(porcine)

Purite Ltd (Oxfordshire, UK): distilled/deionised water

New England Biolabs (Hitchin, UK): Neuraminidase A (Sialidase A, E. coli), Neuraminidase S (Sialidase S, E. coli), rapid Peptide-N-glycosidase F (rapid PNGase F)

QA bio (CA, USA): Endo- β -galactosidase (B. fragilis)

Roche (East Sussex, UK): Peptide-N-glycosidase F(PNGase F, E. coli)

ROMIL (Waterbeach, UK): Acetonitrile (ACN), ammonia, acetic acid, butan-1-ol, methanol (MeOH), chloroform, dimethylsulfoxide (DMSO), formic acid, propan-1-ol (PrOH), sodium hydroxide (NaOH) and trifluoroacetic acid.

Rose Chemicals (London, UK): Sodium chloride (NaCl)

Sigma-Aldrich Corporation (Poole, UK): Ammonium bicarbonate (AMBIC), dithiothreitol (DTT), hexanes, iodoacetic acid (IAA), potassium borohydride, Pur-A-Lyzer Mega Dialysis Kit (3.5 KDa, MWCO), sodium acetate, sodium cholate, sodium borodeuteride, trypsin (porcine pancreas), Endo- β -galactosidase (B. fragilis)

Takara Bio (Saint-Germain-en-Laye, France): Recombinant Endo-glycoceramidase II (rEGCase II, E. coli)

Waters (Herefordshire, UK): Sep-pak Classic C18, Oasis HLB and tC18 (Plus) cartridges

2.2.2 Instruments

API QSTAR Hybrid LC-MS/MS system online connected to a Dionex UltiMate 3000 liquid chromatography (LC) adapted with a Pepmap C18 nano-capillary column (75µm x 150mm)

Applied Biosystems 4800 plus MALDI-TOF/TOF mass spectrometer

Bruker 456-GC/SCION SQ instrument adapted with a BR-5ms column (15m, 0.25mmID, 0.25µm: 5% phenyl 95% dimethyl arylene siloxane)

Waters SYNAPT G2-S mass spectrometer equipped with a nano-flow ESI source on-line connected to a Waters ACQUITY UPLC M-Class system with an ACQUITY UPLC column (HSS T3 1.8µm 75µm x 150mm)

2.2.3 Software

The MALDI TOF and TOF/TOF data were interpreted by using Data Explorer software (version 4.9)

The Q-Star data was interpreted by using an Analyst TF software (version 1.6)

The SYNAPT data was visualised by a MassLynx software (version 4.1)

The GC-MS data was interpreted by using an MS data review software (version 8.0.1)

The Mascot search was performed by ProteinScape (version 3.1) software

The GlycoWorkBench software (version 2.1) was used for annotating glycan structures with standardised cartoons.

2.3 Methods

2.3.1.1 Homogenisation of mice tissues

Murine lung tissues were homogenised with 5ml lysis buffer (25mM Tris, 150mM NaCl, 5mM EDTA and 1% (w/v) CHAPS, pH 7.4). Homogenised sample was transferred into a 3.5KDa MWCO dialysis kit/cassette with a subsequent 48 hr dialysis against regularly changed AMBIC buffers (50mM, pH 7.5). Homogenates were harvested into a clean Falcon Tube and lyophilised.

2.3.1.2 Homogenisation of melanoma cells with lipids extraction

Formerly lyophilised melanoma cells (assumed 0% water content) were homogenised with 4ml ice-cold water determined as the total aqueous volume of homogenate. 1.33 volumes of chloroform and 2.67 volumes of methanol were added into the homogenate to extract glycolipids. The homogenate was well-mixed with these solvents by vortex before a 3000-rpm centrifugation for 10 min. The supernatant was collected and its volume measured. 0.173 volumes of water were added into the supernatant and then vortexed thoroughly. The mixed solution was centrifuged at 3000-rpm for 15 min to separate into two phases: upper polar and lower non-polar phases. These two phases were then subjected to lipid recovery as in **Section 2.3.1.3 &2.3.1.4**. Excess organic solvents were removed by briefly drying under nitrogen and the remaining protein pellet was stored at –80 °C before following procedure 2.3.2.1.

2.3.1.3 Polar lipid recovery

Polar lipid containing samples from **Section 2.3.1.2** were directly loaded onto a Sep-pak tC18 Plus Cartridge pre-conditioned with 5ml of each of methanol, 50% (v/v) methanol, 50% (v/v) methanol/chlorofom, and 15ml of 50% (v/v) methanol. The samples were washed with 15ml of 50% (v/v) methanol prior to successive elution with 5ml of each of methanol and 50% (v/v) methanol/chloroform. The two fractions were combined and dried down under nitrogen.

2.3.1.4 Non-polar lipid recovery

15ml of each of chloroform and water were added into non-polar lipid containing samples from **Section 2.3.1.2.** The mixture was vortexed vigorously and subsequently centrifuged at 3500 rpm for 15min. The aqueous layer was removed, and the water extraction was repeated for the remaining chloroform layer. The chloroform was then dried down under nitrogen.

2.3.2.1 Reduction and Carboxymethylation of cell and mice tissue homogenates

Samples from **Section 2.3.1.1 &2.3.1.2** were dissolved in 1ml of 2mg/ml Dithiothreitol (DTT) in degassed Tris buffer (0.6M, pH 8.5) and incubated at 500C for 2 hr. After addition of 1ml of 12mg/ml

Iodoacetic acid (IAA) in degassed Tris buffer (0.6M, pH 8.5) incubation was continued for 90 min at room temperature in the dark. Reaction was terminated by a 48hr dialysis against AMBIC buffer (50mM, pH 7.5) with regular changes. Dialysed sample was transferred to a clean Falcon tube and lyophilised.

2.3.2.2 Reduction and Carboxymethylation of PSG1 and cryoglobulins

Previously lyophilised samples were re-solubilised into 50 µl of 10mM DTT in degassed Tris Buffer (0.6M, pH 8.5) and then incubated at 37 °C for 60min. 50 µl of 60mM IAA in degassed Tris Buffer (0.6M, pH 8.5) was subsequently added with a further incubation at room temperature in dark environment for 90 min. Reactions were terminated by dialysing as in **Section 2.3.2.1**. Samples were transferred into clean Falcon tubes and lyophilised after dialysis.

2.3.3.1 Tryptic digestion of tissue homogenates

A trypsin solution was made with a final concentration of 1 µg/µl against 50 mM AMBIC buffer (pH 8.4). 1.5 mg of trypsin enzyme was added per gram of tissue with an incubation at 37 °C for 14-16hr. The reaction was terminated by boiling the samples for 3min and adding a drop of acetic acid. Samples were purified by reverse-phase chromatography (Oasis HLB Plus Cartridge) using the propan-1-ol/5% (v/v) acetic acid system.

2.3.3.2 Tryptic digestion of cell homogenate

Trypsin solution was prepared as in **Section 2.3.3.1**. 50 µg of trypsin were applied to per million cell count of initial samples with subsequent incubation at 37 °C for 14-16 hr. This proteolytic digestion was terminated by heating the sample at 100 °C for 3 min following a drop of acetic acid. The resulting peptides were purified by reverse-phase chromatography (Oasis HLB Plus cartridge) employing the propan-1-ol/5% (v/v) acetic acid system.

2.3.3.3 Tryptic digestion of glycoproteins

The trypsin solution was consistent with **Section 2.3.3.1**. This protease was added into samples under a protease: protein ratio of 1:50 followed by an incubation at 37 °C for 14-16 hr. Three drops of acetic acid terminated this reaction and reverse-phase chromatography (Classic C18 cartridge) was employed to purify the proteolysed peptides with the propan-1-ol/5% (v/v) acetic acid system.

2.3.3.4 In solution digestion of glycopeptides for glycoproteomics

A 10µg protein pellet was incubated with 1µl 10mM DTT in 50mM AMBIC buffer (pH 8.4) at 56°C

for 30min. 1μl 55mM IAA in 50mM AMBIC buffer (pH 8.4) was added with a further incubation at room temperature for 30 min in a dark environment. Samples was treated with sequencing grade trypsin (trypsin to sample, 1:50 w/w) and incubated at 37°C for overnight. Proteolytic digestion was terminated by heating the solution to 100°C.

2.3.4 Reverse-phase chromatographic purification (Sep-Pak C18 & Oasis HLB Plus): Propan-1-ol/5% (v/v) acetic acid system

Cartridges were attached to glass syringes and conditioning was sequentially carried out by eluting with 5ml methanol, 5ml 5% (v/v) acetic acid, 5ml propan-1-ol and 15 ml 5% (v/v) acetic acid. Samples were loaded and washed on cartridges with 20ml 5% (v/v) acetic acid (containing potential hydrophilic contamination). Peptides were successively eluted with 4ml of each of 20%, 40% and 100% (v/v) propan-1-ol against 5% (v/v) acetic acid. All fractions were then combined, dried down in a SpeedVac and lyophilised.

2.3.5.1 PNGase F digestion and Sep-pak C18 purification of N-glycans: Propan-1-ol/ 5% (v/v) acetic acid system

Combined propan-1-ol fractions from **Section 2.3.4** were re-solubilised in 200 μl of 50mM AMBIC buffer (pH 8.4). 5-10 U of PNGaseF enzyme was added into the solution and incubated at 37 °C for 12-24 hr. Addition of another 5-10U of enzyme was followed by a further 12-24 hr incubation subject to the assumed quantity of glycopeptides or materials. Lyophilisation was used to terminate this reaction and samples were then resuspended into 5% (v/v) acetic acid prior to loading onto Sep-pak C18 cartridges. N-glycans were eluted within 5ml 5% (v/v) acetic acid following by elution of remaining peptides (including O-linked glycopeptides) with 4ml of 20%, 40% and 100% (v/v) Propan-1-ol in 5% (v/v) acetic acid. The N-glycan containing acetic acid fraction and the combined propan-1-ol fractions were reduced in volume by a SpeedVac and lyophilised.

2.3.5.2 Rapid PNaseF digestion of recombinant IgG Fc and Sep-pak C18 purification of N-glycans: Propan-1-ol/5% (v/v) acetic acid system

50 μg of lyophilised recombinant IgG Fc were re-solubilised into 16 μl of water. 4μl of working buffer (supplied from manufacturer) and 1U of rapid PNGaseF were added into the glycan solutions. The reaction was incubated at 50°C for 1hr and then terminated by lyophilisation. The lyophilised mixtures were re-dissolved in 5% (v/v) acetic acid to pass through a Sep-pak Classic C18 Cartridge as in **Section 2.3.5.1** for purification.

2.3.6 Reductive elimination of O-glycans and Ion-exchange purification

Lyophilised combined propano-1-l fractions from **Section 2.3.5.1** were dissolved in 400µl of 55mg/ml potassium borohydride in 0.1M potassium hydroxide solution and incubated at 45 °C for 18hr. The reaction was quenched with 5 drops of acetic acid and released O-glycans were subsequently purified by the 50W-8C Dowex ion-exchange resin packed in customised columns. The columns were washed by 15ml 5% (v/v) acetic acid and samples were directly loaded into the columns after quenching. O-glycans were eluted with 5ml 5% (v/v) acetic acid and lyophilised. Excess borates were removed by co-evaporation with 2ml 10% (v/v) methanolic acetic acid under a stream of nitrogen.

2.3.7.1 Recombinant endo-glycoceramidase II digestion

Polar/non-polar Glycolipids from **Section 2.3.1.3** &**2.3.1.4** were resuspended into 200 µl to 1ml of 50 mM sodium acetate with 0.2% (w/v) sodium cholate buffer (pH 5.5), subject to the volume of samples. 25mU of reECGase II was initially added into samples with an incubation at 37 °C for 6 hr. Another aliquot of 25mU reECGase II was added into samples with a further incubation at 37 °C for 18 hr. The digestion was diluted by water to 2ml in total volume. 2ml of butan-1-ol were added to extract lipid components with vigorous vortexing. The butan-1-ol layer was removed and the procedure repeated for 3 times. Remaining traces of butan-1-ol were removed by drying under nitrogen for 20 min. The butan-1-ol washed mixtures were passed to a Sep-pak Classic C18 cartridge successively preconditioned with 5ml methanol, 5ml 5% (v/v) acetic acid, 5ml methanol and 15ml 5% (v/v) acetic acid; and the unbound fraction collected. A Hypercarb column sequentially preconditioned with 3 columns of 80% (v/v) acetonitrile in 0.1% (v/v) TFA and water was applied for further purification of the unbound fraction with successive elution of 3 columns of water and 2 columns of 25% (v/v) acetonitrile in 0.05% TFA. The glycan containing acetonitrile fraction was collected, volume reduced by a SpeedVac and lyophilised.

2.3.7.2 Deuteroreduction of lipid-glycans

Lyophilised lipid-glycans from **Section 2.3.7.1** were re-dissolved into 200µl of 10mg/ml sodium boro-deuteride in 2M ammonia solution and incubated at room temperature for 2 hr. The basic mixture was neutralised by adding 5 drops of acetic acid and dried down under a stream of nitrogen. Excess borates were co-evaporated by 10% (v/v) methanolic acetic acid under nitrogen.

2.3.8 Endo- β -galactosidase digestion

An aliquot of lyophilised N-glycans was re-solubilised into 100 µl of sodium acetate (100mM, pH 5.8) buffer. 0.1 U of endo-β-galatosidase (B. fragilis) was added with incubation at 37°C for 24 hr. A second 0.1 U of the enzyme was added with a further 24 hr incubation at 37°C. The digestion was terminated

by lyophilisation, and a Sep-pak C18 cartridge was used to purify glycans as in **Section 2.3.5.1** with the same conditions. Only the glycan-containing 5% (v/v) acetic acid fraction was collected and lyophilised.

2.3.9 Sialidase A digestion

250 µl of 100mM sodium acetate buffer (pH 5.5) was used to dissolve an aliquot of lyophilised N-glycans. 170mU of sialidase A was added, and the reaction was incubated at 37°C for 12hr. A further addition of 170mU of sialidase A was followed by an incubation at 37°C for another 12hr. The digest was then lyophilised and purified by a Sep-pak Classic C18 Cartridge using the same conditions as in **Section 2.3.5.1**. The Glycan-containing fraction was collected and lyophilised.

2.3.10 Sialidase S digestion

Lyophilised N-glycans were dissolved in 250 µl of 100mM sodium acetate buffer (pH 5.5). The digestion was incubated at 37°C with initially 170mU of Sialidase S for 12 hr. Another 170 mU of sialidase S was added with a further incubation at 37°C for 12 hr. The reaction was terminated by lyophilisation, and glycans were purified by a Sep-pak Classic C18 Cartridge employing the same conditions in **Section 2.3.5.1**. The fraction containing glycans was collected and lyophilised.

2.3.11.1 Permethylation

A NaOH-DMSO slurry was made by grinding 5 pellets of NaOH in 3ml of DMSO with a pestle and mortar. 1ml of this slurry was added to lyophilised glycans from **Section 2.3.5.1, 2.3.5.2, 2.3.6. 2.3.7.2, 2.3.8, 2.3.9 and 2.3.10** together with 0.6 ml of methyl iodide. The reactants were mixed vigorously on an automatic vortexer for 30-40 min at room temperature. The reaction was terminated by dropwise quenching with water. Glycans were then extracted into 1ml chloroform, and the glycan-containing chloroform layer was washed with 4 ml of water for 3 times. The chloroform was dried down under a steady flow of nitrogen. The derivatised glycans were purified and fractionated by Sep-pak Classic C18 recruiting aqueous acetonitrile system, Section 2.3.12.

2.3.11.2 Permethylation of sulfated glycans

The same NaOH-DMSO slurry was made, and the same amount of each reagent was added to glycans as in Section 2.3.11.1. The reactants were gently mixed on an automatic vortexer at 4°C for 3hr. The reaction was quenched on ice by a dropwise addition of 5% (v/v) ice-cold acetic acid to reach the final pH value 6. The permethylated non-sulfated glycans were extracted in 1ml of chloroform, and the remaining aqueous solution was directly loaded onto a Sep-pak Classic C18 Cartridge sequentially pre-conditioned with 5ml of each of methanol, water, acetonitrile and 15 ml of water. The glycans

were successively washed with 15ml of water and 3ml of 10% (v/v) acetonitrile, and then sequentially eluted with 3ml of each of 25% (v/v) acetonitrile and 50% (v/v) acetonitrile. The volumes of collected fractions were reduced and then the fractions were lyophilised. The chloroform layer was processed separately as in **Section 2.3.12**.

2.3.12 Sep-pak Classic C18 purification: Aquenous Acetonitrile system

Sep-pak Classic C18 Cartridge was conditioned with successive elution of 5ml of each of methanol, water, acetonitrile and 15ml of water. Permethylated glycans were resuspended in 200 μl of 50% (v/v) methanol and loaded onto the cartridge. 5ml of water wash for the glycans was followed by successive elution with 3ml of each of 15%, 35%, 50% and 75% (v/v) acetonitrile. All fractions were volume-reduced on SpeedVac and lyophilised

.

2.3.13 Preparation of partially methylated alditol acetates for linkage analysis

Permethylated glycans from **Section 2.3.11.1** were incubated in 200μl of 2M TFA at 121°C for 2 hr for hydrolysis to partially methylated monosaccharides. The aqueous mixture was dried down under a stream of nitrogen, and then the monosaccharides were reduced with 200μl of 10 mg/ml sodium borodeuteride in 2 M ammonia solution at room temperature for 2hr. The reaction was neutralised by 5 drops of acetic acid, and the resulting mixture was dried down under nitrogen. Excess borates were removed by co-evaporation with 10% methanolic acetic acid. The dried samples were incubated with 200μl acetic anhydride at 100°C for 1hr and then dried down under nitrogen. The resulting products were extracted by 1ml chloroform following with 3 times of water-wash and dried down. Partially methylated alditol acetates were resuspended in hexanes for GC-MS-linkage analysis.

2.3.14 MALDI-TOF and TOF/TOF

A 4800 MALDI-TOF/TOF was employed to acquire MALDI-TOF MS and TOF/TOF tandem MS spectra. Instrumental calibration and tuning were carried out with the 4700 Calibration standard kit. The collision energy was set to 1kV for TOF/TOF with argon as collision gas. 3,4-diamonoben-zophenone (DABP) was used as the solo assisting matrix, and the resulting concentration of matrix was 20mg/ml in 75% (v/v) acetonitrile solution. 10μl of methanol was used to dissolve permethylated glycans, and 1μl of this solution was mixed with matrix following a ratio of 1:1. 1μl of the mixture was spotted onto plate.

2.3.15 GC-MS Linkage analysis

The method to prepare partially methylated alditol acetates was described in previous **Section 2.3.13**. A Bruker 456-GC/SCION SQ instrument configurated with a BR-5ms column was employed for data acquisition. Samples were injected into the instrument when the column was at 60°C. The column was then heated with a rate of 8°C/min to 300°C in 30 min.

2.3.16 LC-ESI-MS

LC-ESI-MS experiments were conducted on two instruments (the QSTAR and SYNAPT).
The Q-STAR was set with the following LC gradient: Analytes were eluted from column with solutions from 95.5% (v/v) water/acetonitrile to 95.5%(v/v) acetonitrile/water. 0.05% formic acid was in both solutions.

The LC for SYNAPT was constantly running with a flow rate of 0.3μl/min at 50°C. The elution was initialised with 97/3 (v/v) solution A (0.1% formic acid in water)/ solution C (0.1% formic acid in acetonitrile) and the concentration of solution C was increased gradually to 50% over 90min and maintained for 5min. The concentration of solution A was then increased back to initial level within 0.5 min and kept for 14.5 min.

The Q-STAR instrument was kindly operated by Mrs. Dinah Ramen and Dr. Maria Panico; and Dr. Paul Hichen kindly operated SYNAPT instrument for experiments.

Chapter III. Pregnancy Specific Glycoprotein 1

Chapter 3 Pregnancy Specific Glycoprotein 1

3.1 Introduction

Pregnancy specific glycoprotein 1 (PSG1) is an abundant secretory member of the pregnancy-specific glycoprotein family expressed by trophoblastic cells in maternal circulation throughout human pregnancy(Lin et al., 1974). Its expression commences from the first trimester of gestation and reaches the highest level in late pregnancy in maternal serum(Lee et al., 1979). The role of PSG1 has been recognised as a vital modulator to pregnancy success because of its involvement in multi-biological events of immunomodulation and foetal development (Moore and Dveksler, 2014). The immunomodulatory functions of PSG1 have been illustrated to be capable of inducing cellular responses from DC and monocytes/macrophages to produce immunomodulatory cytokines, including IL-6, IL-10 and TGF-β1 (SNYDER et al., 2001). Through the signalling of these cytokines, PSG1 could indirectly influence other immune cells, such as Treg cells, with respect to their population and polarisation (Martinez et al., 2012). However, corresponding cellular receptors have yet to be characterised.

Dveksler and colleagues have found that the N-domain of PSG1 is capable of inducing TGF-β1 secretion on macrophages (Blois et al., 2014); and its B2 domain, which is typically associated with TGF-β1, can activate the latent complex of TGF-β1 (Ballesteros et al., 2015). PSG1 can regulate placental angiogenesis by interacting with heparan and chondroitin sulfate glycosaminoglycans on endothelial cells to trigger tube formation (Lisboa et al., 2011). Treatment with PSG1 could also elevate TGF-β1 dependent secretion of vascular endothelial growth factor (VEGF) A, a key modulator of angiogenesis and vasculogenesis (Ferrara and Davis-Smyth, 1997), on extra-villous trophoblast cell lines and primary monocytes (Ha et al., 2010).

Besides, PSG1 has been suggested to interact with integrins, several of which can modulate trophoblastic activities including migration, invasion and adhesion with endothelial cells (Rout et al., 2004, Harris et al., 2009). There is a Lys/Gly/Asp (KGD) tripeptide on the PSG1 N-domain(Scarborough et al., 1991) analogous to an Arg/Gly/Asp (RGD) integrin recognition motif on snake venom disintegrins (Rutherfurd et al., 1995). Shanley's group reported that PSG1 could bind to integrin αIIbβ3 on platelets thereby triggering subsequent cellular signalling (Shanley et al., 2013). This study indicated that PSG1 could interfere with cell-ECM interactions via integrins similar to that between platelets and fibrinogen. However, other domains of PSG1 could also behave like the N-domain against the same integrin without either the presence of KGD motif or even the entire N-domain (Shanley et al., 2013). This observation indicates that there are other possible mechanisms or motifs involved in the PSG/

integrin interactions.

The primary sequence (PSG1_Human, UniProt Entry:P11464, in Figure 3.1) of PSG1 folds into four domains: one Ig variable region-like domain at its N-terminus (N-domain) followed by three Ig constant region-like domains (A1, A2 and B2). There are seven potential N-glycosylation sites on the PSG1 sequence that distribute through its N (Asn61, 104 and 111), A1 (Asn199) and A2 (Asn259, 268 and 303) domains. Correspondingly, several serine and threonine residues, which could be candidates for O-glycosylation, are present in the primary sequence. However, only limited outcomes have been achieved in elucidating the biochemical composition and site-occupation of PSG1 glycans.

An early study of native PSG1 firstly measured the content of its monosaccharides as 6.2% Neu5Ac, 5.8% Gal, 13%GlcNAc, 6.5% Man, and 1.1% Fuc (Osborne et al., 1982). This physicochemical study provided preliminary information of compositional constitution of PSG1 glycans and provided evidence for N-linked glycans. However, no exact type of glycosylation or structures of possible glycans could be illustrated based on this information. This information also indirectly implied that there was probably no O-glycosylation because there was no detection of the monosaccharide GalNAc which can initialise O-glycosylation on the side chain of serine and threonine residues on a secretory protein sequence(**Section 1.3.2**, Brockhausen and Stanley, 2015).

Two histochemical studies investigated the reactivity of amniotic fluid and serum PSG1 proteins against plant Concanavalin A, which has binding specificity for terminal hexoses (Mannose/Glucose) and affinity to core-mannose residues of bi-antennary complex glycans(Mandal et al., 1994). Both studies concluded that Concanavalin A was interactive with a large proportion of native PSG1(Heikinheimo et al., 1987, Koistinen et al., 1981) from healthy controls. Thus, these identical binding results suggested that N-glycan structures carrying these features (high-mannose, hybrid and bi-antennary complex type) were likely assembled on PSG1 proteins with a considerable quantity. A more recent histochemical study of gestational diabetes mellitus detected enhanced fluorescent staining of patient serum PSG1 binding to Aurantia lectin, which has a binding preference for the α1,6 linked core-fucose of N-glycans (Nagalla et al., 2015, Matsumura et al., 2007). This enhanced imaging in gestational diabetes

Figure 3.1 The primary sequence of PSG1 and domain arrangements.
The N, A1, A2 and B2 domains in the sequence are bolded and coloured as blue, green, red and black, successively. Tri-residue N-glycosylation sequons are underlined; and potential N-glycosylation sites are highlighted by an under star. (Adapted from PSG1_Human, UniProt Entry:P11464)

mellitus indicated that normal PSG1 N-glycans were likely to be partially core-fucosylated.

In addition, site-occupation of N-glycans on PSG1 has been addressed in a few studies. For instance, Li and colleagues firstly identified a trypsinised N-glycopeptide of PSG1 containing two glycosylated sites (Asn104 and 111) in the secretome of human metastatic hepatocellular carcinoma (Li et al., 2013). Albertolle's group also detected this peptide with glycosylation in human saliva(Albertolle et al., 2015). The approach both groups employed to identify the N-glycosylation was the concomitant observation of a mass shift from glycan-linking Asparagine (Asn) to Aspartic Acid (Asp) on peptide backbones when the N-glycans were enzymatically removed. This is because the employed enzyme, PNGaseF, which can remove an entire N-glycan sequence from its asparagine base residue, triggers a deamidation on the linking side chain(Plummer et al., 1984). Nevertheless, even without intended chemical or enzymatical triggering, deamidation can still spontaneously occur on peptide sequences (Hao et al., 2017). Therefore, these site-occupancy studies might provide false-positive results for N-glycosylation site identification.

To study the interacting mechanism between PSG1 N-domain and integrins, a recombinant sequence of the N-domain (Q35 to L144, 109AA) has been produced in our collaborator's laboratory. This recombinant sequence includes three N-glycosylation sites and the KGD motif of the PSG1 N-domain fused with a C-terminal IgG1 Fc tag designated for purification via affinity chromatography. The recombinant protein was produced in CHO-K1 (Chinese hamster ovary) and Expi-CHO cell lines via stable and/or transient transfection. Interestingly, the recombinant products behaved differently against a selected integrin, α5β1, which plays a vital role in trophoblast development(Rout et al., 2004, Harris et al., 2009). The recombinant material produced by stable transfection (Stable CHO-K1) bound to this integrin in a similar manner to the full-length native PSG1. In contrast, the same construct produced via transient transfection in CHO-K1 was unable to interact with the same integrin. Intriguingly, the recombinant material that was transiently produced in the Expi-CHO cell line behaved similarly to that made in the stable CHO-K1 cell line.

A previous study by Ha's group found that N-glycosylation of murine PSG17/19 is required for its interaction with the cellular receptor CD9 (Ha et al., 2008). This binding activity was lost in recombinant murine PSG produced in bacterial cells, *E. coli,* without glycosylation. However, no other study at this stage described a similar effect of glycosylation on interactions relevant to PSG proteins.

Therefore, my objective was to carry out a comprehensive glycomics and glycoproteomics investigation of native PSG1 and recombinant PSG1 N-domains in order to assist our collaborators in their functional studies. My study aimed to profile PSG1 glycosylation in respect of the complete glycome and related site-occupation. An important goal was the elucidation of differences of glycosylation between the recombinant products expressed in CHO-K1 and Expi-CHO cell lines.

3.2 Results

3.2.1 Research strategies

For this project, we were in collaboration with Professor Gabriela Dveksler of Uniformed Services University, Bethesda. They prepared four sets of PSG1 materials: 1) one set of native PSG1 proteins pooled and extracted and pooled from maternal sera of healthy pregnant donors in the United States under strict regulations of sampling and processing; 2) three sets of recombinant PSG1 N-domain with Fc tag expressed under different expressing conditions as stable CHO-K1 (expressed in CHO-K1 cells with stable transfection), transient CHO-K1 (expressed in CHO-K1 cells with transient transfection) and transient Expi-CHO (expressed in Expi-CHO cells with transient transfection). The mass spectrometric strategies for analysing these materials are summarised in Figure 3.2.
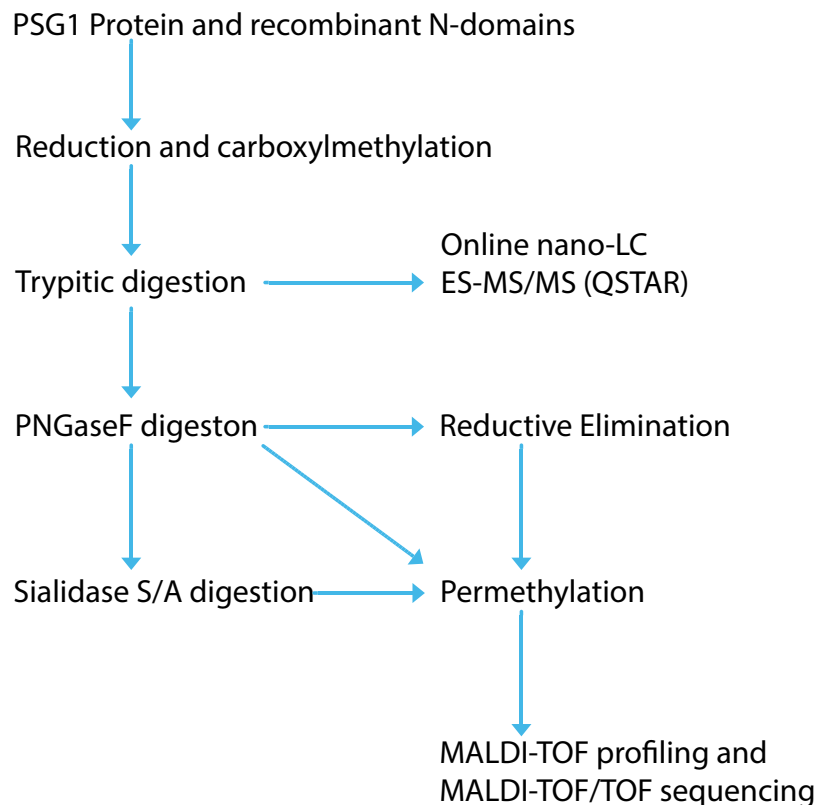
Figure 3.2 Schematic diagram of sample processing and mass spectrometric strategies for glycomics and glycoproteomics analyses of PSG1 materials.
Protein pellets of native PSG1 proteins and recombinant PSG1 N-domain infused with an IgG1 Fc tag (PSG1N-Fc) are analysed.

3.2.2 Data interpretation for MALDI-TOF profiling

MALDI-TOF generally produces singly charged molecular ions ([M+Na]$^+$) with a sodium adduct for permethylated glycans in the positive mode. Molecular ions of glycans present an isotope envelope arising from the natural coexistence of two stable isotopes of carbon, ^{12}C and ^{13}C (99:1). The shift of peak distribution and increasing probability of containing ^{13}C are the consequences of higher carbon content as the molecular weight increases (Figure 3.3A). The lowest m/z peak in an isotope envelope, which contains exclusively ^{12}C, is defined as the monoisotopic peak to assign the composition of a molecular ion unambiguously. The mass to charge ratio (m/z) of a monoisotopic peak is used to interpret the MALDI spectra of glycans as the number and type of monosaccharides can be deduced from the mass of a molecular ion.

As described in the introducing chapter (**Section 1.1**), eight monosaccharides are common constituents of glycans in mammalian systems. Derivatisation of glycans by permethylation confers the following residue mass to each monosaccharide: Fuc (174.1Da); Man, Gal and Glc (Hexose, 204.1Da); GalNAc and GlcNAc (HexNAc, 245.1Da); NeuAc (361.2Da); and NeuGc (391.2Da). The monosac-
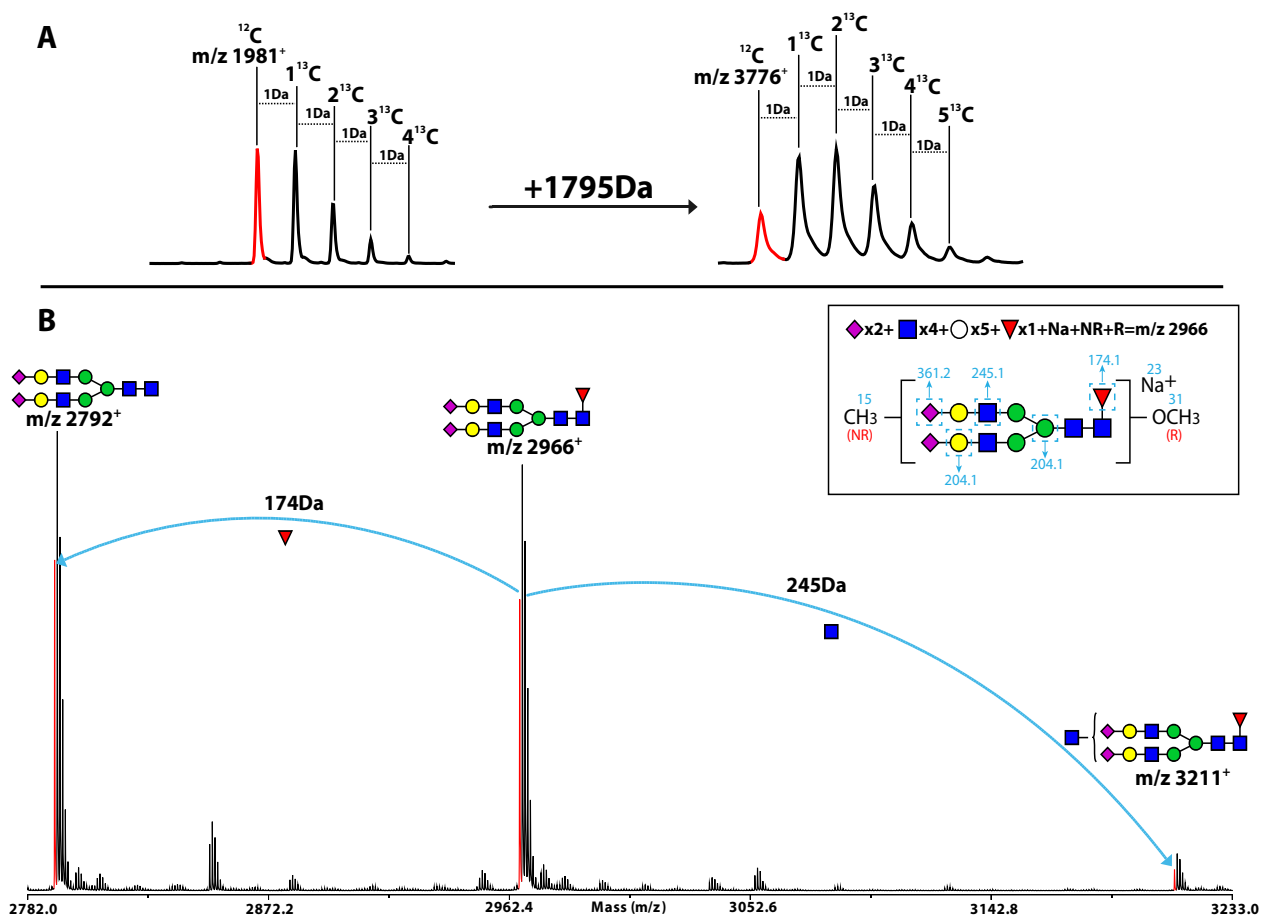
Figure 3.3 Isotopic envelopes of molecular ions and MALDI-TOF spectrum of permethylated N-glycans.
Shift of peak abundance in isotopic envelopes for molecular ions of glycan from m/z 1981 to 3776 (A) and their monoisotopic ions are highlighted as the basis for structural elucidation. A MALDI-TOF spectrum centred monoisotopic peak m/z 2966 (B) and detailed numerical interpretation of m/z 2996 is illustrated in a square at the upper-right corner.

charide composition of each detected glycan can be extrapolated from its molecular ion as follows: the mass of a glycan ion is the total of the residue masses plus the derivatised reducing (31Da), non-reducing (15Da) ends, and the non-covalent sodium (23Da) adduct.

By combining the matched monosaccharide composition and well-established biosynthetic pathways of each type of glycosylation in mammalian systems, elucidation of saccharide sequences on glycans becomes achievable in practice. This is illustrated by Figure 3.3B. Thus the monoisotopic peak of the N-glycan at *m/z* 2966, corresponds to an oligosaccharide consisting of 2NeuAc, 5Hex, 4HexNAc and 1Fuc. Apart from the conserved (3Man2GlcNAc) core on most mammalian N-glycans, the other seven residues could putatively form two sialyl-LacNAc antennae and a core-fucose on the reducing-end chitobiose GlcNAc. Hence the structure of this ion could be assigned as a disialylated bi-antennary complex N-glycan with core-fucosylation. Moreover, due to mammalian glycans usually having conserved core structures, proliferation or modification above the conserved cores can be gleaned by sequential addition of monosaccharides or characterised structural moieties. Technically, assigning a spectrum could also be conducted by calculating the *m/z* increments/decrements between known/near and of interest peaks to elucidate the corresponding composition. Two examples are shown in Figure 3.3B: *m/z* 2792 is a fucose less than *m/z* 2966 and *m/z* 3211 is a GlcNAc higher. Further refinement of putative assignments can be carried out by additional information from MALDI-TOF/TOF MS/MS sequencing experiments (**Section 3.2.2.3**).

3.2.2.1 MALDI-TOF profiling of native PSG1 N-glycans

Figure 3.4 shows a representative spectrum (selected from 3 repeat experiments) of N-glycans recovered from native PSG1. A full list of identified structures is summarised in Table 3.1. The assigned glycan repertoire consists of all three N-glycan classes: High-mannose, hybrid and complex types.

Complex glycans are the dominant class in this glyco-repertoire in respect of structure distribution (ranging from around *m/z* 1980 to 4600) and relative abundance (*e.g.* major signals at *m/z* 2792, 2966, 3602, and 3776). Bi-antennary structures (*e.g. m/z* 2792, 2966 and 3211) are the most abundant members of this glycan family followed by a successive decrease of abundance in tri- (*m/z* 3602, and 3776) mono- (e.g *m/z* 1981 and 2156) and tetra-antennary (*e.g. m/z* 4413 and 4587) structures. Sialylation of the antennary building block, LacNAc, by NeuAc is virtually a global feature observed on most of the complex glycans. The degree of sialylation varies from mono- to tetra-sialylation subject to the availability of antennary LacNAc substrates. Fully sialylated bi-antennary structures *m/z* 2792 and 2966 are the most abundant structures; their counterparts of fully sialylated tri- and tetra-antennary structures are also dominant in their sub-families. About 50% of the complex glycans are core-fucosylated. Bisecting GlcNAc is potentially present on structures such as *m/z* 2850 and 3211 which have a terminal GlcNAc in their composition.
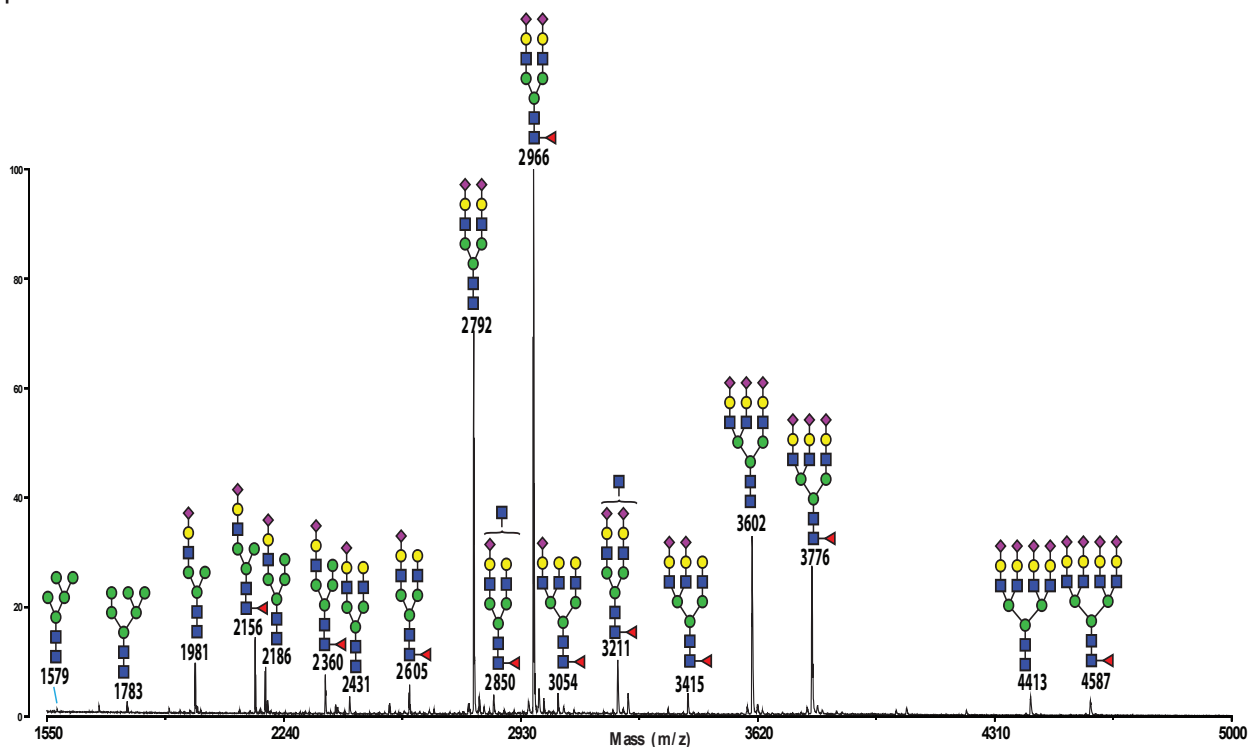
Figure 3.4 MALDI-TOF spectrum of N-glycans of native PSG1.
The 50% acetonitrile/water fraction of permethylated N-glycans from native PSG1 materials was screened by MALDI-TOF after reverse-phase C18 Sep-pak purification. All ions are observed as [M+Na]⁺ and the annotation is based on knowledge of biosynthetic pathway and compositional information.

Hybrid glycans share structural features with complex glycans including antennary sialylation and differentiated core-fucosylation status (*e.g.* 2186 and 2360). The relative abundance of Man4 hybrid glycans (*m/z* 2186 and 2360) surpasses the tetra-antennary family (*m/z* 4413 and 4587). Man5 hybrid glycans are found (*m/z* 2390) in trace amounts. In addition, all five high-mannose structures (Man5 to Man9; *m/z* 1579, 1783, 1987, 2192, and 2396) can be found in spectra of all tested samples, although with relatively low abundance.

3.2.2.2 MALDI-TOF profiling of PSG1 native O-glycans

O-glycans are detected in one of the three analysed batches of native PSG-1 materials (Figure 3.5). In that sample there were two glycan molecular ions at *m/z* 895 and 1256 corresponding to mono- and di-sialylated Core 1 structures. Due to the inconsistency between analysed batches, further work is needed to clarify the cause of this observation.

3.2.2.3 MALDI-TOF/TOF tandem MS/MS sequencing of native PSG1 N-glycans

To verify the putative structures of glycans gleaned from MALDI-TOF spectra, major characteristic structures were validated by tandem MS/MS fragmentation employing positive MALDI-TOF/TOF. As described in chapter I (**Section 1.5.2.3**), targeted molecular ions can be selected and then fragmented via CID inside the tandem mass spectrometer. The resulting cleavages of glycosidic bonds generate a library of B and Y fragment ions which provide vital structural information for defining

Table 3.1 Identified N-glycan structures from native PSG1 materials

N-glycan	Mass to charge ratio (permethylated, *m/z*)
5Man2GlcNAc	1579
6Man2GlcNAc	1783
1NeuAc1LacNAc3Man2GlcNAc	1981
7Man2GlcNAc	1987
1NeuAc1LacNAc3Man1Fuc2GlcNAc	2156
1NeuAc1LacNAc4Man2GlcNAc	2186
8Man2GlcNAc	2192
1NeuAc1LacNAc4Man1Fuc2GlcNAc	2360
1NeuAc1LacNAc5Man2GlcNAc	2390
9Man2GlcNAc	2396
1NeuAc2LacNAc3Man2GlcNAc	2431
1NeuAc2LacNAc3Man1Fuc2GlcNAc	2605
1GlcNAc1NeuAc2LacNAc3Man2GlcNAc	2676
2NeuAc2LacNAc3Man2GlcNAc	2792
1GlcNAc1NeuAc2LacNAc3Man1Fuc2GlcNAc	2850
2NeuAc2LacNAc3Man1Fuc2GlcNAc	2966
1GlcNAc2NeuAc2LacNAc3Man2GlcNAc	3037
1NeuAc3LacNAc3Man2GlcNAc	3054
1GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNAc	3211
2NeuAc3LacNAc3Man2GlcNAc	3241
2NeuAc3LacNAc3Man1Fuc2GlcNAc	3415
3NeuAc3LacNAc3Man2GlcNAc	3602
3NeuAc3LacNAc3Man1Fuc2GlcNAc	3776
1GlcNAc3NeuAc3LacNAc3Man1Fuc2GlcNAc	4022
3NeuAc4LacNAc3Man2GlcNAc	4052
3NeuAc4LacNAc3Man1Fuc2GlcNAc	4226
4NeuAc4LacNAc3Man2GlcNAc	4413
4NeuAc4LacNAc3Man1Fuc2GlcNAc	4587
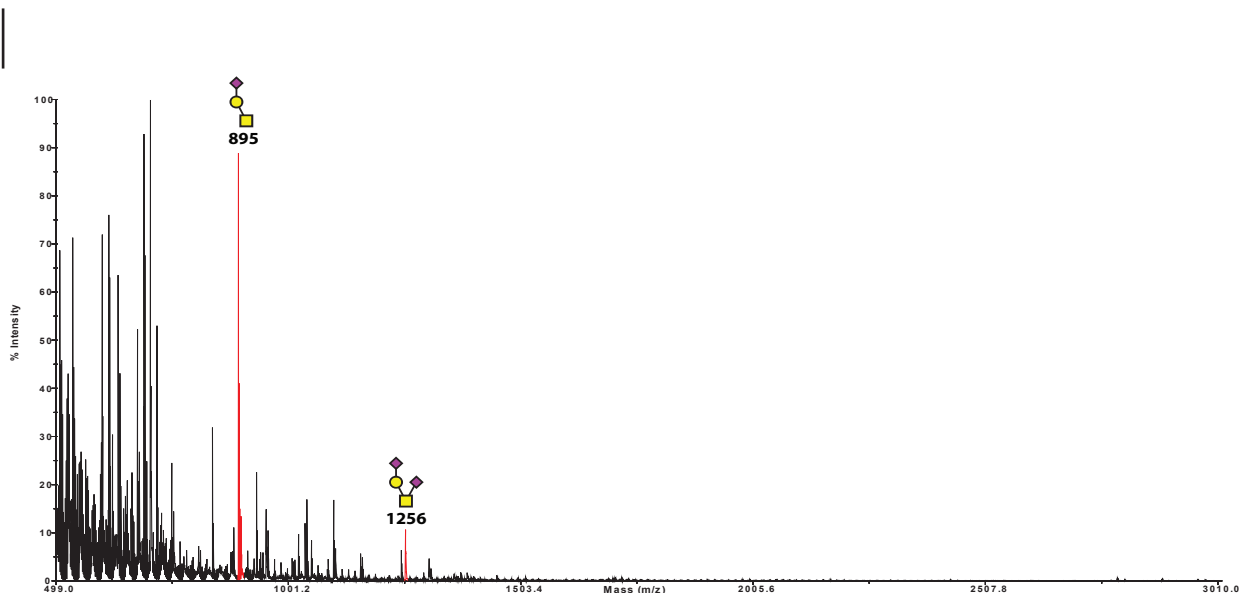
Figure 3.5 MALDI-TOF spectrum of native PSG1 O-glycans.
Data acquisition was performed on the 50%Acetonitrile/water fraction after reversed-phase chromatography. The O-glycans are reduced at reducing end and permethylated. All the ions are observed as [M+Na]⁺. All the structures are assigned upon knowledge of biosynthetic pathway and compositional information.

the sequential arrangements, such as terminal epitopes and antennary sequences, of monosaccharides locally and globally on the selected glycan (Figure 3.6).

To confirm the suggested structural features and identity of N-glycans, MS/MS experiments were generally performed for abundant and putatively characteristic glycans. MS/MS spectra from four molecular ions (*m/z* 1981, 2156, 3211, and 4587) have been selected to represent the large volume of data collected. Their MALDI-TOF/TOF spectra are shown in Figures 3.7 and 3.8.

As a chosen pair of putative mono-antennary glycans, *m/z* 1981 (Figure 3.7A) and 2156 (Figure 3.7B),

NeuAc (361) +Gal (204) +GlcNAc (245)
non-reducing end CH₃ (15) +proton (1)
+reducing end oxygen (16)
+ Na(23) -elminated water (18)

m/z 847 (B-ion)

m/z 474 (Y-ion)

GlcNAc (245) +Fucose (174)
+reducing end OCH3 (31)
+ Na (23)+ proton (1)
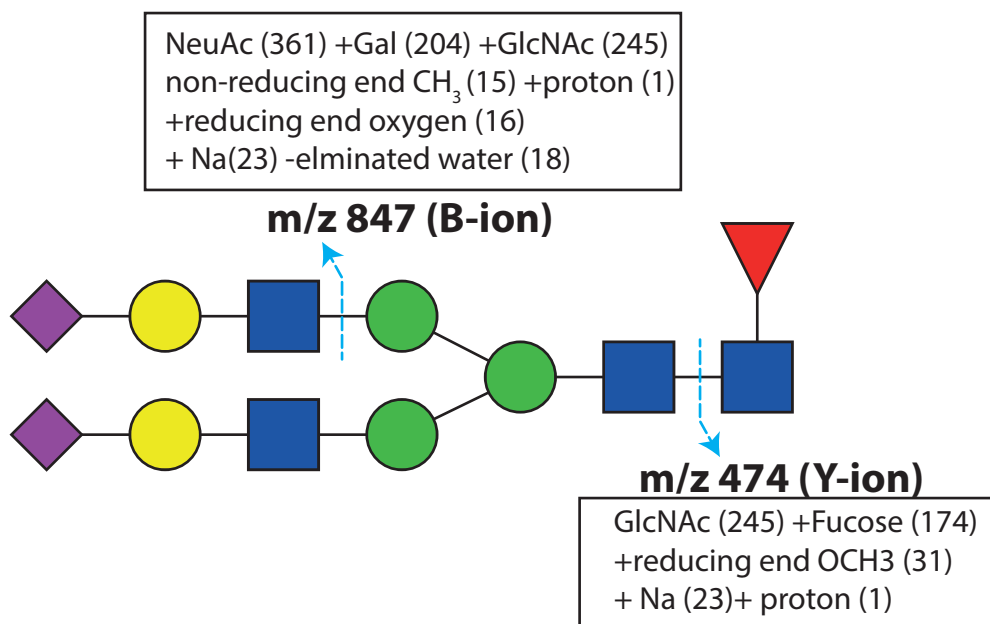
Figure 3.6 Schematic paradigm of CID induced B and Y ions and m/z calculation.

their identities are verified by observing a series of characteristic fragment ions (*m/z* 847, 953, 1157, 1606 and 1704 from parental ion *m/z* 1981; and *m/z* 474, 847, 1127, 1331, 1704 and 1780 from parental ion *m/z* 2156). The presence of a NeuAc sialic acid residue is confirmed by observing the y-ions corresponding to losses of a NeuAc residue (*m/z* 1606 and 1780) for both structures. The presence of *m/z* 847 is indicative for a Sialyl-LacNAc antenna on both glycans, corroborated by the corresponding y-ions from both structures, *m/z* 1157 and 1331, respectively. A linear arrangement of this sialyl-LacNAc tri-saccharide is subsequently elucidated by minor y-ions losing a NeuAc-galactose di-saccharide (*m/z* 1402 and 1576). Co-presence of a pair of y- (*m/z* 474) and b- (*m/z* 1704) ions in the MS/MS spectrum of parental ion *m/z* 2156 validates the putative assignment of core-fucose on observed structures. This observation agrees with the description in Chapter I (**Section 1.3.1.3**) concerning the biosynthetic events leading to core-fucosylation by Fuc-T VIII. The timing of this addition could be as early as
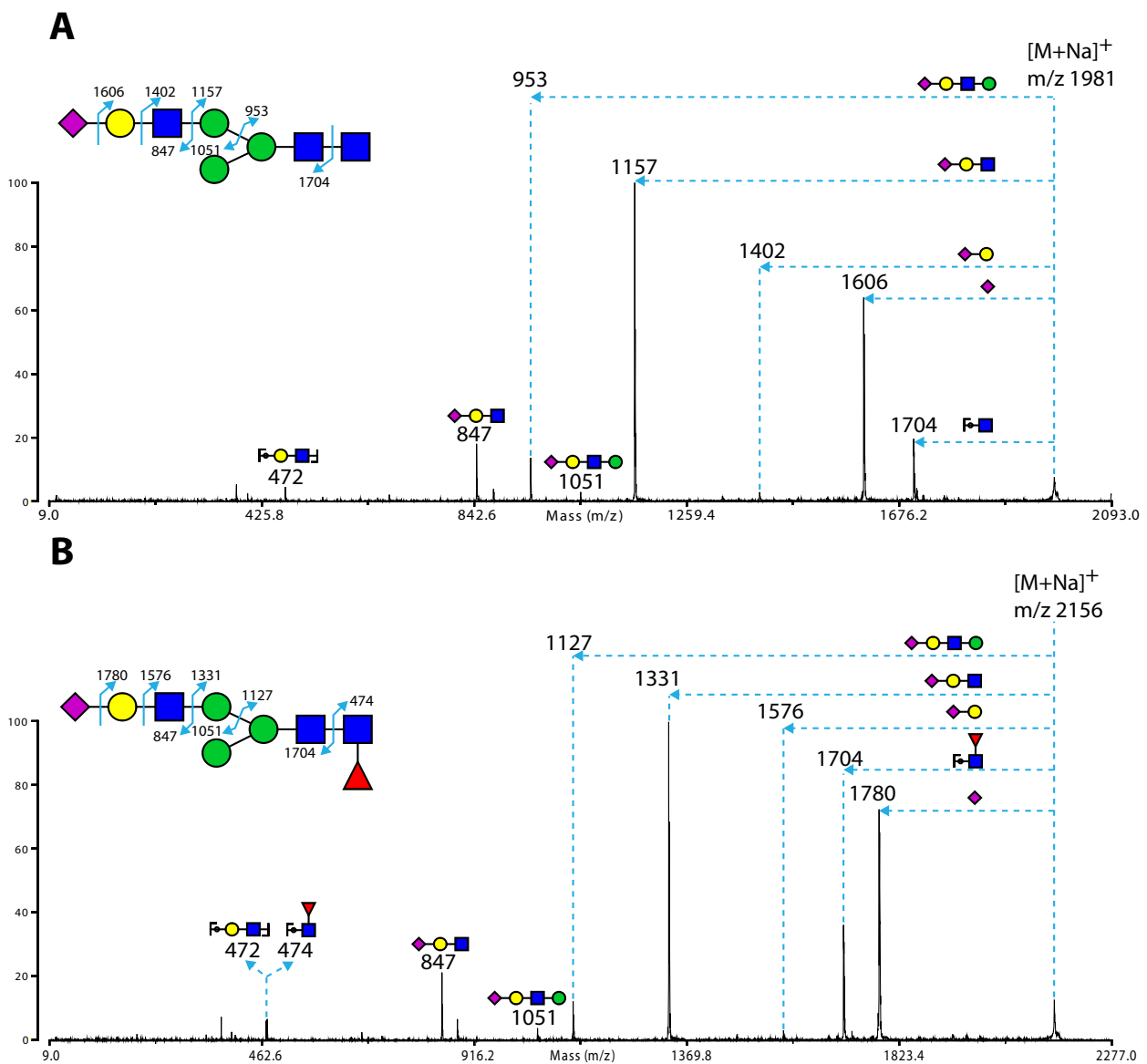
Figure 3.7 MALDI-TOF/TOF spectra of m/z 1981 and 2156 from Native PSG1 N-glycans.
Selected N-glycans m/z 1981(A) and 2156 (B) from 50% acetonitrile/water fraction were fragmented by CID MALDI-TOF/TOF. Fragment ions are detected as [M+Na]⁺. The interpretation and assignments are based on knowledge of fragmentation pathways.

around trimming off the 6 arm mannose residues by mannosidase II since *m/z* 2360 (Figure 3.4) and this structure were simultaneously observed. Therefore, all hybrid and complex glycans in this glycan repertoire can be further classified into two sub-families subject to their core-fucosylation status.

Figure 3.8A shows the MS/MS spectrum of *m/z* 3211. Characteristic b-ion for sialyl-LacNAc (*m/z* 847), y-ion for core-fucosylation (*m/z* 474) and their corresponding ions (*m/z* 2386 and 2760) validate both features for this bi-antennary structure. Besides, a y-ion, *m/z* 2952, was found exclusively corresponding to the loss of a GlcNAc residue from a non-reducing end of this glycan. This fragmentation confirms that a terminal GlcNAc is on this structure, which could potentially be bisected on the central Mannose of the conserved trimannosyl core. However, further experiments are required to
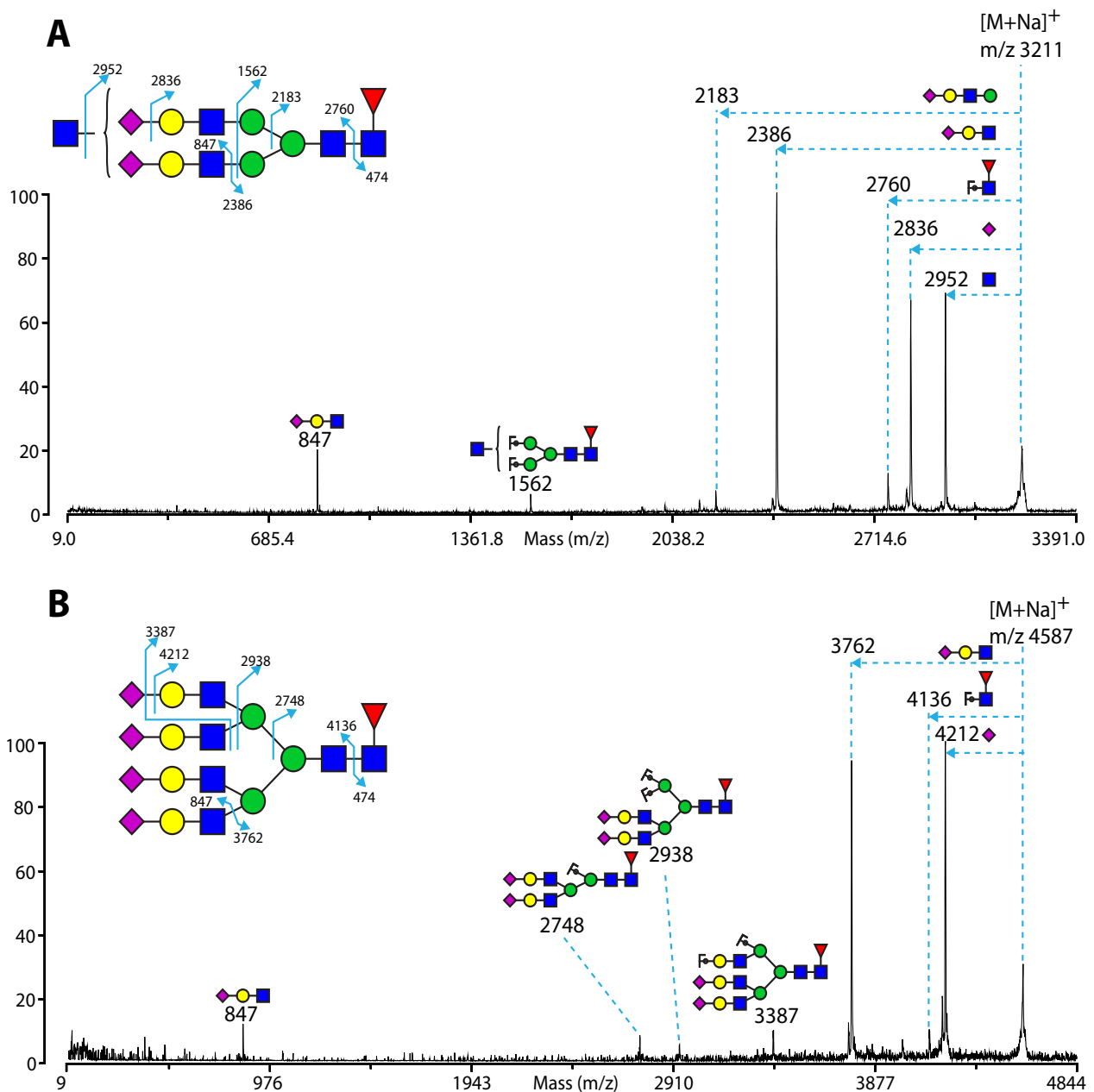
Figure 3.8 MALDI-TOF/TOF spectra of m/z 3211 and 4587 from Native PSG1 N-glycans.
Molecular ions of N-glycan at m/z 3211(A) and 4587 were selected for fragmentation analysis. See the legend of Figure 3.7 for assignments.

validate the actual linkage of this terminal GlcNAc to its underneath mannose residue. This is because glycosidic bonds at the reducing side of GlcNAc residues are preferred for CID induced fragmentation and bisecting GlcNAc and its antennary truncated counterparts are similarly susceptible to cleavage.

The largest structure detected by MALDI-TOF MS, m/z 4587, was fragmented (Figure 3.8B) to investigate the potential presence of poly-LacNAc extension and less-common modifications of sialic acids, such as di-sialyl-LacNAc or poly-sialic acids. However, no fragment ions are found corresponding to a poly-LacNAc antenna or an uncommon sialylation on antennae. Instead, y-ions m/z 3762 and 2938 indicate fragmentation to lose one or two Sialyl-LacNAc antennae, and b-ion m/z 4136 validates the presence of core-fucose. m/z 2748 corresponds to a loss of one arm of the trimannosyl core including two assembled sialyl-LacNAc antennae. The ion m/z 3387 corresponds to losses of a sialyl-LacNAc antenna with an extra Sialic acid by di-fragmentation. Thus, fragmentation analysis for ion m/z 4587 indicate that the structure of this ion is tetra-antennary with terminal sialylation and core-fucosylation.

3.2.2.4 Sialidase S and A digestion of native PSG1 N-glycans

Linkage of sialic acids to their underlying glycans is a vital factor modulating the binding specificity of glycan moieties to lectins. Thus, it is essential to define the glycosidic linkages of sialic acid residues for the native PSG1 N-glycans (**Section 3.2.2.1**). To determine the linkages of NeuAc residues to their carrying structures, Sialidase S and A were used to respectively cleave off sialic acids specifically (cleaving α 2, 3 linkage) and non-specifically on purified N-glycans from native PSG1 materials.

The sialidase S treatment for the N-glycans resulted in substantial but incomplete digestion (shown in Figure 3.9A). Thus the sialic acid content of treated glycans showed a dramatic decrease as most sialylated glycans completely disappeared in the spectrum with the exception of a small group of structures with minor abundance (*e.g.* m/z 2431 and 2792). Correspondingly, asialylated multi-antennary structures dominantly appear in spectrum (*e.g. m/z* 1620, 2244, 1794, 2070, 2489, 2519, 2693, 2968 and 3142) after digestion and m/z 2244 becomes the most abundant structure as base peak. This shift of relative abundance of structures implies that the majority of sialic acids have an α 2, 3 linkage rather than an α 2, 6 linkage on their N-glycan antennae.

In parallel, sialidase A cleavage was performed on another aliquot of the purified N-glycans (Figure 3.9B). Complete removal of sialic acid content was observed as well as a substantial increase of relative abundance of potential GlcNAc bisecting structure, m/z 2489, from 17% to 31% in comparison with base peak between that after Sialidase S and Sialidase A Digestion. A tetra-antennary structure with potential bisecting GlcNAc, m/z 3387, was detected at a trace level after this digestion. In addition, benefiting from improved ionisation by the removal of sialic acids, two putative structures, m/z 3418 and 3592, with 5 LacNAc units in composition appear at trace levels in the spectrum after sialidase A digestion . The observation of these two structures indicates the presence of antennary elongation by
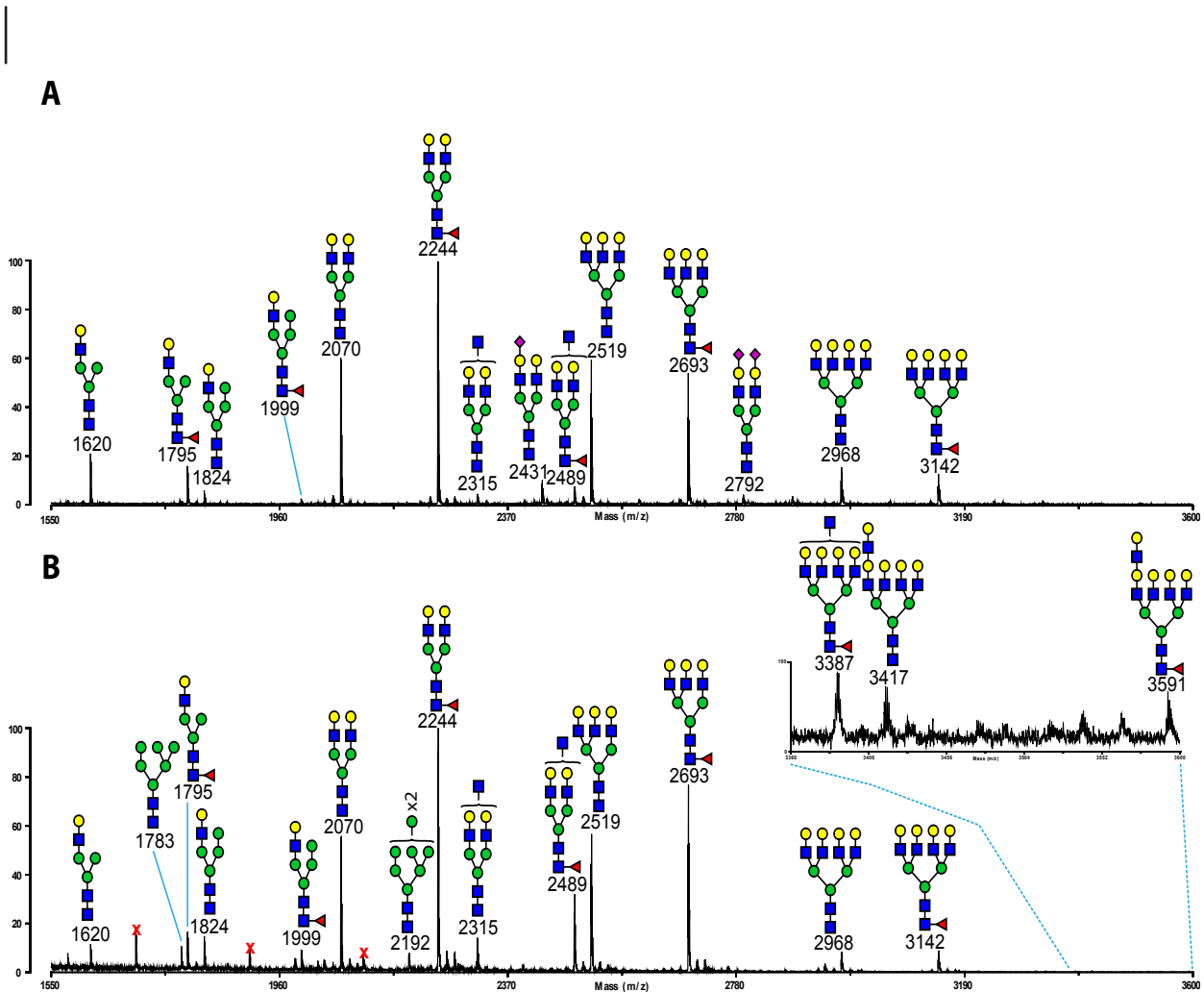
Figure 3.9 MALDI-TOF spectra of native PSG1 N-glycans after Sialidase S and A digestions.
Native PSG1 N-glycans were digested with sialidase S(A) and A(B); and they were then permethylated. High mass range of Sialidase A spectrum is zoomed in and presented (*m/z* 3340 to 3600) at upper right of the panel. Poly-hexose contaminations are indicated by red crosses. See the legend of Figure 3.4 for assignments.

poly-LacNAc on PSG1 glycans.

3.2.2.5 Summary of glycosylation of native PSG1

N-glycans of native PSG1 are dominated by multi-antennary complex structures with heavy sialylation. Most NeuAc sialic acid residues have an α2, 3 glycosidic linkage, and a minor population of α 2, 6 linked NeuAc is on limited complex glycans. There are a few potentially GlcNAc bisected complex glycans with relatively low abundance. Hybrid glycans are dominated by mono-antennary Man4 structures which are exclusively sialylated with α2, 3 linked NeuAc. Core-fucosylation is on nearly 50% of all hybrid and complex glycans with a comparable abundance to their core-afucosylated counterparts. High mannose structures are of low abundance and trace levels of complex glycans with poly-LacNAc extension appear after sialidase A digestion. O-glycosylation was detected on one of three analysed materials indicating inconsistency between different batches of samples.

3.2.2.6 MALDI-TOF profiling of recombinant PSG1N-Fc (Stable CHO-K1, Transient CHO-K1 and Transient Expi-CHO) N-glycans.

As a widely used series of cell lines for industrial production of therapeutic secretory proteins, genetic and glycomics studies for CHO cell lines have been well established (Tejwani et al., 2018, Hossler et al., 2009). Generally, the strain of original CHO cell lines, such as CHO-K1, mainly assembles human-like glycans onto its glycoproteins. Glycomics studies of CHO-K1 cells and their secretory products reveal iconic glycosylation for their N- and O-glycans (North et al., 2010, Lee et al., 2018). The N-glycans of CHO-K1 are conventionally 1) core-fucosylated; 2) poly-LacNAc elongated; 3) sialylated with α2,3 NeuAc sialic acid, and 4) bisected with GlcNAc at a negligible level. Only three O-glycans have been found as neutral, mono- and di-sialylated core 1 structures. It is worth to note that all the recombinant PSG1 N-domain materials were tagged with an IgG1 Fc. The recombinant IgG1 Fc tag can be glycosylated simultaneously. A glycomics analysis for the Fc tag made in transient CHO-K1 system as a negative control for its corresponding material (Figure 3.10) indicates a glycoform which is dominated by bi-antennary core-fucosylated complex structures with varied antenna galactosylation (*m/z* 1836, 2040 and 2244). A Man5 high mannose glycan at *m/z* 1579 and a mono-sialylated bi-antennary structure at *m/z* 2605 are also observed in the spectrum. This glycosylation is similar to human IgG Fc glycosylation (Kiyoshi et al., 2017). Therefore, glycosylation on the Fc tag might also have a contribution to abundance of these structures in the MALDI-TOF profiling of compatible recombinant PSG1N-Fc materials.

Figure 3.11 shows the N-glycan profiles of analysed recombinant N-domain of PSG1 with recombinant C-terminal Fc tag (PSG1N-Fc) from three different expression systems. The recombinant PSG1N-Fc secreted from Stable CHO-K1 (Batch I, Figure 3.11A) and Transient CHO-K1 (Figure 3.11B) shows a
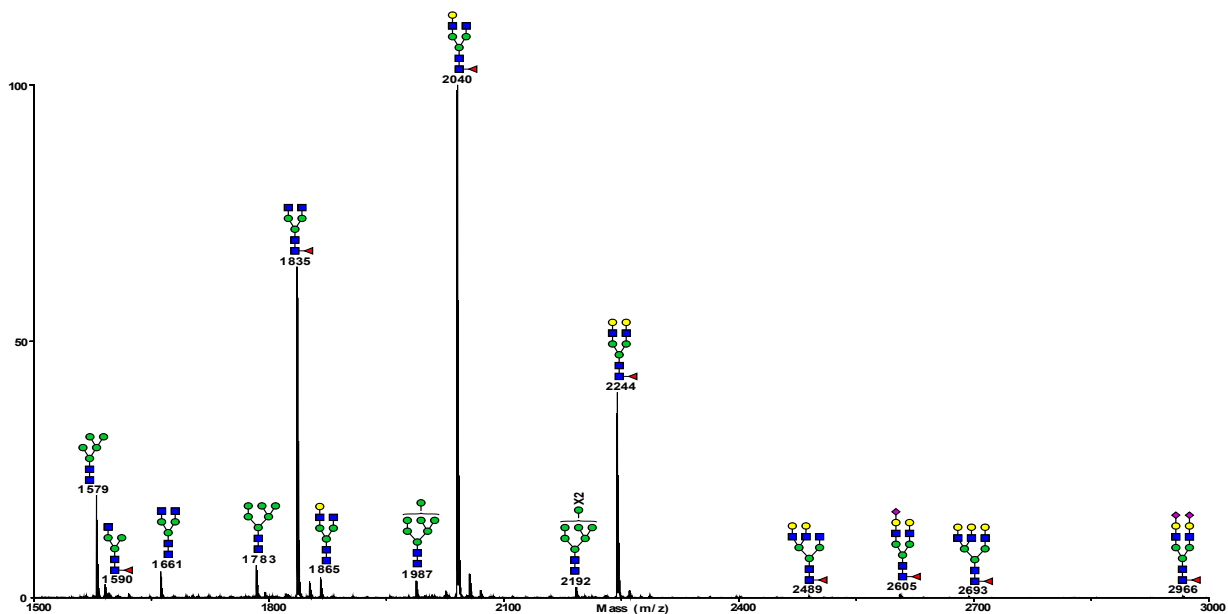
Figure 3.10 MALDI-TOF spectrum of N-glycans of IgG1 Fc tag made in transient CHO-K1 system.
See the legend of Figure 3.4 for assignments.

similar distribution of glycan structures as nearly all complex structures are shared in both materials. Their complex N-glycans well fit with the features of CHO cell glycosylation described above as mainly core-fucosylated, potentially poly-LacNAc extended (*e.g. m/z* 4675) multi-antennary structures.

However, the relative abundance of the different classes of glycans varies depending on the expression system. For example, the transient CHO-K1 material has a higher content of high mannose struc-
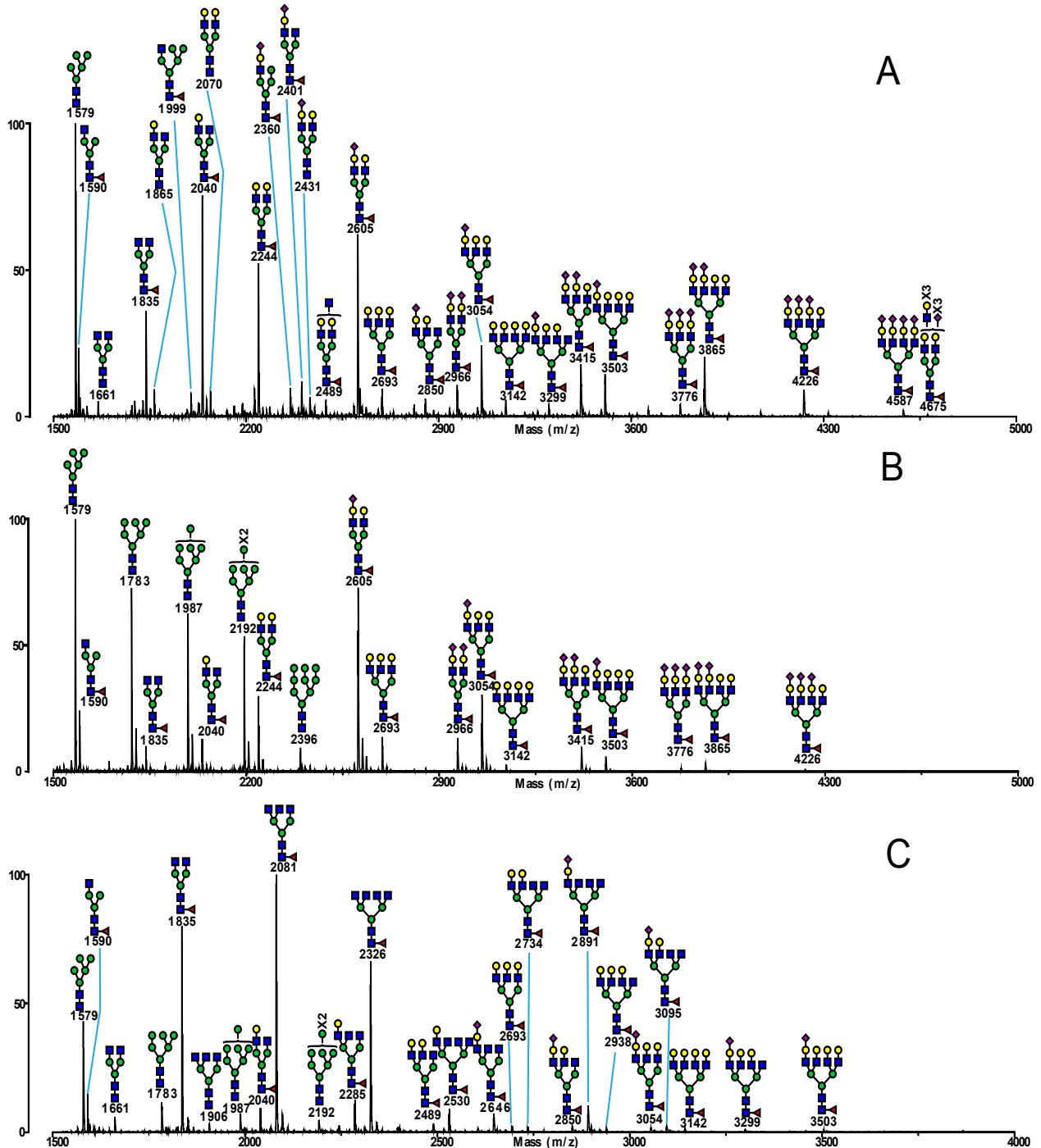
Figure 3.11 MALDI-TOF spectra of N-glycans from recombinant PSG1 N-domains in fusion with a C-terminal IgG1 Fc tag expressed in different CHO systems.
Recombinant N-domains expressed in stable CHO-K1 (A) and transient CHO-K1 (B) and transient Expi-CHO (C) systems. See legend of Figure 3.4 for assignments.
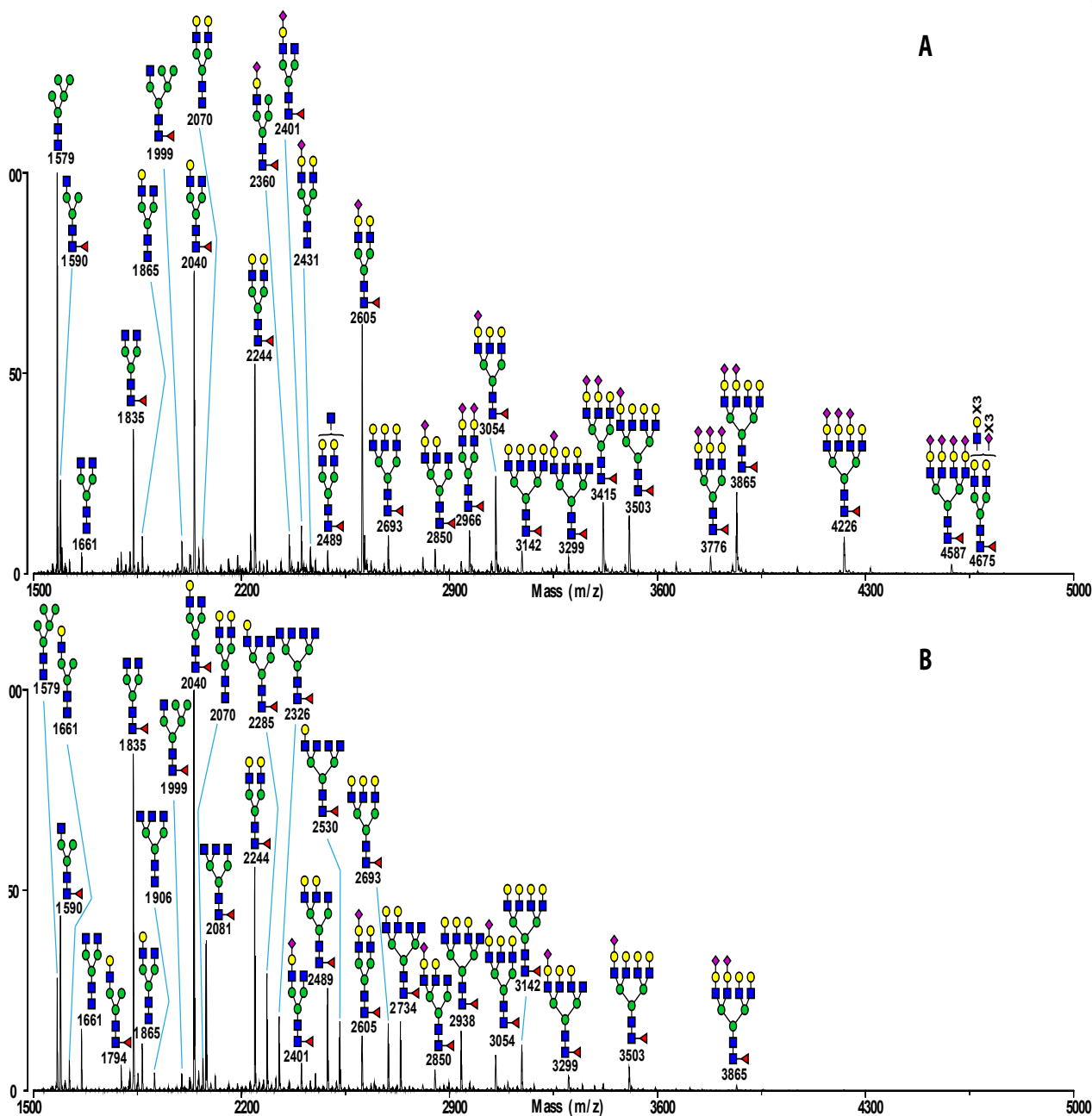
Figure 3.12 MALDI-TOF spectra of N-glycans of two batches of recombinant PSG1N-Fc from stable CHO-K1. Two batches of PSG1N-Fc of stable CHO-K1 were analysed for their N-glycans shown as the first batch(A) and second batch(B). See the legend of Figure 3.4 for assignments.

tures and relatively lower abundance of multi-antennary complex structures compared with stable expression (n>2). The different abundance might reflect relatively less processing of PSG1N-Fc from transient CHO-K1 system in its Golgi apparatus. This reduced processing might be a consequence of hyper-transfection in this system leading to a relative shortage of relevant components in the pathway.

Unlike conventional CHO cell lines, the recombinant PSG1N-Fc produced by the transient Expi-CHO system (Figure 3.11C) exhibits a dominant agalactosylation on complex glycans. Three complex glycans are predominant in this profile at *m/z* 1835, 2081 and 2326 surpassing all the other structures in relative abundance. Their compositions indicate that they possess multiple terminal GlcNAc residues

(n=2-4) above their core-fucosylated pauci-saccharide cores. Core-fucosylation is common on most complex glycans and galactosylation (n=1-4) is present on the complex glycans forming LacNAc antennae (*e.g. m/z* 2285, 2693, 2891 and 3142), but with low abundance. Sialylation is restricted to mono-sialylation on a few structures (*e.g. m/z* 2646, 2850, 2891, 3054 and 3503), in part due to low availability of LacNAc substrates. High mannose glycans were detected with an intermediate abundance (*e.g. m/z* 1579), but no signals for hybrid glycans were detected.

Due to there being insufficient amounts of material of stable CHO-K1(Batch I) PSG1N-Fc (Figure 3.11A&3.12A) to continue subsequent glycorprotemics analysis, a second batch of the stable CHO-K1 (Batch II) PSG1N-Fc (Figure 3.12B) with verified identical integrin reactivity was provided for both glycomics and glycoproteomics experiments. Instead of observing consistent glycomes between these two batches, the stable CHO-K1(Batch II) presents more under- (*e.g. m/z* 2530, 2735 and 2938) and even a-galactosylated (*e.g. m/z* 2081 and 2326) structures than its predecessor. Sialylation and potential poly-LacNAc extension consequently declined or vanished in the batch II. This observation raised concerns of the culture status as a precise hypo-galactosylation which could be a consequence of cells in an unwanted condition such as oxidative stress (Lewis et al., 2016).

3.2.2.7 MADLI-TOF profiling of recombinant PSG1N-Fc O-Glycans

Among all analysed recombinant PSG1N-Fc materials, O-glycans were only found in one batch of sample from stable CHO-K1 (Batch I, Figure 3.13). Two structures are observed as mono- and di-sialylated core 1 structures (*m/z* 895 and 1256). Thereby, O-glycosylation on these recombinant materials requires further validation.

3.2.2.8 MALDI-TOF/TOF sequencing of N-glycans of recombinant PSG1N-Fc

Figure 3.14 shows the fragmentation analysis for three molecular ions of the N-glycans from stable
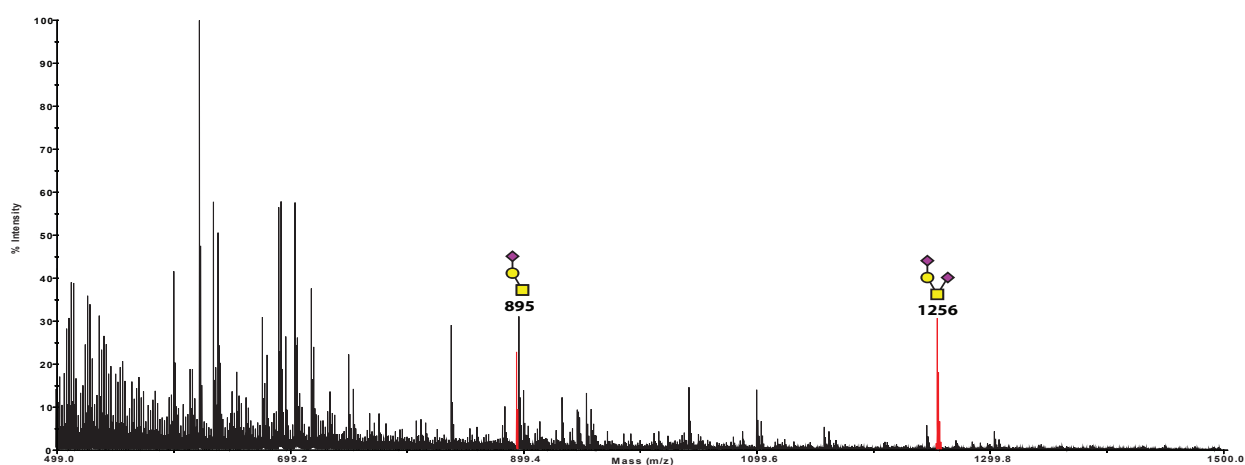
Figure 3.13 MALDI-TOF spectrum of O-glycans of recombinant PSG1N-Fc from stable CHO-K1 system.
See the legend of Figure 3.5 for assignments. Peaks corresponding to glycan structures are highlighted in red.

CHO-K1, transient CHO-K1 and transient Expi-CHO systems, respectively. To verify the consistency of glycosylation between transient and stable CHO-K1 systems for PSG1N-Fc, m/z 3504 (Figure 3.14 A&B) was selected from both N-glycan pools. This ion could be potentially either elongated or branched on their trimannosyl cores as isomers might exist. Instead of observing structural difference, the MS/MS fragmentation for these two ions shows a high degree of similarity. Major fragment ions as m/z 2230, 2591 and 2679 correspond to loss of a terminal sialyl-LacNAc, LacNAc and sialic acid residue, respectively. This validates a consistent presence of these determinants at the non-reducing end of this glycan from both systems. Double-cleavage of antennae is observed on the glycan from stable CHO-K1 system at m/z 1766 corresponding to loss a sialyl-LacNAc and a LacNAc antennae. A possible signal for this fragment ion on the m/z 3054 from transient CHO-K1 system is also visible indicating a possible quantitative difference, but no fragment ions were observed corresponding to an antennary elongation. The signal for core-fucosylation m/z 2603 is consistently low for both glycans, but there is also no detected signal corresponding to branch fucosylation. Therefore, it is likely that the glycosylation machinery in both systems has no precise alteration.
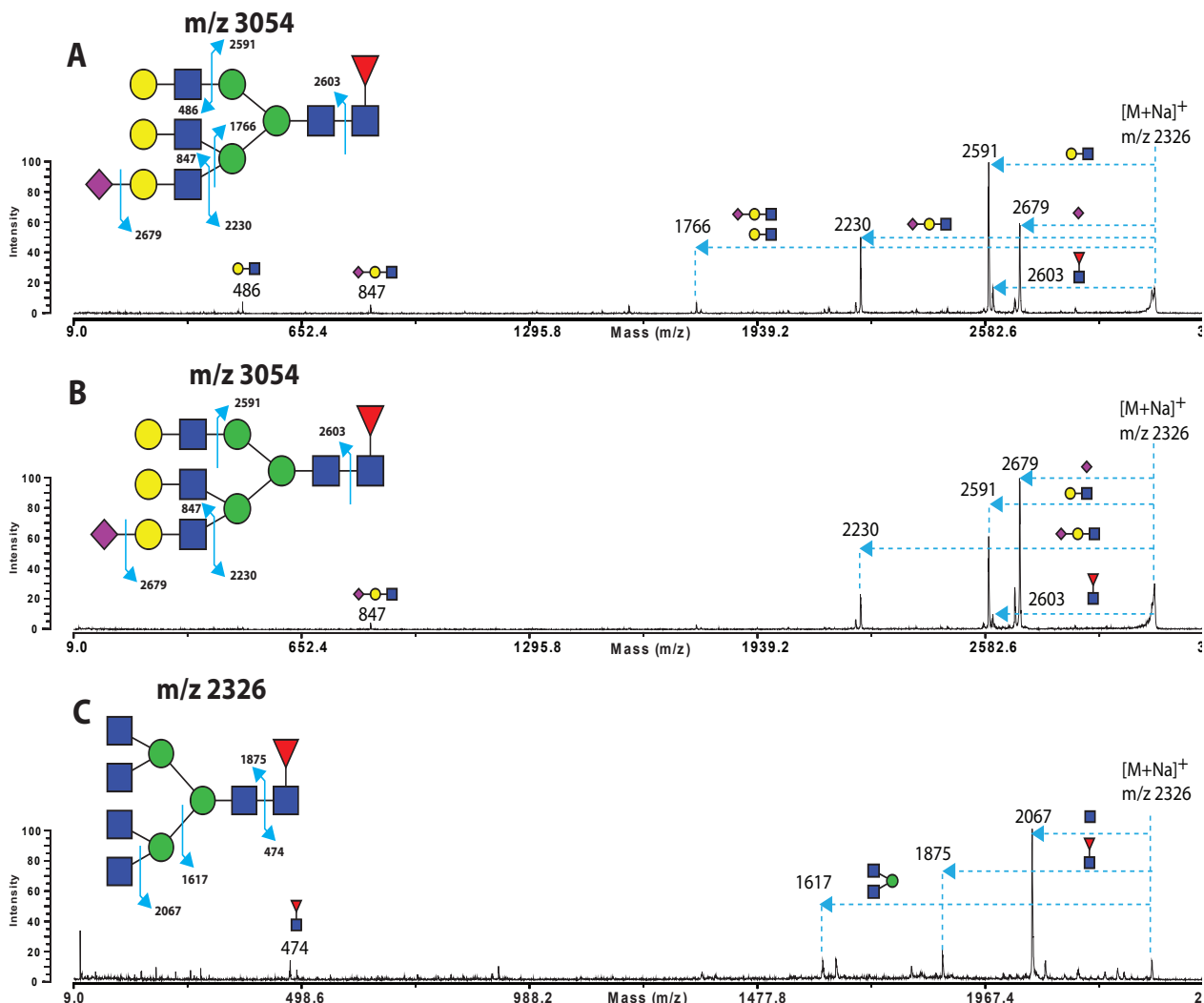
Figure 3.14 MALDI-TOF/TOF spectra of selected N-glycan ions of PSG1N-Fc from CHO cell systems.
There are parental ions from stable CHO-K1 system (m/z 3054, A), transient CHO-K1 system (m/z 3054, B) and transient Expi-CHO system (m/z 2326, C) for fragementation analysis. See the legend of Figure 3.7 for assignments.

In addition, a putative truncated structure, *m/z* 2326, was selected for MS/MS analysis from the N-glycan pool of PSG1N-Fc made in the transient Expi-CHO system (Figure 3.14 C). The presence of the y-ion at *m/z* 2067 indicates the loss of terminal GlcNAc from a non-reducing end as the predominant fragment ion in this spectrum. The y-ion at *m/z* 1617 indicates a cleavage at one arm of the trimannosyl core, and core-fucosylation is validated by a pair of by-ions as *m/z* 474 and 1875. Due to there being no detection of potential fragment ions for LacdiNAc determinants, it is very likely that all the HexNAcs are connected to the core-mannoses as shown in Figure 3.14C rather than being tandemly connected together.

3.2.3 Glycoproteomics of native PSG1 and recombinant PSG1 N-domain

To define the site-specific glycoforms on PSG sequences, LC-ESI-MS glycoproteomics experiments were carried out for native PSG1 proteins and recombinant PSG1 N-domains. Briefly, PSG1 proteins and recombinant N-domains were denatured by carboxymethylation, and the denatured sequences were then digested by sequencing grade trypsin. Their tryptic products were directly loaded onto an Applied Biosystems QSTAR LC-MS/MS system with set LC programmes. Mascot search was subsequently performed for each material to validate sequence identity and coverage against its established proteomic database (Tissot et al., 2009).

3.2.3.1 Identification of Glycopeptides

Since there was no specific enrichment of glycopeptides before loading onto the mass spectrometer, the system acquires MS and MS/MS data for glycopeptides and other peptides/contaminants eluted, simultaneously. Therefore, identification of glycopeptides mainly relied on manual search for the presence of indicative fragment ions of saccharides in the low mass range of MS/MS spectra. There is a targeted pair of indicative saccharide fragment ions at *m/z* 204 (HexNAc) and 366 (LacNAc), which are most commonly observed as an indication for glycosylation. The fragment ion of Sialic acid residue, such as *m/z* 292 (NeuAc), can also be detected as an indicator of the presence of sialylation. Dehydration of these saccharides is commonly observable as the detection of *m/z* 168 (HexNAc-2H$_2$O), 186 (HexNAc-H$_2$O) and 274 (NeuAc-H$_2$O).

All MS/MS spectra including these indicative fragment ions and their matching MS spectra were thoroughly scrutinised for the identification of glycopeptides. The previous MALDI-TOF profiles were employed along with predicted tryptic products of glycopeptides to facilitate this validation. It is worth to note that there was no derivatisation for monosaccharides in this experiment, thus the detected mass of each monosaccharide, including 146Da (Fuc), 162Da (Hex), 203Da (HexNAc), and 291Da (NeuAc) is consistent with their natural molecular weights.

3.2.3.2 Site occupation of native PSG1 and recombinant PSG1 N-domain

Tryptic products of PSG1 materials were screened by the QSTAR LC-MS/MS system with a reversed-phase chromatographic column. Glycans were found on four of the seven N-glycosylation sites on native PSG1 via the observation of five peptide backbones carrying N-glycosylated sites Asn61, 199, 268 and 303, respectively (see Table 3.2 for peptide sequences). Two digestion products were observed for site Asn61 corresponding to complete digestion (peptide mass: 2398Da) and a miss-cleavage at its N-terminus (peptide mass: 2898Da). Double miss-cleavages occurred on peptides for site Asn303 which has mis-cleavages at both N- and C-terminals (peptide mass: 2629Da). Complete tryptic digestion was observed on peptides for sites Asn199 (peptide mass: 718Da) and 268 (peptide mass: 2409Da).

As confirmed with our collaborators, there was no artificial alteration of the engineered PSG1 N-domain sequence in and around the regions of expected tryptic products carrying N-glycosylation sites. Therefore, the tryptic products of recombinant PSG1 N-domains are comparable to that of native PSG1 N-domain. Two digestion products (peptide masses: 2398Da and 2898Da) for site Asn61 are found in all recombinant materials resulting from the same miss-cleavage observed in native PSG1. Glycosylation of sites Asn104 and 111 was exclusively detected in the recombinant N-domain from transient CHO-K1 (peptide mass: 1779Da, Table 3.2). Additionally, a glycopeptide from the engineered IgG1 Fc domain (EEQYNSTYR; peptide mass: 1188Da) was also detected in all recombinant materials.

3.2.3.3 Glycosylation at site Asn61

ESI generates multiple charges ($[M+nH]^{n+}$ or $[M-nH]^{n-}$) for each ionised molecule in both MS and MS/MS modes. The charge states of each molecule can be used to facilitate its identification. The charge state of ions can be identified by calculating the m/z difference between isotopic peaks. This is 1 for a single charge, 0.5 for a double charge, 0.33 for triple charge and 0.25 for a quadruple charge. The sequential composition of peptides and glycans can be revealed by comparing mass intervals consistent with amino acid and/or saccharide residues between ions with the same charge state. For relatively

Table 3.2 Observed PSG1 N-glycopeptides in tryptic digest

Peptide Sequence	Peptide Molecular Weight	Native PSG1	T. Expi-CHO	T. CHO-K1	S. CHO-K1
DVLLLVHNLPQNLTGYIWYK	2398 Da	X	X	X	X*
VSEGKDVLLLVHNLPQNLTGYIWYK	2898 Da	X	X	X	
ETAYSNASLLIQNVTR	1779 Da			X	
LSETNR	718 Da	X			
SENYTYIWWLNGQSLPVSPR	2409 Da	X			
ILILPSVTRNETGPYQCEIRDR	2629 Da	X			
EEQYNSTYR	1188 Da		X	X	X

*Relatively less confident assignment due to observing limited peptide fragmentation. The crosses in table indicate the presence of glycopeptide in relevant materials
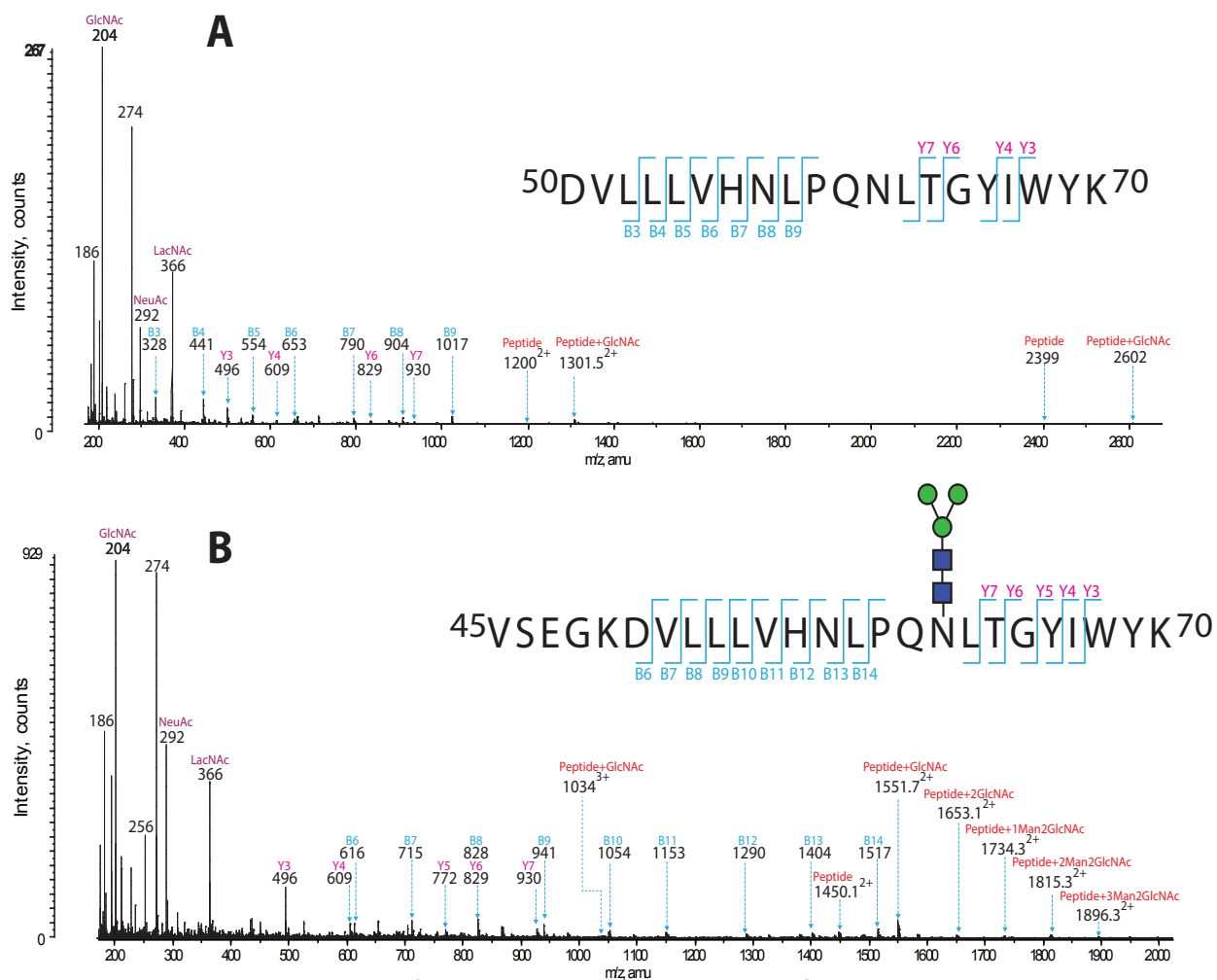
Figure 3.15 LC-ESI-MS/MS spectra of observed tryptic Product I and II for N-glycosylation site Asn61. Fragmentation analysis was conducted for tryptic product I(A) and II(B) for glycopeptide identification. Mono- and di-saccharide fragments are highlighted by name; dehydrated saccharide fragments are labelled in numbers. B and Y ions are highlighted in blue and purple capital letters; intact peptides are indicated in red letters with their remaining monosaccharide residues. Most fragments form a [M+H]$^+$ protonated ion, intact backbones and mis-cleaved peptides can form [M+nH]n+ ions, the number of proton adducts depends upon the numbers of remaining basic residues. An observed putative structure is also assigned on the upper right sequence in panel (B). All the assignments are based on knowledge of peptide fragmentation and glycan fragmentation pathways as well as the biosynthetic pathways along with available compositional information.

low molecular weight analytes as digested glycopeptides, the charge state of their MS/MS fragments is mainly determined by protonation and deprotonation of peptides (Konermann et al., 2013). Therefore, basic or acidic amino acid residues will influence the formation of multiple charge states in the MS/ MS data.

As described in **Section 3.2.3.2**, there are two tryptic digestion products for site Asn61. The identity of these two products was firstly validated via mapping fragment ions of their peptide backbone and saccharides, respectively, in MS/MS spectra. An example of each spectrum from native PSG1 proteins is shown in Figure 3.14. Intact backbone of fully digested Asn61 glycopeptide (^{50}DVLLLVHNLPQN-LTGYIWYK^{70}) was detected as a singly charged ion at *m/z* 2399 (Product I, Figure 3.15A). The doubly charged ion for this backbone (*m/z* 1200^{2+}) was also detected. Two basic residues (histidine and lysine)

contributed the double charge state. The identity of this peptide was verified by mapping series of singly charged N-terminal b-ions (m/z 328, 441, 554, 790, 904 and 1017) and C-terminal y-ions (m/z 496, 609, 829 and 930). Its glycosylation status is evident by the presence of saccharide fragments (m/z 186, 204, 274, 292 and 366) and singly charged peptide backbone carrying a GlcNAc residue (m/z 2602).

The miss-cleaved Asn61 glycopeptide (singly charged m/z 2899, [45]VSEGKDVLLLV-HNLPQNLTGY-IWYK[70], Product II, Figure 3.15B) was discerned by the presence of a group of doubly charged fragment ions including m/z 1450.1^{2+}, 1551.7^{2+}, 1653.2^{2+}, 1734.3^{2+}, 1815.3^{2+} and 1896.3^{2+}. The mass intervals calculated between these doubly charged ions indicate a consecutive loss of monosaccharides from a peptide with pauci-saccharide N-glycan core ($Man_3GlcNAc_2$) m/z 1896.3^{2+} to a naked backbone m/z 1450.1^{2+}. A triply charged peptide backbone with a GlcNAc residue was also detected at m/z 1034.4^{3+} indicating an additional basic residue in the sequence. A series of fragment b-ions of this peptide (m/z 616, 715, 828, 941, 1054, 1153, 1290, 1404 and 1517) indicates a mis-cleavage occurred at its N-terminal penta-amino acid sequence VSEGK. Y-ions from the C-terminus, including m/z 496, 609, 772, 829 and 930 provide complementary peptide identification.

3.2.3.3.1 N-glycosylation on Asn61 of native PSG1

There were three elution fractions (Figure 3.16) for the two Asn61 glycopeptides from native PSG1. Two fractions for Product II (2898Da) eluted at 61.5-62.5min (Fraction A) and 71.3-75.4min (Fraction B); one fraction for Product I (2398Da) eluted at 80-82.7min (Fraction C), respectively.

In the fraction A of Product II, twelve N-glycans (Table 3.3) were identified on eluted glycopeptides by manual examination. Abundant N-glycans in this fraction are annotated on its MS spectrum shown in Figure 3.16A. There are mono- to tetra-antennary structures with and without core-fucosylation. Sialylation is a common modification on all N-glycans to form tri-saccharide sialyl-LacNAc moieties as characterised in the glycomics experiments (**Section 3.2.2.1** and **Section 3.2.2.2**). Up to three sialyl-LacNAc moieties can be found on a single structure as m/z 1441^{4+} and 1477^{4+}, which are tri-antennary with predominant abundance. Mono- and dominant bi-antennary structures are fully sialylated on their LacNAc antennae. Unmodified LacNAc moieties are also observable on bi- to tetra-antennary structures with mono- to di-sialylation (m/z 1404^{4+} and 1149^{5+}). A mono-sialylated tetra-antennary structure, m/z 1437^{4+}, was detected with a terminal GlcNAc which could potentially be a bisecting one. Bisecting GlcNAc could equally be on a minor bi-antennary glycan (m/z 1364^{4+}) .

Glycopeptides eluted in fraction B contain tri- and tetra-antennary glycan structures, exclusively. Core-fucosylation consistently occurred on about half of all assigned structures (Figure 3.16B). The level of sialylation on these structures is generally higher than that of the structures in fraction A as mainly with tri- and tetra-sialylation. This observation might suggest that higher content of sialic acid residues could lead to a later elution of glycopeptides under the set LC programme. Tetra-antennary
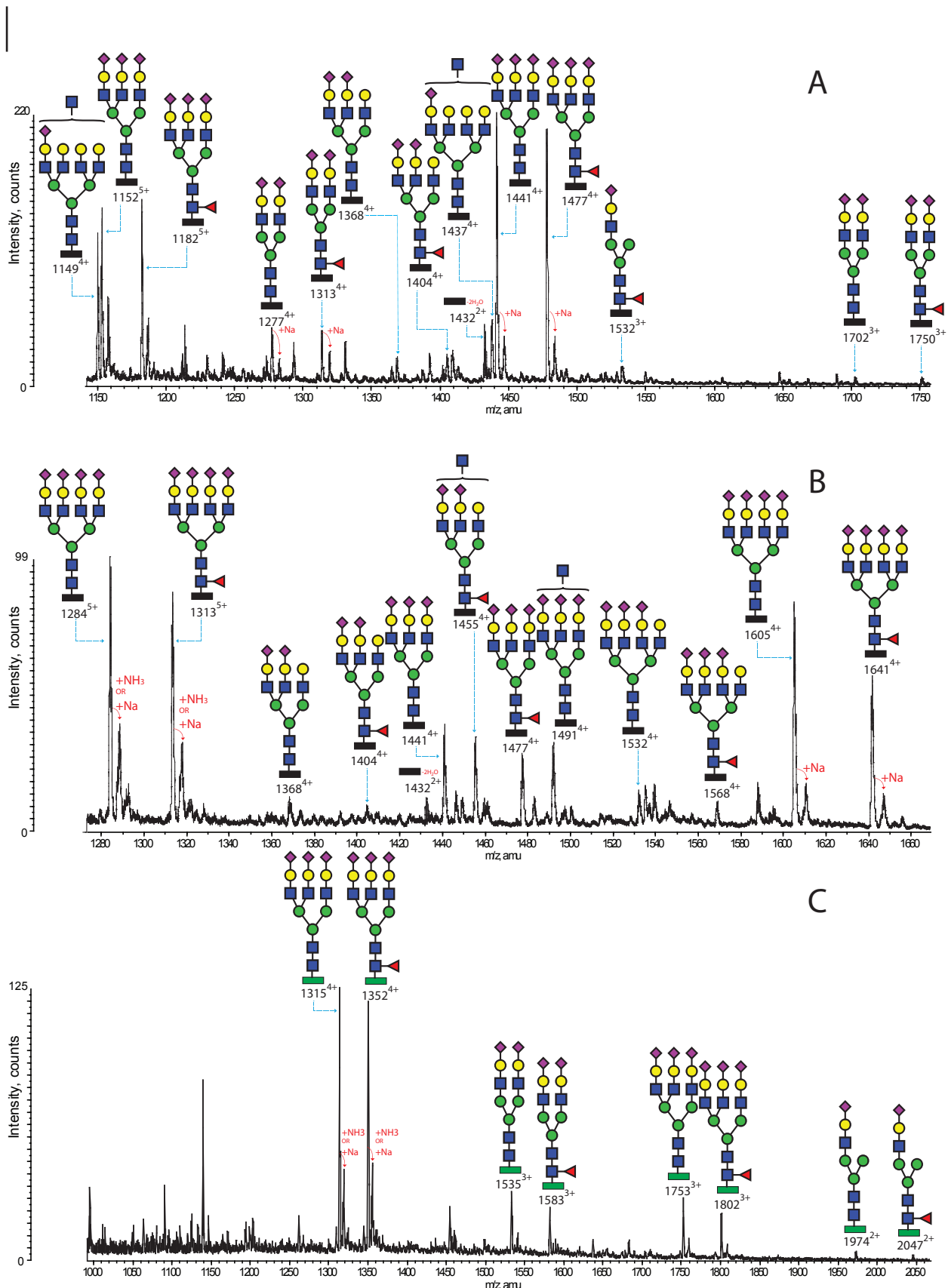
Figure 3.16 Representative portions of LC-ESI-MS spectra for digested products of Asn61 from native PSG1. Fraction A of Product II (A), Fraction B of Product II (B) and Product I (C) for N-glycosylation site Asn61 of native PSG1. Interpretable ions corresponding to glycopeptides are annotated with N-glycan structure and an under bar for peptide backbones. Product I and II are represented by the green and black bars, respectively. Major ions are observed as [M+nH]n+. Additional ammonia or sodium adducts for peaks are also labelled in red characters.

structures with four sialyl-LacNAc antennae are predominant at m/z 1605[4+](m/z 1284[5+]) and 1641[4+] (m/z 1313[5+]). Potentially bisecting GlcNAc is identified on two tri-antennary glycans as m/z 1455[4+] and 1491[4+].

Moreover, identification of the naked peptide backbone, m/z 1432[2+], with double dehydration in the MS spectra of both fractions could be indicative for the occurrence of an in-source fragmentation of eluted glycopeptides. The observed mono- and di-sialylation on tri- and tetra-antennary structures in both fractions could be at least partially contributed by the loss of NeuAc residues from their fully sialylated counterparts. This is because sialic acid residues are labile to loss within the ion-source (Kailemia et al., 2014).

Glycosylated peptides of the Product I were eluted much later into fraction C (Figure 3.16C, Table 3.4) than their single mis-cleaved counterpart, Product II, even though the predominant glycans (m/z 1315[4+] and 1352[4+]) in this fraction are shared with that in fraction A. No structure carrying terminal GlcNAc was detected in this fraction.

In addition, the relative abundance of each glycoform in eluted fractions could not be considered as a reliable reference for accurate quantification in this study. This is because a variety of factors through-out the entire experimental procedure affect the detection and distribution of each glycopeptide. For example, biochemical properties of amino acid sequences could have impacts on the efficiencies of

Table 3.3 N-glycosylation of fraction A &B of Asn61 from native PSG1

N-glycan	Molecular Weight (Da)	Fraction A			Fraction B		
		3+	4+	5+	3+	4+	5+
1NeuAc2LacNAc3Man2GlcNAc	4812	1605					
1NeuAc1LacNAc5Man1Fuc2GlcNAc	4917			1230			
1NeuAc2LacNAc3Man1Fuc2GlcNAc	4958		1653				
1GlcNAc3LacNAc3Man2GlcNAc	5089			1273			
2NeuAc2LacNAc3Man2GlcNAc	5103	1702	1277				
2NeuAc2LacNAc3Man1Fuc2GlcNAc	5249	1750	1313				
1GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNAc	5452		1364				
2NeuAc3LacNAc3Man2GlcNAc	5468		1368			1368	
2NeuAc3LacNAc3Man1Fuc2GlcNAc	5614		1404			1404	
1GlcNAc1NeuAc4LacNAc3Man2GlcNAc	5745		1437	1150			
3NeuAc3LacNAc3Man2GlcNAc	5759	1920	1441	1153		1441	
1GlcNAc2NeuAc3LacNAc3Man1Fuc2GlcNAc	5817					1455	1164
3NeuAc3LacNAc3Man1Fuc2GlcNAc	5905	1969	1477	1182		1477	
1GlcNAc3NeuAc3LacNAc3Man2GlcNAc	5962					1491	
3NeuAc4LacNAc3Man2GlcNAc	6124					1532	
3NeuAc4LacNAc3Man1Fuc2GlcNAc	6270					1568	
4NeuAc4LacNAc3Man2GlcNAc	6415				2139	1605	1284
4NeuAc4LacNAc3Man1Fuc2GlcNAc	6561				2188	1641	1313

Molecular weight corresponds to the designated backbones plus the composition of assigned N-glycans

Table 3.4 N-glycosylation of fraction C of Asn 61 from native PSG1

N-glycan	Molecular Weight (Da)	Fraction C		
		2+	3+	4+
1NeuAc1LacNAc3Man2GlcNAc	3946	1974		
1NeuAc1LacNAc3Man1Fuc2GlcNAc	4092	2047		
1NeuAc2LacNAc3Man2GlcNAc	4311	2157		
1NeuAc2LacNAc3Man1Fuc2GlcNAc	4457	2230		
2NeuAc2LacNAc3Man2GlcNAc	4602	2302	1535	
2NeuAc2LacNAc3Man1Fuc2GlcNAc	4748	2375	1584	
3NeuAc3LacNAc3Man2GlcNAc	5258		1754	1316
3NeuAc3LacNAc3Man1Fuc2GlcNAc	5404		1802	1352

Molecular weight corresponds to the designated backbones plus the composition of assigned N-glycans

proteolytic digestions and peptide ionisation (Mallick et al., 2007). In-source fragmentation on glycans could lead to underestimation or even no detection of relatively labile monosaccharides such as sialic acids. Moreover, various charge states and observable non-covalent adducts (NH_3-17Da and Na-22Da) of glycopeptides could also distribute their signal intensity to a broader range of corresponding peaks causing difficulties for accurate quantitative estimation.

In summary, the N-glycosylation at site Asn61 of native PSG1 exhibits a nearly complete collection of N-glycans identified in previous glycomics experiments (**Section 3.2.2.1**). It has a range of mono- to tetra-antennary complex glycans with substantial hyper-sialylation, and has core-fucosylation on nearly half of identified structures. Besides, bisecting GlcNAc is probably present at low levels on several multi-antennary complex glycans.

3.2.3.3.2 N-glycosylation of Asn61 on recombinant PSG1 N-domains

Comparable to native PSG1, Product I and II for N-glycosylation site Asn61 were also observed in recombinant PSG1N-Fc materials from three different expression systems: Transient CHO-K1, Stable CHO-K1 and Transient Expi-CHO. Glycan structures eluted in each chromatographic fraction for each product are listed in Table 3.5. There are 4 fractions of Product I for both transient Expi-CHO and CHO-K1, but only one fraction for stable CHO-K1. Glycopeptides of Product I from the former two expression systems exhibited an orderly elution with glycans from high-mannose to complex type (from mono- to tetra-sialylation).

Figure 3.17 shows MS spectra of a representative fraction of Product I from each of the transient CHO-K1 and Expi-CHO expression systems. The fraction III (57.2-57.9min) from transient Expi-CHO (Figure 3.17A) exhibits a typical glyco-repertoire with antennary truncation. Major structures in this fraction contain at least one probable terminal GlcNAc residue in their composition (m/z 1433[3+] and 1500[3+]) as well as core-fucosylation. Up to four terminal GlcNAc residues can be found on a single glycan (m/z 2052[2+]). Presence of these structures is evidence for substantial hypo-galactosylation.
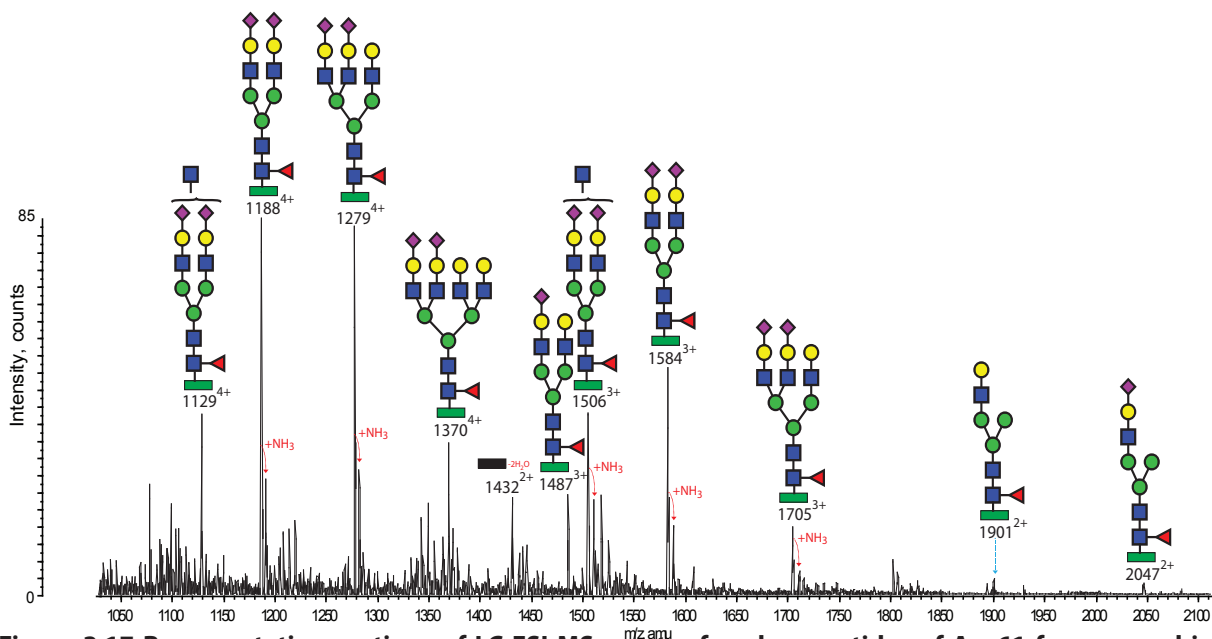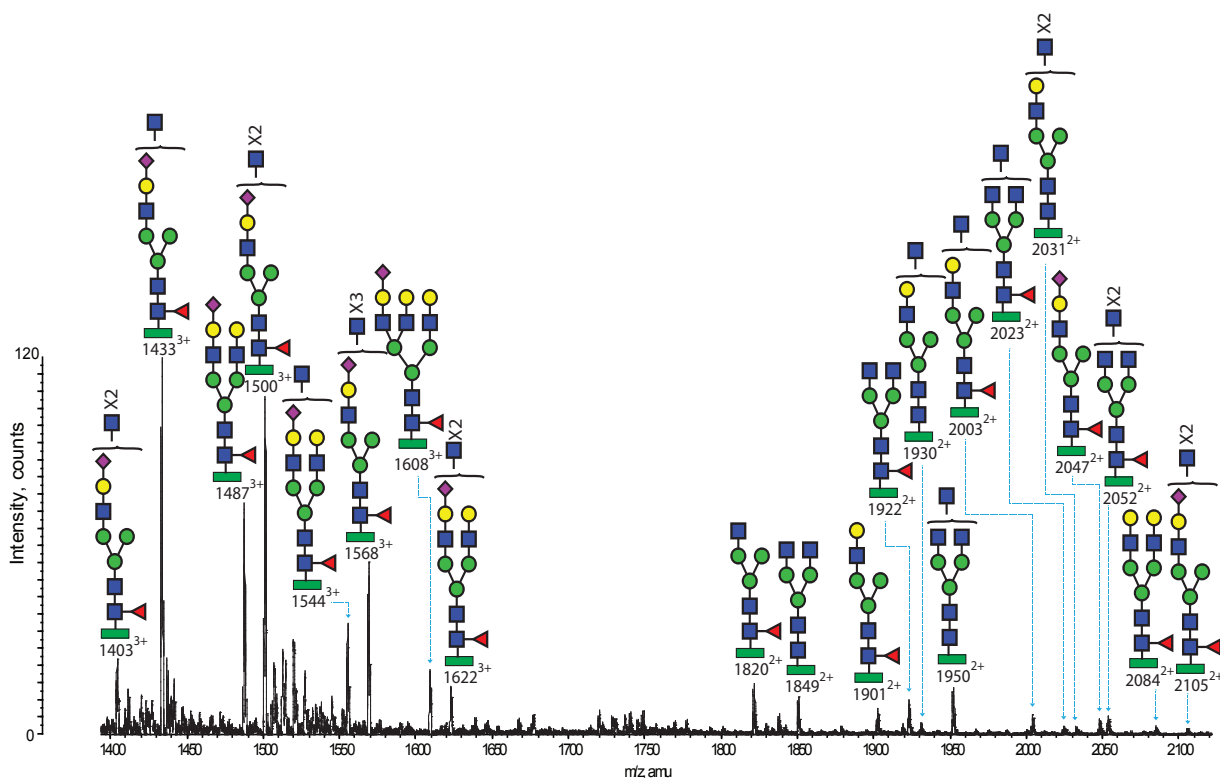
Figure 3.17 Representative portions of LC-ESI-MS spectra for glycopeptides of Asn61 from recombinant PSG1N-Fc expressed in transient CHO cell systems.
Glycopeptides carrying site Asn61 are shown as (A) from transient Expi-CHO system and (B) from transient CHO-K1 system. See the legend of Figure 3.16 for assignments.

Nevertheless, fully galactosylated multi-antennary structures (m/z 1487^{3+} and 1608^{3+}) with mono-sialylation are still present at an intermediate abundance. As the Expi-CHO cell line is designed for producing secretory proteins with high titers, it is likely that there could be a relative shortage of galactosyltransferases and/or UDP-monosaccharide substrates in its glycosylation machinery resulting

Table 3.5 Eluted fractions of Product I and its linking N-glycans from recombinant expression systems

N-glycan	Molecular Weight (Da)	Expi-CHO.T				CHO-K1.T				CHO-K1.S
		I	II	III	IV	I	II	III	IV	I
		25.9'-26.7'	29.0'-29.3'	57.2'-57.9'	76.4'-84.2'	55.6'-56.5'	59.8'-60.7'	64.0'-65.7'	72.1'-74.0'	76.4'-84.2'
1GlcNAc3Man2GlcNAc	3494	x	x				x			
5Man2GlcNAc	3615		x			x				
1GlcNAc3Man1Fuc2GlcNAc	3640	x	x	x		x	x	x		
1LacNAc3Man2GlcNAc	3656					x	x			
2GlcNAc3Man2GlcNAc	3697	x	x	x						
6Man2GlcNAc	3777	x	x	x		x				
1LacNAc3Man1Fuc2GlcNAc	3802			x		x	x	x		
2GlcNAc3Man1Fuc2GlcNAc	3843	x	x	x		x				
1GlcNAc1LacNAc3Man2GlcNAc	3859			x				x		
3GlcNAc3Man2GlcNAc	3900	x	x	x						
7Man2GlcNAc	3939	x	x			x				
1GlcNAc1LacNAc3Man1Fuc2GlcNAc	4005	x		x	x	x	x			
2LacNAc3Man2GlcNAc	4021					x	x			
3GlcNAc3Man1Fuc2GlcNAc	4046	x	x	x						
2GlcNAc1LacNAc3Man2GlcNAc	4062			x						
1NeuAc1LacNAc3Man1Fuc2GlcNAc	4093			x	x		x	x		
8Man2GlcNAc	4101	x	x			x				
4GlcNAc3Man2GlcNAc	4103			x						
2LacNAc3Man1Fuc2GlcNAc	4167	x		x	x	x	x	x		
2GlcNAc1LacNAc3Man1Fuc2GlcNAc	4208	x		x						
1GlcNAc2LacNAc3Man2GlcNAc	4224						x	x		
4GlcNAc3Man1Fuc2GlcNAc	4249	x								

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples.

Table 3.5 Eluted fractions of Product I and its linking N-glycans from recombinant expression systems (continued 1)

N-glycan	Molecular Weight (Da)	Expi-CHO. T				CHO-K1.T				CHO-K1.S
		I 25.9'-26.7'	II 29.0'-29.3'	III 57.2'-57.9'	IV 76.4'-84.2'	I 55.6'-56.5'	II 59.8'-60.7'	III 64.0'-65.7'	IV 72.1'-74.0'	I 76.4'-84.2'
9Man2GlcNAc	4263	X				X				
1GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	4296			X	X		X			
1NeuAc2LacNAc3Man2GlcNAc	4312						X			
1GlcNAc2LacNAc3Man1Fuc2GlcNAc	4370					X		X		
1NeuAc2LacNAc3Man1Fuc2GlcNAc	4458			X	X			X		
2GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	4499			X		X	X			
1GlcNAc1NeuAc2LacNAc3Man2GlcNAc	4515							X		
3LacNAc3Man1Fuc2GlcNAc	4532					X				
2NeuAc2LacNAc3Man2GlcNAc	4603									X
1GlcNAc1NeuAc2LacNAc3Man1Fuc2GlcNAc	4661			X						
3GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	4702			X						
2NeuAc2LacNAc3Man1Fuc2GlcNAc	4749				X			X		X
4LacNAc3Man2GlcNAc	4751								X	
1GlcNAc2NeuAc2LacNAc3Man2GlcNAc	4806									
1NeuAc3LacNAc3Man1Fuc2GlcNAc	4823			X	X		X			
2GlcNAc1NeuAc2LacNAc3Man1Fuc2GlcNAc	4864			X						
1GlcNAc1NeuAc1LacNAc3LacNAc3Man2GlcNAc	4880							X		
1GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNAc	4952				X					
2NeuAc3LacNAc3Man1Fuc2GlcNAc	5114				X			X	X	
5LacNAc3Man2GlcNAc	5116				X					
1NeuAc4LacNAc3Man1Fuc2GlcNAc	5188						X			
3NeuAc3LacNAc3Man2GlcNAc	5259									X

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples.

Table 3.5 Eluted fractions of Product I and its linking N-glycans from recombinant expression systems (continued 2)

N-glycan	Molecular Weight (Da)	Expi-CHO. T				CHO-K1.T				CHO-K1.S
		I 25.9'-26.7'	II 29.0'-29.3'	III 57.2'-57.9'	IV 76.4'-84.2'	I 55.6'-56.5'	II 59.8'-60.7'	III 64.0'-65.7'	IV 72.1'-74.0'	I 76.4'-84.2'
3NeuAc3LacNAc3Man1Fuc2GlcNAc	5405									x
1NeuAc5LacNAc3Man2GlcNAc	5407				x					
2NeuAc4LacNAc3Man1Fuc2GlcNAc	5479							x		
1GlcNAc3NeuAc3LacNAc3Man1Fuc2GlcNAc	5608				x					
3NeuAc4LacNAc3Man1Fuc2GlcNAc	5771				x				x	
2NeuAc5LacNAc3Man1Fuc2GlcNAc	5845								x	

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples.

in observed hypo-galactosylation. These results are consistent with a recent study that Expi-CHO cell lines produced less galactosylated N-glycans than usual on recombinant IgG and increased availabilities of GT and monosaccharide substrate could reversibly improve the level of galactosylation and sialylation (Zhong et al., 2018).

In contrast, the glycopeptides from transient CHO-K1 (Figure 3.17B) are dominated by the fully galactosylated N-glycans with sialylation and core-fucosylation. Major structures are bi- and tri-antennary, and one dominant bi-antennary glycan (m/z 1129^{4+} and 1584^{3+}) could potentially have a bisecting GlcNAc. In addition, a poly-LacNAc extension could be identified on structures in the latter fractions of both materials as containing 5 LacNAc di-saccharide units (Table. 3.5, 5408Da for Epxi-CHO.T and 5846Da for CHO-K1.T).

Elution of glycopeptides of Product II followed a succession similar to that of Product I, whereas more sialylated glycans were eluted with this product (Table 3.6). Tetra-antennary structures with four sialyl-LacNAc are found in both materials as the largest glycan structure (6563Da) linked to the observed tryptic peptide. Several core-afucosylated structures were also detected on this glycopeptide (*e.g.* 5673Da).

Molecular ions for the Asn61 glycopeptides for stable CHO-K1 were of low abundance and only four interpretable structures were observed (Table 3.5 and Figure 3.18A). The poor data quality raised a concern of under-glycosylation of peptides. The MS/MS spectrum for this fraction was scrutinised, whereas only two ions were found correlating to the peptide backbone with trace amounts as B9 m/z 1017 and intact peptide backbone with a GlcNAc residue, m/z 2602 (Figure 3.18B).

Glycosylation of the recombinant Fc tag was subsequently evaluated. The MS spectrum of its Fc N-glycosylation site sequence indicates a glycosylation pattern dominated by core-fucosylated mono- and bi-antennary structures with varied galactosylation levels (Figure 3.19A). Minor fragmentation in the ion source might also occur because m/z 1538^+ is observed, which correlates to the peptide backbone carrying a Fuc-GlcNAc disaccharide.

There are two XIC (Extracted-ion Chromatogram) peaks for 366.1 between 33 to 36 min in the MS/MS data (black dash lined square Figure 3.19B) which are potential LacNAc fragment ions. However, there is no corresponding HexNAc fragment ion (m/z 204) indicating that these peaks are unlikely to be derived from glycopeptides. In addition, a few putative structures were annotated as glycopeptides for Product I and II (Table 3.7, Figure 3.20) after manual scrutiny of their MS spectra. Although their m/z values numerically matched with digested peptides carrying putative glycans, and their m/z intervals to nearby peaks with the same charge state also corresponded to saccharide residues, there were no corroborating fragment ions observed in the MS/MS data. Therefore further experiments are needed to confirm their identities once a larger amount of material is available.

A

B

$$^{50}DVLLLVHNL|PQNLTGYIWYK^{70}$$

B9
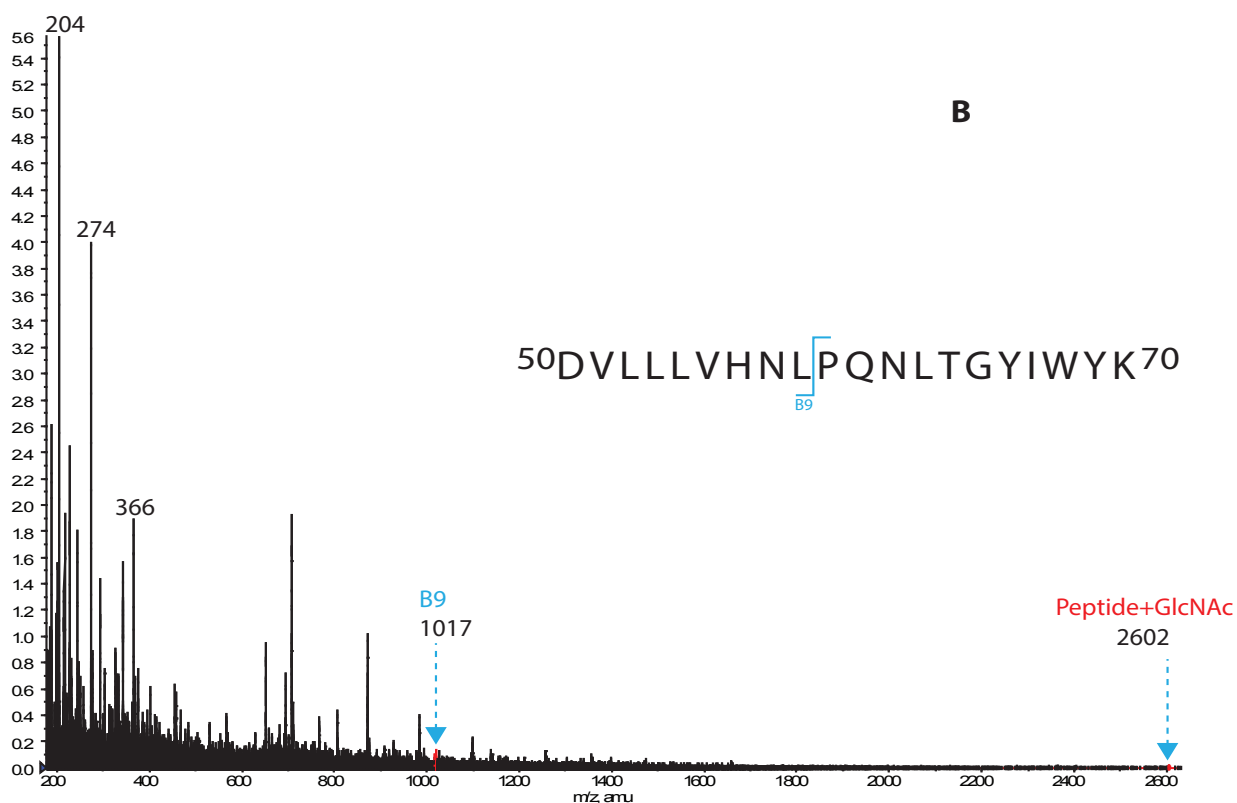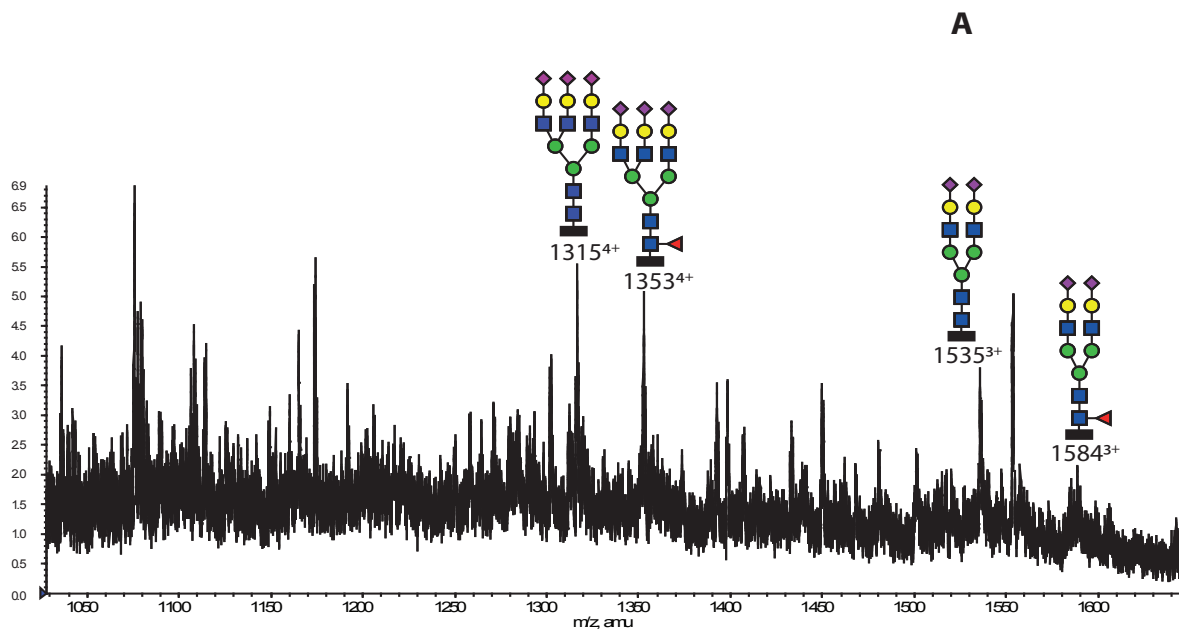
Figure 3.18 Representative portion of ESI-MS/MS and MS spectra of Product I (2398Da) for N-glycosylation site Asn61 of PSG1N-Fc eluted at 76.4-84.2 min from stable CHO-K1 system.

In the MS spectrum (A), four interpretable peaks are assigned with structures based upon knowledge of the biosynthetic pathway and compositional information of glycans and peptide backbone. In the MS/MS spectrum (B), Peaks for saccharide fragment ions are labelled with numbers, one B-ion of the peptide fragment is named in blue letters. The observed intact peptide backbone with a GlcNAc residue is labelled in red letters and the sequence of this glycopeptide Product is at the up-right corner of the panel.
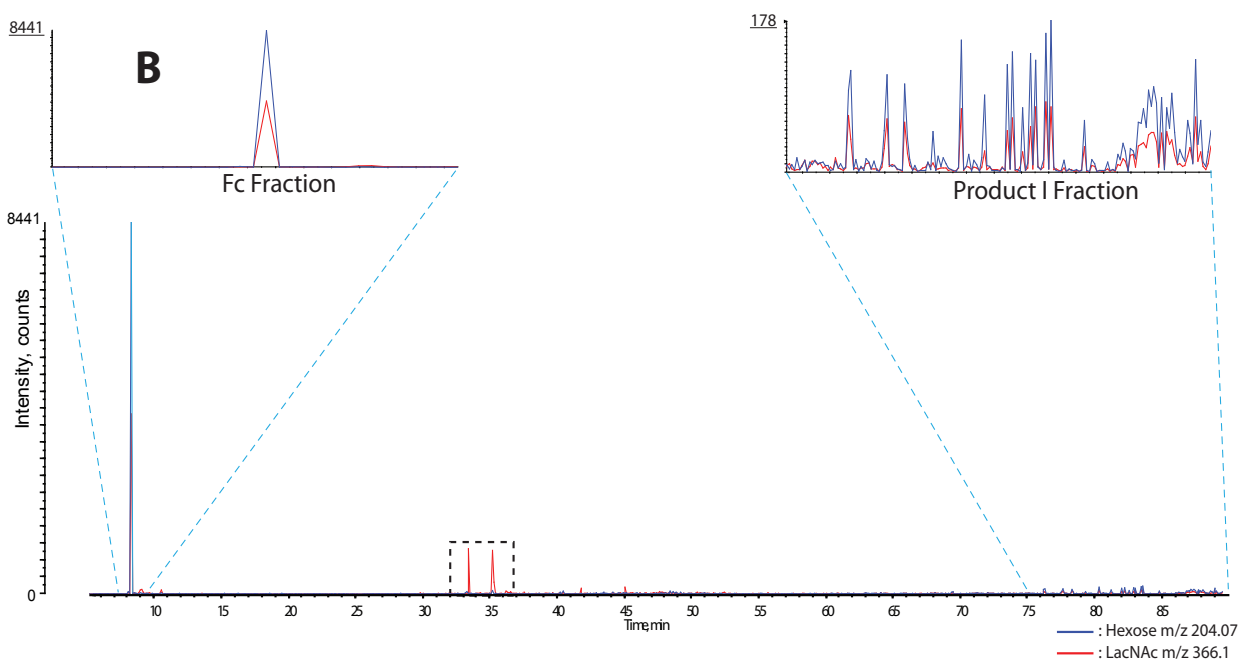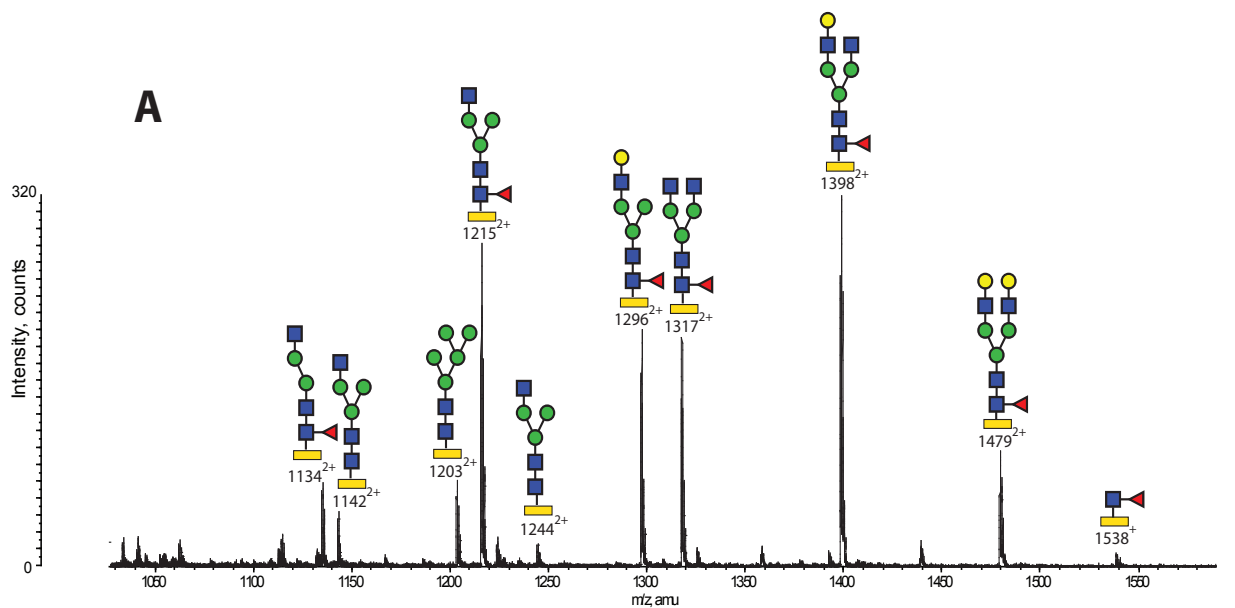
Figure 3.19 Representative portion of ESI-MS spectra of the N-glycopeptides from fused IgG1 Fc-tag of recombinant PSG1N-Fc from stable CHO-K1 system and corresponding overlapped XIC chromatograms of Hex and LacNAc fragments from ESI MS/MS.

The N-glycopeptides eluted at 8.6-8.8 min in the LC system(A) and XIC chromatograms (B) were extracted and overlapped from the MS/MS chromatograms of PSG1-Fc of stable CHO-K1 system. In the MS spectrum (A), the Fc glycopeptide is represented by a yellow bar, assignments are based the on knowledge of biosynthetic pathway and compositional information of IgG peptide and glycans. In the overlapped chromatograms, the region for Fc elution and Asn61 peptide (Product I) elution are zoomed in separate sub-spectra with their ion count labelled on the up-right of their y-axis. The blue line in the chromatograph represents the diagnostic ion 204.06 which corresponds to HexNAc fragment ions, and the red line represents the diagnostic ion 366.1 which correlated to HexHexNAc fragment ions. The dashed black square shows peaks for 366.1 in the absence of 204.07 indicating that the former fragment ion is unlikely to be glycan-derived.

Table 3.6 Eluted fractions of Product II and its linking N-glycans from recombinant expression systems

N-glycan	Molecular Weight (Da)	Expi-CHO.T			CHO-K1.T				
		I	II	III	I	II	III	IV	V
		56.1'-57.1'	58.7'-59.0'	73.5'-75.1'	52.4'-52.7'	54.5'-55.5'	57.7'-59.7'	62.3'-63.3'	68.7'-69.5'
4Man2GlcNAc	3953				X				
5Man2GlcNAc	4115				X				
1GlcNAc3Man1Fuc2GlcNAc	4140					X			
2GlcNac3Man2GlcNac	4197					X	X		
6Man2GlcNAc	4277				X				
1LacNAc3Man1Fuc2GlcNAc	4302					X	X		
1GlcNac1LacNAc3Man2GlcNac	4359					X	X		
7Man2GlcNAc	4439				X				
1NeuAc1LacNAc3Man1Fuc2GlcNAc	4593					X	X		
8Man2GlcNAc	4601				X				
1GlcNac1NeuAc1LacNAc3Man2GlcNac	4650					X	X		
2LacNAc3Man1Fuc2GlcNAc	4667				X	X	X		
1GlcNac2LacNAc3Man2GlcNac	4724				X	X	X		
9Man2GlcNAc	4763				X				
1GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	4796	X				X			
1NeuAc2LacNAc3Man2GlcNac	4812					X			
2GlcNac1NeuAc1LacNAc3Man2GlcNac	4853	X	X				X		
1GlcNAc2LacNAc3Man1Fuc2GlcNAc	4870						X		
1NeuAc2LacNAc3Man1Fuc2GlcNac	4958	X	X			X	X	X	
1GlcNac1NeuAc2LacNAc3Man2GlcNac	5015	X	X	X			X	X	
3GlcNac1NeuAc1LacNAc3Man2GlcNac	5056	X					X		
2GlcNAc2LacNAc3Man1Fuc2GlcNAc	5073							X	
1GlcNac3LacNAc3Man2GlcNac	5089					X			
2NeuAc2LacNAc3Man2GlcNac	5103						X		

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples.

Table 3.6 Eluted fractions of Product II and its linking N-glycans from recombinant expression systems (continued)

N-glycan	Molecular Weight (Da)	Expi-CHO.T			CHO-K1.T				
		I 56.1'-57.1'	II 58.7'-59.0'	III 73.5'-75.1'	I 52.4'-52.7'	II 54.5'-55.5'	III 57.7'-59.7'	IV 62.3'-63.3'	V 68.7'-69.5'
1GlcNAc1NeuAc2LacNAc3Man1Fuc2GlcNAc	5161		x						
1NeuAc3LacNAc3Man2GlcNac	5177					x			
2GlcNac1NeuAc2LacNAc3Man2GlcNac	5218	x	x						
2NeuAc2LacNAc3Man1Fuc2GlcNAc	5249	x	x				x		x
4GlcNac1NeuAc1LacNAc3Man2GlcNac	5259	x							
1GlcNac2NeuAc2LacNAc3Man2GlcNac	5306			x				x	
1NeuAc3LacNAc3Man1Fuc2GlcNac	5323					x			
1GlcNac1NeuAc3LacNAc3Man2GlcNac	5380	x					x		
1GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNac	5453	x	x				x		
2GlcNac2NeuAc2LacNAc3Man2GlcNac	5510		x						
2NeuAc3LacNAc3Man1Fuc2GlcNAc	5615	x	x	x			x	x	
2GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNAc	5656	x	x						
1GlcNac2NeuAc3LacNAc3Man2GlcNac	5672		x	x			x	x	
1NeuAc4LacNAc3Man1Fuc2GlcNac	5689					x			
3GlcNac2NeuAc2LacNAc3Man2GlcNac	5713		x						
1GlcNAc2NeuAc3LacNAc3Man1Fuc2GlcNAc	5818	x	x						
2GlcNac2NeuAc3LacNAc3Man2GlcNac	5875		x						
3NeuAc3LacNAc3Man1Fuc2GlcNac	5906		x	x				x	
1GlcNac3NeuAc3LacNAc3Man2GlcNac	5963			x					x
2NeuAc4LacNAc3Man1Fuc2GlcNac	5980	x	x	x			x		
1GlcNac2NeuAc4LacNAc3Man2GlcNac	6037							x	
1GlcNAc3NeuAc3LacNAc3Man1Fuc2GlcNac	6109		x						
2GlcNac3NeuAc3LacNAc3Man2GlcNac	6166			x					
3NeuAc4LacNAc3Man1Fuc2GlcNAc	6271		x	x				x	
1GlcNac3NeuAc4LacNAc3Man2GlcNac	6328			x					
4NeuAc4LacNAc3Man1Fuc2GlcNAc	6562			x					x

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples .
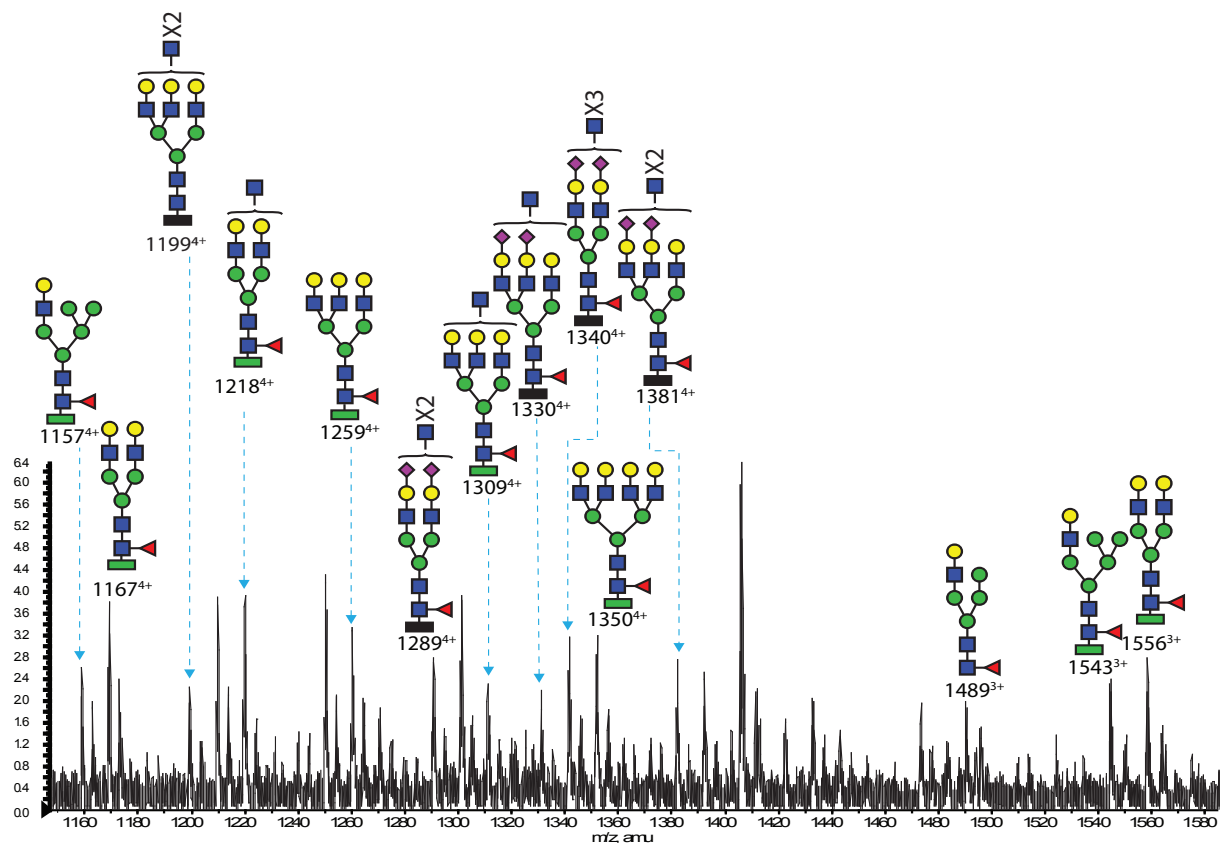
Figure 3.20 ESI-MS spectra of Asn61 containing glycopeptides eluted at 33-36min of PSG1N-Fc from stable CHO-K1 system.

The coloured bar under assigned structures represent different corresponding peptide Products as Product I for black bar (2398Da) and Product II for green bar (2898Da). The mathematically matched structures range from hybrid, bi-, tri- to tetra-antennary complex structures. Core-fucosylation is universal and terminal GlcNAc is placed above brackets to indicate its undefined linkage. See the legend of Figure 3.16 for assignments.

Table 3.7 Putative N-glycan structures at Asn61 of PSG1N-Fc from stable CHO-K1 eluted at 33-36min

N-glycan	Molecular Weight (Da)	Product I	Product II
1LacNAc4Man1Fuc2GlcNac	4463		x
1LacNAc5Man1Fuc2GlcNac	4625		x
2LacNAc3Man1Fuc2GlcNAc	4666		x
2GlcNAc3LacNAc3Man2GlcNAc	4791	x	
1GlcNAc2LacNAc3Man1Fuc2GlcNAc	4869		x
3LacNAc3Man1Fuc2GlcNAc	5031		x
2GlcNAc2LacNAc3Man1Fuc2GlcNAc	5072		x
2GlcNAc2Sia2LacNAc3Man1Fuc2GlcNAc	5154	x	
1GlcNAc3LacNAc3Man1Fuc2GlcNAc	5234		x
1GlcNAc2Sia3LacNAc3Man1Fuc2GlcNAc	5316	x	
3GlcNAc2Sia2LacNAc3Man1Fuc2GlcNAc	5357	x	
4LacNAc3Man1Fuc2GlcNAc	5396		x
2GlcNAc2Sia3LacNAc3Man1Fuc2GlcNAc	5519	x	
5LacNAc3Man1Fuc2GlcNAc	5761		x

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and crosses represent the presence of N-glycans in the designated samples.

3.2.3.4 Glycosylation on Asn104 and 111 of transient CHO-K1

N-glycosylation on sites Asn104 and 111 of PSG1 N-domain was exclusively found on the recombinant PSG1N-Fc materials from transient CHO-K1 system. The tryptic digestion generates a sixteen amino acids backbone product ([99]ETAYSNASLLIQNVTR[114], 1779Da) retaining these two N-glycosylation sites and their N-glycans. There were four fractions (I, II, III and IV) eluted for this Product as at 37.2-38.4min, 41.9-42.6min, 47.0-47.8min and 53.8-54min (Fraction I in Figure 3.21A and Fraction II in 3.21B). Due to there being two potential sites on this Product for N-glycosylation, it is likely to observe singly and/or doubly glycosylated eluents as an occurrence of macro-heterogeneity of glycosylation. A small population of singly glycosylated peptides was eluted along with doubly glycosylated peptides in I and II fractions, especially at fraction II, exhibiting a single high mannose or complex type structures on its backbone (*e.g. m/z* 1161^{3+} and 1304^{3+}). A bi-antennary structure (m/z 1183^{3+}) with core-fucosylation is predominant in this sub-population. Moreover, intensive manual scrutiny enables a detailed listing of all probable structures. Other singly glycosylated structures (Table 3.8) are present in the listing with potential poly-LacNAc as carrying 5 (4787Da) or 6 (5152Da) LacNAc units in their composition.

However, co-elution of singly and doubly glycosylated peptides leads to an unselective fragmentation of mixed parental ions. Thus, observed glycosylated fragment ions as y5 and y7-9 (shown in Figure 3.20C) with a signle GlcNAc residue could not be definitive evidence to suggest a site-specific occupancy.

Doubly glycosylated peptides were differentially eluted into four fractions based on their glycosylation. High-mannose linked peptides were mostly eluted into fraction I at 37.2-38.4min (Figure 3.21A). A series of triply charged ions are observed with an m/z interval of 54, and this m/z interval corresponds to a 162Da Hexose residue. There are up to 18 Hexoses on two putative N-glycans on the peptides, m/z 1836^{3+}, which could correspond to two Man9 high mannose structures (Table 3.8), indicating limited processing of these glycans on these two sites. However, complex glycans were still eluted in the remaining fractions.

CID can induce intensive cleavage of glycans on glycopeptides rather than their peptide backbones. Thereby, no specific antennary arrangements or chitobiose core-modification could unambiguously be assigned for the structure on a specific N-glycosylation site eluted in the III and IV fraction as shown in Figure 3.22. For the doubly glycosylated peptide, maximal five LacNAc units and three sialic acid residues plus two fucoses can be found in the list on two N-glycans as 6554Da (m/z 1639^{3+}, Figure 3.22B, Table 3.8).
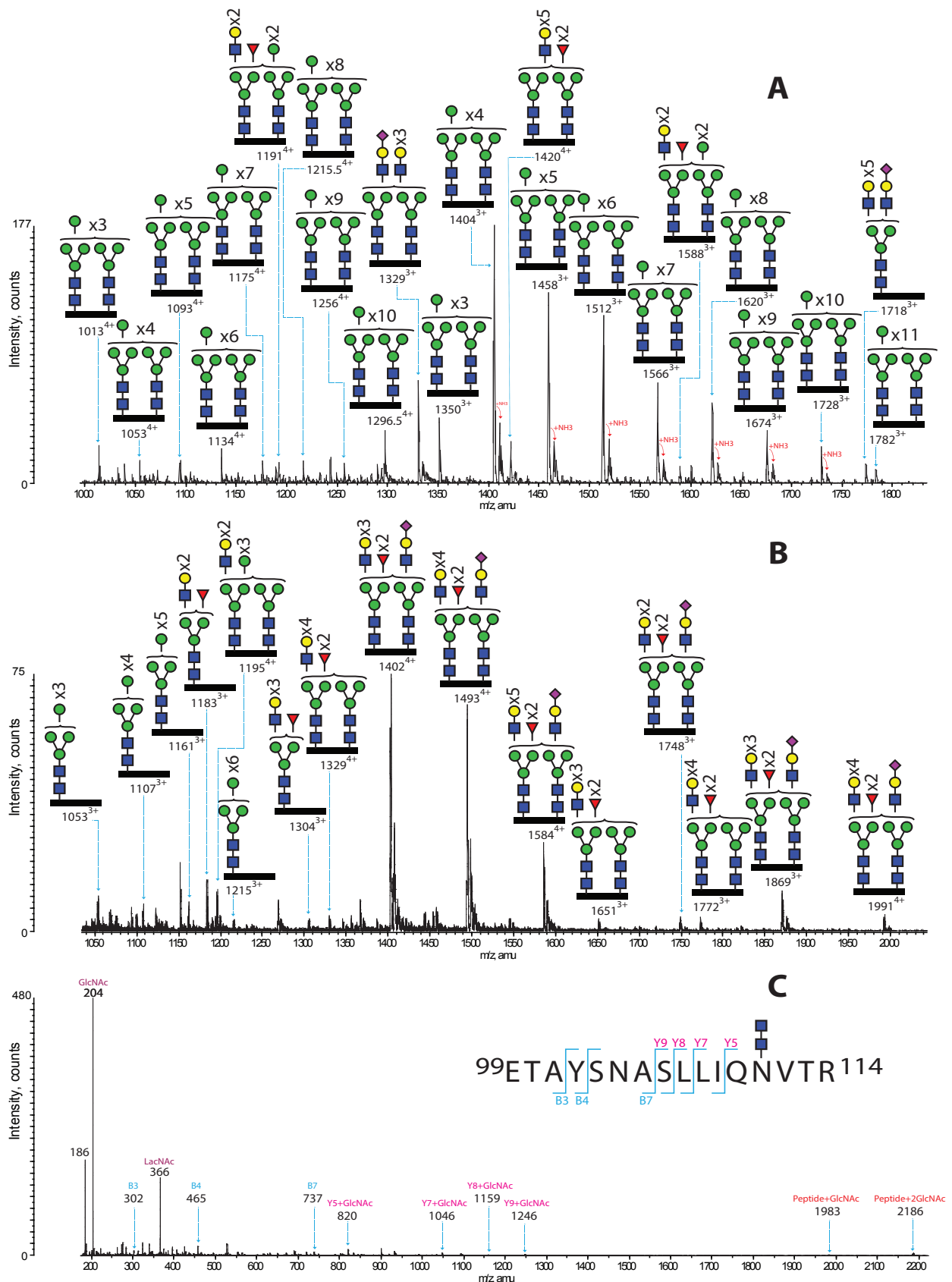
Figure 3.21 Representative portion of ESI-MS spectra of two early fractions for glycopeptides of Asn 104 and 111 and a representative ESI-MS/MS spectrum of the glycopeptide from transient CHO-K1.
The first fraction of N-glycopeptides for Asn 104 and 111 was eluted at 37.2-38.4min(A), and the second fraction eluted at 41.8-42.6min (B). The glycopeptide was fragmentated for peptide identification (C). See the legends of Figure 3.15 &3.16 for assignments.
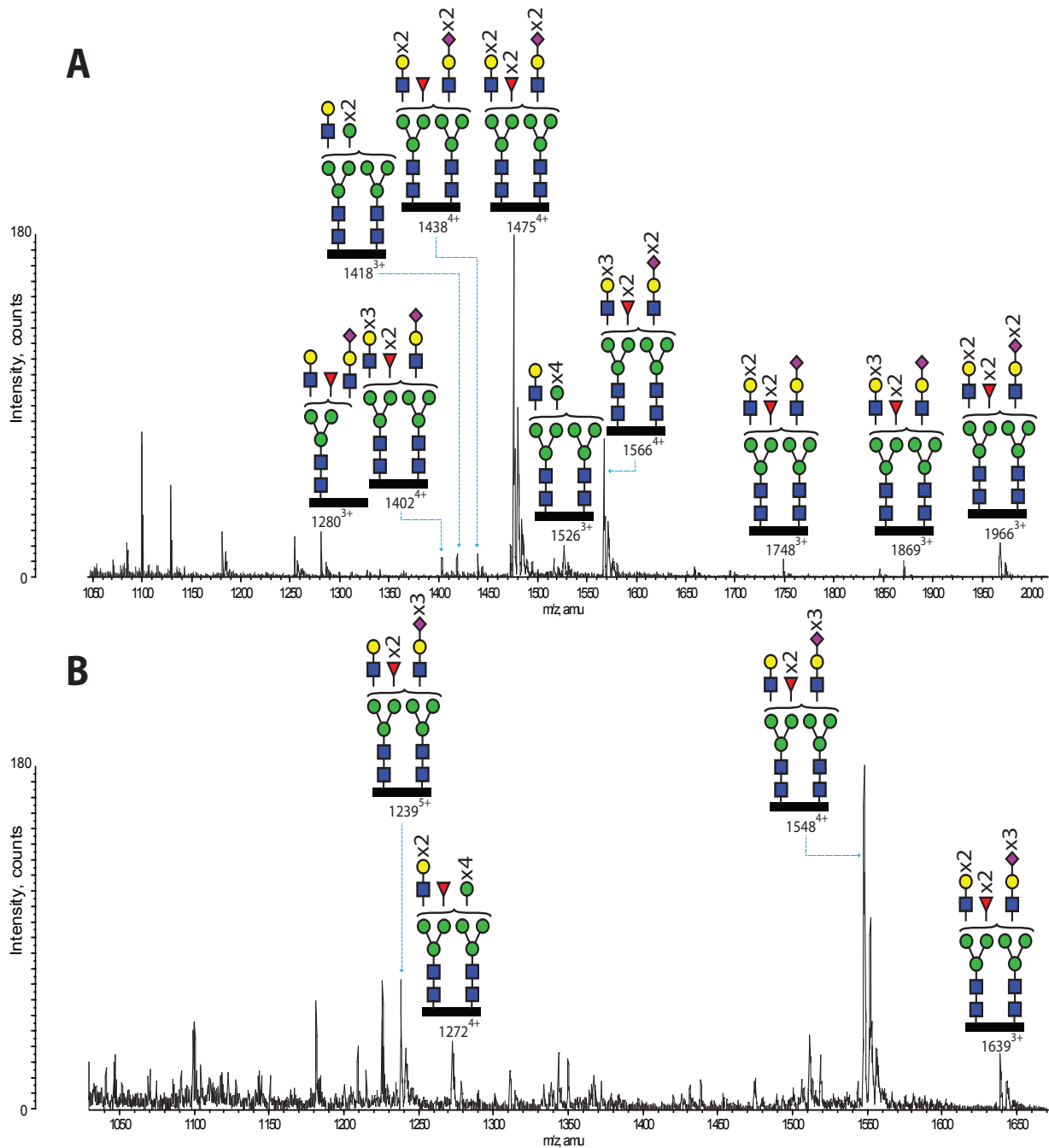
Figure 3.22 ESI-MS spectra of two latter fractions of glycopeptides for Asn 104 and 111 of PSG1N-Fc from transient CHO-K1.

These two fractions of the glycopeptides were sequentially eluted at 47.5-47.8min(A) and 53.8-54.0min(B). See the legend of Figure 3.15 for assignments.

Table 3.8 Fractions of glycopeptides for Asn104 and Asn111 with putative linking N-glycans

1st N-glycan	2nd N-glycan	Molecular weight (Da)	I 37.2'-38.4'	II 41.8'-42.6'	III 47.5'-47.8'	IV 53.8'-54.0'
5Man2GlcNAc	-	2994		x		
6Man2GlcNAc	-	3156		x		
7Man2GlcNAc	-	3319		x		
8Man2GlcNAc	-	3481		x		
2LacNAc3Man1Fuc2GlcNAc	-	3547		x		
9Man2GlcNAc	-	3643		x		
1NeuAc2LacNAc3Man1Fuc2GlcNAc	-	3838			x	
3LacNAc3Man1Fuc2GlcNAc	-	3912		x		
5Man2GlcNAc	3Man2GlcNAc	3886	x			
6Man2GlcNAc	3Man2GlcNAc	4048	x			
5Man2GlcNAc	5Man2GlcNAc	4211	x			
5Man2GlcNAc	1LacNAc3Man2GlcNac	4252			x	
5Man2GlcNAc	6Man2GlcNAc	4373	x			
1LacNAc3Man2GlcNac	1LacNAc3Man1Fuc2GlcNAc	4439		x		
6Man2GlcNAc	6Man2GlcNAc	4535	x			
1LacNAc3Man2GlcNac	7Man2GlcNAc	4576			x	
1LacNAc3Man1Fuc2GlcNAc	1LacNAc3Man1Fuc2GlcNAc	4585	x			
5Man2GlcNAc	2LacNAc3Man2GlcNac	4617	x			
6Man2GlcNAc	7Man2GlcNAc	4697	x			
5Man2GlcNAc	2LacNAc3Man1Fuc2GlcNAc	4763	x			
6Man2GlcNAc	2LacNAc3Man2GlcNac	4779		x		
1NeuAc5LacNAc3Man2GlcNac	-	4787		x		
7Man2GlcNAc	7Man2GlcNAc	4859	x			
6Man2GlcNAc	2LacNAc3Man1Fuc2GlcNAc	4925	x			
1LacNAc3Man1Fuc2GlcNAc	2LacNAc3Man1Fuc2GlcNAc	4950		x		

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples.

Table 3.8 Fractions of glycopeptides for Asn104 and Asn111 with putative linking N-glycans (continued)

1st N-glycan	2nd N-glycan	Molecular weight (Da)	I 37.2'-38.4'	II 41.8'-42.6'	III 47.5'-47.8'	IV 53.8'-54.0'
7Man2GlcNAc	8Man2GlcNAc	5021	x			
7Man2GlcNAc	2LacNAc3Man1Fuc2GlcNAc	5087	x			x
1NeuAc6LacNAc3Man2GlcNac	-	5152	x	x		
8Man2GlcNAc	8Man2GlcNAc	5183	x			
1LacNAc3Man1Fuc2GlcNAc	1NeuAc2LacNAc3Man1Fuc2GlcNAc	5241		x	x	
2LacNAc3Man2GlcNac	1NeuAc2LacNAc3Man2GlcNac	5314	x			
2LacNAc3Man1Fuc2GlcNAc	2LacNAc3Man1Fuc2GlcNAc	5315		x		
8Man2GlcNAc	9Man2GlcNAc	5345	x			
7Man2GlcNAc	1NeuAc2LacNAc3Man1Fuc2GlcNAc	5378		x		
2LacNAc3Man2GlcNac	1NeuAc2LacNAc3Man1Fuc2GlcNAc	5460		x		
9Man2GlcNAc	9Man2GlcNAc	5507	x			
1LacNAc3Man1Fuc2GlcNAc	4LacNAc3Man2GlcNac	5534				x
2LacNAc3Man1Fuc2GlcNAc	1NeuAc2LacNAc3Man1Fuc2GlcNAc	5606		x	x	
2LacNAc3Man1Fuc2GlcNAc	3LacNAc3Man1Fuc2GlcNAc	5680	x			
1NeuAc2LacNAc3Man2GlcNac	1NeuAc2LacNAc3Man1Fuc2GlcNAc	5751			x	
1NeuAc1LacNAc4Man1Fuc2GlcNAc	3LacNAc3Man1Fuc2GlcNAc	5768		x		
1NeuAc2LacNAc3Man2GlcNac	3LacNAc3Man1Fuc2GlcNAc	5825		x		
2LacNAc3Man1Fuc2GlcNAc	2NeuAc2LacNAc3Man1Fuc2GlcNAc	5898			x	x
1LacNAc3Man1Fuc2GlcNAc	1NeuAc4LacNAc3Man1Fuc2GlcNAc	5972		x		
1NeuAc1LacNAc3Man1Fuc2GlcNAc	2NeuAc3LacNAc3Man1Fuc2GlcNAc	6189				x
3LacNAc3Man1Fuc2GlcNAc	2NeuAc2LacNAc3Man1Fuc2GlcNAc	6263			x	
3LacNAc3Man1Fuc2GlcNAc	1NeuAc3LacNAc3Man1Fuc2GlcNAc	6337		x		
2NeuAc2LacNAc3Man1Fuc2GlcNAc	1NeuAc3LacNAc3Man1Fuc2GlcNAc	6554				x

The molecular weight corresponds to the designated peptide plus assigned N-glycans, and the crosses represent the presence of N-glycans in the designated samples.

3.2.3.5 Glycosylation on Asn199 for native PSG1

The N-glycosylation site Asn199 is located in the A1 domain of full-length native PSG1. Tryptic digestion gives a hexapeptide backbone ([194]LESTNR[200], 718Da) and its glycoforms eluted at 11.2-11.7min (Figure 3.23A). The N-glycans on this peptide are exclusively mono- and bi-antennary structures and major components are sialylated (*e.g.*1023^{3+}, 1134^{2+} and 1462^{2+}). Core-fucosylation is found on two dominant structures (*m/z* 1207^{2+} and 1535^{2+}). The structure of *m/z* 1235^{2+} is putatively assigned as bi-antennary with a truncated antenna, and other minor structures are also listed in Table 3.9. As a representative example, the MS/MS spectrum of *m/z* 1976 is shown in Figure3.23B. A sequential decrement from a putative hepta-saccharide (LacNAc3Man2GlcNAc, *m/z* 1976) to a single GlcNAc (*m/z* 922) is the major fragmentation observed. Lack of peptide fragmentation is to be expected because a singly charged ion was selected for MS/MS.
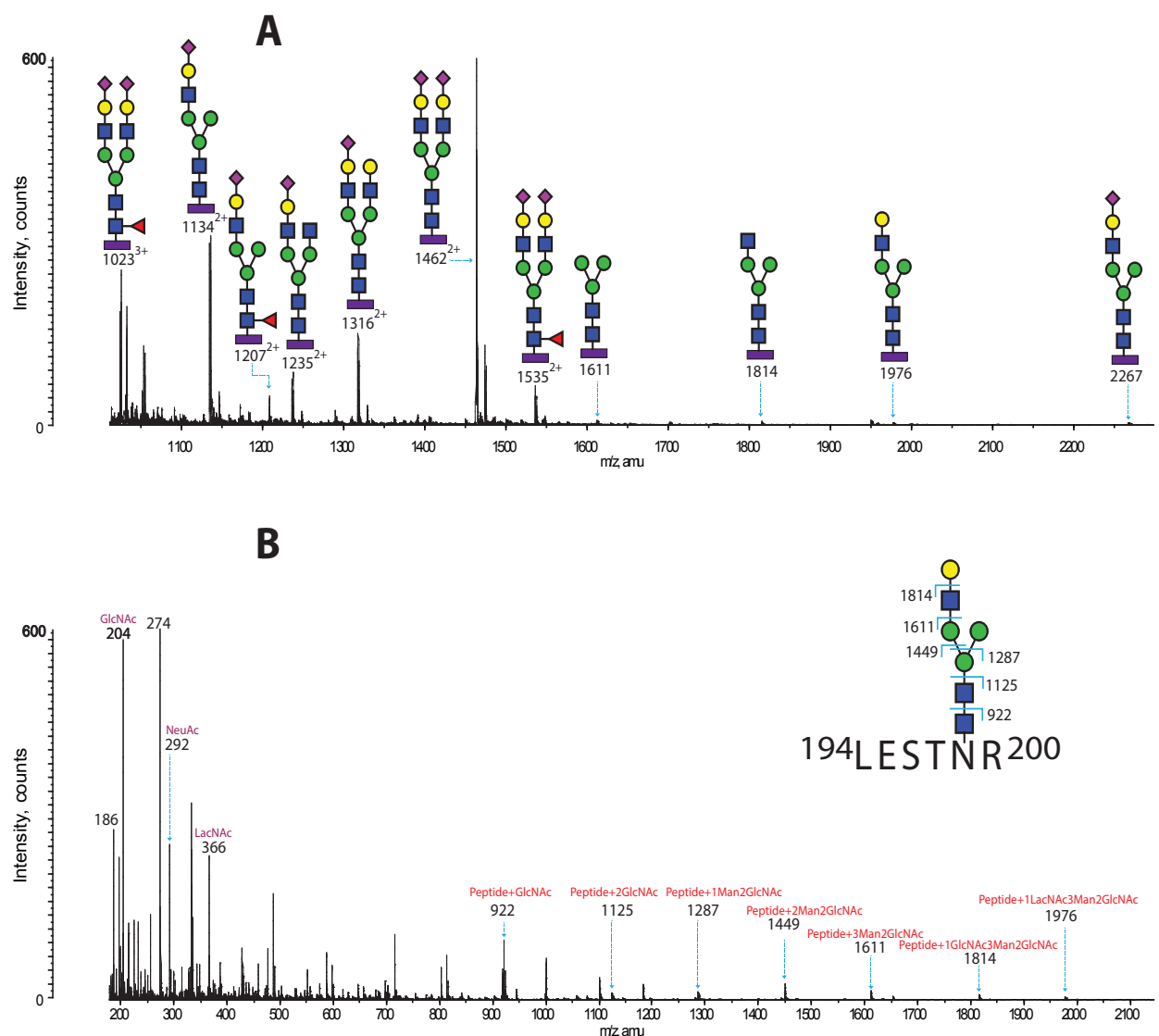
Figure 3.23 ESI-MS spectrum for the glycopeptides of Asn199 on native PSG1 and corresponding ESI-MS/ MS spectrum.
The glycopeptides eluted at 11.2-11.7min as one fraction(A) and fragmentation analysis was conducted for glycopeptide identification(B). See the legends of Figure 3.14 and 3.15 for assignments.

Table 3.9 N-glycan structures at Asn199 of native PSG1

N-glycan	Molecular weight (Da)	1[+]	2[+]	3[+]
3Man2GlcNAc	1610	1611		
1GlcNAc3Man2GlcNAc	1813	1814	908	
1LacNAc3Man2GlcNAc	1975	1976		
1GlcNAc1LacNAc3Man2GlcNAc	2178	2179	1090	
1NeuAc1LacNAc3Man2GlcNAc	2266		1134	
2LacNAc3Man2GlcNAc	2340	2341	1171	
1NeuAc1LacNAc3Man1Fuc2GlcNAc	2412	2413	1207	
1GlcNAc1NeuAc1LacNAc3Man2GlcNAc	2469	2470	1236	
1GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	2615	2616	1309	
1NeuAc2LacNAc3Man2GlcNAc	2631	2632	1317	
1NeuAc2LacNAc3Man1Fuc2GlcNAc	2777	2778	1390	
2NeuAc2LacNAc3Man2GlcNAc	2922	2923	1462	
2NeuAc2LacNAc3Man1Fuc2GlcNAc	3068	3069	1535	1024

The molecular weight corresponds to the designated peptide plus assigned N-glycan

3.2.3.6 Glycosylation at Asn268 for native PSG1

This twenty amino-acid peptide was cleaved from the A2 domain of full-length native PSG1 protein sequence by tryptic digestion ([266]SENYTYIWWLNGQSLPVSPR[285], 2409Da). Two fractions of this glycopeptide were eluted at 62-70min and 83-90min (Figure 3.24 A&B). Glycans eluted in the former fraction are exclusively mono- and bi-antennary structures. Their LacNAc antennae are generally capped by terminal sialic acids, and core-fucosylation is on about half of detected structures in this fraction (*e.g. m/z* 1191[4+] and 1369[3+]). Glycans of the latter fraction share these features and two more fully sialylated tri-antennary structures were identified at *m/z* 1318[3+] and 1354[3+] shown in Figure 3.24B as well as in Table 3.10. In addition, potentially GlcNAc bisecting bi-antennary structures are also found in these two fractions, such as *m/z* 1605[3+] and 1655[3+]. Another potentially GlcNAc bisecting structure is also perceivable in the latter fraction at *m/z* 1558[3+]. However, this mono-sialylated structure was only eluted in the latter fraction, and as described previously it is expected that well-sialylated structures should elute later. Thus, it is likely that this structure could be a product of in-source fragmentation from its bi-sialylated counterpart at *m/z* 1655[3+]. In addition, a representative MS/MS spectrum is shown in Figure 3.24C. Dominant fragmentation y5 and y9 were observed as *m/z* 555 and 940; intact peptide with a GlcNAc residue was also detected as *m/z* 2613. Besides, deamidation on asparagine residue 277 (N[277] to D[277]) was identified by observation of a mass shift for 1Da on y10 to y13 ions.
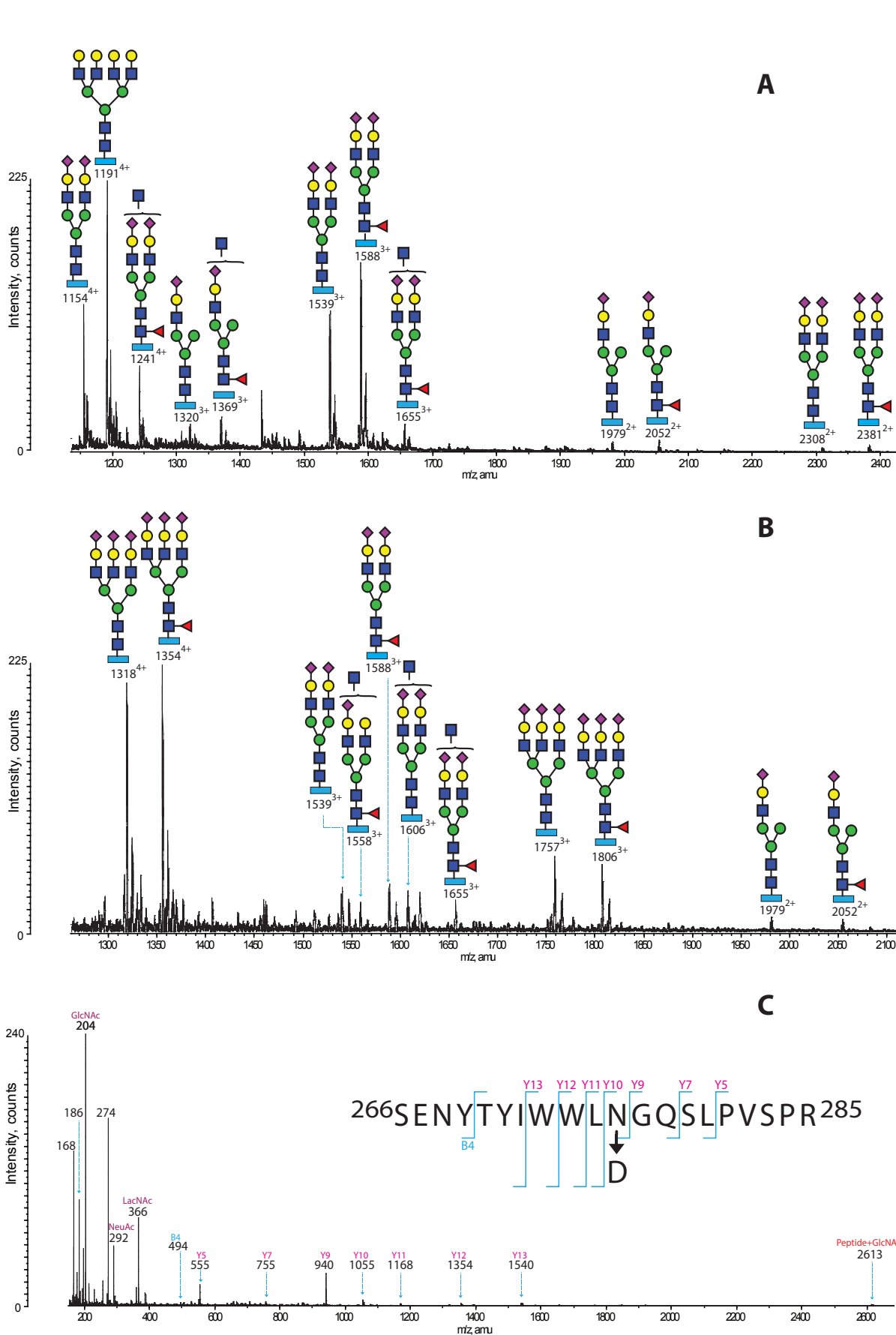
Figure 3.24 ESI-MS spectra of Asn268 N-glycopeptides and a corresponding ESI-MS/MS spectrum from native PSG1.

N-glycopeptides of Asn268 eluted at 62-70min(A) and 83-90min(B); fragmentation was conducted for glycopeptides identification as shown in a representative spectrum(C). See the legends of Figure 3.14 and 3.15 for assignments

Table 3.10 N-glycosylation at Asn 268 of Native PSG1

N-glycan	Molecular Weight (Da)	62-70min			83-90min		
		2+	3+	4+	2+	3+	4+
1NeuAc1LacNAc3Man2GlcNAc	3957	1980	1320		1980		
1NeuAc1LacNAc3Man1Fuc2GlcNAc	4103	2053	1369		2053		
1GlcNAc1NeuAc1LacNAc3Man2GlcNAc	4160				2081		
1GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	4306	2154					
1NeuAc2LacNAc3Man1Fuc2GlcNAc	4468	2235	1490				
2NeuAc2LacNAc3Man2GlcNAc	4613	2308	1539	1154	2308	1539	
1GlcNAc1NeuAc2LacNAc3Man1Fuc2GlcNAc	4671					1558	
2NeuAc2LacNAc3Man1Fuc2GlcNAc	4761	2382	1588	1191			
1GlcNAc2NeuAc2LacNAc3Man2GlcNAc	4816		1606	1205		1606	
1GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNAc	4962	2482	1655	1242			
2NeuAc3LacNAc3Man1Fuc2GlcNAc	5124		1709	1282			
1GlcNAc2NeuAc3LacNAc3Man2GlcNAc	5181						1296
3NeuAc3LacNAc3Man2GlcNAc	5269				2636	1757	1318
1GlcNAc2NeuAc3LacNAc3Man1Fuc2GlcNAc	5327					1777	1333
3NeuAc3LacNAc3Man1Fuc2GlcNAc	5415				2709	1806	1355

The molecular weight corresponds to the designated peptide plus assigned N-glycan

3.2.3.7 Glycosylation at Asn303 for native PSG1

Asn303 is located in the A2 domain of full-length native PSG1. A double miss-cleavage resulted in the observed peptide backbone [294]ILILPSVTRNETGPYQCEIRDR[315] (2629Da). There was one fraction of glycopeptides which eluted at 43.3-43.5min corresponding to this peptide backbone (Figure 3.25A). Only mono- and bi-antennary structures are found in this fraction. Consistently, core-afucosylation is found on a small group of these glycans while sialylation is universal on all structures. A potential GlcNAc bisecting structure is observed as *m/z* 1296[4+]; antennary truncation is also observable as *m/z* 1510[3+]. All identified structures are listed in Table 3.11. In addition, the fragmentation pattern of this peptide (2629Da, Figure 3.25B) is distinctive among observed MS/MS spectra as it has a collection of singly, doubly and triply charged fragment ions, including *m/z* 559, 698[2+] and 994[3+]. This feature is evident to that there are multiple basic amino acid residues, likely lysine or arginine, on this peptide. Attentive scrutiny of this MS/MS spectrum matching with annotated PSG1 sequence from UniProt (P11464) confirmed the identity of this peptide and its double mis-cleavage status. The observed peptide resulted from mis-cleavages of both N- and C-terminal of Product [303]NETGPYQ-CEIR[313] (Cysteine was carboxymethylated, +58Da). Typical b and y ions are assigned on Figure 3.24C to partially map out the entire sequence.

Table 3.11 N-glycosylation on Asn303 of Native PSG1

N-glycan	Molecular Weight (Da)	3+	4+
1LacNAc3Man1Fuc2GlcNAc	4033	1345	
1NeuAc1LacNAc3Man2GlcNAc	4178	1394	
1GlcNAc1LacNAc3Man1Fuc2GlcNAc	4236	1413	
2LacNAc3Man2GlcNAc	4252	1418	
1NeuAc1LacNAc3Man1Fuc2GlcNAc	4324	1442	
1GlcNAc1NeuAc1LacNAc3Man2GlcNAc	4381	1461	
2LacNAc3Man1Fuc2GlcNAc	4398	1467	
1GlcNAc1NeuAc1LacNAc3Man1Fuc2GlcNAc	4527	1510	
1NeuAc2LacNAc3Man2GlcNAc	4543	1515	1137
1NeuAc2LacNAc3Man1Fuc2GlcNAc	4689	1564	1173
2NeuAc2LacNAc3Man2GlcNAc	4834	1612	1210
2NeuAc2LacNAc3Man1Fuc2GlcNAc	4982	1662	1247
1GlcNAc2NeuAc2LacNAc3Man1Fuc2GlcNAc	5183	1729	1297

The molecular weight corresponds to the designated peptide plus assigned N-glycan
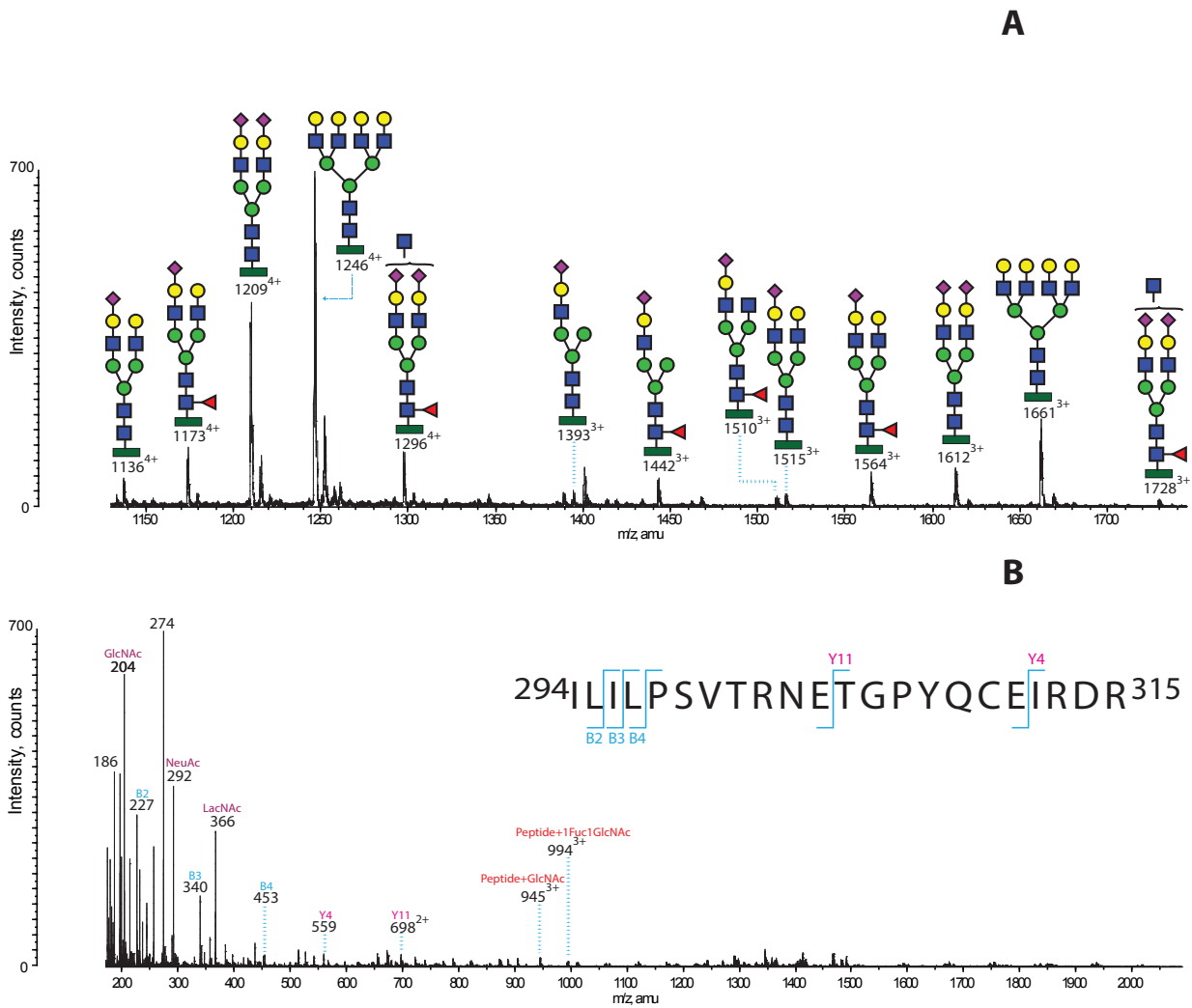
Figure 3.25 ESI-MS spectrum for glycopeptides of Asn303 from native PSG1 and corresponding ESI-MS/MS spectrum

The glycopeptides were eluted at 43.3-43.5min(A) and corresponding fragments were analysed for glycopeptide identification as shonw in (B). See the legends of Figure 3.14 and 3.15 for assignments.

3.3 Discussion

In this chapter, we firstly report a complete N-glycan repertoire of native PSG1 (pooled from healthy donors, **Section 3.2.2.1**) and preliminary N-glycosylation site mapping for native PSG1 (**Section 3.2.3.3.1; 3.2.3.5; 3.2.3.6** and **3.2.3.7**) by mass spectrometry. The recombinant N-domains of this protein were also analysed for N-glycan profiling and site mapping (**Section 3.2.3.3.2** and **3.2.3.4**). This study combined glycomics and glycoproteomics techniques utilising systematic mass spectrometric strategies based on MALDI-TOF/TOF and LC-ESI-MS systems. Our mass spectrometric data demonstrates that the systematic mass spectrometric methodologies are still advantageous in structural and sequential analyses of highly heterogeneous macromolecules such as glycans and glycoproteins.

3.3.1 Glycosylation of native PSG1

3.3.1.1 N-glycosylation

Our data illustrates that native PSG1 N-glycans are dominated by complex structures ranging from mono- to tetra-antennary. As agreed with previous findings in the literature (**Section 3.1**), mono- and bi-antennary structures are characteristic in the overall population; and nearly half of all structures have core-fucosylation.

Antennary modifications are highlighted as sialylation which is virtually universal on N-glycans with mainly α2,3 linkages. Potential GlcNAc bisecting is on a minor group of N-glycans; poly-LacNAc extension is also observed on several glycans after sialidase A digestion. Precedent glycomics studies of mammalian placental materials indicated that these antennary modifications were generally consistent on placental N-glycans across species (Chen et al., 2016, Jones et al., 2007). GlcNAc bisecting and poly-LacNAc extension were verified on N-glycans from human cyto-(CTB) and syncytio-trophoblast (STB) cells by GC-MS linkage and end-β-galactosidase experiments in a previous study (Chen et al., 2016). As PSG1 proteins are naturally secreted by these trophoblastic cells, it is logical to reason that the same modifications could be retained on PSG1 N-glycans as illustrated in the result section.

However, identified high-mannose structures in the glycomics of these native materials could be contaminations from other serum glycoproteins because there was no detection of high-mannose linking glycopeptides from PSG1 via glycoproteomics (**Section 3.2.3**). Simultaneously, Mascot search for this material identified other serum glycoproteins, such as IgM (Table 3.12). The presence of high-mannose structures could be attributed to the N-glycosylation of IgM (Arnold et al., 2005) and other presented glycoproteins in the list.

Table 3.12 Mascot proteomics search result list for native PSG1 (the batch presented for N-glycan analysis)

prot_accession	prot_description	Protein score
PSG1_HUMAN	Pregnancy-specific beta-1-glycoprotein 1 OS=Homo sapiens GN=PSG1 PE=2 SV=1	474
PSG3_HUMAN	Pregnancy-specific beta-1-glycoprotein 3 OS=Homo sapiens GN=PSG3 PE=2 SV=2	420
PSG4_HUMAN	Pregnancy-specific beta-1-glycoprotein 4 OS=Homo sapiens GN=PSG4 PE=2 SV=3	366
PSG2_HUMAN	Pregnancy-specific beta-1-glycoprotein 2 OS=Homo sapiens GN=PSG2 PE=2 SV=2	341
TRFE_HUMAN	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=3	341
PSG11_HUMAN	Pregnancy-specific beta-1-glycoprotein 11 OS=Homo sapiens GN=PSG11 PE=2 SV=3	291
PSG8_HUMAN	Pregnancy-specific beta-1-glycoprotein 8 OS=Homo sapiens GN=PSG8 PE=2 SV=2	267
PSG5_HUMAN	Pregnancy-specific beta-1-glycoprotein 5 OS=Homo sapiens GN=PSG5 PE=1 SV=3	243
FETUA_HUMAN	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	222
PSG7_HUMAN	Putative pregnancy-specific beta-1-glycoprotein 7 OS=Homo sapiens GN=PSG7 PE=5 SV=2	189
HPT_HUMAN	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	181
IGHG3_HUMAN	Ig gamma-3 chain C region OS=Homo sapiens GN=IGHG3 PE=1 SV=2	148
IGHA1_HUMAN	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	128
MUCB_HUMAN	Ig mu heavy chain disease protein OS=Homo sapiens PE=1 SV=1	123
IGHA2_HUMAN	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	90
IGHG1_HUMAN	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1	90
IGHM_HUMAN	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	87
PSG9_HUMAN	Pregnancy-specific beta-1-glycoprotein 9 OS=Homo sapiens GN=PSG9 PE=2 SV=2	84
A1AG1_HUMAN	Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1	62
LAC1_HUMAN	Ig lambda-1 chain C regions OS=Homo sapiens GN=IGLC1 PE=1 SV=1	51
IGKC_HUMAN	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	50
PSG6_HUMAN	Pregnancy-specific beta-1-glycoprotein 6 OS=Homo sapiens GN=PSG6 PE=2 SV=1	44
IGHG2_HUMAN	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2	40
KV113_HUMAN	Ig kappa chain V-I region Lay OS=Homo sapiens PE=1 SV=1	37

3.3.1.2 O-glycosylation

The status of O-glycosylation of native PSG1 is unclear due to the inconsistent observations of O-glycosylation between different batches of native PSG1 materials. These native PSG1 materials were collected and pooled from donors. Natural variations of O-glycosylation could exist in these materials from different individuals due to the possible heterogeneity of their genetic backgrounds. Besides, various trophoblast cells can secrete PSG1 proteins during gestation. CTB cells are capable of producing hyper-glycosylated secretory proteins with heavy N- and O-glycosylation as hyperglycosylated human chorionic gonadotropin (H-hCG) at the first trimester of gestation (Evans et al., 2015, Handschuh et al., 2007, Camolotto et al., 2010). As CTB is also a source of PSG1, it might assemble O-glycans onto PSG1 proteins as H-hCG during the same period of pregnancy recruiting a shared glycosylation machinery. Glycoproteomics experiments for the latest batch of samples (without detection of O-glycans

in Glycomics) helped to confirm that there is no glycopeptide from native PSG1 proteins carrying O-glycan. Because there is neither a fragmentation ion corresponding to serine/threonine glycosylation in MS/MS nor a molecluar ion corresponding to a potentially co-presence of N- and O-glycosylation in MS. However, this approach for verifying the O-glycosylation of the other analysed native PSG1 materials is unachievable, due to no availability of previous batches of samples for this experiment.

In future, another glycoproteomics experiment employing electron transfer dissociation (ETD) as the fragmenting technique for further native PSG1 preparations could provide another layer of information to address the presence of O-glycosylation (Mechref, 2012). Briefly, ETD MS/MS offers a characteristic fragmentation of glycopeptides producing CZ ions. Instead of a dominant fragmentation of glycans in CID, ETD mainly leads to fragmentation of peptide backbones leaving glycan sequences mostly intact. Both N- and O-glycosylations could subsequently be preserved and detected on fragmentated glycopeptides. Simultaneously, the site-occupancy for both N- and O-glycosylations could also be elucidated.

3.3.1.3 Site-specific N-glycosylation of PSG1

Site-specific glycosylation on the native PSG1 sequence is preliminarily elucidated from the results of glycoproteomics experiments (**Section 3.2.3**). Characterised PSG1 N-glycans are not uniformly distributed on N-glycosylation sites. Generally, all the N-glycosylated sites (Asn 61, 199, 268 and 303) have mono- and bi-antennary structures, whereas further branched tri- and tetra-antennary structures are found on two of the four N-glycosylation sites. N-glycosylation sites Asn 61 and 268 on N and A2 domains exhibit tri-antennary structures, respectively; tetra-antennary structures are exclusively found on site Asn61. This site-specific glycosylation of multi-antennary structures could correlate to differentiated accessibility of each N-glycosylation site for glycosyltransferases and monosaccharide substrates. Interestingly, excluding site Asn61 on the N-domain, all other three N-glycosylation sites have a nearby cysteine residue around their conserved sequons, especially sites Asn 268 (C^{262}) and 303 (C^{310}). Some early studies suggested that cysteine residues near to N-glycosylation sequons could be capable of impairing or even inhibiting N-glycosylation processing via introducing conformational constraints by local disulfide bonding (Bause et al., 1982, Hasegawa et al., 1992, Mellquist et al., 1998). The relatively restricted branching of N-glycans on these three sites (Asn199, 268 and 303) could be a consequence of this hindering effect.

Bisected GlcNAc is also observed on N-glycans at sites Asn 61, 268 and 303 while not Asn199. The lack of such modification on Asn199 could indicate a more severe conformational restriction on this site as bisected GlcNAc is sequentially assembled on N-glycans after the formation of N-glycan di-antennae (**Section 1.3.1.3**).

Unlike the site-specific N-glycan branching, partial core-fucosylation is generally uniform on N-gly-

cans at all glycosylated sites. The level of core-fucosylation of PSG1 might associate with physiological conditions. For instance, patience with gestational diabetes mellitus was found to have PSG1 with elevated level of core-fucosylation (Ngalla et al., 2015). Elevated core-fucosylation was found on the α subunit of hCG (hCGα), which is a protein likely to be secreted from the same trophoblast cells as PSG1, around 14-17 weeks after the last menstrual period (Nemansky et al., 1998). Moreover, the β subunit of hCG (hCGβ) from patients with testicular cancer or mole invasion exhibits even a site-specific elevation of core-fucosylation on N-glycans at its Asn13 N-glycosylation site (Valmu et al., 2006).

3.3.2 Functional implication of native PSG1 N-glycan structures

As summarised in the previous **Section 3.3.1**, sialylation, partial core-fucosylation, GlcNAc bisection and mono-antennary structures are highlighted as characteristic for the N-glycosylation of native PSG1. These features of PSG1 N-glycans might correlate with the potential functions of PSG1 in placental developments and immune regulation to ensure a successful pregnancy.

3.3.2.1 Potential functions of NeuAc sialylation

Sialylation is beneficial for secretory glycoproteins in circulation as it can help to extend their retention time in serum (Fukuda et al., 1989). In addition, most immune cells express siglec family receptors to monitor environmental sialic acid ligands (Macauley et al., 2014). Ligation with siglec receptors enables a "self" recognition of ligand conjugates to maintain an immunomodulatory effect. For instance, induction of siglec 1 and 9 on macrophages resulted in increased production of immunosuppressive cytokines such as TGF-β1 and IL-10 (Wu et al., 2016, Matsumoto et al., 2016). Notably, these two siglec receptors have a shared reactivity to α 2,3-linked sialyl-LacNAc moieties (Blixt et al., 2003, Crocker et al., 2007). As native PSG1 proteins have a high content of such moieties, it is likely that the PSG1 induced productions of these cytokines during pregnancy could employ these siglec signalling mechanisms (Blois et al., 2014). Additionally, PSG1 could also activate DC to produce TGF-β and IL-10 for immunomodulation (Martinez et al., 2012). As DCs also share the expression of siglec 9 with macrophages, it is likely that the same signalling mechanism might also be recruited by macrophages for immunomodulation.

Moreover, CTB and STB cells have no presentation of class I and II human leukocyte antigens (HLA) on their plasma membrane(Clark and Schust, 2013, Hunt et al., 1987, Apps et al., 2009). It is an exclusive strategy recruited by trophoblastic cells to evade adaptive immune surveillance by blocking the initiation of a classic "self" and "non-self" discrimination. Theoretically, without HLA antigens, trophoblastic cells should become more susceptible to being attacked by the innate immune system, such as NK cells. In particular, their local subset, decidual natural killer (dNK) cells, constitute 70% of immune cells at the maternal-foetal interface(King et al., 1996). In reality, trophoblastic cells maintain a harmonious relationship with them. This relationship might not only rely on the local dNK cells

as a special subset lacking expression of CD16 receptor (Koopman et al., 2003), but also benefit from the predominant expression of inhibitory receptors on NK cells, siglec 7 and 9 (Macauley et al., 2014, Nicoll et al., 1999). The binding affinity of siglec 7 is also nearly equivalent toward α2,3 and α2,6-linked terminal sialic acids (Yamaji et al., 2002). Thus, PSG1 could be a candidate to bind both siglec receptors on NK cells and ensure a desirable immunomodulatory effect.

3.3.2.2 Potential functions of LacNAc moiety

The expression of galectins in human trophoblast was initially reported in the 1980s as human placental tissue protein 13 (PP13 and Galectin 13) and human placenta β-galactoside-binding lectin (Galectin 1), although the protein family had yet been named galectin (Bohn et al., 1983, Hirabayashi and Kasai, 1984). During the mid-1990s, the importance of placental galectins was recognised after knowing that galectin 1 plays a vital role in regulating T cell apoptosis (Perillo et al., 1995).

Currently, there are five galectin family proteins recorded as highly expressed members at the maternal-foetal interface: galectin 1, 3, 8, 9 and 13 (Blois and Barrientos, 2014). They may generally synergise together in gestation for facilitating multiple events at this interface including immune modulation, embryo implantation and angiogenesis (Blois and Barrientos, 2014). Galectins primarily function through cross-linking glycoproteins via recognition of their LacNAc moieties (Figure 1.26). The formation of a galectin 3 lattice is a classic example for the cross-linking found on the cell surface in mice to block immune responses (Demetriou et al., 2001). However, the functions and formation of cross-linking between secretory glycoproteins or secretory/membrane glycoproteins via galectins remain largely unclear. As the N-glycans of PSG1 have multi-antennary and poly-LacNAc elongated moieties with mainly 2, 3-linked sialic acid terminals, it is highly likely that this human protein would be a ligand to some studied galectins (**Section 1.6.4.1**).

β integrins have been preliminarily shown in our collaborator's laboratory to mediate cell adhesion of endothelial cells and a trophoblast cell line (HTR-8) on PSG1 coated wells (Moore and Dveksler, 2014). Although there is an RGD-analogue, KGD motif, on the PSG1 N-domain, so far there is no evidence to support a direct interaction between this domain and integrins (Shanley et al., 2013). Instead, the PSG1-integrin interaction might be through an indirect association via galectin cross-linking. An early work suggested that galectin 8 could directly bind to β integrins on trabecular meshwork cells to mediate their cell adhesion on the galectin 8-coated wells (Diskin et al., 2009). Addition of lactose could abolish this cell adhesion, and cell spreading on the galectin-coated wells could even be inhibited by supplying NeuAcα2-3Galβ1-4GlcNAc tri-saccharide. Recently, galectin 1 was also shown to bind β1 integrin on human trophoblast via a glycan-lectin interaction *in vitro* (Bojic-Trbojevic et al., 2018). Major galectins have no possession of RGD or analogous motifs, so that integrin glycans could likely be the sole determinant for these galectin-integrin interactions. Moreover, galectin 1 was shown to bind a single β1 integrin but did not cross-link between β1 integrins (Moiseeva et al., 2003). This

ligand selectivity would favour the formation of heterogenous cross-linkings between ECM and membrane glycoproteins. A more recent preliminary study in our collaborator's laboratory validated the interaction between galectin 1 and recombinant PSG1 N-domain from CHO-K1 cell lines indicating the availability of PSG1 N-glycans to this lectin-glycan interaction. Other studies testing integrin-fibronectin-mediated cell adhesion also reported that disruption of N-glycan maturation or complete de-glycosylation led to reduced cell adhesion (Hsiao et al., 2017, von Lampe et al., 1993, Duksin and Bornstein, 1977, Oz et al., 1989). Taken together, PSG1 N-glycans could possibly facilitate galectins to connect with β integrins on cells as a complementary approach for manipulating cell adhesion and migration, such as trophoblast.

In addition, it is very likely that PSG1 N-glycans could engage with galectins to indirectly connect PSG1 proteins with receptors/signalling molecules on immune cells. For instance, T-cell immunoglobulin mucin 3 (TIM-3)/ galectin 9 ligation on macrophages has recently been shown to alleviate induced preeclampsia-like impairments in rats (Li et al., 2018). The signalling after this ligation induced the production of IL-10 and TGF-β cytokines and switched polarisation of decidual macrophages to the M2 subtype. Although the ligation between galectin 9 and TIM-3 might not be related to TIM-3 N-glycosylation (Lee et al., 2009), it would still be interesting to validate whether the TIM-3 ligated Galectin 9 could cross-link with PSG1 N-glycans simultaneously to tune the oberseved immunomodulatory effect.

3.3.2.3 Potential functions of bisecting GlcNAc

Bisecting GlcNAc is believed to play a role in regulating the functions of NK cells in immunity. In a seminal experiment, a human chronic myelogenous leukemia cell line, K562, was transfected with MGAT-3 gene which encodes GlcNAc-T III transferase responsible for GlcNAc bisecting. The GlcNAc-T III transfected cells showed an increased level of bisecting-GlcNAc and, importantly, were found to be resistant to killing by NK cells. However, the underling mechanism of suppression of NK cytotoxicity was not established (Yoshimura et al., 1996). As PSG1 has a small proportion of putative GlcNAc-bisecting structures, it will be of interest in future work to investigate whether its bisected glycoforms are capable of suppressing the cytotoxicity of NK cells for trophoblasts.

3.3.3 Glycosylation of recombinant PSG1N-Fc

The glycosylation of recombinant PSG1 constructs produced in three different CHO expression systems was evaluated in this study. As shown in **Section 3.2.2.5**, stable and transient transfected CHO cells can equivalently produce the same level of structural complexity, forming tetra-antennae and even poly-LacNAc chains on the N-glycans of PSG1N-Fc. This similarity indicates that the glycosylation machinery was not largely changed or affected in terms of functionality in both systems. Thereby, recruiting both transfection approaches could technically generate equivalent protein products

in terms of the quality of post-translational modifications. However, a principally semi-quantitative comparison of the structures characterised in both systems indicates that there is a higher content of complex glycans assembled onto the proteins produced in CHO cell system with stable transfection. Stable transfection normally introduces controllable copies of gene into cells, whereas the transient transfection can introduce numerous copies of gene into cells. Thereby, the difference in the processed glycan content could possibly be caused by a relative shortage of relevant components for glycosylation in the transient transfection system. In addition, the second batch of PSG1N-Fc from stable CHO-K1 system exhibited a glycoform with more truncated N-glycan structures. However, this feature might not reflect a conventional glycosylation of CHO-K1. Updated with our collaborator, the quality of glycosylation for this batch of materials might be compromised by technical issues in our collaborator's laboratory, and the causes are under inspection to clarify.

A comparison of glycosylation of recombinant samples from Epxi-CHO and CHO-K1 cell lines reveals some differences. We found that the CHO-K1 cell system can assemble multi-antennary complex glycans with antennary elongation by poly-LacNAc moieties as well as high levels of sialylation and galactosylation. On the contrary, the Expi-CHO system generally produced a much limited amount of complex glycans with galactosylation and sialylation. It is dominant in producing highly truncated structures to expose more terminal GlcNAc residues as confirmed by MS/MS analysis. Generally speaking, the Epxi-CHO cell line would be for secretory protein production with less requirements of glycosylation, in particular galactosylation and sialylation.

Connecting these results with the performed integrin binding in our collaborator's laboratory, it is difficult to suggest structural features of glycans on the PSG1N-Fc made in transient CHO-K1 that can explain suggested interactions. The reasons are obvious: 1) the protein from transient transfected CHO-K1 system has a similar N-glycan repertoire to its stable transfected counterparts; and 2) even the materials which have similar interactions against tested integrin were found to have significant variations in their glycomes. For example, there are differences of glycomes between different batches of PSG1N-Fc made in stable CHO-K1 system as well as between PSG1N-Fc from stable CHO-K1 and Expi-CHO.

In contrast, the occupancy of N-glycosylation sites examined by glycoproteomics shows some degree of consistency with integrin binding outcomes among experimental groups. As described in **Section 3.2.3.2**, all recombinant (PSG1N-Fc) and even native (PSG1) materials have N-glycosylation at site Asn61 on their N-domains. The examined stable CHO-K1 material as an exception that shows a predominant glycosylation for its IgG1 Fc-tag while a reduced glycosylation for Asn61. Putative N-glycan structures on Asn61 are summarised in Table 3.7 matching with the characterised higher mass range structures from glycomic profiling (Figure 3.9, **Section 3.2.2.5**).

Glycosylation of N-glycosylation sites Asn104 and/or 111 on PSG1 N-domain was exclusively found on transient CHO-K1 material. The glycoproteomic study indicates that N-glycans on these two sites are mainly immature oligo-mannose structures. Steric constraint around these N-glycosylation sites could be the main cause of limited processing as close sequon arrangements could be inhibitive for the development of N-glycans. Ph. D. work of Dr. O'Riordan included a mutagenesis study for the N-glycosylation site-occupancy of PSG1 N-domain with a C-terminal histone tag (PSG1N-His, O'Riordan, 2014). In that work, Asn61 and Asn111 were suggested as conserved glycosylation sites on the sequence which was transiently expressed in a Freestyle human embryonic kidney 293 (HEK-293) cell system. The suggested Asn111 glycosylation is compatible with that of the transient CHO-K1 material since we observed single-occupancy of the tryptic peptide spanning the Asn104 and 111 consensus sequences. However, as the glycosylation for the sites Asn104 and/or 111 were only found in transiently transfected cell systems, it is likely that the glycosylation on these two sites could be an artefact of the expression system.

Finally, tested PSG1 materials are generally interactive to integrin α5β1, but with an exception for transient CHO-K1 material. An early study assessed the impacts of glycosylation on integrin-ligand binding (Diamond et al., 1991). They carried out site-specific mutagenesis of the Mac1 ligand ICAM 1, to disrupt its N-glycosylation sequons. It was found that the binding between ICAM 1 and Mac 1 was enhanced after silencing N-glycosylation on the third domain of ICAM1 (Diamond et al., 1991). Recently, our collaborator preliminarily generated a recombinant PSG1N-Fc with mutation of all three N-glycosylation sites. Except for a significantly low production, the "naked" backbone peptides were still reactive with α5β1 protein. Thus, the N-glycosylation of site Asn104 and/or 111 on PSG1 N-domain could be responsible for interfering in its binding to integrin α5β1.

Chapter IV. Canine Melanoma

Chapter 4 Canine melanoma

4.1 Introduction

Melanoma is a life threating and mostly untreatable disease which is estimated to cause near 55,500 cases of mortality worldwide (Schadendorf et al., 2018). The disease is generally caused by the transformation of rarely dividing melanin-producing melanocytes to melanoma with high load gene mutations and uncontrollable proliferation under intrinsic and environmental triggers (Jimbow et al., 1975, Shain et al., 2015). It becomes fatal and untreatable once it reaches the metastatic stage. Melanoma can be morphologically categorised by the appearance of tumour lesions at primary sites: cutaneous, acral, mucosal, conjunctival and uveal melanomas. The pathogenic mechanism of this lethal disease is still not fully understood. Current genomic and genetic studies on mouse models and human cell lines facilitate the identification of key gene mutations and cell signalling transductions which are highly correlative to the development of malignant melanoma (Hodis et al., 2012, Cancer Genome Atlas, 2015, Testa et al., 2017). Inhibitory medication for identified signalling transductions can improve therapy response and survival rate (Schadendorf et al., 2018). However, drug resistance, persistent side effects and high variation of active gene mutation between different melanoma subtypes restrict the eligible population to receive this medication (Cancer Genome Atlas, 2015, Gencler and Gonul, 2016, Hayward et al., 2017). Hence, immunotherapy with high specificity to individual tumours has become an alternative approach to pursuing a potential cure.

A compatible in vivo animal model is required to underpin immunotherapy studies of melanoma. There are three fundamental preconditions for compatible animal models: 1) high load gene mutation; 2) morphologically and physiologically consistent tumour development; 3) immunocompetence. However, widely applied mouse animal models cannot fully fulfil these criteria. Importantly, mouse is not a vulnerable species for melanoma since it normally has a very limited population of melanocytes in its epidermis at restricted locations such as hair bulb and tail (Walker et al., 2011). Consequently, induced or spontaneous murine melanoma does not generally result in epidermal changes and metastasis (Damsky and Bosenberg, 2010, Walker et al., 2011, Stei et al., 2016). Engineered immunocompetent mouse models can mimic several features of human melanoma, but their limitations restrict their usefulness. For example, a transgenic mouse model can exhibit consistent melanoma development with immunocompetence, but its gene mutation load is lower than the natural melanoma (Walker et al., 2011). Syngeneic injection of melanoma cells can induce melanoma, but metastasis is not generally developed in this model (de Lange et al., 2012, Grossniklaus et al., 1995). Moreover, carcinogens can also induce murine melanoma, how-

ever, it is generally uncontrollable with no promising outcomes (Dithmar et al., 2000). Practically, mouse models also require a stringent housing environment with diet requirements (Park et al., 2016).

In contrast to the mouse, dogs appear to be a more ideal models for human melanoma. Canine melanoma mainly occurs as spontaneous mucosal melanoma in the tongue cavity with high incidence rate. The primary tumour of canine melanoma is normally visible and has high metastatic capability. The mutation distribution of canine melanoma is similar to its human counterpart, and its tumour progression is also consistent (Hernandez et al., 2018, Khanna et al., 2006). Moreover, as a companion animal, it is very likely that canines were exposed to very similar environmental triggers as their human counterparts (Nishiya et al., 2016). In a study, whole-cell vaccination of human gp100 transfected canine melanoma cell lines to naturally occurring canine melanoma patients was well tolerated, and resulted in an extended life-span and even regression in six patients (6 of 34) (Alexander et al., 2006). This study highlights the compatibility and efficiency of canine melanoma models to preclinical research of immunotherapy. More recently, our collaborators at the University of Missouri have innovated a whole-cell vaccination strategy which is based on inactivating tumour-associated glycosylation. A highly promising pre-clinical outcome was obtained on this novel canine melanoma model. For example, complete remission was reached on 35% of canine patients without adverse effects (Dr Gary Clark, personal communication of unpublished work). This is an unprecedented outcome. Hence, it is of great interest to investigate the equivalence of the biochemical properties of canine melanoma, especially glycosylation, with its human counterpart to validate whether the exciting outcomes on canines could be translated to humans.

Cancer glycosylation is dominantly different from that of most normal cells. Glycomics studies of melanoma have been mainly carried out on human cell lines and a spontaneous murine melanoma cell line named B16 and its variants (Overwijk and Restifo, 2001). The N-glycosylation of human melanoma was consistently portrayed as having progressively increasing β 1,6 branching on the trimannosyl core and varying sialylation between different cell lines, whereas the level of high mannose was unchanged (Litynska et al., 2001, Ochwat et al., 2004, Przybylo et al., 2008, Kolasinska et al., 2016). Uveal melanoma was found to have a higher content of β 1,6 branched structures than cutaneous melanoma and the higher content of β1,6 branching was correlated to superior cell mobility (Przybylo et al., 2008, Pochec et al., 2015). An early screening for Selectin ligands on human melanoma cells suggested that sialyl LewisX/A epitopes were not on glycoproteins so are likely to be carried by glycolipids (Miller et al., 1996). A later study examined the expression level of FUT1 and FUT4 (Fuc-T I and IV) which increased in human melanoma cell lines from primary to metastatic states, thereby an increased fucosylation was suggested to be associated with the progression of melanoma (Ciolczyk-Wierzbicka et al., 2007). However, a more recent study into the signalling regulation of melanoma found that the fucosylation on glycoproteins was negatively regulated by fucokinase along with the cancer progression (Lau et al., 2015). The latter study is more consistent with recent reports on the

N-glycosylation of human melanoma cell lines. These studies showed that core, but not antennary fucosylation is abundantly expressed during cancer progression. This was consistently observed on N-glycans among all tested human cell lines (Sweeney et al., 2018, Magalhaes et al., 2017, Hoti et al., 2018). There is one study of Bowes melanoma which proposed the presence of sulphation on the core-fucosylated chitobiose (Zamze et al., 2001). In addition, Poly-LacNAc extension on N-glycans was also observed in recent studies and the ratio of linear to branched poly-LacNAc chains on N-glycan antennae was substantially higher in melanoma cell lines than controls(Sweeney et al., 2018, Kinoshita et al., 2014). The N-glycosylation on murine melanoma was suggested to be generally consistent with the observations in humans including poly-LacNAc extension and branching (Srinivasan et al., 2009, Kawano et al., 1991). The sialylation on murine melanoma was found to be generally α2,3-linked and the expression of these linkage specific sialic acid residues positively correlated to the metastatic potential of murine melanoma (Chang et al., 2006).

O-glycosylation has been less studied on melanoma. There is one report involving antibody detection which suggests that less than 30% of human cutaneous tumours express T and Tn antigens. These antigens were not detected on non-epithelial melanoma (Kanitakis et al., 1998). Melanoma glycolipids have been more rigorously investigated than O-glycans. Thus, studies of the gangliosides of human or murine melanoma suggested that GM and GD families of glycolipids were prominently expressed on melanoma cell surfaces. GM2 and GD2 were observed as the most complex GSL glycans in major examined human cell lines. GM2 and GD2 had high abundance but the level of GD2 was dependent on that of GD3 (Yamashiro et al., 1993, Ruan and Lloyd, 1992). Because the B16 murine melanoma cell line can efficiently convert the short chain ceramide to GM3, this glycolipid predominates in this murine cell model (Komori and Ito, 1995).

Until now, no glycomics investigations have been undertaken for canine melanoma. Therefore, in this study, my objective was to fully characterise the N-, O- and lipid- linked glycans expressed by melanoma cells derived from a canine patient that had been successfully vaccinated at the University of Missouri. We aimed to enhance the understanding of canine melanoma from a glycobiological viewpoint and to verify the compatibility of this animal model for human melanoma research. Hopefully, we would also provide vital evidence to assist investigations of the underlying mechanisms of glycan-related immunomodulation.

4.2 Results

4.2.1 Research strategy

We are collaborating with Professor Gary Clark and his colleagues at the University of Missouri in the United States. Our collaborators prepared and provided the primary cells of canine melanoma from a stage III patient for my comprehensive glycomics study. The overall research strategy is summarised in Figure 4.1.

4.2.2 N-glycosylation of canine melanoma

The MALDI-TOF spectrum of canine melanoma showed a sophisticated N-glycan repertoire, which consisted of high mannose and complex glycans (Figure 4.2). The high mannose glycans have the overall significance of high relative abundance. Some studies in other cancers, such as breast cancer, reported this feature of tumour N-glycans (Sindhura et al., 2017). However, it is important to bear in mind that the N-glycome was directly recovered from the homogenates of cell lysates and will include glycans derived from the secretory pathway. Thus, detection of the high mannose precursor *m/z* 2600 with an un-cleaved Glc is indicative for the presence of glycans from glycoproteins in the ER compart-
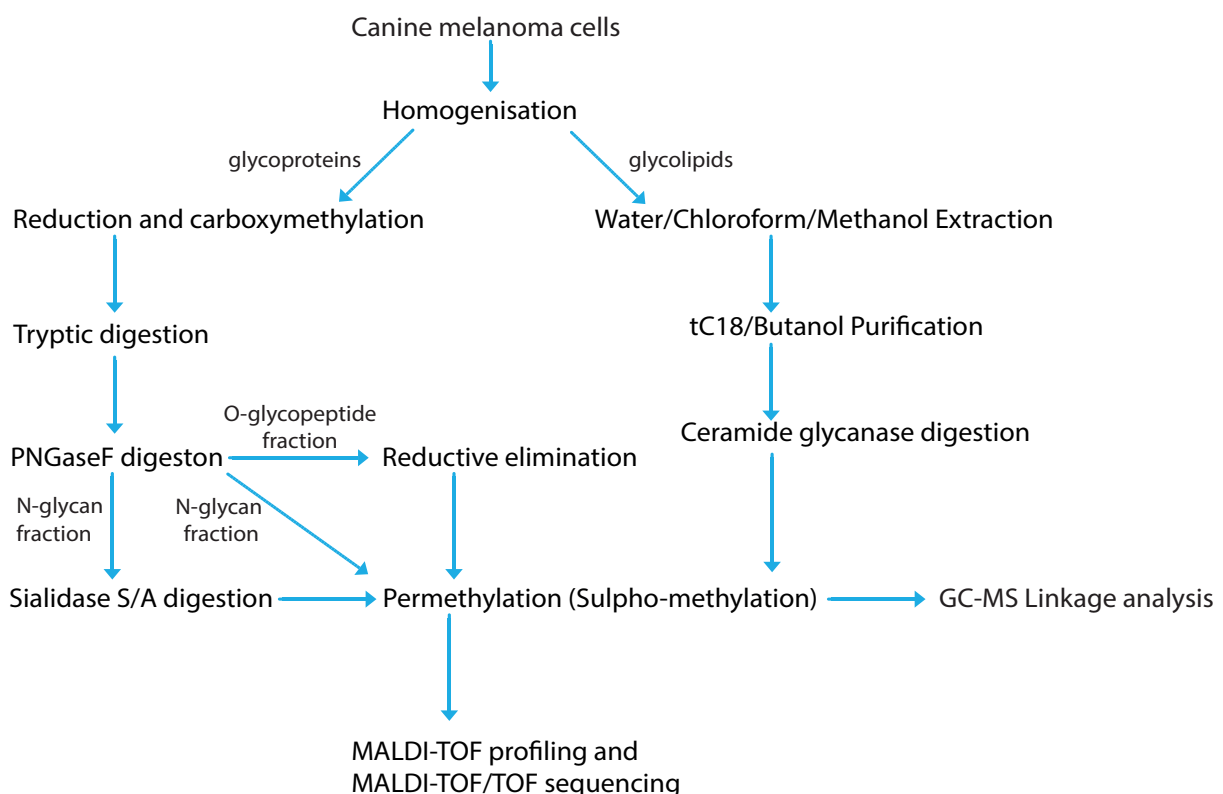
Figure 4.1 Schematic paradigm of research strategies for canine melanoma cells
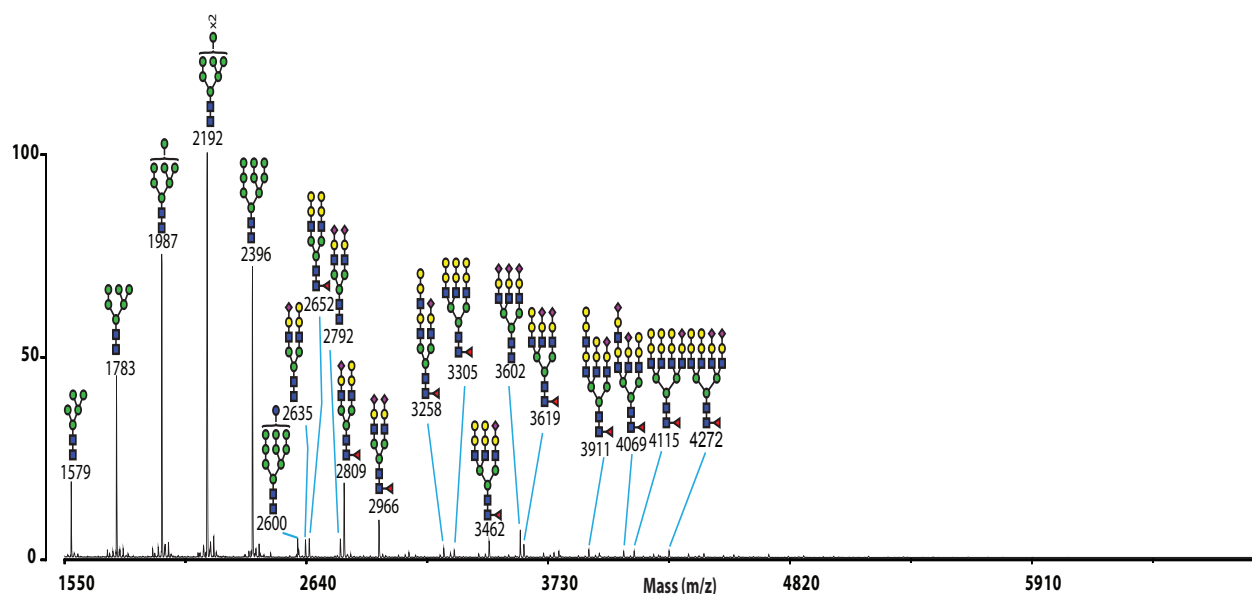
Figure 4.2 Overview of the MALDI-TOF spectrum of canine melanoma N-glycans.
For more details see Fig 4.3. The MS data were acquired from the 50%Acetonitrile fraction of permethylated N-glycans purified by C18 reversed-phase chromatography. Molecular ions are detected as [M+Na]+. The structural assignments of all the peaks are based on knowledge of biosynthetic pathway and compositional information.

ment which have not yet passed the quality control for folding. It is possible that the exceptionally high abundance of high mannose glycans might not be surface-associated.

The dominant peaks for complex glycans correspond to bi-antennary structures with NeuAc sialic acid and putative α-Gal on LacNAc forming sialyl-LacNAc (*e.g. m/z* 2966) and Gal-α-Gal (*e.g. m/z* 2652) terminal moieties. The Gal-α-Gal moiety widely exists in mammals with the exception of human beings because the gene for the biosynthesising enzyme, α1,3 galactosyltransferase, is only partially transcribed in humans (Lanteri et al., 2002). Tri- and tetra-antennary structures are also observed with the shared terminal modifications (*e.g. m/z* 3305 and 3602) and potential poly-LacNAc extension (*e.g. m/z* 3911 and *m/z* 4069).

Higher *m/z* signals corresponding to complex glycans were detected reaching the screened upper-limit of *m/z* 7000. To more effectively present the complexity of canine melanoma N-glycans, the higher mass range of spectrum is spread out from *m/z* 2900 to 6900 (Figure 4.3). There is an extensive poly-LacNAc elongation on antennae of the higher mass complex glycans ranging from bi-, tri- to tetra-antennary. There are up to seven internal LacNAc units found on a N-glycan sequence excluding terminal epitopes (*e.g. m/z* 6450). The predominant population of complex glycans is core-fucosylated, and the termini of antennary LacNAc are generally occupied by either sialic acid or α-Gal residues.
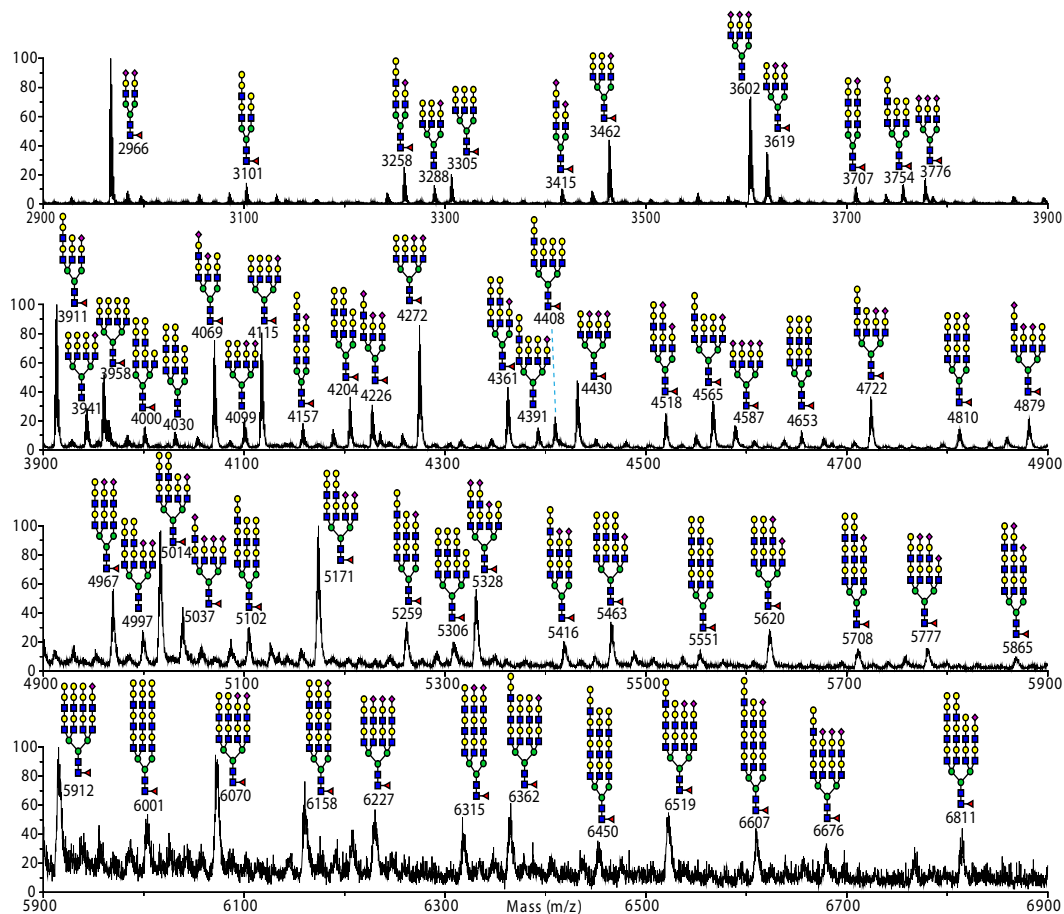
Figure 4.3 The MALDI-TOF spectrum of canine melanoma N-glycans from m/z 2900-6900.
The spectrum is displayed in four panels, each normalised to the most abundant peak in the mass range defined by the panel: *m/z* 2900-3900 (A), 3900-4900 (B), 4900-5900 (C) and 5900-6900 (D). . See the legend of Figure 4.2 for assignments.

4.2.2.1 Fragmentation analysis for melanoma N-glycans

To validate the presence of suggested terminal epitopes and poly-LacNAc extension, fragmentation analyses were conducted on three selected N-glycan ions (*m/z* 3258, 4069 and 5463) by MALDI-TOF/TOF (Figure 4.4).

The ion *m/z* 3258 (Figure 4.4A) could putatively be assigned as either a bi- or tri-antennary structure as the LacNAc moiety could be terminal or internal with two suggested terminal moieties. Its single fucose reisdue is validated as core-fucose on the reducing end GlcNAc by the b-ion *m/z* 2807. The suggested terminal moieties are firstly validated as the co-presence of pairing b and y ions for sialyl-LacNAc (b-ion *m/z* 847 and y-ion 2434) and Gal-α-Gal (b-ion *m/z* 690 and y-ion 2591). Internal LacNAc is identified on antenna with either terminal moiety via the observation of y-ions at *m/z* 1984 and 2142. The y-ion at *m/z* 2795, corresponding to loss of a terminal LacNAc, validates the existence of a tri-antennary structure. Hence, three structural variants contribute to this molecular ion. These are

depicted in the insets in Fig 4.4A. Interestingly, *m/z* 1984, 2142 and 2795 have a similar relative abundance in this MS/MS spectrum. This similarity might imply that there is a nearly equivalent probability for the structures to be branched or extended without a special preference for terminal moieties.

The ion at *m/z* 4069 (Figure 4.4B) was also subjected to fragmentation analysis for verifying the presence of LacNAc extensions. It is not surprising that three fragment y-ions are found corresponding to terminal LacNAc (*m/z* 3605) and internal LacNAc associated with the Sialyl-LacNAc (*m/z* 2796) or Gal-α-Gal (*m/z* 2952). This co-presence of internal and terminal LacNAc validates the presence of
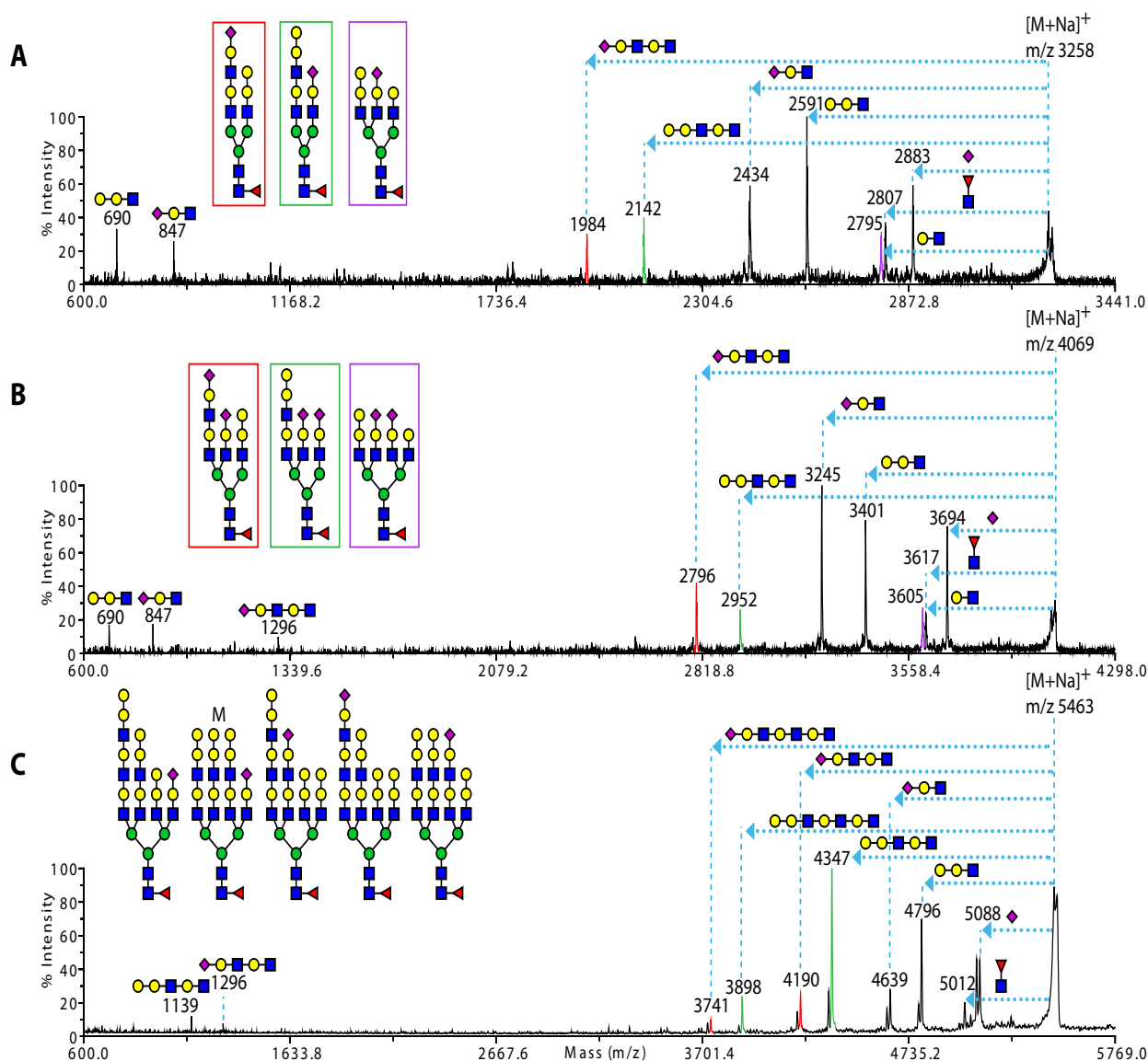
Figure 4.4 MALDI-TOF/TOF spectra for m/z 3258, 4069 and 5463 from the N-glycans of canine melanoma.
Three ions m/z 3258(A), 4069(B) and 5463(C) were selected to validate the terminal epitopes and antennary arrangements. All the ions are observed as [M+Na]+. The fragmentation ions are annotated based on knowledge of fragmentation pathways. All possible structures are presented at the up-left of each panel and the major structure is marked with an "M" above the structure. Peaks for fragment y-ions containing terminal sialyl-LacNAc, Gal-α-Gal and LacNAc are separately highlighted in red, green and purple.

the tri- and tetra-antennary structures shown in the insets on panel B. Besides, the relative abundance of m/z 2952 and 3605 is nearly equivalent but that of m/z 2796 is relatively higher. This difference of relative abundance corresponds to the number of each putative terminal moiety and is consistent with the suggestion above of an equivalent distribution of terminal or internal LacNAc.

The ion m/z 5463 (Figure 4.4C) was fragmentated to further validate the distribution of internal LacNAc units on antennae. The composition of this molecular ion ($NeuAc_1Hex_{13}HexNAc_9Fuc_1$) suggests tetra-antennary glycans containing one sialylated antenna and three α-Gal capped antennae. In addition the composition predicts three internal LacNAc moieties. The abundant y-ions at m/z 4796, 4347 and 3898 are consistent with the presence of three Gal-α-Gal terminals on antennae that have up to two LacNAc repeating units. Sialyl-LacNAc with one internal LacNAc is confirmed by the pair of ions m/z 1296 and 4190. The latter has an abundance less than one third of that of its Gal-α-Gal counterpart (m/z 4347), which is consistent with there being only one sialylated antenna in these tetra-antennary glycans. Sialyl-LacNAc with two internal LacNAc units is also observed as a minor y-ion at m/z 3741. However, no signal was detected corresponding to an antenna with three internal LacNAc units. Hence, five structures are assigned for this molecular ion with LacNAc extensions on multiple antennae (structures are shown on Panel C in Fig 4.4). In addition, the abundance distribution of antenna fragments with from one to two internal units follows a successive decrease as expected for the fragmentation of linear poly-LacNAc sequences (*e.g.* m/z 4347 to 3898 and m/z 4190 to 3741).

Taken together, these MS/MS analyses verified the suggested structural features as core-fucosylation, poly-LacNAc extension, sialyl-LacNAc and Gal-α-Gal. With the exception of Gal-α-Gal, all the other structural features are compatible with the conclusion of recent glycomics work on human melanoma cell lines in our laboratory (Sweeney et al., 2018).

4.2.2.2 sialidase A and S digestion

To verify the linkage of sialic acids, sialidase S was conducted on an aliquot of native canine melanoma N-glycans. The glycome after Sialidase S digestion is shown in Figure 4.5B &4.6B in comparison with the untreated control (Figure 4.5A &4.6A). There is a nearly complete digestion of terminal sialic acids, with an exception of a minor signal m/z 2809 which has one sialyl-LacNAc antenna (Figure 4.5B). This digestion outcome shows that the vast majority of sialic acids on canine melanoma N-glycans are α2,3-linked. Besides, benefiting from the near complete removal of sialic acids which enhanced our ability to detect low abundance high mass ions , the signals for desialylated poly-LacNAc carrying structures reach a ceiling up to m/z 6042 which corresponds to a structure carrying ten LacNAc moieties (Figure 4.6B).

In addition, Sialidase A digestion (Figure 4.5C &4.6C) was carried out to improve the detection for poly-LacNAc extended structures by eliminating sialic acids with all type of linkages. The spectrum
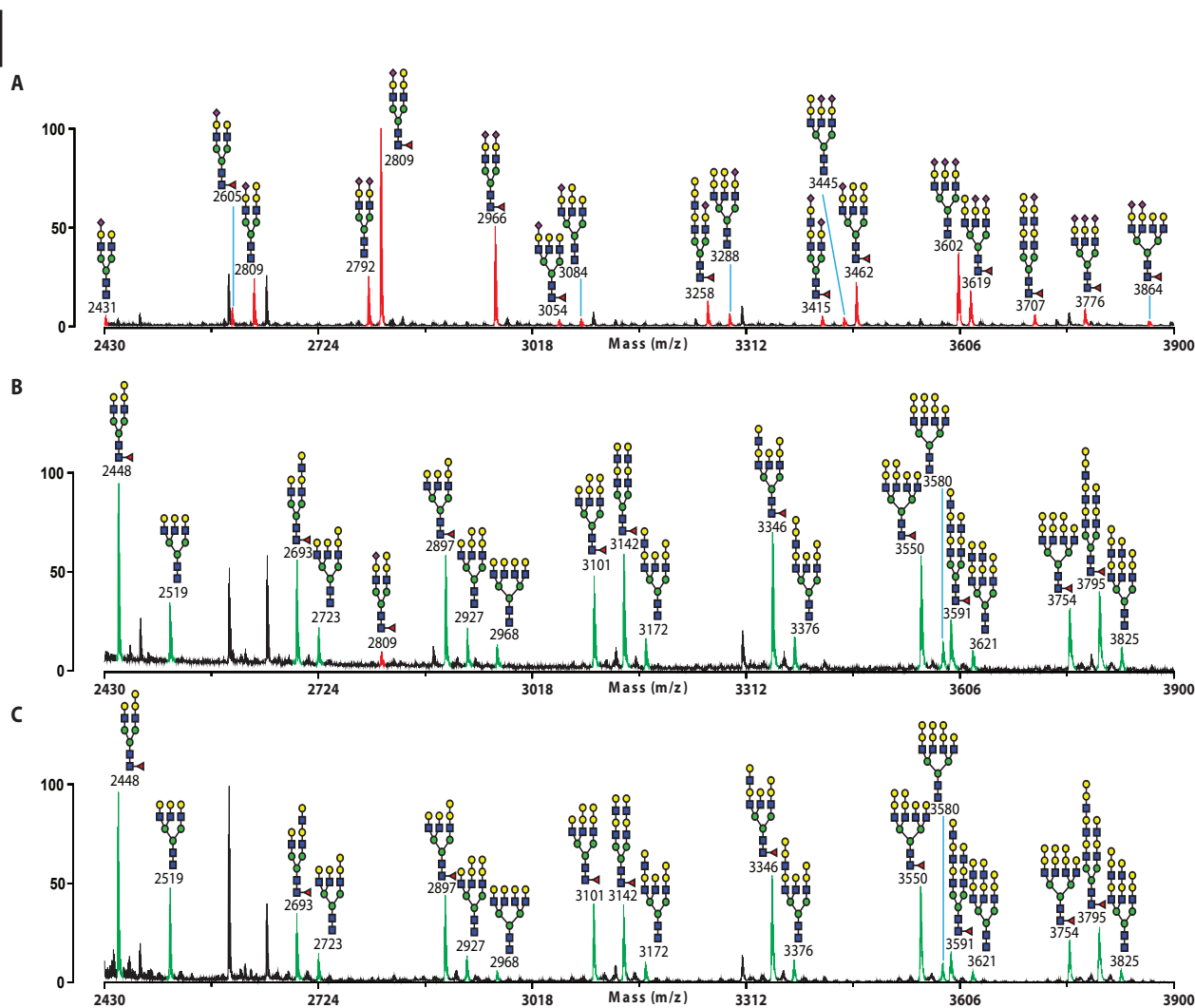
Figure 4.5 Low mass range comparison of MALDI-TOF spectra of untreated, Sialidase S and A treated canine melanoma N-glycans.
Permethylated N-glycans of canine melanoma are shown as untreated (A), sialidase S treated(B) and Sialidase A treated(C) in different panels.Peaks for sialic acid containing structures are highlighted in red. Peaks for digested product structures are highlighted in green. See the legend of Figure 4.2 for assignments. In addition, the assignments of Gal-α-Gal or Sialyl-LacNAc on extended antennae based on general understandings of biosynthetic pathway, but other localisations of these eptiopes are still possible.

of sialidase A digestion is generally identical to that of sialidase S, although an improved digestion is noticeable by the disappearance of *m/z* 2809 (Figure 4.5C). Detection for the highest *m/z* signal is also slightly improved in this experiment. Thus, a weak signal for *m/z* 6246 (Figure. 4.6C) becomes visible corresponding to a structure with one more Gal-α-Gal than the *m/z* 6042 component in the sialidase S digestion.

In summary the sialidase experiments showed that the major sialic acid residues on canine melanoma N-glycans are α2,3-linked and the detectable largest number of LacNAc moieties is ten in an individual N-glycan structure.
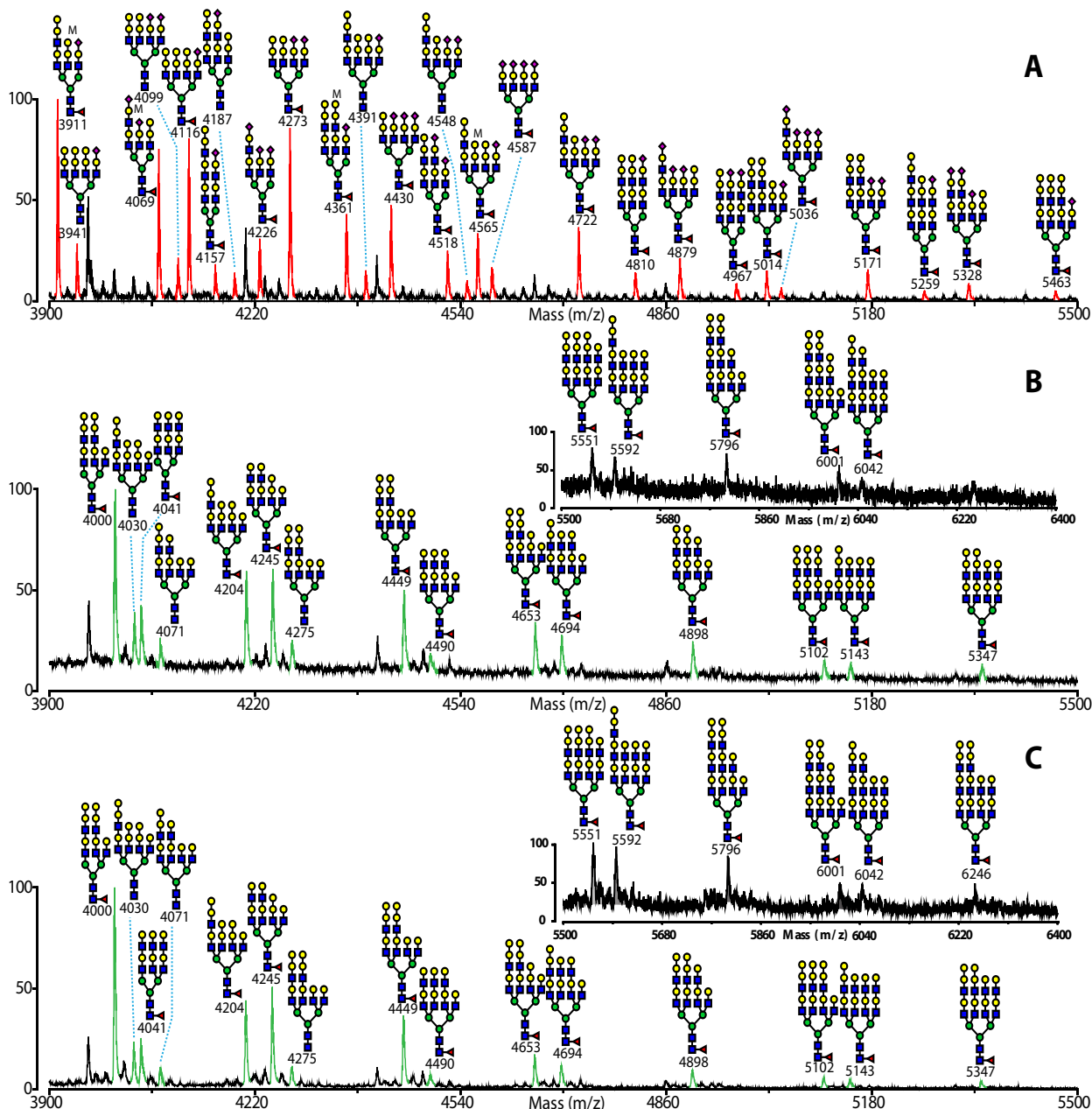
Figure 4.6 High mass range of MALDI-TOF spectra of untreated, Sialidase S and A treated canine melanoma N-glycans.

Permethylated N-glycans of canine melanoma are shown as untreated (A), sialidase S treated(B) and Sialidase A treated(C) in different panels. Higher mass range of spectra for Sialidase S and A are presented on the up-right corners of panel B and C, respectively. See the legend of Figure 4.2 for assignments. In addition, the assignments of Gal-α-Gal or Sialyl-LacNAc on extended antennae based on general understandings of biosynthetic pathway, but other localisations are still possible.

4.2.2.3 Endo-β-galatcosidase digestion

Endo-β-galatcosidase digestion was carried out to determine whether there is branching of the internal LacNAc units and to provide information on the number of extended antennae in multi-antennary glycans. This enzyme should cleave the internal β1,4 galactoside on unbranched Poly-LacNAc chains. Branched poly-LacNAc sequences are resistant to cleavage. The enzyme activity would pro-
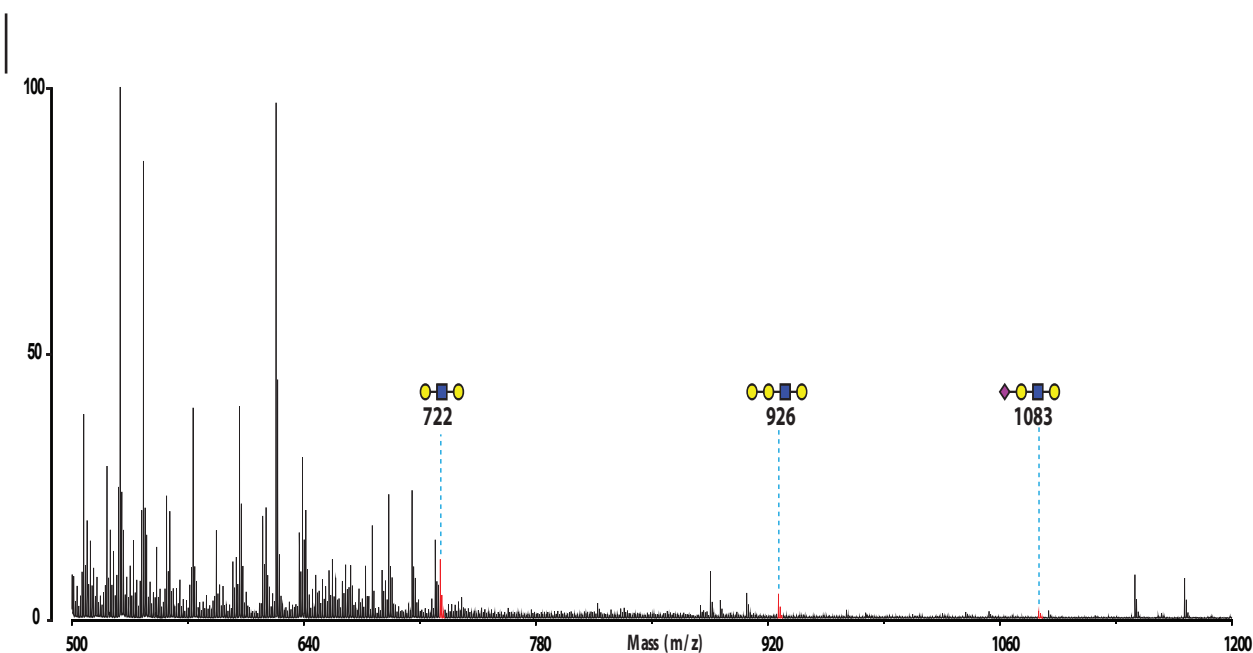
Figure 4.7 Low mass range of MALDI-TOF spectrum of canine melanoma N-glycans after endo-β-galactosidase digestion.
Peaks for digested products are highlighted in read and interpreted. The assignments are based on the enzyme specificity and expectable digestion products.

duce a series of characteristic low molecular weight digestion products from the non-reducing end terminal and internal LacNAc units. The product from cleavage of internal LacNAc is a di-saccharide, GlcNAc-Gal, with m/z 518 in the MALDI-TOF. Due to the presence of the three types of terminal epitopes suggested above, there are predicted to be three terminal products: 1) Gal-GlcNAc-Gal, m/z 722; 2) Gal-Gal-GlcNAc-Gal, m/z 926; and 3) NeuAc-Gal-GlcNAc-Gal, m/z 1083.

Figure 4.7 shows the low mass range of canine melanoma N-glycans after this enzyme digestion. The terminal products m/z 722, 926 and 1083 are discerned with a limited abundance, whereas no signal for internal digestion, m/z 518, is detectable above the chemical background which makes detection of low levels of low molecular weight glycans difficult. The quality of the data suggested that very little digestion had occurred.

Examination of the higher mass range confirmed this tentative conclusion. Although no signals above m/z 6000 were detectable (Figure 4.8B) after digestion, poly-LacNAc extended structures are still present (*e.g.* m/z 4653 and 4810). Surprisingly, there are unexpected changes in the abundance for m/z 3602 and 3305 in comparison with a reference ion m/z 3462 since all the three structures are neither poly-LacNAc extended nor expected to be reactive to the enzyme. It is possible that the enzyme preparation could have other catalytic activities to these structures, such as de-sialylation, to cause the shift of relative abundance.

The spectrum ranging from m/z 1820 to 2890 was also assessed for the abundance of products corresponding to complete digestion (Figure 4.9). The relative abundance of the putative bi-antennary fully digested product, m/z 1835, is slightly increased after digestion in comparison with a high mannose
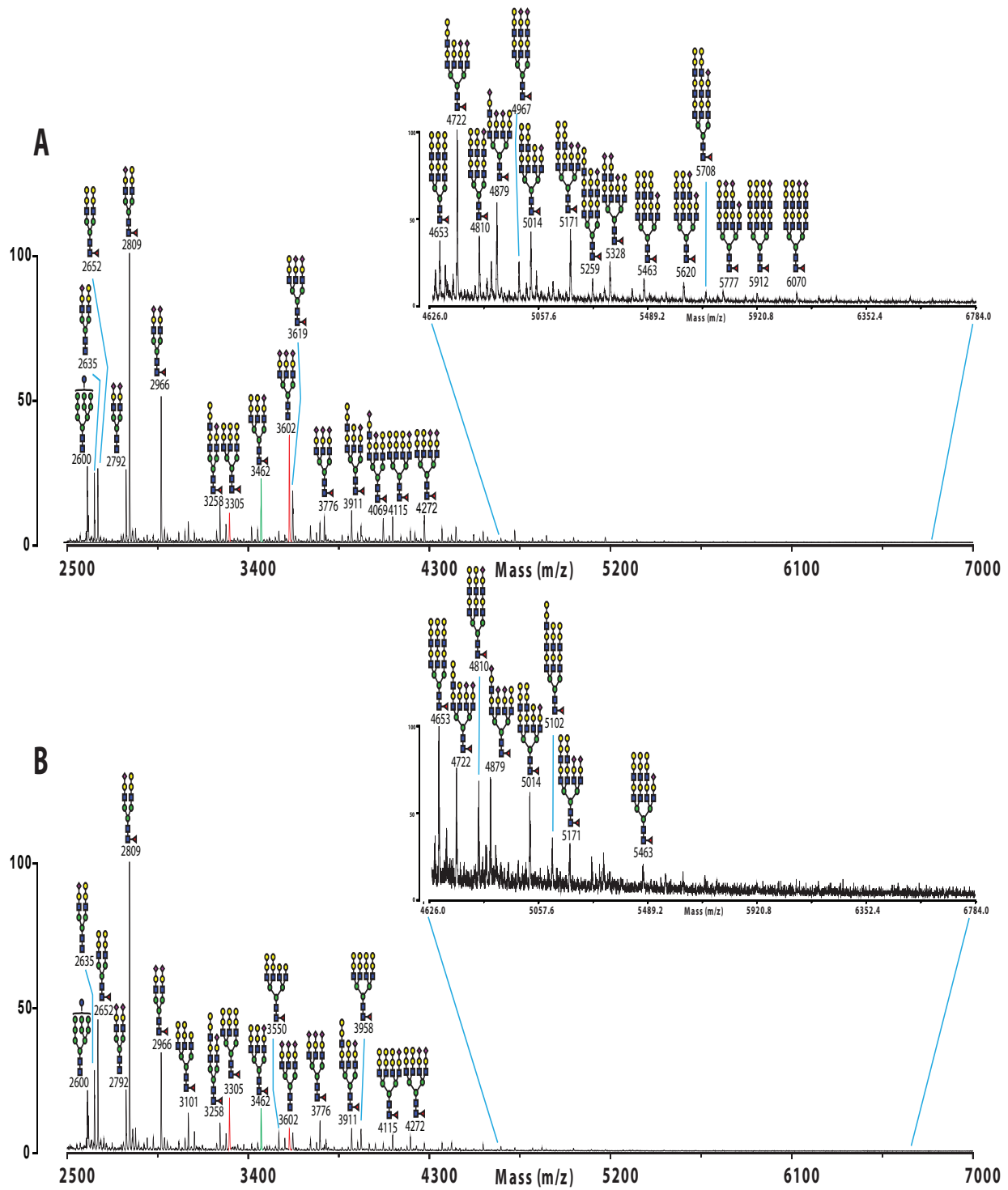
Figure 4.8 High mass range of MALDI-TOF spectra of untreated and endo-β-galactosidase treated canine melanoma N-glycans.

Permethylated N-glycans of canine melanoma are shown as untreated (A) and endo-β-galatosidase treated (B) in two panels. An expansion of the higher range of the spectrum is displayed in the upper-right corner of each panel. The peaks for the comparison of relative abundance are highlighted in red and a reference peak is highlighted in green. See the legend of Figure 4.2 for assignments.

structure, *m/z* 2192, as a reference. However, this shift of relative abundance may not be significant because the abundances of high mannose and complex-type glycans are known to be differently affected by sample complexity. Also two other putative digest products (*m/z* 2080 and 2326) are found in both spectra without a significant increase of relative abundance.

Ions whose masses corresponded to compositions predicted to be derived from digestion of poly-LacNAc-containing glycans were subjected to MS/MS analysis. However, instead of observing a fragment ion that corresponds to loss of a GlcNAc from the non-reducing end of the predicted structure (*e.g. m/z* 3707, Fig 4.10A), the MS/MS data identified two non-digested glycans whose structures are shown in Fig 4.10B. Similar results were obtained from all ions that had been selected for MS/MS analysis on the basis that their compositions might correspond to digested products. Hence the endo-β-galactosidase was not degrading the sample as expected. This could be due to either an inactive enzyme for the poly-LacNAc sequences or the poly-LacNAc sequences themselves that are resistant to the digestion.

To investigate which of these explanations was more likely, the N-glycans of murine lung, which have a rich population of poly-LacNAc antennae known to be sensitive to endo-β-galactosidase (Bern et al., 2013), were purified and subjected to digestion with three commercially available endo-β-
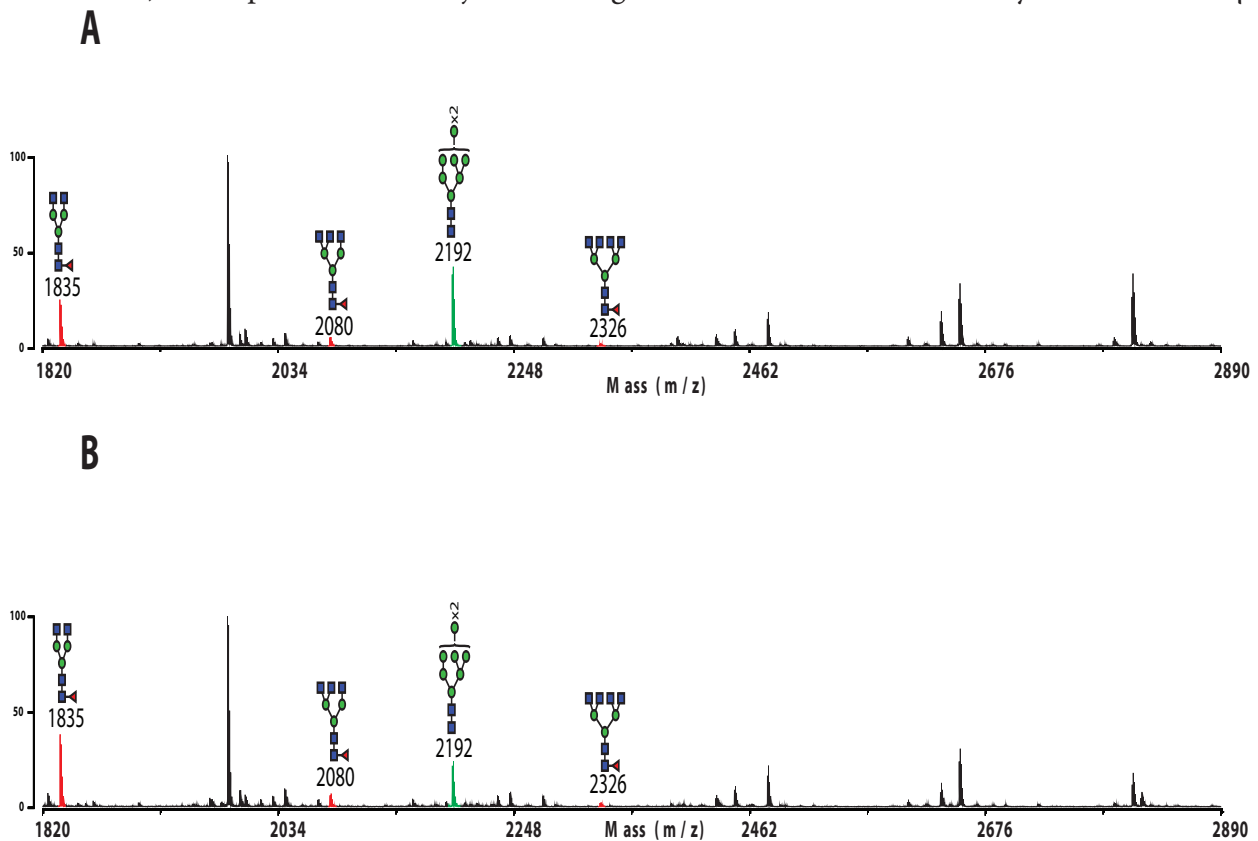
Figure 4.9 MALDI-TOF spectra of canine melanoma and endo-β-galactosidase treated N-glycans ranging from m/z 1820 to 2890.
Permethylated N-glycans of canine melanoma are shown as untreated (A) and endo-β-galatosidase treated (B) in two panels. Peaks for putative fully-digested products are highlight in red and a peak for reference as a high mannose structure is highlighted in green. See the legend of Figure 4.2 for assignments.
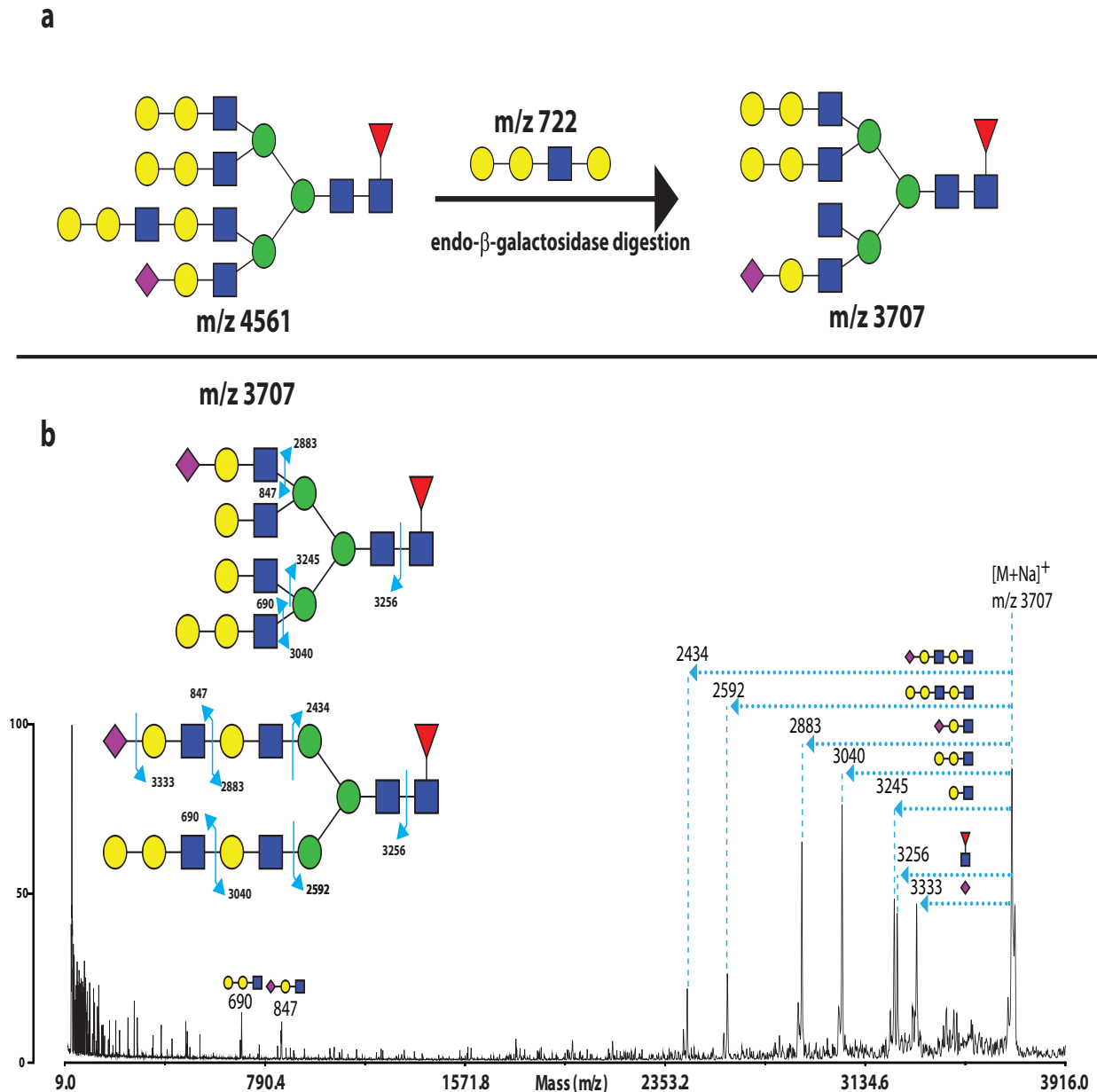
Figure 4.10 A putative digestion of endo-β-galactosidase and MALDI-TOF/TOF spectrum of m/z 3707 from canine melanoma N-glycans treated with endo-β-galactosidase.
A putative digestion is shown from a precursor m/z 4561 to product m/z 3707 after enzyme treatment(A) and possible product ion was subsequently fragmentated for verification (B). See the legend of Figure 4.4 for assignments.

galactosidases. Disappointingly, no low molecular weight digestion products were found in the N-glycan spectra of mice lung tissue after digestion by all enzymes (Figure 4.11). In addition, poly-LacNAc-containing glycans were observed at high mass after digestion (Figure 4.12). Therefore we concluded that, at this time, commercial endo-β-galactosidase preparations were not reliable for characterising poly-LacNAc structures.

In summary, the observation of minor quantities of cleaved terminal products confirmed the presence of linear poly-LacNAc but the limited activity of the commercial endo-β-galactosidase precluded further investigation. Therefore GC-MS linkage analysis was attempted as an alternative means of assessing whether the poly-LacNAc antennae are branched.
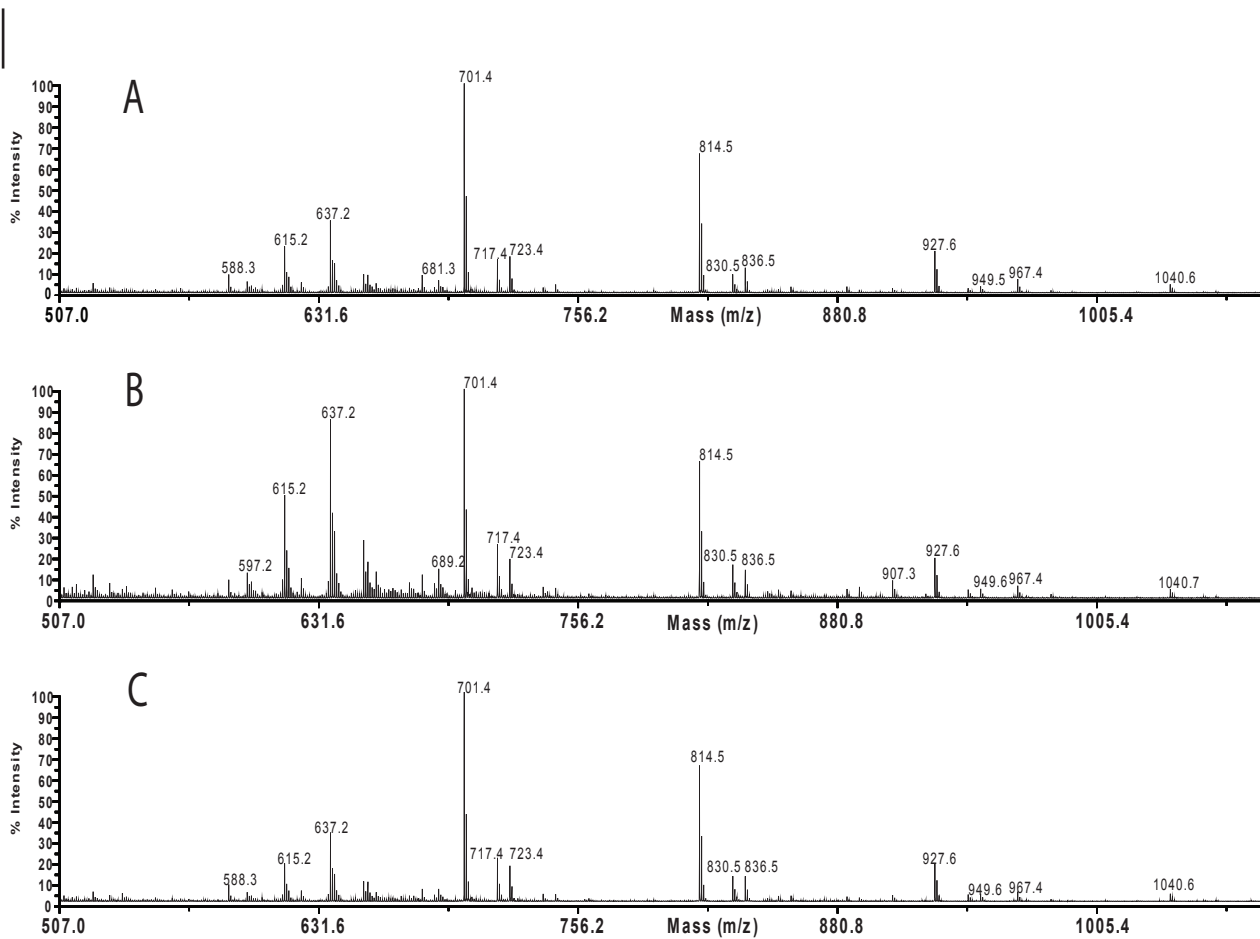
Figure 4.11 Low mass range of MALDI-TOF spectra of murine lung N-glycans treated with endo-β-galactosidases from commercial sources under consistent condition.

Untreated murine lung N-glycans are shown in (A). Three enzymes were obtained from Calbiochem (B), Sigma (C) and QA (D) for treatment. No corresponding digestion product is perceivable.
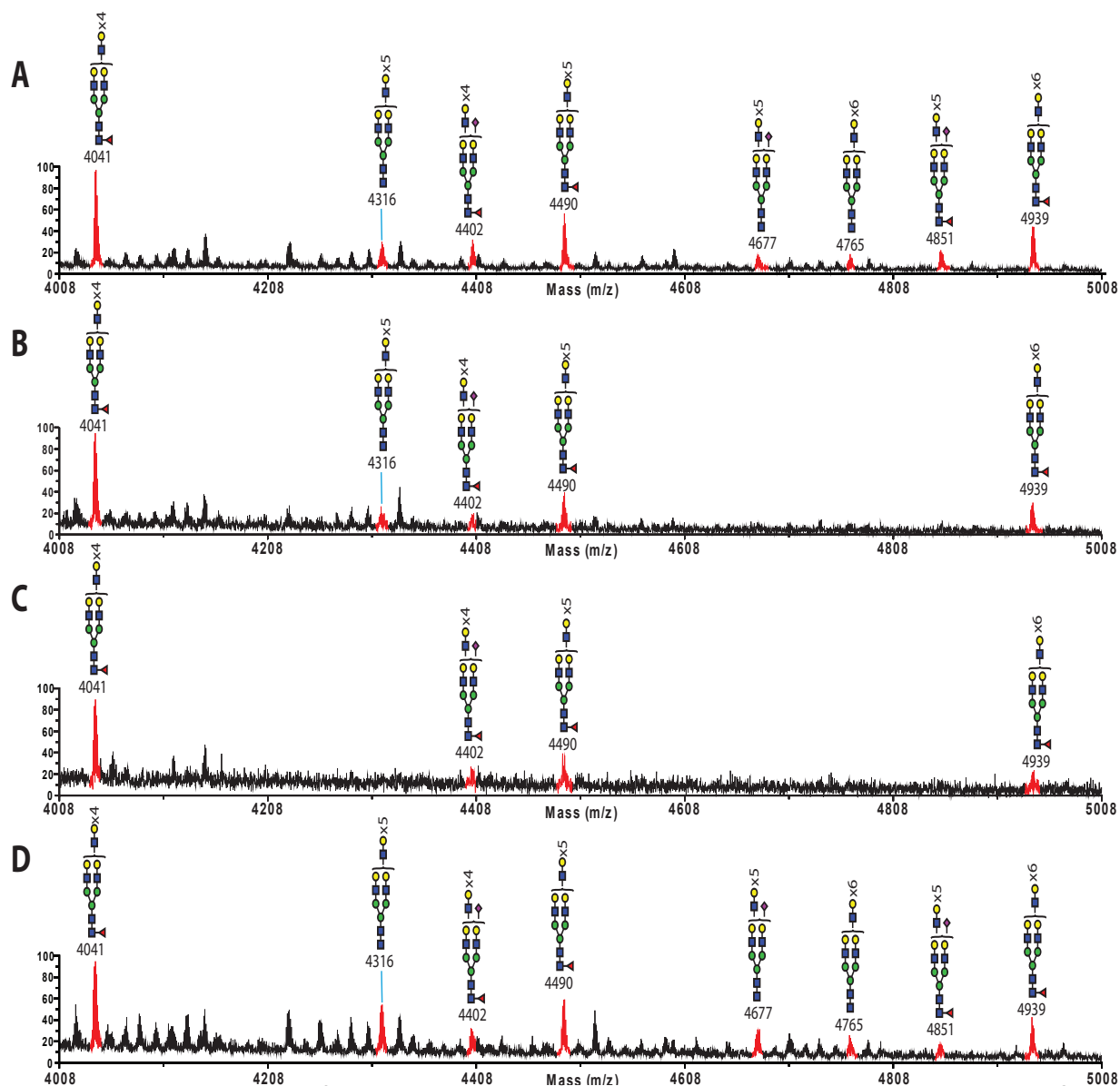
Figure 4.12 High mass range of MALDI-TOF spectra of murine lung N-glycans and treated with endo-β galactosidases from different commerical sources.

Untreated murine lung N-glycans are shown in (A). Three enzymes were obtained from Calbiochem (B), Sigma (C) and QA (D) for treatment. Peaks corresponding to putative structures with Poly-LacNAc are highlight in read and interpreted. See the legend of Figure 4.2 for assignments.

4.2.2.4 GC-MS linkage analysis for canine melanoma N-glycans

To collectively identify the linkage of each monosaccharide from the N-glycan repertoire, GC-MS analysis was conducted for canine melanoma N-glycans. Briefly, all N-glycans were hydrolysed and the monosaccharide products were chemically converted to partially methylated alditol acetates (PMAAs) for GC-MS analysis. Linkage information can be deduced from the characteristic fragmentation patterns of the modified molecules in combination with gas-phase chromatography which retains each monosaccharide differently in time. The linkage information would help to confirm the linkage related details, such as the type of LacNAc linkage or the presence of bisecting GlcNAc. In addition, sialic acids would normally be hydrolysed without detection.

The elution and characteristic fragment ions of each detected PMMA are listed in Table 4.1, and relative abundance is given by normalising the signal intensity of 2-linked Man to 1. The presence of 4-linked instead of 3-linked GlcNAc verifies the existence of type II poly-LacNAc extension. The absence of 3,6-linked Gal confirms the linear arrangement of poly-LacNAc sequences on antennae. The relatively high abundance of 3-linked Gal correlates to the presence of α-Gal, α2,3-linked sialic acid and the LacNAc units. 6-linked Gal is also detected at a trace level which is consistent with the described results of sialidase digestion (**Section 4.2.2.2**). The presence of 2,4- and 2,6-linked Man confirms the detection of tri- and tetra-antennary structures and core-fucosylation is verified by the presence of 4,6-linked GlcNAc. The absence of 3,4,6-linked Man rules out the presence of bisected GlcNAc. This is consistent with the glycomics data described earlier.

Table 4.1 Monosaccharide elution and characteristic fragment ions from canine melanoma N-glycans

Elution Time/min	Charateristic Fragment Ions	Assignment	Relative Abundance
12.774	89+102+115+118+131+162+175	Terminal Fuc	0.144
14.223	87+102+118+129+145+161+162+205	Terminal Man	0.856
14.486	87+102+118+129+145+161+162+205	Terminal Gal	0.178
15.327	87+88+101+129+130+161+190+205+234	2-linked Man	1
15.65	101+118+129+161+202+217+234	3-linked Gal	0.347
16.129	87+102+118+129+162+189+233	6-linked Gal	0.034
16.455	87+99+113+130+233	2,4-linked Man	0.178
16.852	87+99+129+190	2,6-linked Man	0.195
17.038	87+118+129+189+234	3,6-linked Man	0.44
18.579	117+129+143+159+189+346	4-linked GlcNAc	0.678
19.876	117+159+261	4,6-linked GlcNAc	0.11

4.2.3 O-glycosylation of canine melanoma

There are three core 1 O-glycans (Figure 4.13) found in canine melanoma as neutral (T-antigen), mono-(sialyl-T antigen) and di-sialylated Core 1 structures (*m/z* 534, 895 and 1256). The di-sialylated T antigen is the most abundant O-glycan in the spectrum. They are expected O-glycan sequences on tumour cells in cancer progression that have been reported from previous studies (Fu et al., 2016).

MS/MS fragmentation analysis was performed for these structures to verify their identities. Figure 4.14 shows two representative MALDI-TOF/TOF spectra for *m/z* 895 and 1256. For the fragmentation of *m/z* 895 (Figure 4.14A), β-elimination gives the pair of ions at *m/z* 298 and 620. These ions are diagnostic for the β 1,3 linkage between Core 1 Gal and GalNAcitol. A pair of b and y ions at *m/z* 398 and 520 validate the present of sialic acid on this tri-saccharide structure.
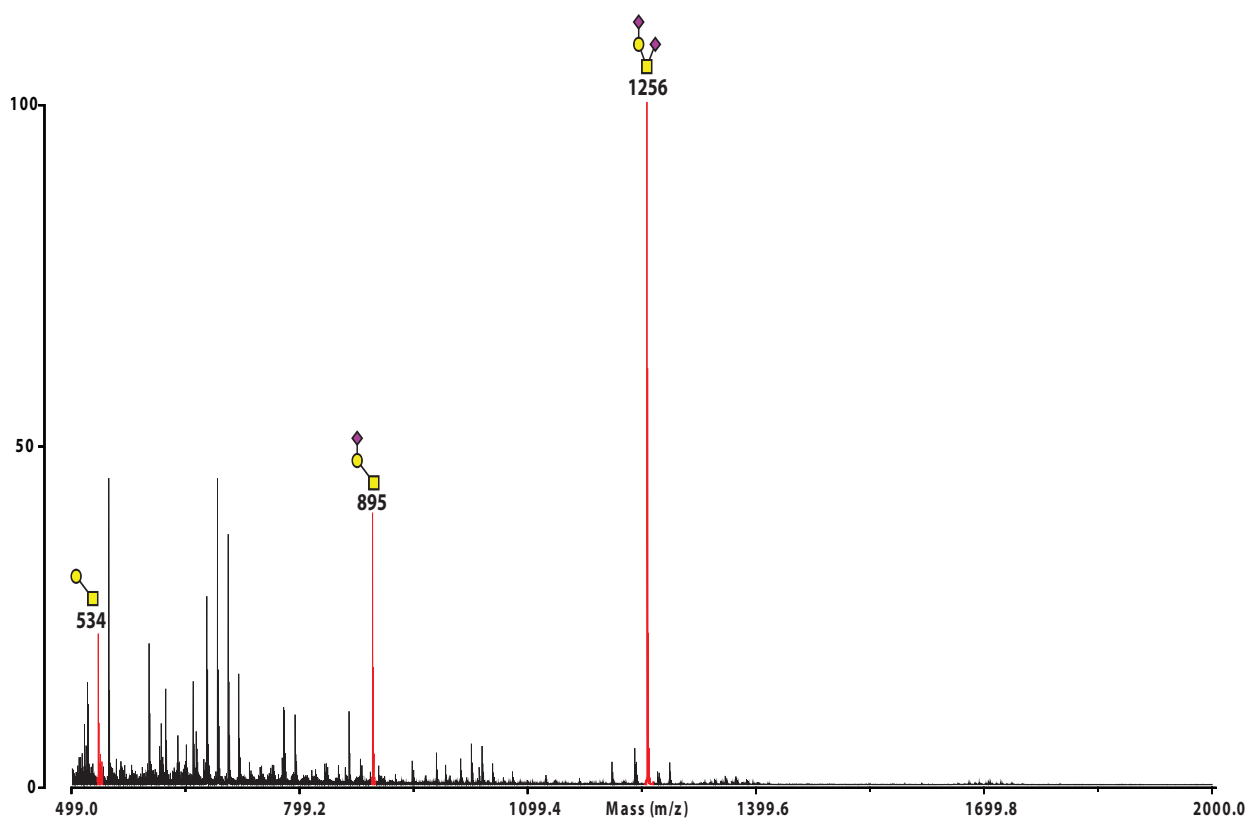
Figure 4.13 MALDI-TOF spectrum of canine melanoma O-glycans.
35% Acetonitrile fraction after purification by C18 reversed phase chromatography. O-glycans were reduced and per-methylated. Signals are detected in the form as [M+Na]+. Structures are assigned based knowledge of biosynthetic pathway and compositional information. Peaks corresponding to glycan structures are highlighted in read and annotated.
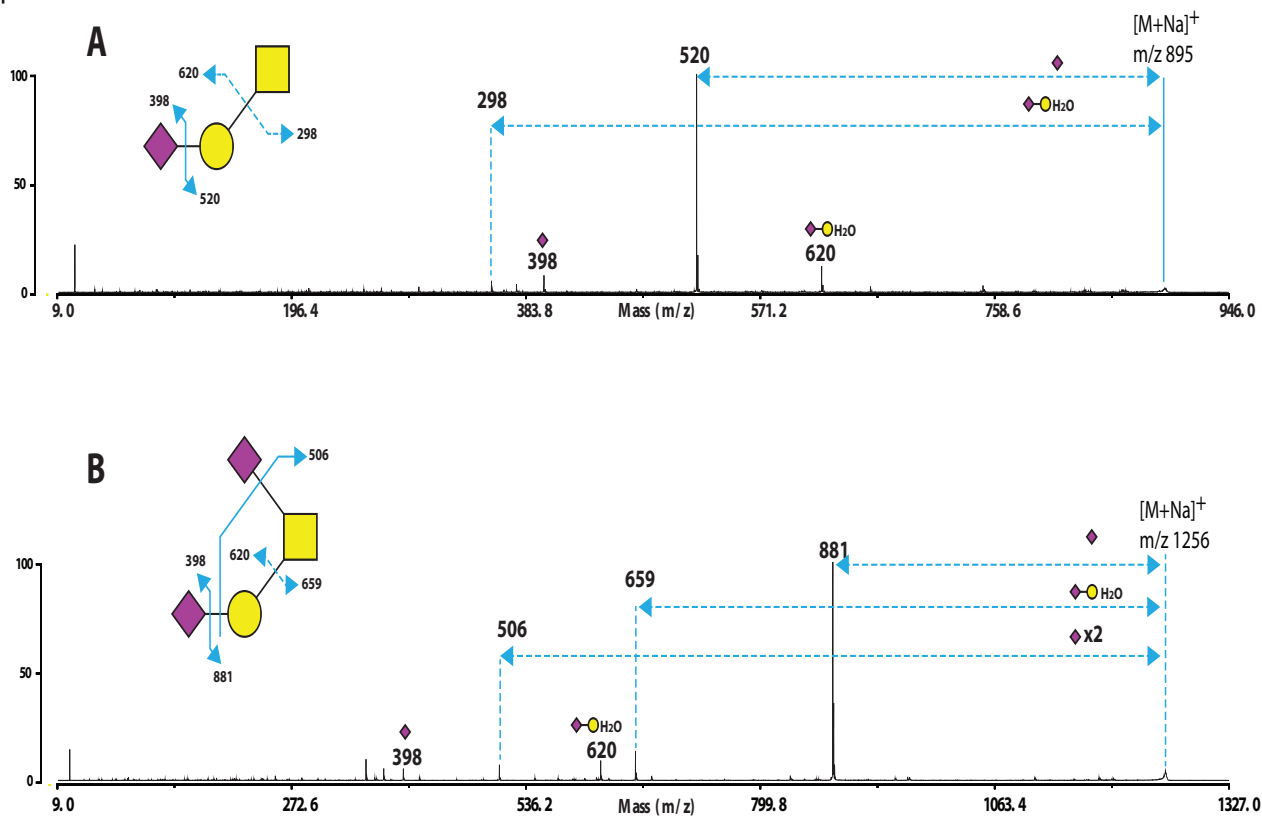
Figure 4.14 MALDI-TOF/TOF spectra of canine melanoma O-glycans for selected ions m/z 895 and 1256.
O-glycans m/z 895 (A) and 1256 (B) are selected for structure identification. All the fragment ions are observed as [M+Na]+. β-elimination on monosaccharide to lose an extra water molecule is labelled as H2O with fragment ions, and corresponding bond cleavage on the structure is represented by dashed lines with arrows.

MS/MS fragmentation of m/z 1256 (Figure 4.14B) generated five fragment ions from the putative O-glycan structure. The pair of b and y ions at m/z 398 and 881 indicates the presence of sialic acid. The y-ion at m/z 506 corresponds to loss of two sialic acids from different non-reducing ends because it has an extra 14Da decrement when compared with the single-cleaved core 1 y-ion at m/z 520 in Fig 4.14A. A β-elimination of the Core 1 arm gives the pair of ions at m/z 620 and 659 which correspond to NeuAc-Gal and NeuAc-GalNAcitol fragments, respectively.

4.2.4 Glycosylation of melanoma glycolipids

Glycolipids were physiochemically extracted from the cell lysates of canine melanoma, and further fractioned by their polarity and enzymatically digested to release the glycans from their lipid carriers. The released glycans were tagged by a deuterium (+2Da) at the C-1 of their reducing end Glc after reduction to assist data interpretation.

There are three prominent structures shared in two recovered fractions (Fraction A: Figure 4.15 and Fraction B: 4.16) of lipid-glycans. These are monosialylated glycans derived from GM1, GM2 and GM3 (*m/z* 1304, 1100 and 855, respectively). The glycan from GM3 is the most abundant component in both spectra. Di-sialylated glycans GD1, GD2 and GD3 were detected in fraction A at *m/z* 1665, 1461, and 1216 (Figure 4.15). In fraction B, there are three exclusive structures which are putatively extended from the asialo-GM1 precursor with additional LacNAc, sialyl-LacNAC and Gal-α-Gal epitopes (*m/z* 1392, 1596 and 1753; Figure 4.16).

MS/MS fragmentation analysis was performed on structures which might have variants in sequential arrangement. MS/MS spectra for *m/z* 1304, 1665 and 1753 are shown in Figure 4.17. The presence of two variants for *m/z* 1304, as GM1a and GM1b, is confirmed by y ions *m/z* 480 and 841 (Figure 4.17A). They correspond to loss of a Gal-GalNAc or NeuAc-Gal-GalNAc saccharide fragment. Comparison of their relative abundance indicates a majority of GM1a in the glyco-repertoire. The ion *m/z* 466 corresponds to a Gal-Glc di-saccharide fragment with double-cleavage which is diagnostic for the identification of GM1a. Besides, MS/MS analysis for *m/z* 1665 verifies the presence of GD1a rather than other variants (Figure 4.17B). The presence of a y-ion *m/z* 841 correlates to the loss of a NeuAc-Gal-GalNAc fragment from the non-reducing end. The ion *m/z* 466 is indicative for double-cleavage which should correspond to the losses of a Neu-Gal-GalNAc tri-saccharide and a NeuAc from two non-reducing ends.

In addition, *m/z* 1753 is validated as an extension of asialo-GM1 due to the presence of b-ion *m/z* 847 and y-ion *m/z* 929 which correspond to loss of a sialyl-LacNAc from the non-reducing end. The presence of pairing b and y ions *m/z* 480 and 1296 also validates the linear extension and the presence of the novel asialo-GM1 based structure.
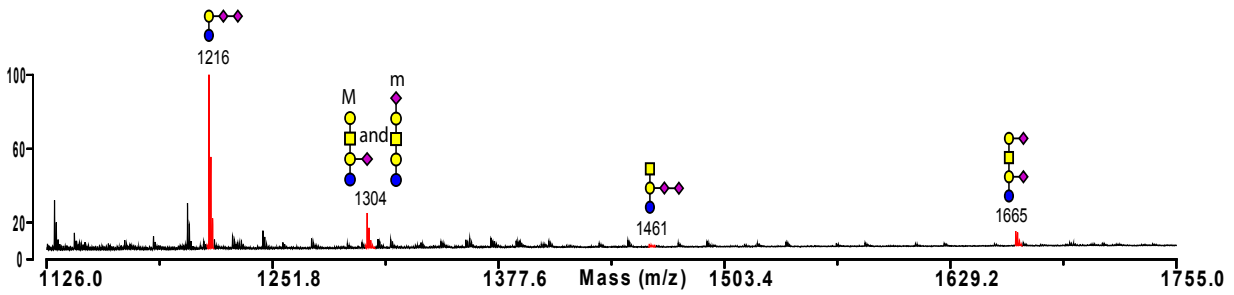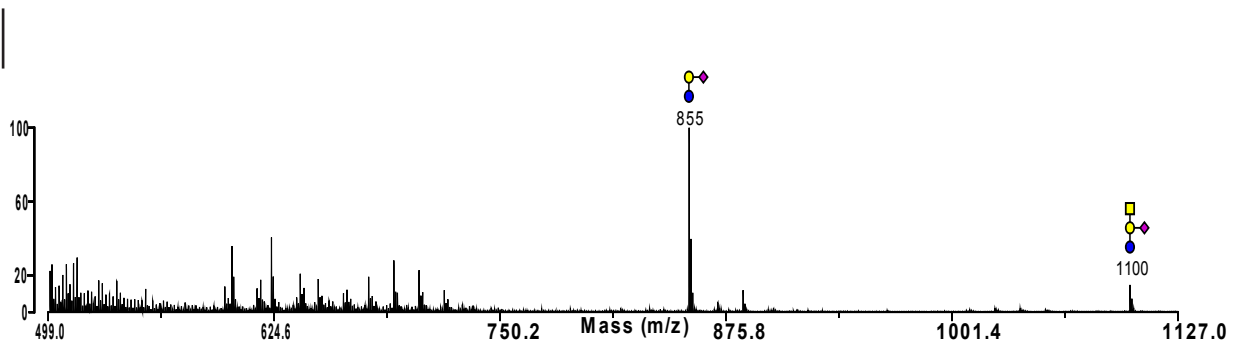
Figure 4.15 Split MALDI-TOF spectra of canine melanoma lipid-glycans from fraction A (Non-Polar).
35%Acetonitrile after purification by C18 reversed phase chromatography which was cleaned up by butanol extraction. Structures are reduced and labelled with deuterium at reducing end. All ions are observed as [M+Na]+. Peaks for lipid-glycans are highlighted in red and interpreted based on knowledge of biosynthetic pathway and compositional information
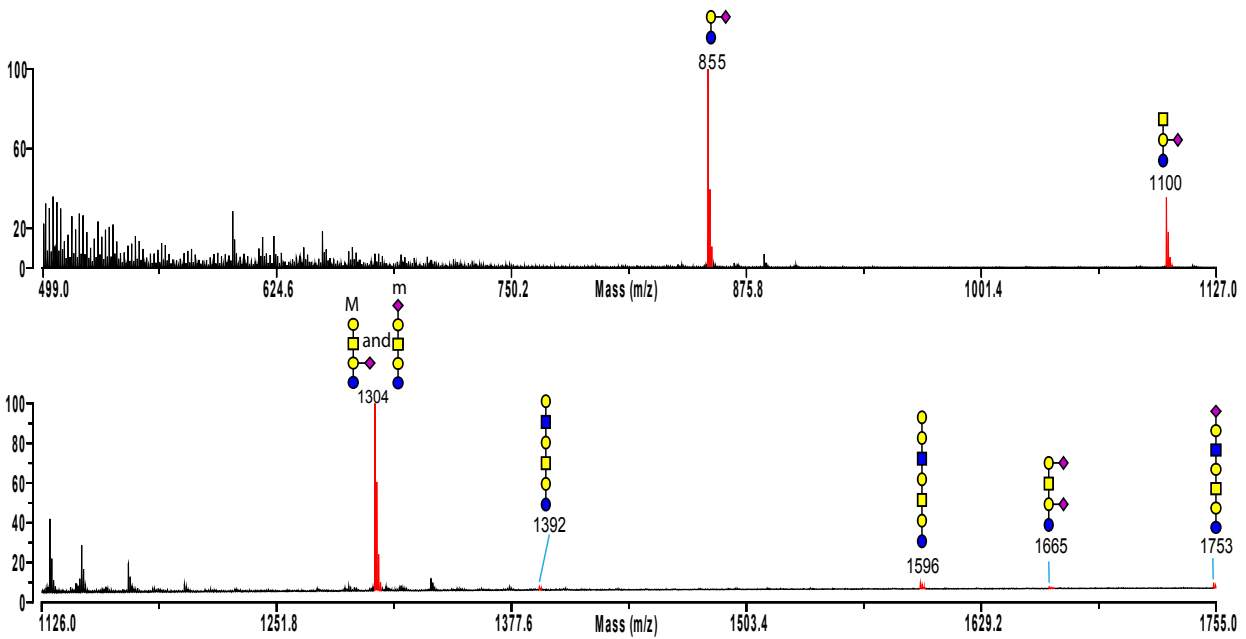
Figure 4.16 Split MALDI-TOF spectra of canine melanoma lipid-glycans from fraction B (Polar)
The glycans were cleaned up by tC18 chromatography before enzymatic release of glycans. Peaks corresponding to glycans are labelled in red and interpreted. See the legend of Figure 4.15 for assignments.
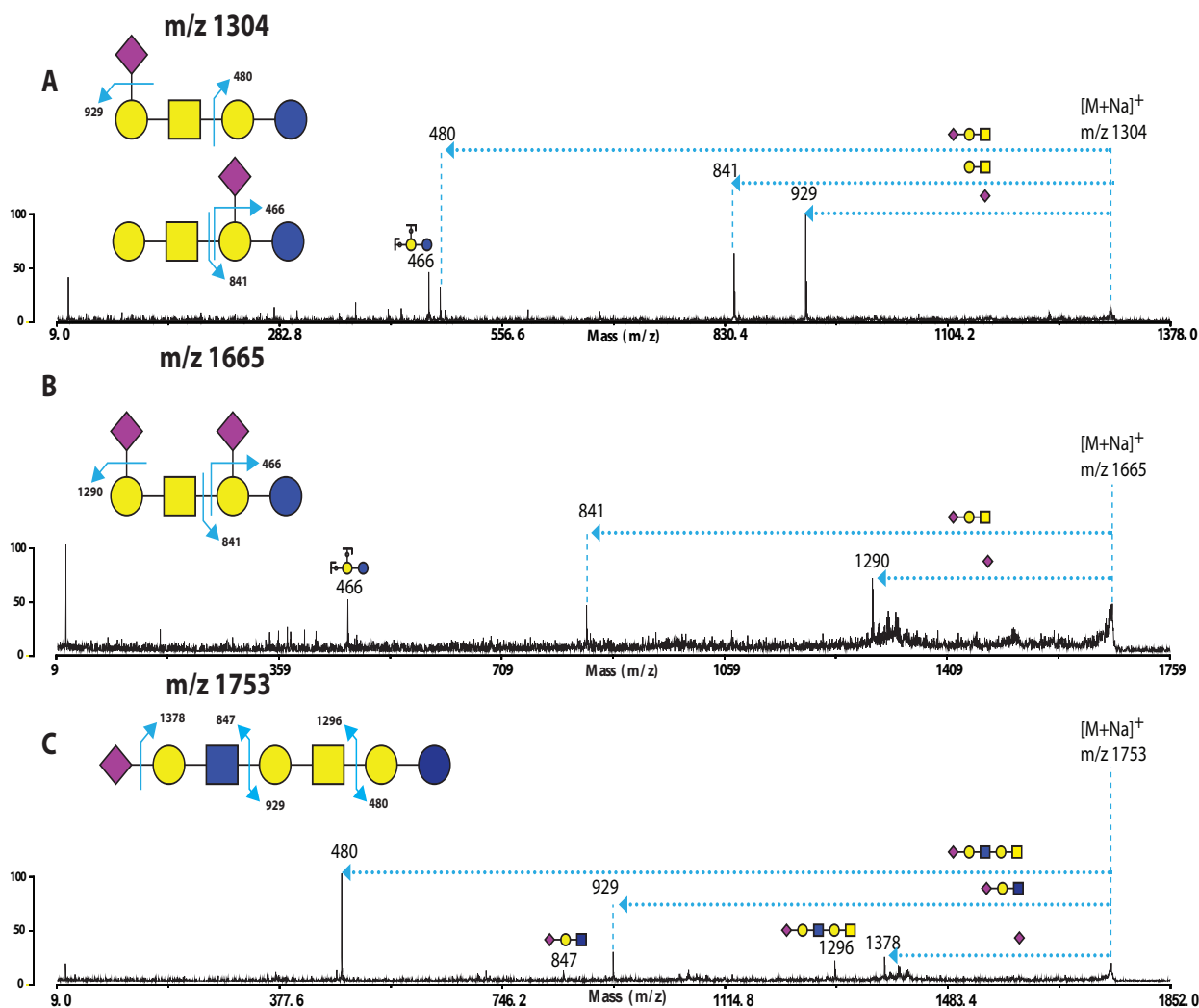
Figure 4.17 MALDI-TOF/TOF spectra of canine melanoma lipid-glycans for m/z 1304 , 1665 and 1753 from fraction B.

Three ions, 1304(A), 1665(B) and 1753(C), are selected to verify the position of their sialylation. See the legend of 4.14 for assignments.

4.2.5 Sulpho-glycomics analysis of canine melanoma glycans does not detect sulphation

To investigate the possible existence of sulphation on the glycans of canine melanoma, we conducted sulphoglycomics experiments on canine melanoma glycans and on murine small intestine O-glycans (positive control) following a previously established protocol (Yu et al., 2009, Holmen Larsson et al., 2013). The successfully permethylated sulphated glycans would have an increment of 88Da for each sulphate group when compared with its non-sulphated counterpart. This is because the sulphate group (80Da) and charge-balancing sodium cation (23Da) is observed instead of the methyl group (15Da).

In the parallel experiments, we successively identified mono-sulphated O-glycans from the positive control at *m/z* 867, 1071 and 1275 (shown in Figure 4.18). The ion *m/z* 867 corresponds to the mono-sulphated Core 2 structure. Substitution of this sulphated core structure by Gal and a further α-Gal gives the mature Core 2 structures at *m/z* 1071 and 1275.
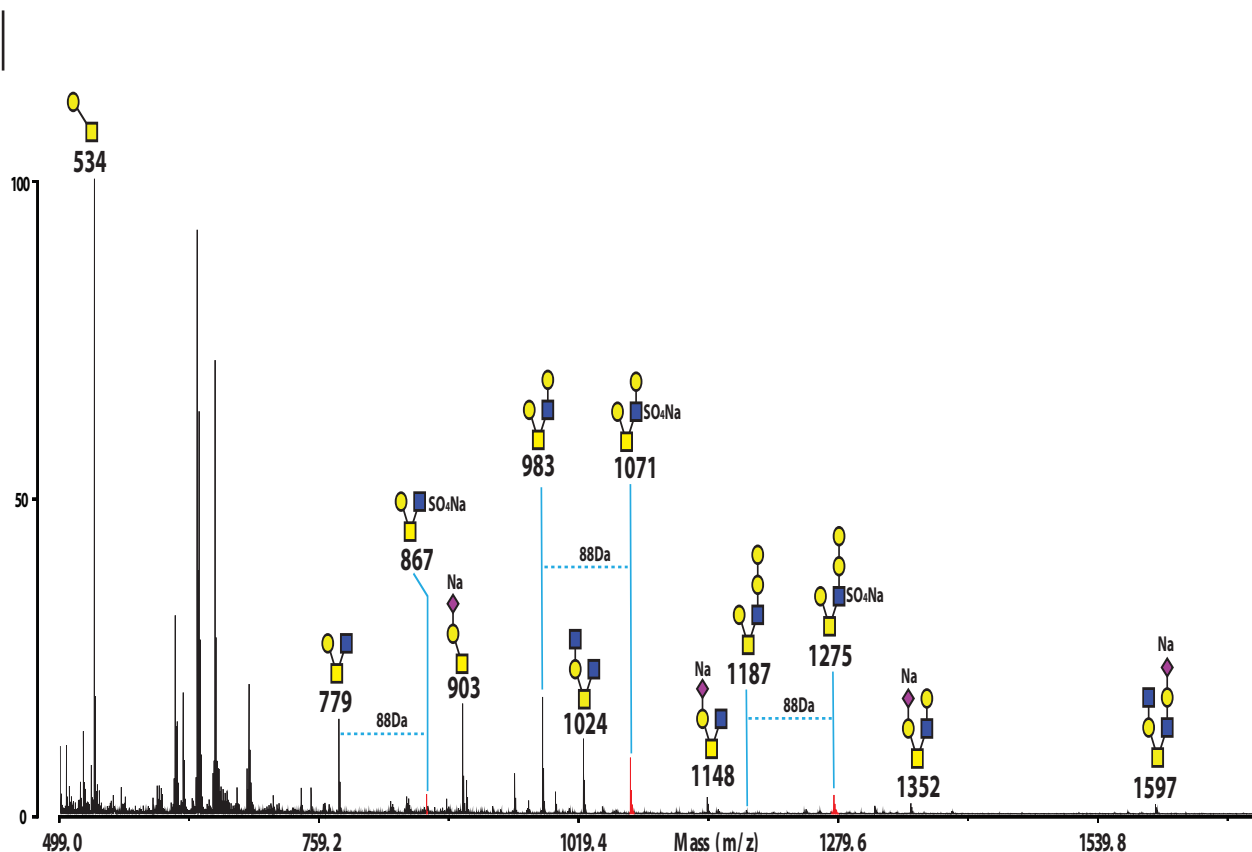
Figure 4.18 MALDI-TOF spectrum of sulphated O-glycans from murine small intestine.
The data acquisition was performed on 25% Acetonitrile fraction after purification by C18 reversed-phase chromatography. The 88Da mass shift from non-sulphated to sulphated glycans are labelled and the peaks for identified sulphated glycans are highlighted in red. The sulphation on glycans is presented as SO4Na reflecting the detected composition. All the ions are observed as [M+Na]+ and structures are assigned based the knowledge biosynthetic pathway and compositional information. The Na on sialic acid indicate a counter-balance sodium cation for its carboxyl group.

MS/MS for these structures validated their identities. Figure 4.19A (MS/MS of m/z 867) shows that cleavage at the 6-arm of GalNAcitol produced the b-ion m/z 370 and the y-ion m/z 520. The b-ion corresponds to a GlcNAc residue plus a charge balanced sulphated group indicating the presence of a sulphate group on this GlcNAc residue. A β-elimination on the 3-arm is observed via the presence of the ion at m/z 631.

MS/MS of m/z 1071 further confirms the presence of the sulphate group (Figure 4.19B) on the 6-arm antenna via the b-ion m/z 574 and the y-ion m/z 520. The b-ion at m/z 574 has an increment of 204 from the b-ion at m/z 370 derived from m/z 867. This reflects an addition of a methylated Gal residue consistent with the established pathway of core 2 O-glycan biosynthesis.

In contrast, there was no detection of sulphation in all types of examined canine melanoma glycans following the same procedure in parallel. An example is shown in Figure 4.20 as the spectrum of canine melanoma N-glycans after sulphate-permethylation. Structures with incompletely permethylated sialic acid are found in this spectrum while there is no detection of any structure corresponding to sulphated N-glycans. However, it is hard to draw a negative conclusion merely from these results without multi-dimensional verifications. Further efforts are required to optimise the methodology
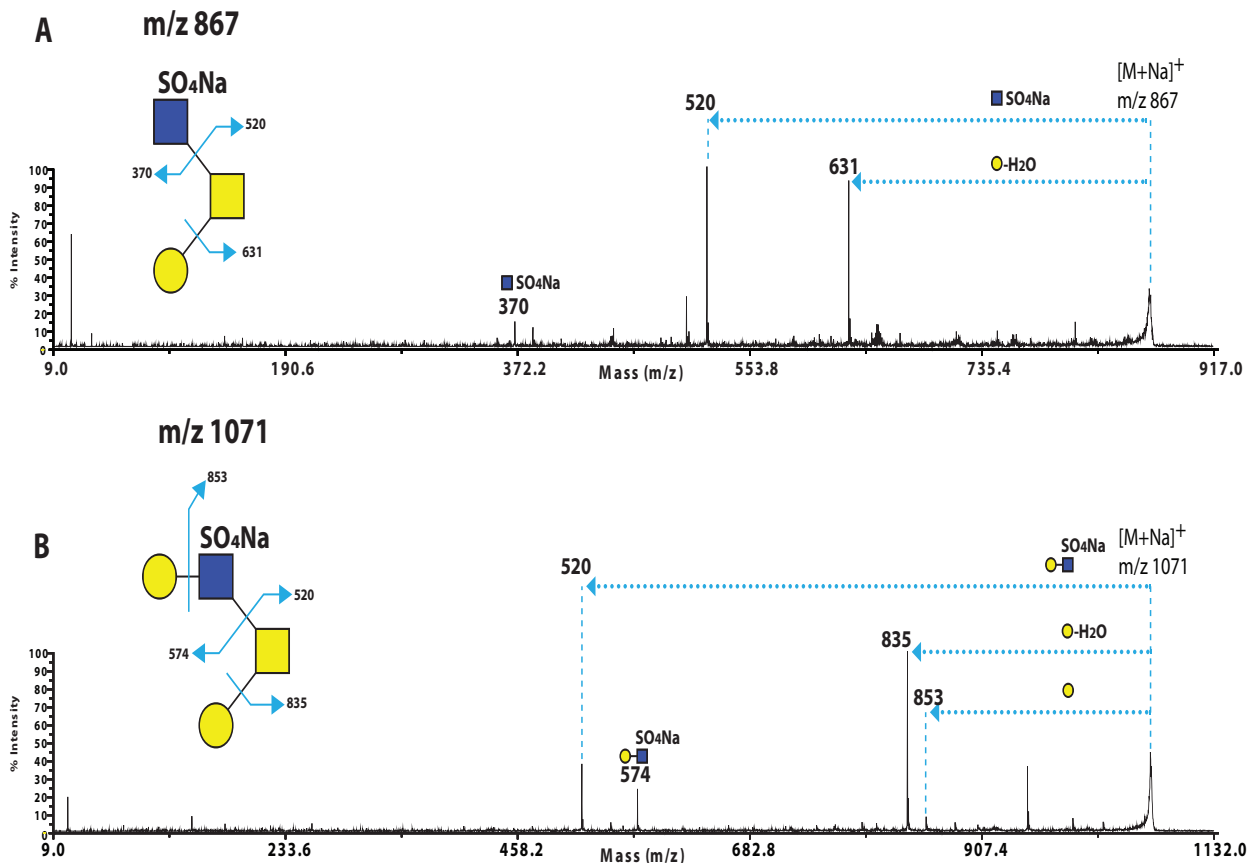
Figure 4.19 MALDI-TOF/TOF spectra of sulphated O-glycans m/z 867 and 1071 from murine small intestine. Selected putative sulphated O-glycans m/z 867(A) and 1071(B) from murine small intestine were fragmentated to verify their sulphation status. See the legend of 4.14 for assignments.

for permethylation of sulphated glycans and to provide a comprehensive detection strategy for this modification.

4.3 Discussion

In this chapter, we firstly characterised the glycomes of canine melanoma including N-linked, O-linked and lipid-glycans from a stage III patient. Detailed structural definitions were carried out for glycan sequences, glycosidic linkages and absence of sulphation via multiple biochemical and mass spectrometric technologies. The comprehensive structural characterisation enables a direct comparison of human and canine melanoma models from a glycobiological viewpoint. The results of this work will help to enhance our understandings of the consistent alterations of glycosylation in cancer progression. Moreover, this work should be immediately beneficial for the studies of veterinary medication to help improve the medical design for targeting canine melanoma with vaccines based on global inactivation of immunosuppressive carbohydrate sequences on the tumour cells(Clark, 2017).

Our results demonstrated that the canine melanoma cells have a large repertoire of bi- to tetra-antennary and elongated N-glycans with terminal sialylation or Gal-α-Gal. There is abundant core fucosylation but no antenna fucosylation was detected. MS/MS fragmentation analysis confirmed the
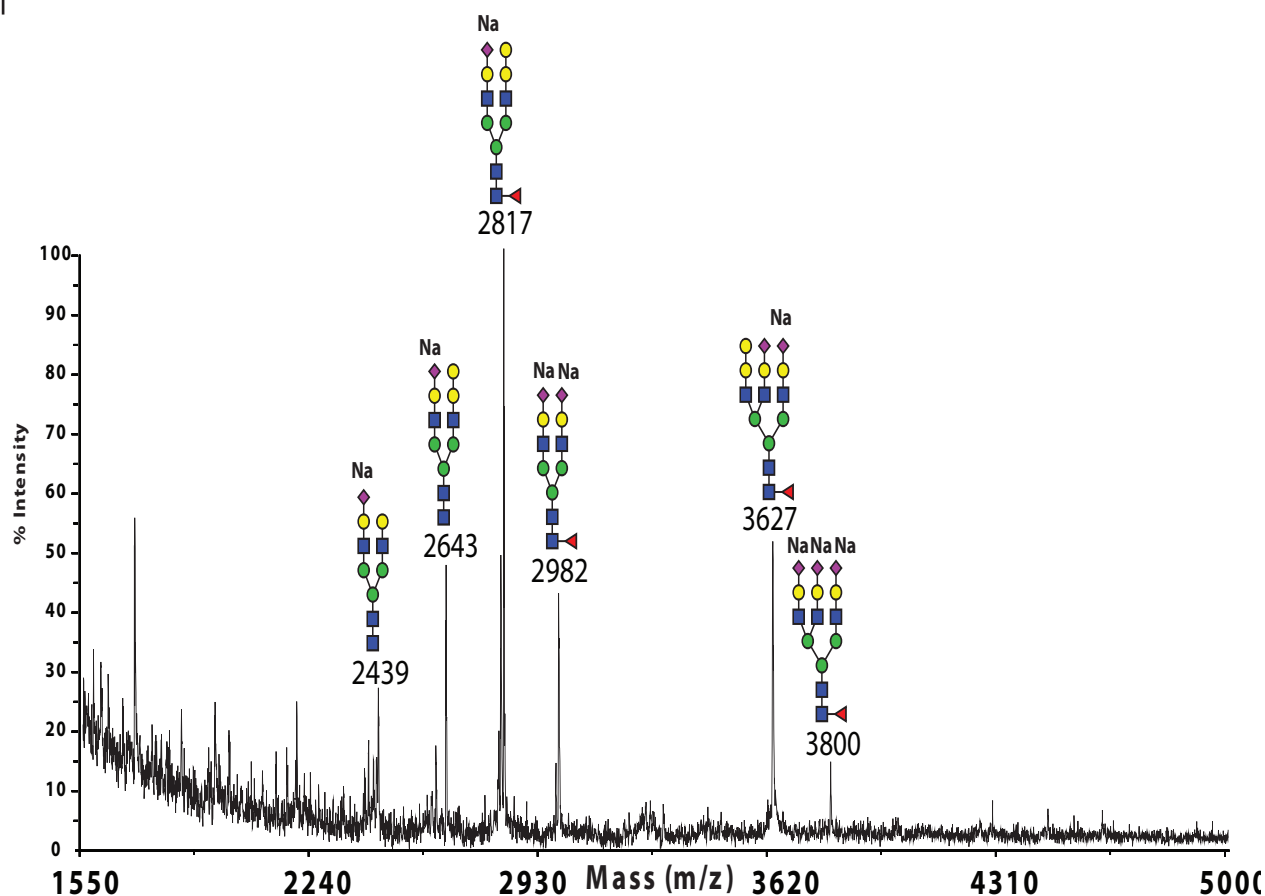
Figure 4.20 MALDI-TOF/TOF spectrum of canine melanoma N-glycans after sulphate-permethylation. Na is labelled for the counter-balance sodium cations on imcompletely permethylated NeuAc residues. See the legend of Figure 4.2 for assignments.

exclusive existence of NeuAc sialic acids; sialidase digestions and linkage analysis verified that the NeuAc sialic acids on N-glycans are mainly α2,3-linked at LacNAc terminals and there is only a trace level of α2,6-linked sialic acid. A β 1,6 branching on the 6-arm of tri-mannosyl core is validated by the presence of 2,6-linked Man in GC-MS linkage analysis. Type II poly-LacNAc is indirectly validated by the exclusive presence of 4-linked GlcNAc in GC-MS linkage analysis as well as the partial digestion by endo-β-galactosidase. Moreover, no detection of 3,6-linked GlcNAc in the GC-MS linkage analysis verifies that there is no further branching on the poly-LacNAc sequences. Besides, there is a small population of N-glycans without core-fucosylation. Due to the primary cells being cultured to expand the population *in vitro*, this minor group of N-glycans might come from the cell culture. The trace amount of α2,6-linked sialic acid could also be from the N-glycans of glycoproteins in the cell culture. Taken together, apart from the species-specific expression of the Gal-α-Gal, the characterised canine melanoma N-glycans are highly comparable to those found in human melanoma cell lines (see **Section 4.1**). This similarity would imply that the canine model being developed at the University of Missouri has promise for translating to humans.

In a recent study of human melanoma, an identical mass spectrometric analysis was conducted for

its N-glycans (Sweeney et al., 2018). This enables a direct and detailed comparison of our results of canine melanoma N-glycans to its human counterpart. In that study, two melanoma cell lines (G361 and A375) have been analysed. The de-sialylated glycome of each cell line is publicly available for comparison. In the two human cell lines, core-fucosylated structures with poly-LacNAc antennae are dominated in high mass range (*m/z* 4600-6500). Up to five LacNAc motifs have been assigned on a single antenna of a N-glycan structure, and a poly-LacNAc chain with three LacNAc moieties (two internal LacNAc) is the major species assembled onto the N-glycans of human melanoma. Moreover, bi- and tri-antennary structures are dominant within the high mass range, it could indicate that forming a long linear poly-LacNAc chain could be preferable in human melanoma than forming the fourth antenna on the N-glycans. In our characterised N-glycans of canine melanoma, the assigned poly-LacNAc chains (**Section 4.2.2.2**) are relatively shorter than that of human melanoma as observed in an equivalent desialylated batch. The major species of the poly-LacNAc chains contains mainly two LacNAc moieties (one internal LacNAc), and most structures in the high mass range (*m/z* 3900-6400) developed the fourth antenna. Moreover, due to the precisely low abundance of several large structures (*e.g. m/z* 6042) with multiple LacNAc in composition in high mass range, no sufficient information can be gained from the MS/MS experiments to fully define the maximal length of its poly-LacNAc antennae. It still possible that the longest poly-LacNAc chains on canine melanoma could have similar length to that of human counterpart. Besides, it is not surprising that the length of the N-glycan poly-LacNAc chain of canine melanoma could be shorter than that of human melanoma. Because the α1,3 galactosyltransferase expressed in canine melanoma could serve as the second type of suppressor for the poly-LacNAc extension after sialylaltransferases. Both glycosyltransferases can result in capping the Gal residues of terminal LacNAc by a sialic acids/Gal residue on N-glycans antennae to terminate atennary elongation, so that the biosynthesis of poly-LacNAc chains on N-glycans from canine melanoma could have a higher probability to be terminated earlier than that in human melanoma. In addition, due to there is no data publicly available about the sialidase-untreated N-glycans from the same study of human melanoma, the difference of sialylation levels between human and canine melanoma cannot be comparatively evaluated. However, it is likely that the sialylation level of human melanoma could be higher in some extent than that of canine melanoma. Because the α1,3 galactosyltransferase also competes with sialyltransferases for the terminal LacNAc acceptor. This is also reflected on the high mass range structures that even after sialidase digestion, the vast majority of high mass range of structures carry at least one terminal Gal-α-Gal. The functional impact of these structural features could be reflected on the reactivity of galectins. For instance, galectin-1 would have a greater binding on canine melanoma cells rather than human melanoma cells because of its binding specificities for terminal Gal and branched structures. Conversely, galectin-3 could have a superior binding on human melanoma cells rather than their canine counterparts due to the binding preference of galectin-3 to internal Gal on a longer poly-LacNAc chain. Thereby, it could be interesting to evaluate the reactivity and expression of galectins in the tumour microenvironment of canine melanoma. This could possibly enhance our understandings of the cancer progression in this animal model and also facilitate

detailed evaluation of the compatibility of this animal model and potential differences.

Because a study of tissue plasminogen activator from Bowes melanoma cell lines revealed that some of its N-glycans have sulphation on the chitobiose GlcNAc (Zamze et al., 2001) we investigated whether sulphate is found on the canine melanoma N-glycans. None was detected. This is perhaps not surprising because the glycosylation of tissue plasminogen activator from Bowes melanoma is very different from that reported for other human melanoma cell lines. For example, it carries lacdiNAc glycans that are typical of neuro-endocrine glycoproteins (Chan et al., 1991). Such glycans have not been found on other human melanoma cell lines (Sweeney et al., 2018). Nevertheless, further work will be required to establish whether sulphation could be present in the canine melanoma cells at levels not currently detectable by our sulphoglycomics methodologies (Clark, 2017).

Although we were able to verify the presence of type II poly-LacNAc extension by indirect linkage analysis, concerns are still drawn for the enzyme activities of commercially available endo-β-galactosidases. In our laboratory, we have historical records of using a well-established condition (sodium acetate buffer) to perform this digestion with the recombinant enzyme from *Escherichia freundii* which is commercially discontinued (Choo et al., 2017, Walther et al., 2013). In this study, we examined all the other available recombinant enzymes from *Bacteroides Fragilis*. It is very surprising that there was nearly no activity of all tested enzymes; and even some of them might have other emzymatic effects on tested glycans such as de-sialylation. An early study examined the efficiency of this form of recombinant enzymes from the *B. Fragilis* origin, but no compromise of biological activities was suggested to be caused by applying the sodium acetate condition (Scudder et al., 1983). Therefore, it might be likely that all the enzymes were inactive in this study; future studies utilising these enzymes from *B. Fragilis* should have control sets or verification of their biological activities prior to investigation.

The O-glycosylation of canine melanoma generally matches the expectation of cancer associated glycosylation as there is abundant expression of T antigen and its sialylated forms (Fu et al., 2016). Some studies suggested the reactivity of T antigen and possibly its α2,3 monosialylated form to galectin 3 (Glinskii et al., 2012). The predominantly expressed mono- and di-sialylated T antigens would be reactive to siglec receptors such as siglec-4 that has a remarkably high affinity for both sialylated T antigens to facilitate cancer invasion (Swanson et al., 2007). However, these suggestions and the functionalities of O-glycans are still in need to be evaluated for melanoma.

Our results illustrate a lipid glycan repertoire which is dominated by GM3 and GM2 structures with identification of less-dominant GD2, GD1, GD3 and GM1 structures. All these structures are associated with the malignancy of tumours as they can affect cell signalling and mobility to favour cancer progression (Kannagi et al., 2018). Moreover, GD2 and GD3 have also been shown to be highly reactive with siglec-7, which is an inhibitory lectin receptor commonly expressed on NK cells to modulate their cell cytotoxicity (Jandus et al., 2014). In addition, the structural and quantitative significances of

the other identified lipid-glycan structures in this study might require a further validation via a comparison with that of original canine melanocytes.

Overall, the glycosylation of canine melanoma is highly consistent with previous glycomics studies of melanoma on human cell lines (summarised in **Section 4.1**) except for the expression of the Gal-α-Gal moiety. This exception is a common case in most mammalian systems, and the presence of this terminal moiety might not have a significant impact on glycan-lectin related interactions in canine models because there is no characterised lectin to specifically recognise Gal-α-Gal. It is very likely that the underlying LacNAc moiety is still reactive to galectins such as galectin 1 since the α1,3 Gal might be equivalent to α2,3-linked sialyl-LacNAc which galectin 1 can bind to but not to α2,6-linked sialyl-LacNAc (Zhuo and Bellis, 2011). Thus, the lectin receptor engagements of canine melanoma should be generally similar to human counterparts as the identical structural features are shared and the canine melanoma model should be capable to serve as a representative pre-clinical model (Pinho and Reis, 2015, Sweeney et al., 2018).

Chapter V. Fc Fusion Protein

Chapter 5 Fc Fusion Protein

5.1 Introduction

The fragment crystallisable region (Fc) is a part of the sequence of constant region of the Ig heavy chains which retains most biological effector activities of an antibody, such as interactions with Fc Receptors (FcRs), but without antigenic recognition. In 1989, this sequence was firstly combined with a CD4 membrane protein sequence to specifically inhibit viral infection of T cells by human immune deficiency virus (Capon et al., 1989). This technique was then broadly recruited for pharmaceutical applications to treat various immune relevant diseases utilising the advantages of the Fc region as well as its jointed sequence. The basic structure of Fc fusion proteins are to combine the Fc with an additional peptide or to modify the Fc structure alone (Czajkowsky et al., 2012). This enables the entire sequence to have a longer half-life in circulation due to altered interaction with the salvage neonatal FcR (Roopenian and Akilesh, 2007) as well as a slow-down of renal clearance (Kontermann, 2011). With the retention of this Fc region, Fc fusion molecules would be interactive with FcRs expressed on immune cells to enhance effector function or suppress unwanted immune responses. For instance, the major approved therapeutical antibodies are the IgG class. IgG1-Fc can be structurally modified to improve its binding to FcγRIIIA on NK cells for enhancing NK directed ADCC (Strohl, 2009). On the other side, IgG1-Fc fused with CD19 could dramatically increase the binding affinity of this hybrid molecule to inhibitory FcgammaRIIb which can lead to suppression of immune responses (Chu et al., 2008). In addition, due to the varied binding affinity of Fc's to cellular receptors, mainly FcRs, polymerisation could be employed to form Fc immune complexes for improving avidity and potency (Bruhns et al., 2009). Particularly, major protective FcRs including FcγRIIA, FcγRIIB and FcRL5 require to interact with high avidity immune complexes of antibodies or Fc-fusion proteins for inducing their protective immune responses (Wilson et al., 2012, Nimmerjahn and Ravetch, 2008).

N-glycosylation on human IgG is conserved at the N297 site in its Fc sequence, this N-glycosylation site generally accommodates a glycoform that is dominated by core-fucosylated bi-antennary complex structures with varied levels of galactosylation and subsequent terminal sialylation. The N-glycosylation on this site is required to enable the effector functions of IgGs and Fc fusion proteins (Jefferis, 2009). For instance, the core-fucosylation of N-glycans on IgG-Fc can reduce its binding affinity to FcγRIIIA leading to suppression of ADCC activities (Ferrara et al., 2011), and subsequent removal of this monosaccharide residue can reverse the effect to enhance the binding and activation of FcγRIIIA in vitro and in vivo (Ferrara et al., 2011, Umana et al., 1999). Moreover, lack of galactosylation and subsequent sialylation has long been associated with rheumatoid arthritis (Parekh et al., 1985), and engineered Fc fusion proteins with full galacosylation might minimise their interaction with C-type

lectins on DCs as well as subsequent internalisation and antigen presentation (Carrillo-Conde et al., 2011). In addition, sialylation, in particular α2,6-linked, on the N-glycans of recombinant Fc-fusion proteins can interact with DC-SIGN on macrophages to induce its expression of FcγRIIb receptor and subsequently suppress arthritic inflammation (Anthony et al., 2011). Thus, the N-glycosylation on Fc-fusion protein is a vital target for pharmaceutical design to optimise diverse immunological functions.

To utilise the immune suppressive role of Fc fusion protein, attempts to alter the sialylation of N-glycans on Fc fusion proteins have been attempted. The natural content of sialylated structures at N297 is less than 2% on Fc sequence expressed in a CHO-K1 cell system (Blundell et al., 2017), and the assembled N-glycans are largely buried in the cavity formed by the Fc CH2 and CH3 domains (Subedi et al., 2014, Frank et al., 2014). Chemical modifications of glycans and mutagenesis programs for

Table 5.1 IgG1-Fc constructs
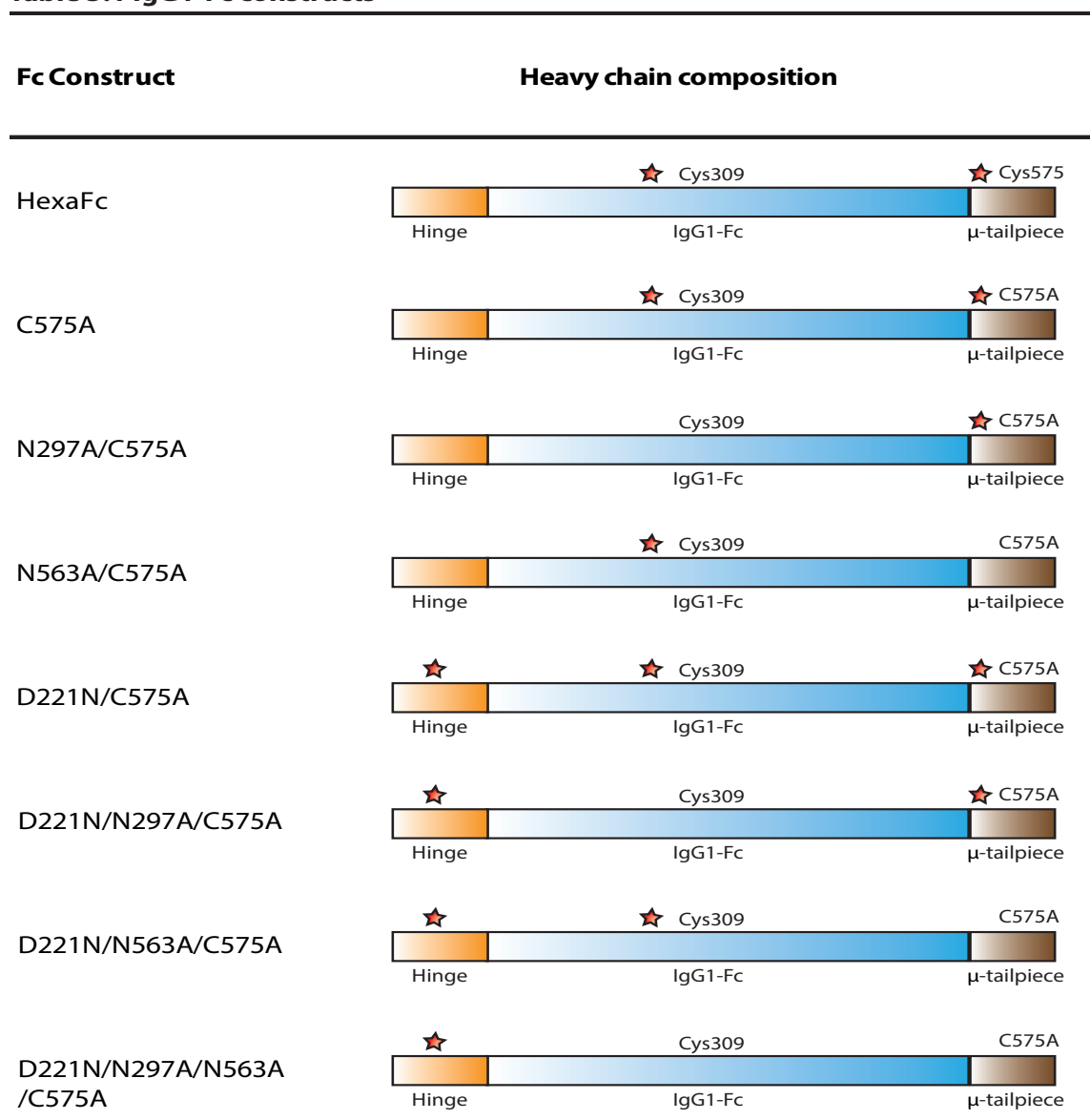
Fc Construct	Heavy chain composition
HexaFc	Hinge — IgG1-Fc (★ Cys309) — μ-tailpiece (★ Cys575)
C575A	Hinge — IgG1-Fc (★ Cys309) — μ-tailpiece (★ C575A)
N297A/C575A	Hinge — IgG1-Fc (Cys309) — μ-tailpiece (★ C575A)
N563A/C575A	Hinge — IgG1-Fc (★ Cys309) — μ-tailpiece (C575A)
D221N/C575A	Hinge (★) — IgG1-Fc (★ Cys309) — μ-tailpiece (★ C575A)
D221N/N297A/C575A	Hinge (★) — IgG1-Fc (Cys309) — μ-tailpiece (★ C575A)
D221N/N563A/C575A	Hinge (★) — IgG1-Fc (★ Cys309) — μ-tailpiece (C575A)
D221N/N297A/N563A/C575A	Hinge (★) — IgG1-Fc (Cys309) — μ-tailpiece (C575A)

Stars represent N-glycosites (N221, N297 and N563) on recombinant IgG1-Fc mutants. The yellow bar represents the sequence of hinge region, the blue bar represents the sequence of IgG1-Fc and the brown bar represents the sequence of attached IgM tailpiece. Cys575 is mutated to creat the C575A panel of mutants. Adapted from (Blundell et al., 2019)

Table 5.1 IgG1-Fc constructs (continued)
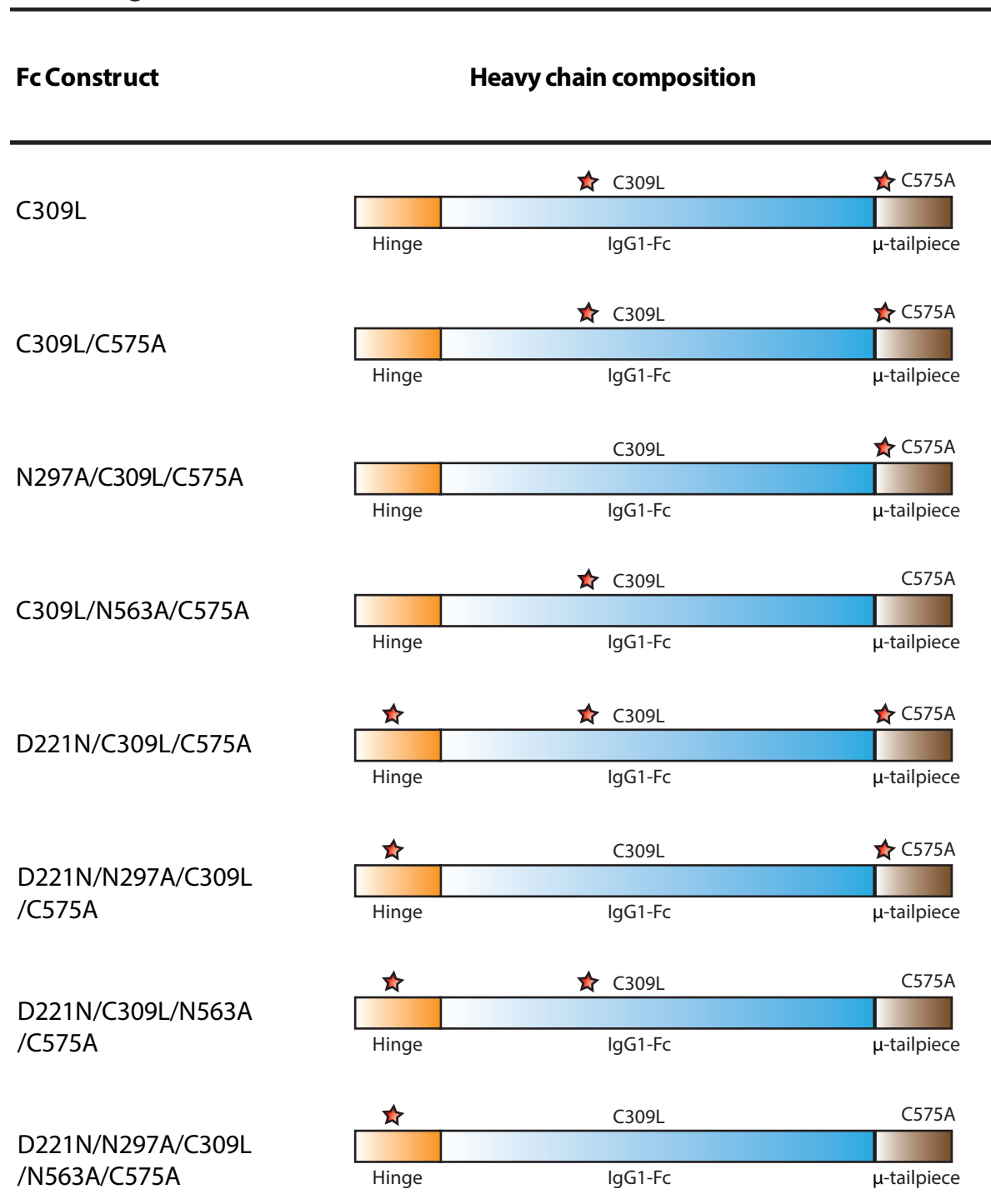
Fc Construct	Heavy chain composition

C309L

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

C309L/C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

N297A/C309L/C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

C309L/N563A/C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

D221N/C309L/C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

D221N/N297A/C309L /C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

D221N/C309L/N563A /C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

D221N/N297A/C309L /N563A/C575A

C309L · C575A
Hinge · IgG1-Fc · µ-tailpiece

Stars represent N-glycosites (N221, N297 and N563) on recombinant IgG1-Fc mutants. The yellow bar represents the sequence of hinge region, the blue bar represents the sequence of IgG1-Fc and the brown bar represents the sequence of attached IgM tailpiece. Cys309 and Leu310 are mutated on the C575A panel to Leucine and histidine to create C309L/C575A panel of mutant. Adapted from (Blundell et al., 2019).

Table 5.2 Polymeric status of engineered IgG1-Fc construct made in CHO cell system assessed by SDS-PAGE and SEC-HPLC

Fc-construct	Polymeric Status	
	SDS-PAGE	SEC-HPLC
Hexa-Fc	barrel	barrel
C575A	monomer >dimer >trimer	monomer >dimer
N297A/C575A	monomer >dimer >trimer	monomer >dimer
N563A/C575A	ladder	monomer, multimer
D221N/C575A	monomer >dimer	monomer >dimer
D221N/N297A/C575A	monomer >dimer	monomer
D221N/N563A/C575A	ladder	multimer
D221N/N297A/N563A/C575A	ladder	multimer
C309L	ladder	multimer
N297A/C309L/C575A	monomer	monomer
C309L/N563A/C575A	monomer	monomer
N297A/C309L/N563A/C575A	monomer	multimer
D221N/C309L/C575A	monomer	multimer
D221N/N297A/C309L/C575A	monomer	monomer
D221N/C309L/N563A/C575A	monomer >dimer	multimer
D221N/N297A/C309L/N563A/C575A	monomer	monomer, multimer >dimer

Adapted from (Blundell et al., 2019)

altering the protein carbohydrate surface have been used to improve the overall sialylation content on pre-existing glycans. However, neither approaches have been shown to be efficient and cost-effective (Dekkers et al., 2016, Yu et al., 2013).

Recently, our collaborator introduced two alternative modifications to improve the sialic acid content of Fc glycans and polymerisation of Fc regions. The first is to add a tailpiece from IgM (18aa) to the C-terminal of the Fc sequence, the second is to add an extra N-glycosylation site onto the N-terminal part of the Fc at position 221 (Blundell et al., 2017). This approach resulted in a radical increase of sialylation on glycans to above 75% on mono- and multi-mers of Fc molecules. IgG-Fc with an increased sialic acid content can bind to glycan receptors including Siglec 1 and myelin-associated glycoprotein, which are functionally implicated in manipulating neuropathology (Wong et al., 2016).

The previously established Fc sequence contains two Cysteine residues (C309 and C575) and two N-glycosylation sites (N297 and N563), which can form homo-hexamers of Fc between Fc sequences, named Hexa-Fc (Blundell et al., 2017, Czajkowsky et al., 2015) to improve their binding affinity with FcγRIIb. In this study, two panels of mutants were established on this sequence, generally named IgG1-Fc, listed in Table 5.1. The two panels of IgG1-Fc differed in relation to the presence of cysteine residue, C309; and the insertion of a possible third N-glycosylation site (D221N). These different mutants are associated with varied polymerisation status, electrophoretic and chromatographic migrations (Table 5.2) as well as divergent binding capabilities to Fc and glycan receptors. Hence, it raised a

great interest to define the N-glycosylation of each mutant structurally and semi-quantitatively.

In this chapter, we analysed these two sets of IgG1-Fc mutants stably produced from CHO and HEK293 cell expression systems. We aim to understand the impact of these mutations on the N-glycosylation of the Fc sequence and subsequent functional consequences.

5.2 Results

5.2.1 Research strategy

The project was undertaken as a collaboration with Prof Richard Pleass from The Liverpool School of Tropical Medicine. They generated and provided two sets of recombinant Fc fragments stably expressed in either CHO or HEK293 cell systems. The recombinant sequences of the Fc fragments are summarized in table 5.1.1 and detailed research methodology are described in chapter 2. Briefly, provided IgG1-Fc material was lyophilised and their N-glycans were directly released by rapid PNGase F. Released N-glycans were purified and permethylated for MALDI-TOF analysis.

5.2.2 N-glycosylation of IgG1-Fc mutants from CHO cell system.

The general N-glycosylation of CHO cell system was described in **Chapter 3**. Briefly, N-glycans assembled on glycoproteins by this system feature as core-fucosylated, well-branched and elongated N-glycan structures with exclusive α2,3 linked NeuAc sialylation. Minimal levels of non-human glycans such as NeuGc or Gal-α-Gal in CHO cell lines have been noted. Therefore, this cell system is still a practical model for the production of human-like glycoprotein (North et al., 2010, Canis et al., 2018).

5.2.2.1 N-glycosylation at N221, N563 and N297 of IgG1-Fc mutants from CHO cell system

The N-glycosylation of N563 from the IgM tailpiece and N221 by insertion of the (D221N) hinge onto IgG1-Fc is characterised from multimeric mutants N297A/C575A (Figure 5.1A) and D221N/N297A/N563A/C575A (Figure 5.1B). They both have a single N-glycosylation site available. Both recombinant sequences share prominent sialylated bi-antennary structures (*m/z* 2605 and 2966) as well as bi-antennary structures with mono- and di-galactosylation on antennae (*m/z* 2040 and 2244). Tri- and tetra-antennary structures with sialylation (*e.g. m/z* 3776 and 4226) and even antennary elongation (*e.g. m/z* 4314 and 4675) are also detectable in the spectra. Mutant N563A/C575A was also analysed for N297 glycosylation (Figure 5.1C). A human-like Fc N-glycosylation pattern recovered from N297
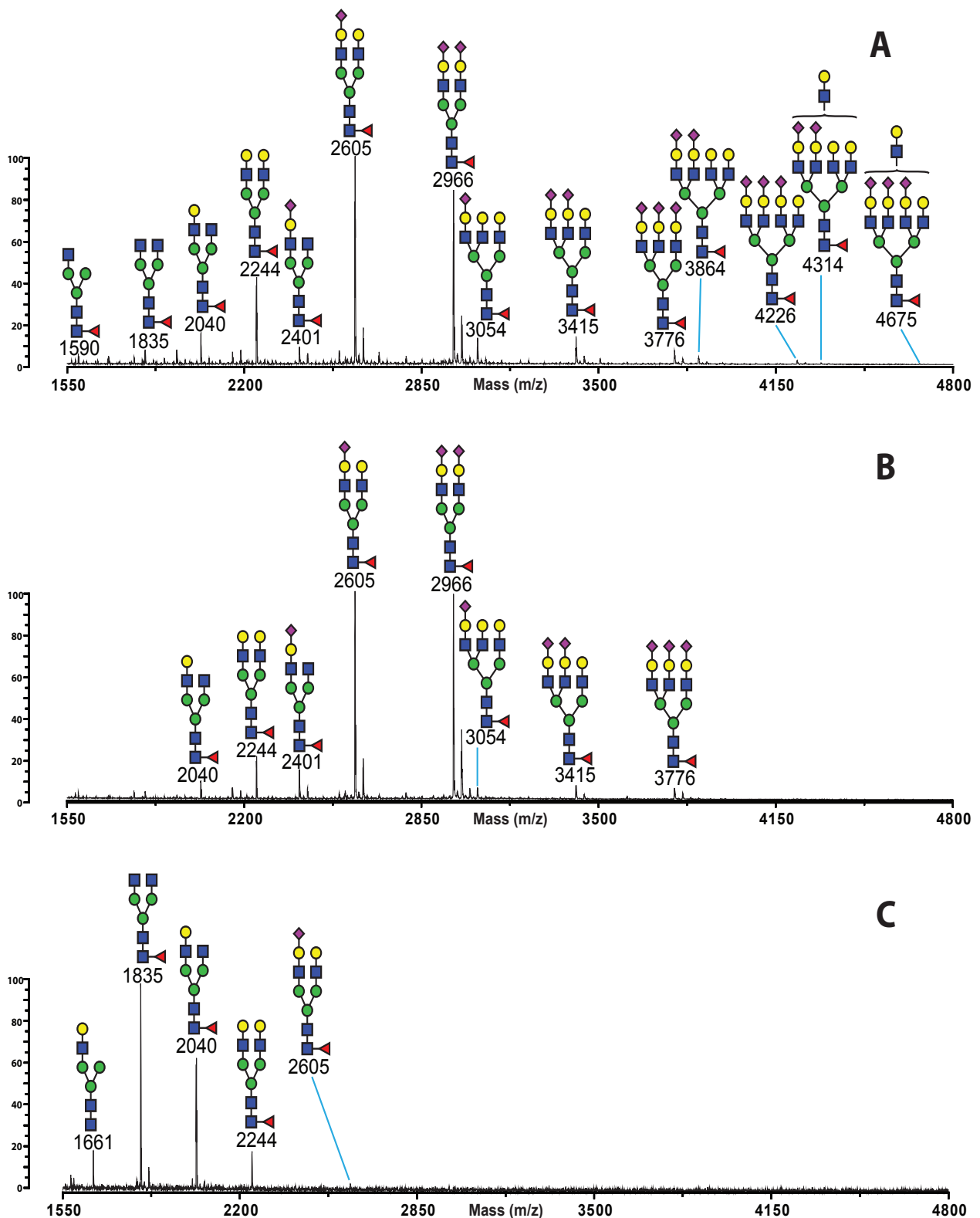
Figure 5.1 MALDI-TOF spectra of IgG1-Fc mutants made in CHO cell system with diverse single N-glycosylation site availability on sequence.
N297A/C575A mutant for N563 (A), D221N/N297A/N563A/C575A mutant for N221 (B) and N563A/C575A mutant for N297 (C). Data acquisition was conducted in the positive mode of MS and all ions are observed as $[M+Na]^+$. The structures are assigned upon knowledge of biosynthetic pathway of N-glycans as well as compositional information. Saccharides placed above bracket indicate their undetermined antennary linkages.

Multimeric

Monomeric

N297A/C575A

A

■ 74.29% Sialylated
■ 25.71% Netural

N297A/C309L/C575A

D

■ 52.79% Sialylated
■ 47.21% Netural

N563A/C575A

B

■ 1.56% Sialylated
■ 98.44% Netural

C309L/N563A/C575A

E

■ 3.55% Sialylated
■ 96.45% Netural

D221N/N297A/N563A/C575A

C

■ 81.82% Sialylated
■ 18.18% Netural

D221A/N297A/C309L/N563A/C575A

F

■ 24.25% Sialylated
■ 75.75% Netural
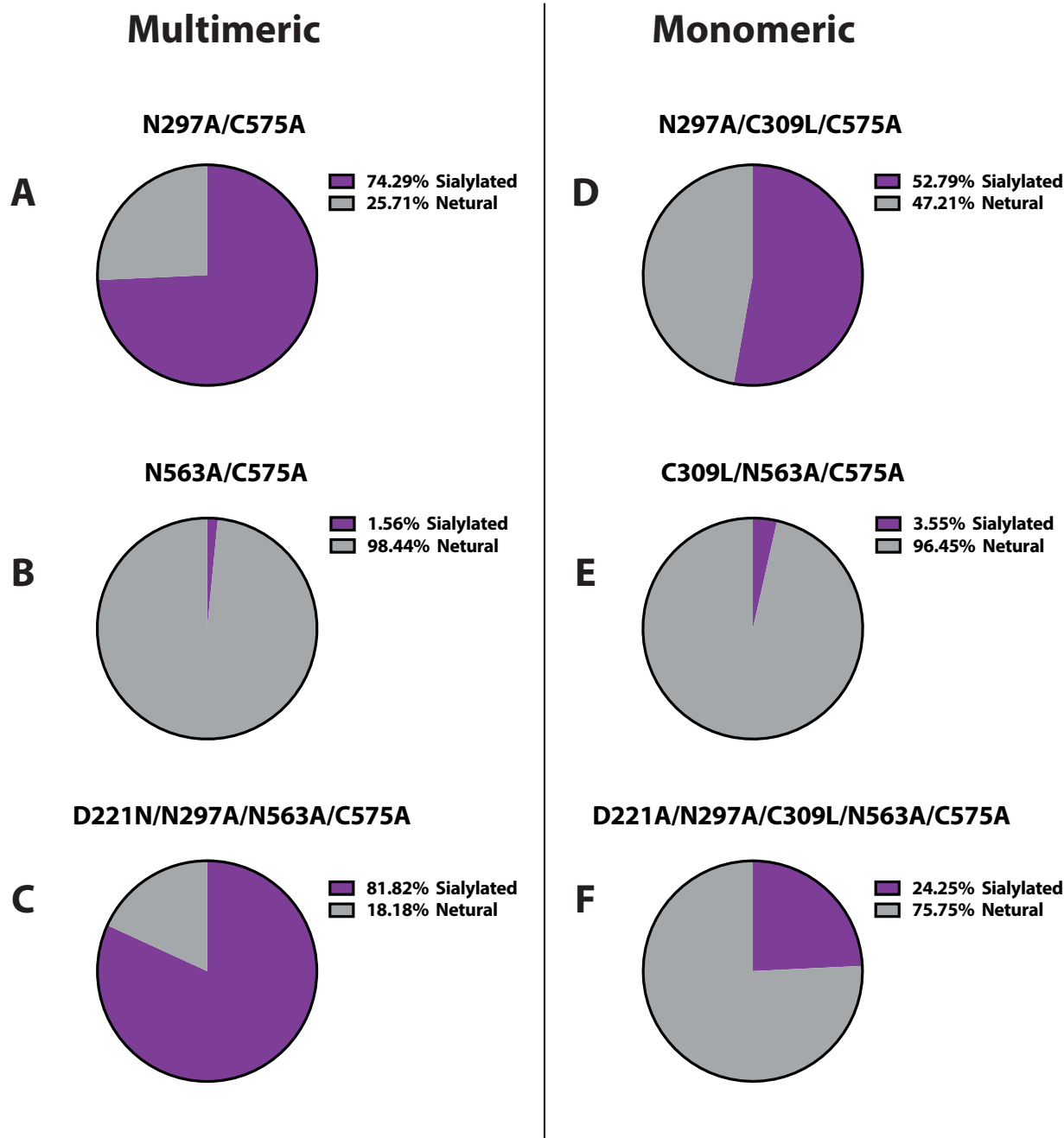
Figure 5.2 Semi-quantitative estimation of sialylation levels of IgG1-Fc mutants made in CHO cell system with single available N-glycosylation site in multimeric and monomeric statuses.
The levels of sialylated (purple) and netural (grey) glycans are reflected in percentage from their summed intensity.

indicating the compatibility of CHO cell system for antibody glycosylation on this conserved N-glycosylation site (Kiyoshi et al., 2017, Blundell et al., 2017). The observed glycome on this site is dominated by the core-fucosylated bi-antennary structures (*m/z* 1835, 2040 and 2244), and trace amounts of core-afucosylated mono antennary (*m/z* 1661) as well as mono-sialylated bi-antennary structures (*m/z* 2605).

The levels of sialylation on glycans from each mutant were principally and semi-quantitatively analysed by comparing the ratio of summed intensity of sialylated structures to the total intensity of all

observed glycan structures. Percentages of sialylated and neutral structures were calculated and visualised as pie charts in Figure 5.2. Mutants N297A/C575A and D221N/N297A/C575A have 74.29% and 81.82% of their intensity contributed from sialylated structures, respectively. This predominance of intensity contribution shows their high levels of sialylation. In contrast, only 1.56% of the overall intensity of N297 N-glycans (N563A/C575A) is from sialylated structures which matches with previous suggestion that less than 2% of glycans structures were sialylated on this site on Fc expressed in CHO cell systems(Blundell et al., 2017).

To assess the possible impacts of disulphide bonding on multimerization of Fc sequences and N-glycosylation of these individual sites, a set of monomeric mutants with a substitution of CL309-310LH (C309L) was also analysed to validate the impacts on the described site-specific N-glycosylation. MALDI-TOF spectra of the set of mutants (D221N/N297A/C309L/N563A/C575A, N297A/C309L/C575A and C309L/N563A/C575A) are shown in Figure 5.3. The N-glycosylation of N563 (N297A/
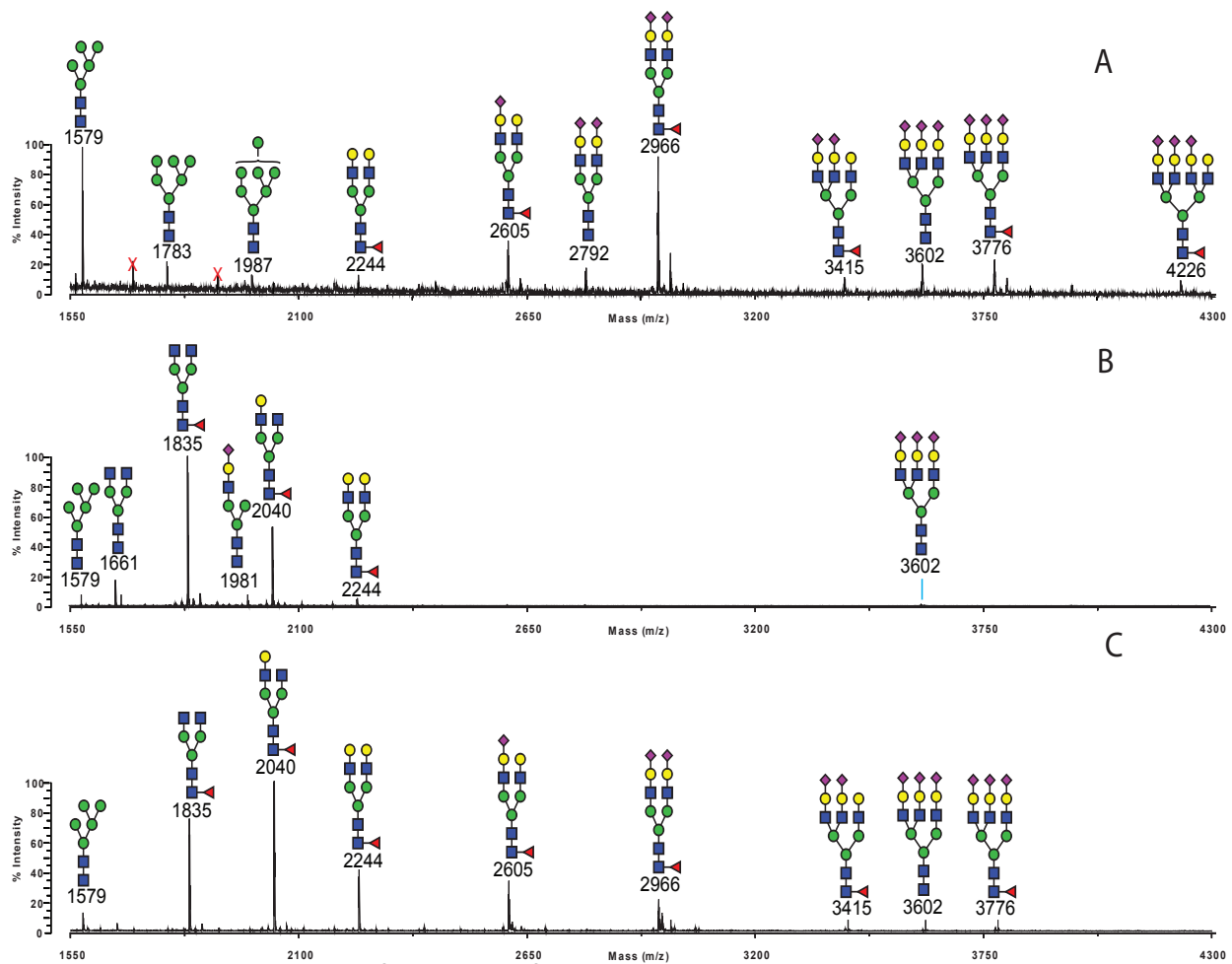
Figure 5.3 MALDI-TOF spectra of N-glycans of IgG1-Fc mutants with monomeric status and single N-glycosylation site on sequence made in CHO cell system.
The recombinant constructs are N297A/C309L/C575A (A), D221N/N297A/C309L/N563A/C575A (B) and C309L/N563A/C575A. See the legend of Figure 5.1 for assignment. Poly-Hexose contamination is highlighted by red crosses in spectra.

C309L/C575A, Figure 5.3A) in the monomeric status still retains a substantial presence of sialylated complex glycans. Bi-, tri- and tetra-antennary N-glycans with varied sialylation levels are the major members of complex class glycans (*e.g. m/z* 2966, 3776 and 4226). Two core-afucosylated structures were also identified as *m/z* 2792 and 3602; and Man5 high mannose structure is also prominent in spectrum indicating a population of relatively less processed N-glycans in this monomeric status. The N-glycome of N221 site on monomeric mutant D221/N297A/C309L/N563A/C575A (Figure 5.3B) is dominated by three core-fucosylated bi-antennary structures as *m/z* 1835, 2040 and 2244 with varied galactosylation levels. Other complex structures are mainly sialylated and core-fucosylated bi- or tri-antennary (*e.g. m/z* 2605, 2966 and 3776), with an exception of *m/z* 3602 which is core-afucosylated. The relative abundance of sialylated structures is generally lower than the predominant three bi-antennary structures. This might indicate a relatively less sialylation and galactosylation in comparison with its multimeric counterparts; Man5 high mannose structure, *m/z* 1579, was also detected with a minor abundance. In addition, the N-glycans observed on N297 (Figure 5.3C) in the monomeric mutant C309L/N563A/C575A are dominant by three core-fucosylated bi-antennary structures as *m/z* 1835, 2040 and 2244. Two sialylated structures are found as mono- (*m/z* 1981) and tri-antennary (*m/z* 3602) with minor amounts. A high mannose structure, *m/z* 1579 as Man5, and core-afucosylated truncated bi-antennary structures are also observable.

The level of their sialylation was also assessed for these three mutants as shown in Figure 5.2. Sialylated N-glycans contribute 52.79% (Figure 5.2D) and 24.25% (Figure 5.2F) of the overall intensity of glycans on mutants D221N/N297A/C309L/C575A and N297A/C309L/C575A, respectively. However, there is only 3.55% of overall intensity on mutant C309L/N563A/C575A contributed from sialylated structures (Figure 5.1.2E). A higher level of sialylations on sites N221 and N563 is observed by comparing to that on N297. In addition, a cross-set comparison of these multi- and monomeric IgG1-Fc mutants shows a significant decrease of sialylation on site N221 (81.82% to 24.25%) and N563 (74.29% to 52.79%) while a slight increase on N297 (1.56% to 3.55%) from multimeric to monomeric.

5.2.2.2 The impact of multiple N-glycosylation sites on IgG1-Fc glycosylation from the CHO cell system

To investigate the impacts of interplay among N-glycosylation sites on the overall N-glycosylation, four mutants were established as C575A, D221N/N563A/C575A, D221N/N297A/C575A, and D221N/C575A in multimeric status. They differ in the number of potential N-glycosylation sites, for example, N297 & N563, N221 & N297, N221 & N563 and N221 & N297 & N563. Their overall glyco-profiles are shown in Figure 5.4. The glycosylation of sites N297 &N563 (Figure 5.4A) is dominated by three core-fucosylated bi-antennary structures with a-, mono- and di-galactosylation (*m/z* 1835, 2040 and 2244). There is only one sialylated structure found as *m/z* 2605, whereas no tri- or tetra-antennary structures were detected; a core-afucosylated bi-antennary structure, *m/z* 1661, with truncated antennae was also detected with trace amounts. On the multimeric mutant D221N/N563A/C575A (Figure
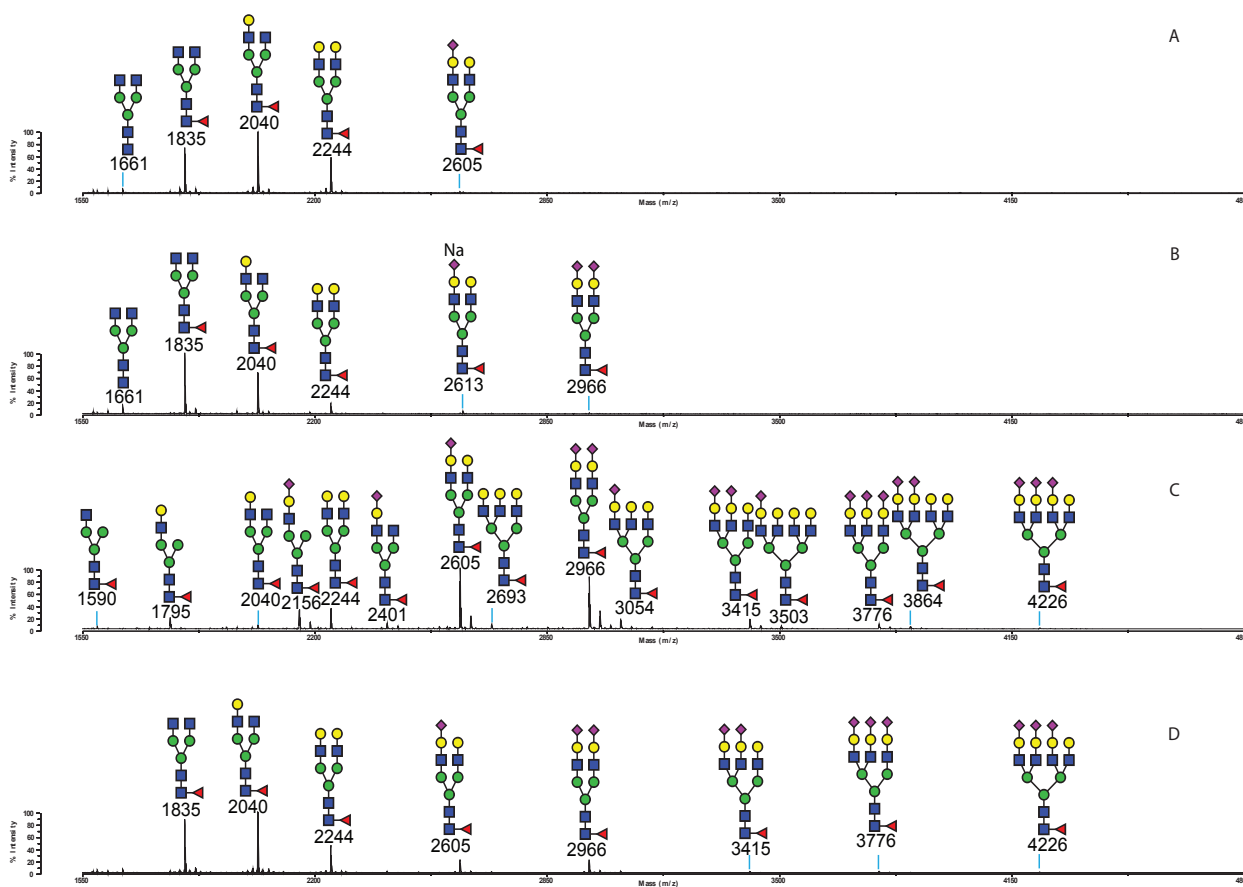
Figure 5.4 MALDI-TOF spectra of N-glycans of IgG1-Fc mutants with multimeric status and two or three available N-glycosylation sites on each sequence.
The recombinant IgG1-Fc constructs are C575A (A), D221N/N563A/C575A (B), D221N/N297A/C575A (C) and D221N/C575A (D). See the legend of Figure 5.1 for assignments. Na labelled above a sialic acid residue indicates the presence of a counter-balance sodium which exists on the detected ions indicating an incomplete permethylation for sialic acid residues.

5.4B), the major glycans are core-fucosylated bi-antennary structures as *m/z* 1835, 2040 and 2244; core-afucosylated bi-antenanry structure with truncated terminal is also observed as *m/z* 1661. Two weak signals were detected corresponding to sialylated bi-antennary structures with core-fucosylation as *m/z* 2613 and 2966.

N-glycosylation of sites N221 &N563 (Figure 5.4C) exhibits a high degree of sialylation as *m/z* 2605 and 2966 are the most predominant pair of structures in relative abundance. Major N-glycan structures ranging from mono- to tetra-antennary are core-fucosylated and sialylated to varied levels (*e.g. m/z* 2156, 2605, 3776 and 4226). Structures without sialylation are also observable as mono- or bi-antennary structures with mono- or di-galactosylation (*m/z* 1590, 1795, 2040 and 2244). In addition, mutant D221N/C575A (Figure 5.4D) exhibits a glycan repertoire which has only complex glycans with predominant abundance of three bi-antennary structures at *m/z* 1835, 2040 and 2244. The other complex glycans are all sialylated bi-, tri- and tetra-antennary.

Multimeric

C575A

A

■ 1.84% Sialylated
■ 98.16% Netural

D221N/N563A/C575A

B

■ 4.54% Sialylated
■ 95.46% Netural

D221N/N297A/C575A

C

■ 78.65% Sialylated
■ 21.35% Netural

D221N/C575A

D

■ 15.66% Sialylated
■ 84.34% Netural

Monomeric

C309L/C575A

E

■ 3.21% Sialylated
■ 96.79% Netural

D221N/C309L/N563A/C575A

F

■ 5.61% Sialylated
■ 94.39% Netural

D221N/N297A/C309L/C575A

G

■ 78.26% Sialylated
■ 21.74% Netural

D221N/C309L/C575A

H

■ 6.96% Sialylated
■ 93.04% Netural

Figure 5.5 Percentage of sialylation of N-glycans on IgG1-Fc mutants with multiple N-glycosylation sites. The recombinant constructs are indicated above these pie chart.
See the legend of Figure 5.2 for detailed interpretation.

The levels of sialylation on these multimeric mutants were subsequently evaluated as described above (**Section 5.2.2.1**) and their pie charts are shown in Figure 5.5. N-glycosylation of N297 &N563 (Figure 5.5A) and N221 &N297 (Figure 5.5B) was largely restricted to neutral glycans with a limited contribution to the overall intensity from sialylated glycans at 1.84% and 4.54%, respectively. Mutant D221N/C575A (Figure 5.5D) with all three sites available for N-glycosylation has 15.66% of intensity contribution from sialylated glycans. Nevertheless, sialylation is predominant on the mutant with N221& N563 (Figure 5.5C) on sequence as its contribution reaches 78.65%. Therefore, the presence of N-glycosylation site N297 might negatively correlate with the overall sialylation on available N-glycosylation sites in these constructed Fc fusion proteins.

The monomeric mutants for these N-glycosylation site combinations are also analysed as shown in Figure 5.6. N-glycosylation of monomeric mutant C309L/C575A (Figure 5.6A) and D221N/C309L/ N563A/C575A (Figure 5.6B) shows an identical pattern to their multimeric counterparts (Figure 5.4A&B). Mutant D221N/N297A/C309L/C575A(Fgiure 5.6C) exhibits a similar glycosylation pattern to its multimeric counterpart (Figure 5.4C) but with loss of signals for tetra-antennary structures as well as truncated mono- or bi-antennary structures. On mutant D221N/C309L/C575A, signals for tri- and tetra-antennary structures are not visible, and three core-fucosylated bi-antennary structures are still predominant in the spectrum at *m/z* 1835, 2040 and 2244 (Figure 5.6D).
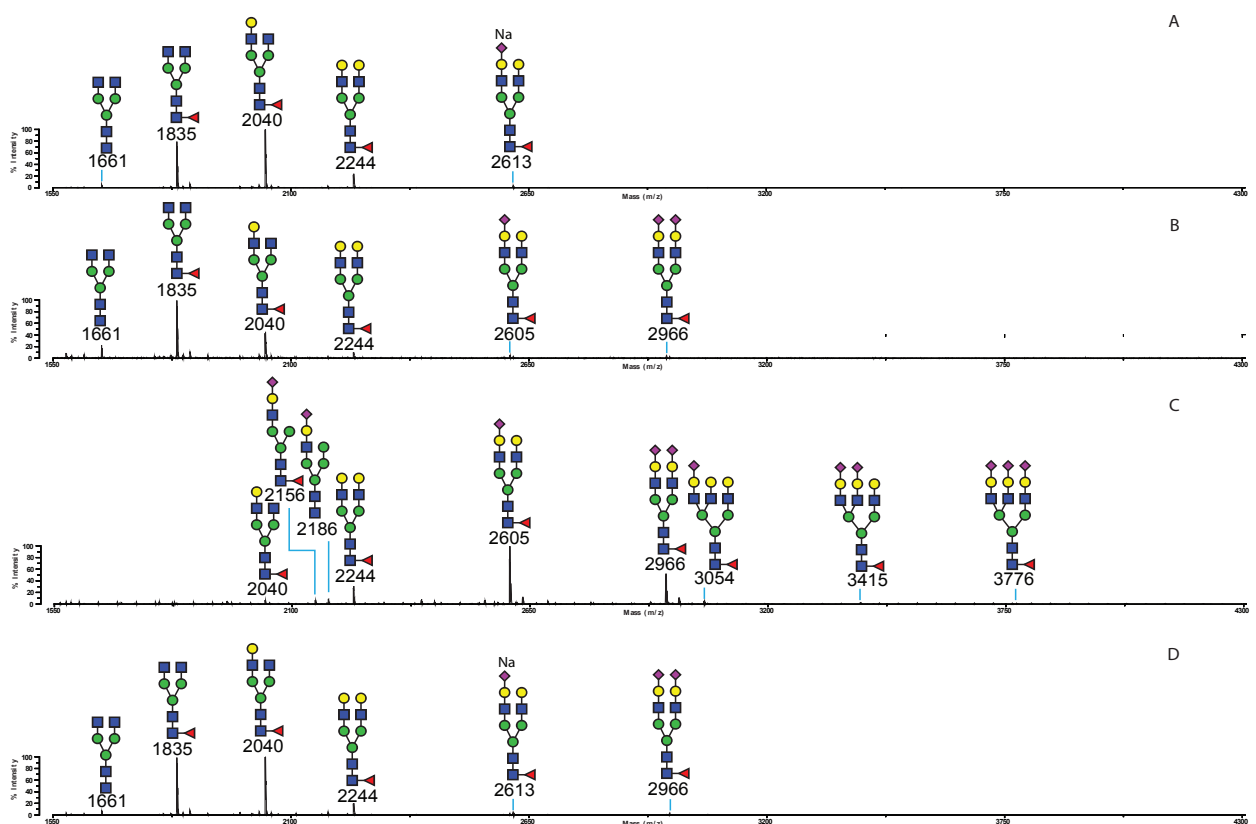
Figure 5.6 MALDI-TOF spectra of N-glycans of recombinant IgG1-Fc mutants with monomeric status and multiple available N-glycosylation sites.
The recombinant constructs are C309L/C575A (A), D221N/C309L/N563A/C575A (B), D221N/N297A/C309L/C575A (C) and D221N/C309L/C575A (D). See the legend of Figure 5.1 for assignments.

The sialylation of these monomeric mutants was also examined via the semi-quantitative apporach (Figure 5.5E, F, G and H), and consistent patterns of sialylation are also found on their matching monomeric mutants. There is a nearly equivalent level of sialylation on mutants D221N/N297A/C575A (Figure 5.5C) and D221N/N297A/C309L/C575A (Figure 5.5G). Besides, a slight increase of sialylation is observed on mutants C309L/C575A (Figure 5.5E) and D221N/C309L/N563A/C575A (Figure 5.5F) in comparison with their multimeric counterparts (Figure 5.5A and 5.5B). However, a decrease of sialylation presents on mutant D221N/C309L/C575A (Figure 5.5H) from its counterpart D221N/C575A (Figure 5.5D).

5.2.2.3 Impacts of cysteine residues on the N-glycosylation of IgG-Fc from CHO cell system

The impacts of two cysteine residues, C309 and C575, on the Fc N-glycosylation was evaluated on four selected mutants with N-glycosylation site N297 &N563 available on sequence as Hexa-Fc, C309L, C575A and C309L/C575A. The N-glycomes of these mutants are shown in Figure 5.7. Hexa-Fc (Figure 5.7A) has a glycome mainly consisted of core fucosylated bi-antennary complex glycans. Tri- and tetra-antennary structures as well as high mannose structures have a minor abundance. Sialylation is observed on three complex structures at *m/z* 2605, 3054 and 3503 ranging from bi-antennary to tetra-antennary. On the mutant C575A, three core-fucosylated bi-antennary structures are still dominant (Figure 5.7B), only one sialylated (*m/z* 2605) and one tri-antennary (*m/z* 2693) structure were detected as well as a core-afucosylated bi-antennary structure with terminal truncation (*m/z* 1661). In mutant C309L (Figure 5.7C), bi-, tri- and tetra-antennary N-glycan structures are found having varied sialylation levels with limited abundance; core-fucosylated bi-antennary structures *m/z* 1835 and 2040 are still dominant in the spectrum. In addition, mutant C309L/C575A (Figure 5.7D) exhibits a glycome with six structures, three of them are the core-fucosylated bi-antennary at *m/z* 1835, 2040 and 2244. Two sialylated structures are identified as mono-sialylated bi- or tri-antenanry structures at *m/z* 2613 (2605) and 3062 (3054).

The levels of sialylation were compared between these mutants as shown in Figure 5.8. Sialylation of Hexa-Fc (11.48%) and C309L (11.23%) are generally equivalent, whereas the sialylation for C575A (1.84%) is nearly one tenth of that for Hexa-Fc or C309L. Mutant C309L/575A with knockouts of both
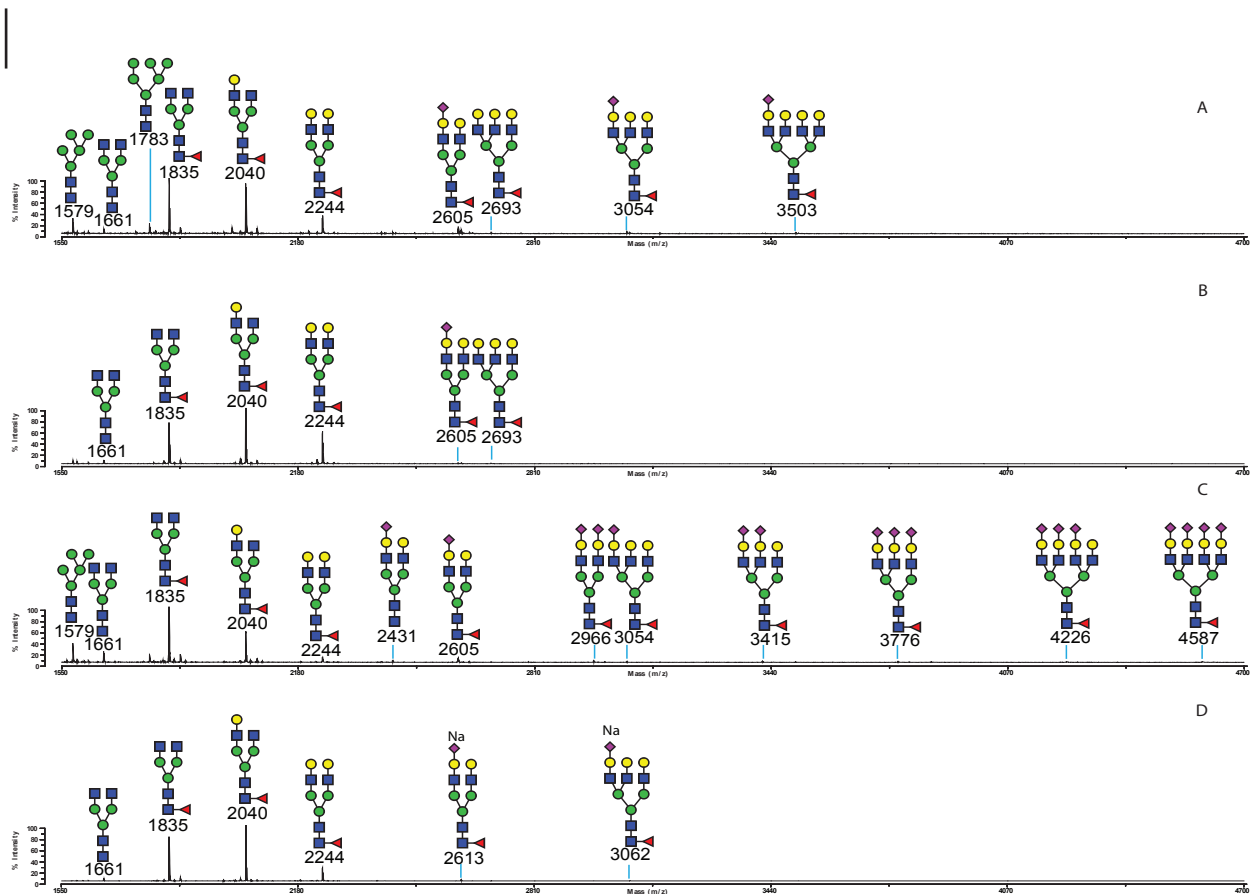
Figure 5.7 MALDI-TOF spectra of N-glycans of IgG1-Fc mutants with available N-glycosylation site N297 &N563.
The recombinant constructs are unmodified Hexa-Fc (A), C575A (B), C309L (C) and C575A/C309L. See the legend of Figure 3.1 for assignments.
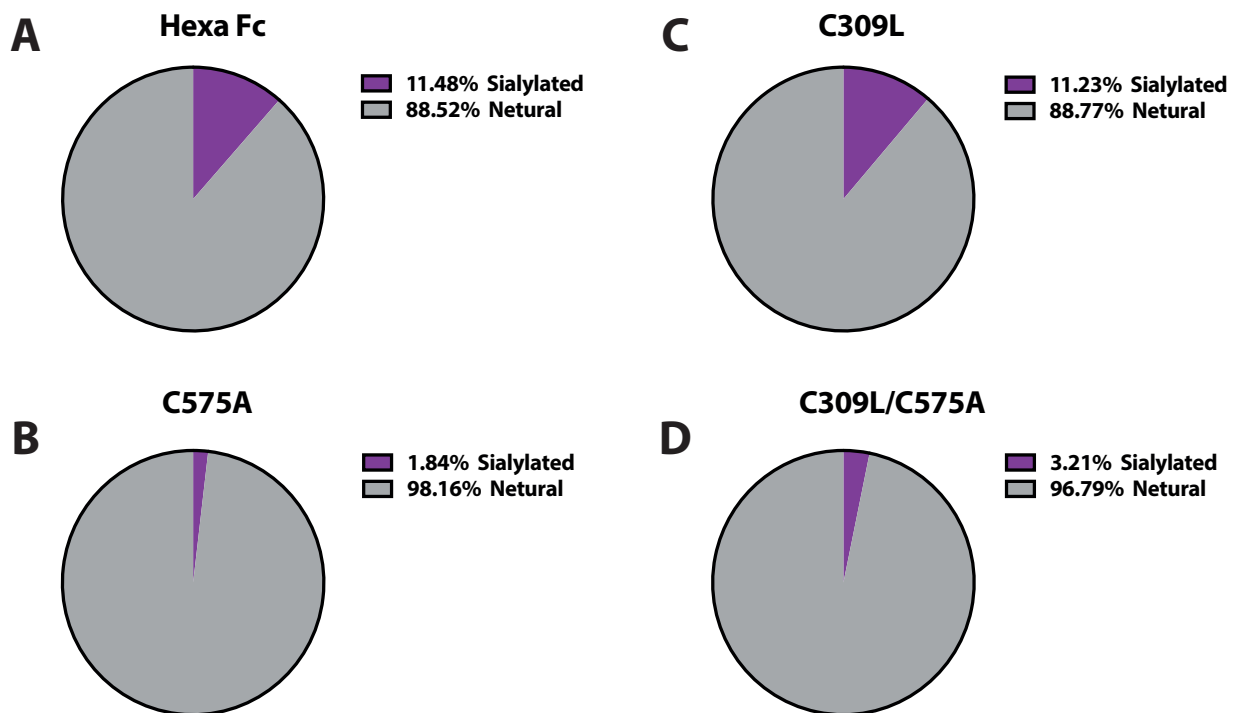
Figure 5.8 Semi-quantitative estimation of sialylation level of IgG1-Fc with variable availabilities of cysteine residues.
Multimeric Hexa-Fc with two cysteines (A); multimeric IgG1-Fc with one cysteine at C309 (B); multimeric IgG1-Fc with one cysteine at C575 (C); monomeric IgG1-Fc without C309 and C575 cysteine residues (D)

cysteine residues has a slightly higher sialylation than mutant C575A as 3.21%. Taken together, these results might suggest that the presence of disulfate bonding could have impacts on protein glycosylation.

5.2.3 N-glycosylation of HEK293 cells

Unlike the relatively simple glycosylation of the CHO cell systems, HEK cells are capable of producing more complex N-glycan structures on their glycoproteins. The N-glycosylation of the secretory products from HEK 293 cells have been previously investigated in our laboratory (Lo et al., 2013). Their N-glycans are featured with core-fucosylation as well as diverse antennary arrangements and modifications such as LacdiNAc (GalNAcβ1-4GlcNAc), fucosylation and sialylation. Moreover, as a human cell line, HEK cells can express α2,6 linked sialic acids on glycans. This linkage of sialic acid has been demonstrated to contribute to the anti-inflammatory effect of IgGs in vivo (Anthony et al., 2011).

5.2.3.1 N-glycosylation at N221, N297 and N563 on IgG1-Fc from HEK293 cells

To validate the N-glycosylaiton of each individual site in the HEK cell system with available C309, mutants N297A/C575A (N563), D221N/N297A/N563A/C575A (N221) and N563A/C575A (N297) were analysed as shown in Figure 5.9. A core-fucosylated bi-atnneary structure without antennary galactosylation, *m/z* 1835, is the base peak of spectra from all IgG1-Fc mutants produced in this system. N-glycosylation of N297 (Figure 5.9B) is dominated by three core-fucosylated bi-antennary glycans (*m/z* 1835, 2040 and 2244) with varied galactosylation levels, and a Man5, *m/z* 1579, high mannose structure is also observed in this spectrum. Besides, glycosylation on N-glycosylation site N563 (Figure5.9A) is dominated by putative truncated structures at *m/z* 2081 and 2285 which have potential antennary GlcNAc or terminal GalNAc. Branch fucosylation is also identified on structures as *m/z* 2820 which could assemble a sialylated LacNAc, Sialyl-Lewis x/a antigen or a sialylated LacdiNAc. The presence of *m/z* 2674 confirmed the existence of fucosylated LacdiNAc moiety on glycans assembled onto this site. The glycosylation in this system for N563 is different to that in the CHO cell system as CHO cells mainly assemble more sialyl-LacNAc as the terminal moiety, whereas this system assembles more diverse structures and is relatively less efficient for sialylation. In addition, the N-glycosylation site, N221, as mutant D221N/N297A/N563A/C575A (Figure 5.9C) exhibits a N-glycome with mainly bi-antennary complex structures. Excluding the base peak *m/z* 1835, four structures are identified as potentially truncated complex structures as *m/z* 2081, 2285, 2459 and 2646; and they could all possibly form LacdiNAc based moiety due to the presence of uncapped GlcNAc and sialic acid and/or fucose residues.

The levels of their sialylation were also semi-quantitatively assessed as shown in Figure 5.10. There is only 0.8% and 2.88% of overall intensity contributed from sialylated glycans for N297 (Figure 5.10B) and N221(Figure 5.10C), respectively. N563 (Figure 5.10A) has the highest content of sialic acid among these three mutants at 46.74%.
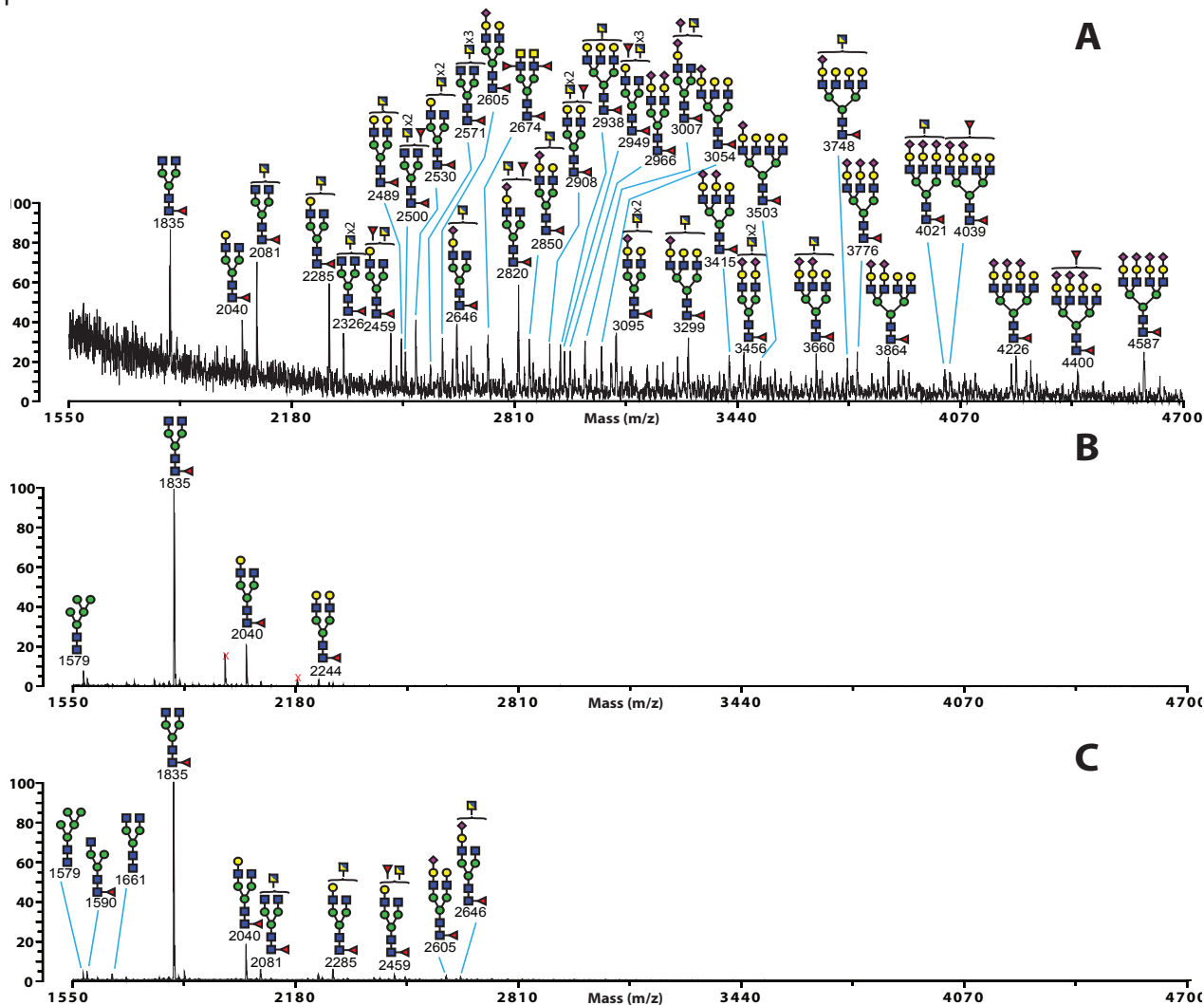
Figure 5.9 MALDI-TOF spectra of N-glycan repertoires of IgG1-Fc mutants with diverse single N-glycosylation site and C309 residue.

The recombinant constructs are N297A/C575A (A), N563A/C575A (B) and D221N/N297A/N563A/C575A (C). Linkage undetermined monosaccharides are positioned above the bracket on a structure. Poly-hexoses contaminants are highlighted with red crosses. See the legend of Figure 3.1 for assignments.

Information about the polymeric status of these Fc proteins has not been determined. For the set of C309 absent mutants, N563 (C309L/N297A/C575A, Figure 5.11A) exhibits a similar N-glycome to that of its counterpart with C309 available. The glycome is dominated by the abundance of potential LacdiNAc or truncated structures and with structural diversity for the antennary modifications. N297 shows a restricted glycan repertoire (mutant C309L/N563A/C575A, Figure 5.11B) the most abundant glycan is a bi-antennary structure with mono-galactosylation and core-fucosylation (*m/z* 2040), and there is no detection of sialylated structures. In addition, the N-glycosylation of N221 (D221N/N297A/N563A/C575A; Figure 5.11C) shows that a dominance of putative truncated or LacdiNAc forming structures (*e.g. m/z* 2081, 2285 and 2459). Minor sialylation could be found on peaks close to the

Cys309

C309L

A N297A/C575A

46.74% Sialylated
53.26% Neutral

D N297A/C309L/C575A

23.50% Sialylated
76.50% Neutral

B N563A/C575A

0.80% Sialylated
99.20% Neutral

E C309L/N563A/C575A

100.00% Neutral

C D221N/N297A/N563A/C575A

2.88% Sialylated
97.12% Neutral

F D221N/N297A/C309L/N563A/C575A

1.13% Sialylated
98.87% Neutral
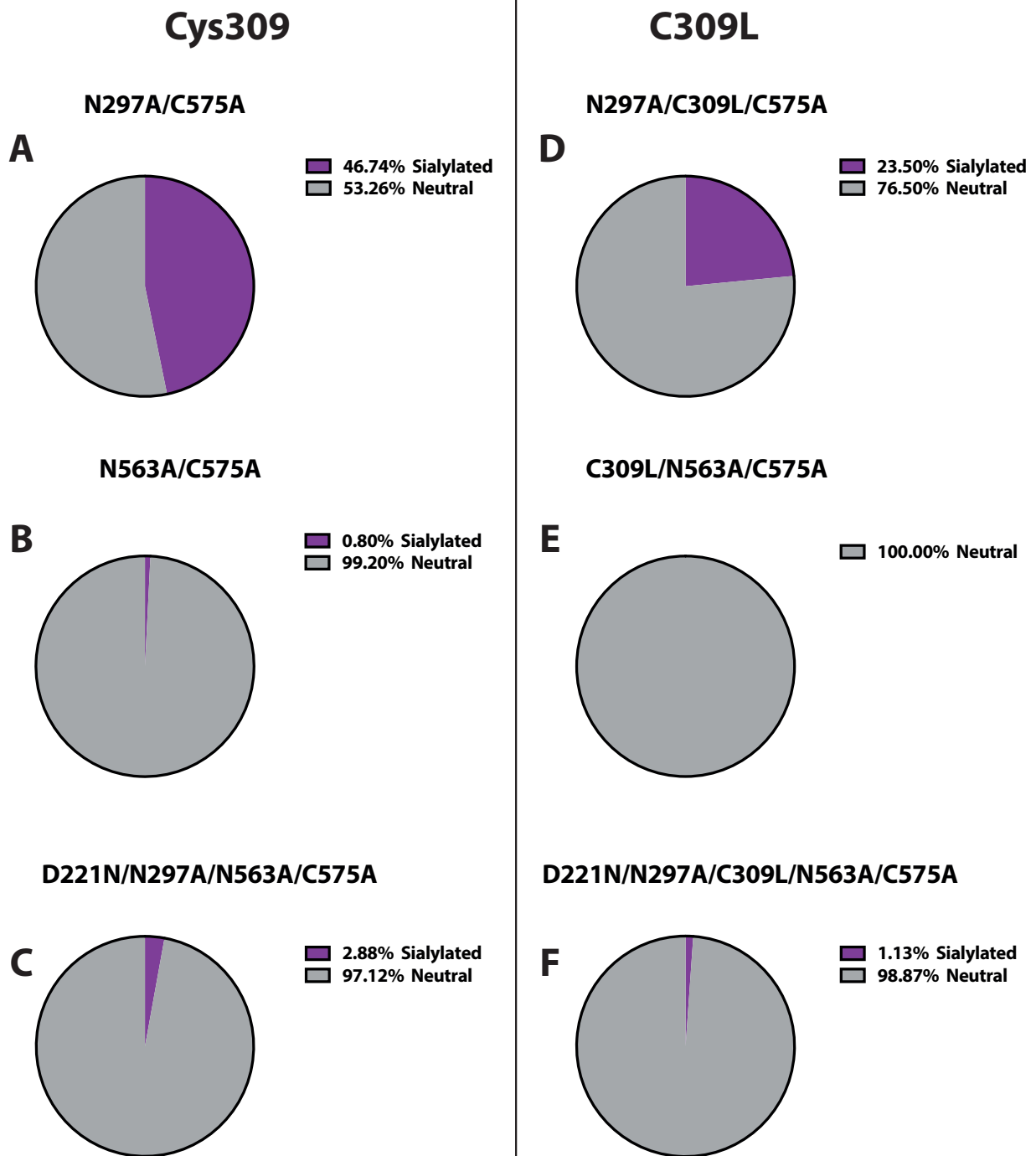
Figure 5.10 Estimation of sialylation level of single N-glycosylation site IgG1-Fc mutants from HEK293 cell system semi-quantitatively.
The recombinant constructs are indicated above these pie chart. See the legend of Figure 5.2 for detailed interpretation.
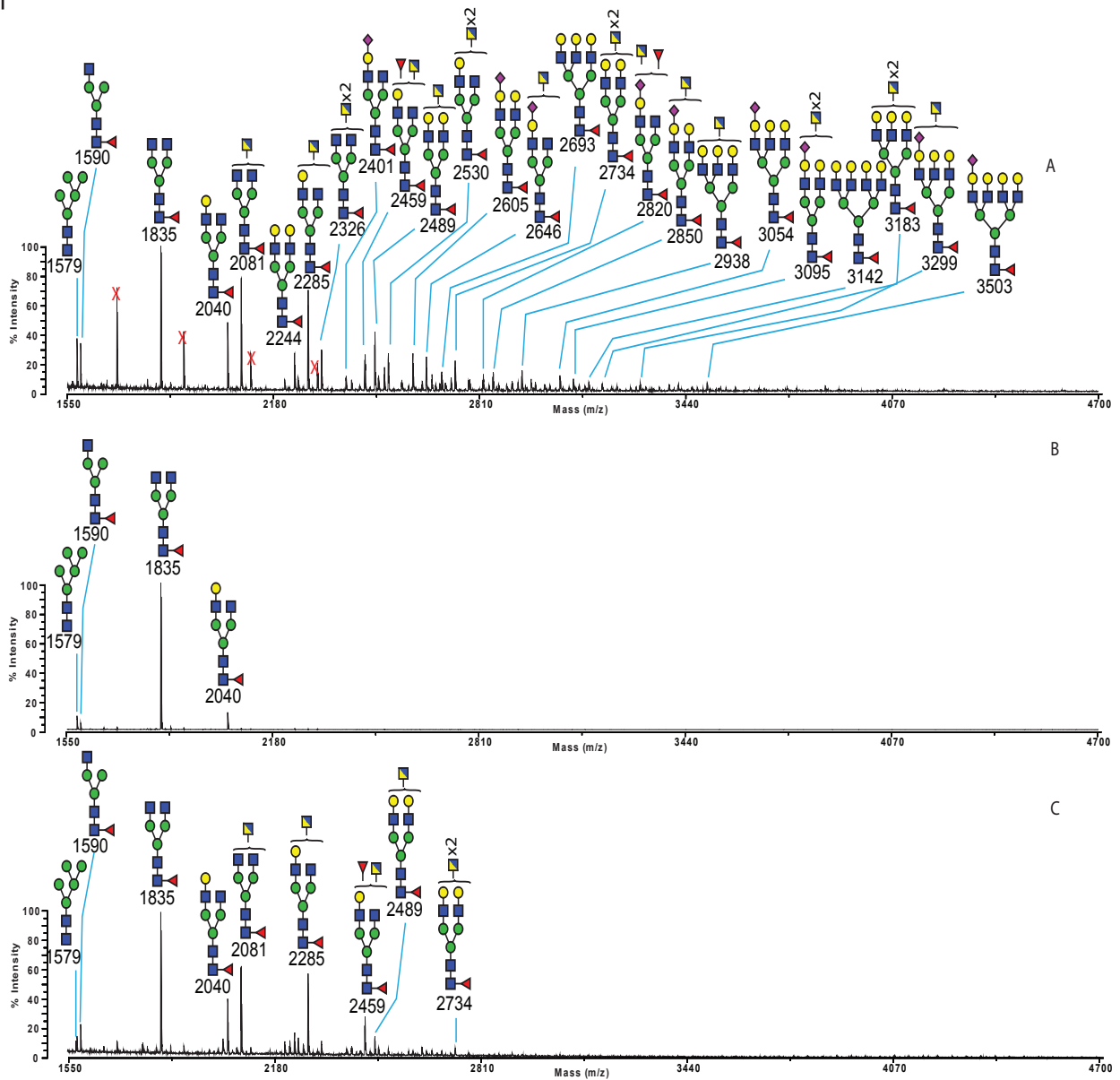
Figure 5.11 MALDI-TOF spectra of N-glycans of IgG1-Fc mutants with single N-glycosylation site but no C309 residue.

The recombinant constructs are C309L/N297A/C575A (A), C309L/N563A/C575A(B) and D221N/N297A/C309L/N563A/C575A (C). See the legend of Figure 5.1 for assignments.

baseline. Hence, the sialylation of these C309 absent Fc mutants was also evaluated (Figure 5.10), and sialylation of these mutants is 1.13% for N221 (Figure 5.10F), 0% for N297 (Figure 5.10E) and 23.50% for N563 (Figure 5.10D). A decline of sialylation is observed between the pair of mutant N297A/C575A (Figure 5.10A) and N297/C309L/C575A (Figure 5.10D) as the percentage of sialylation drops from 46.74% to 23.50%. The other two mutants also have a minor decline of sialylation percentage from their C309 available counterparts.

5.2.3.2 The impact of multiple N-glycosylation sites on IgG1-Fc glycosylation from the HEK293 cell system

There are abundance shifts of sialylation on glycans from mutants with multiple N-glycosylation sites expressed in the CHO cell system. Comparably, the levels of sialylation on mutants with multiple N-glycosylation sites of constructs from the HEK cell system are shown in Figure 5.12.

Consistent with the CHO cell system, the presence of N221 on sequences enhanced the expression of sialic acids on mutants D221N/N563A/C575A and D221N/N297A/C575A (Figure 5.12B and C). However, the synergistic effect of N221 and N563 for sialylation observed in the CHO cell system is not present on mutant D221N/C575A in this system (Figure 5.12D). The levels of sialylation is also affected once the mutation for C309L is included (Figure 5.12E, F, G and H). Sialylation is enhanced on mutants C309L/C575A than C575A (Figure 5.12 A and E; Figure 5.16 B and D), and this generally agrees with the observation in the CHO cell system for the same mutants. A relatively equivalent level of sialylation on D221N/N297A/C575A and D221N/N297A/C309L/C575A in this system is also consistent with the observation of these two mutants in the CHO cell system. However, the changes of sialylation on mutants from D221N/C575A to D221N/C309L/C575A and D221N/N563A/C575A to D221N/C309L/N563A/C575A are reversed in this system from that observed in the CHO cell system.

To understand the absence of synergistic effect for sialylation on mutant D221N/C575A, a comparison of MALDI-TOF spectra of this mutant to D221N/N563A/C575A is present in Figure 5.13. A similar glycan distribution in the low mass range of these two mutants is observed. At higher mass D221N/C575A exhibits a higher content of asialylated multi-antennary structures with potential LacdiNAc antennae, truncated GlcNAc and fucosylation (*e.g. m/z* 2867, 2908 and 2938; Figure 5.13A). D221N/N563A/C575A presents mainly sialylated bi- and tri-antenanry structures with relatively lower level of fucosylation (*e.g. m/z* 2966 and 3054; Figure 5.13B). Furthermore, D221N/N297A/C575A is included in comparison as the mutant without the proposed suppressive effect of N297 for sialylation (Figure 5.13C). There is an increase of the abundance of structures with potential LacdiNAc epitopes, truncation and branch fucosylation (*e.g.* 2081, 2285 and 2459) in this mutant, which are more abundant than sialylated structures. This may imply a competing advantage of these antennary modifications to sialylation in this system.
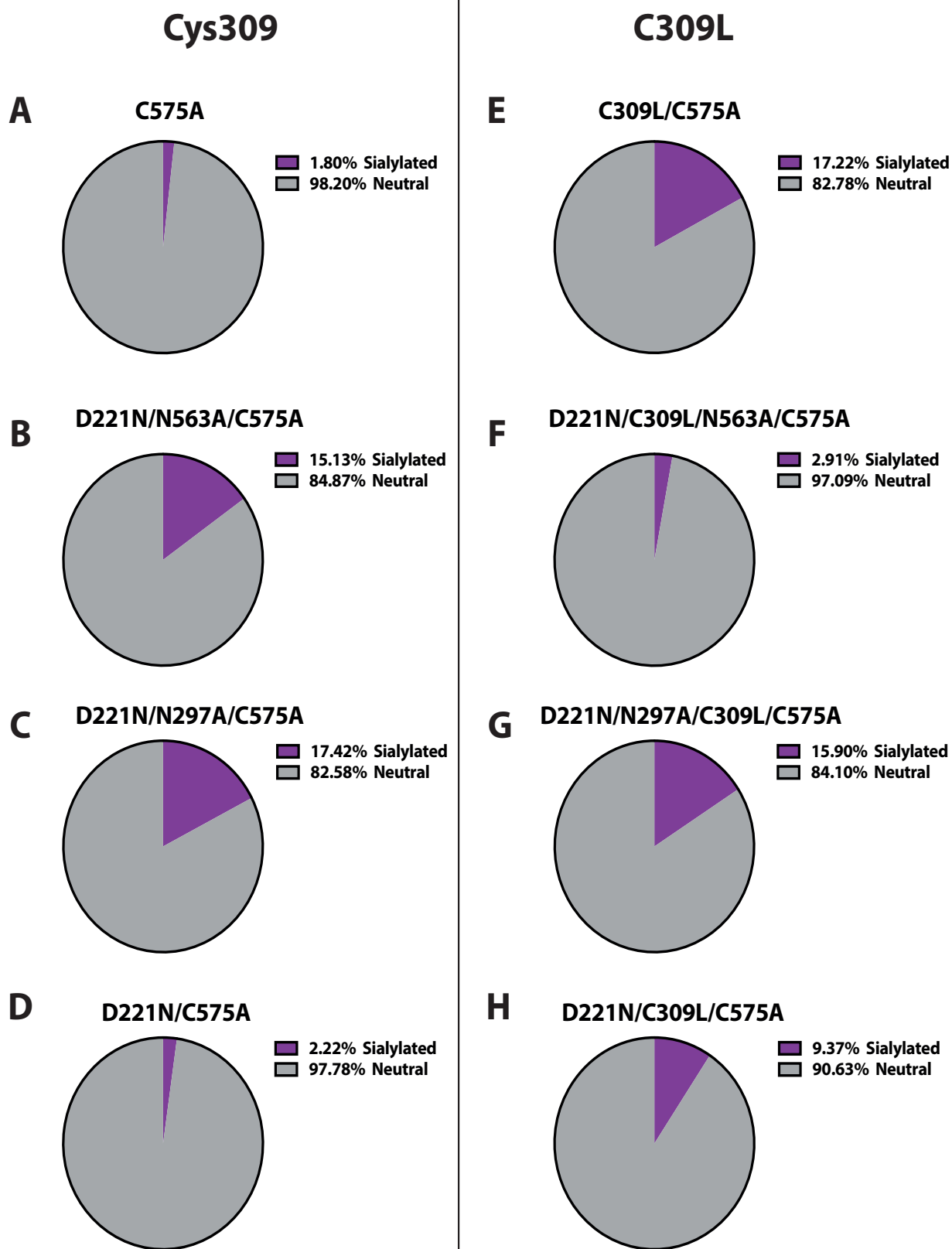
Figure 5.12 Sialylation level of multiple N-glycosylation site IgG1-Fc mutants from HEK293 cell system.
The recombinant constructs are indicated above these pie chart. See the legend of Figure 5.2 for detailed interpretation.
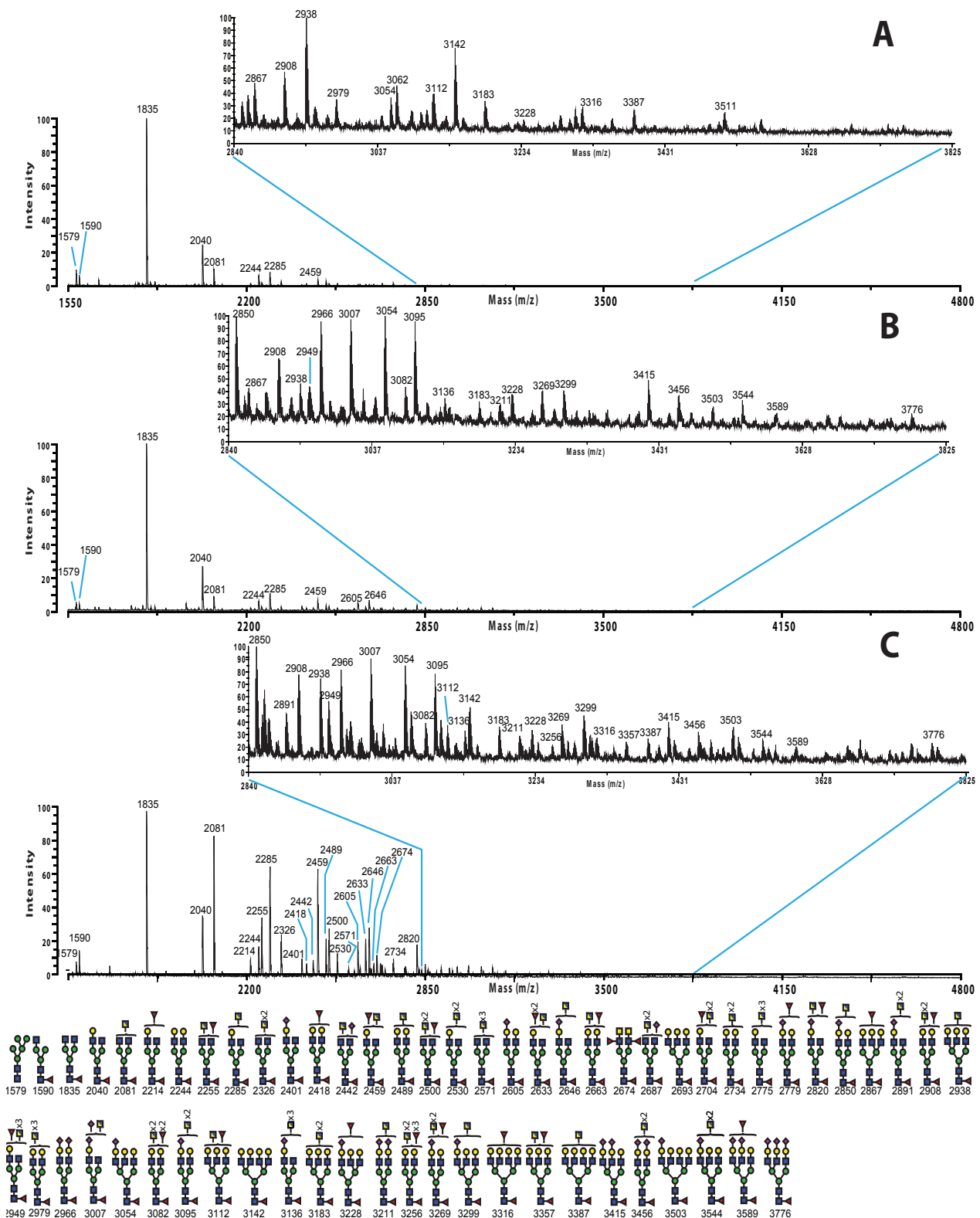
Figure 5.13 MALDI-TOF Spectra comparison of IgG1-Fc mutants with multiple N-glycosylation site with C309.

D221N/C575A (A), D221N/N563A/C575A (B) and D221N/N297A/C575A (C). *m/z* 2840 to 3825 was zoomed in to visualise. *m/z* is assigned on spectra and corresponding cartoons are summarised in the bottom. See the legend of Figure 5.1 for assignments.

This suggested competition was also assessed in the set of C309 absent mutants as shown in Figure 5.15. In mutants D221N/C309L/C575A (Figure 5.14A) and D221N/N297A/C309L/C575A (Figure 5.14C), sialylation is low in abundance as a dominant appearance of various structures carrying potential LacdiNAc epitopes or truncated antennae. Mutant D221N/C309L/N563A/C575A (Figure 5.14B) shows a limited glycan repertoire with mainly truncated bi-antennary complex structures with high mannose and mono-antennary structures indicating even less sialylation than its counterpart with C309 available.
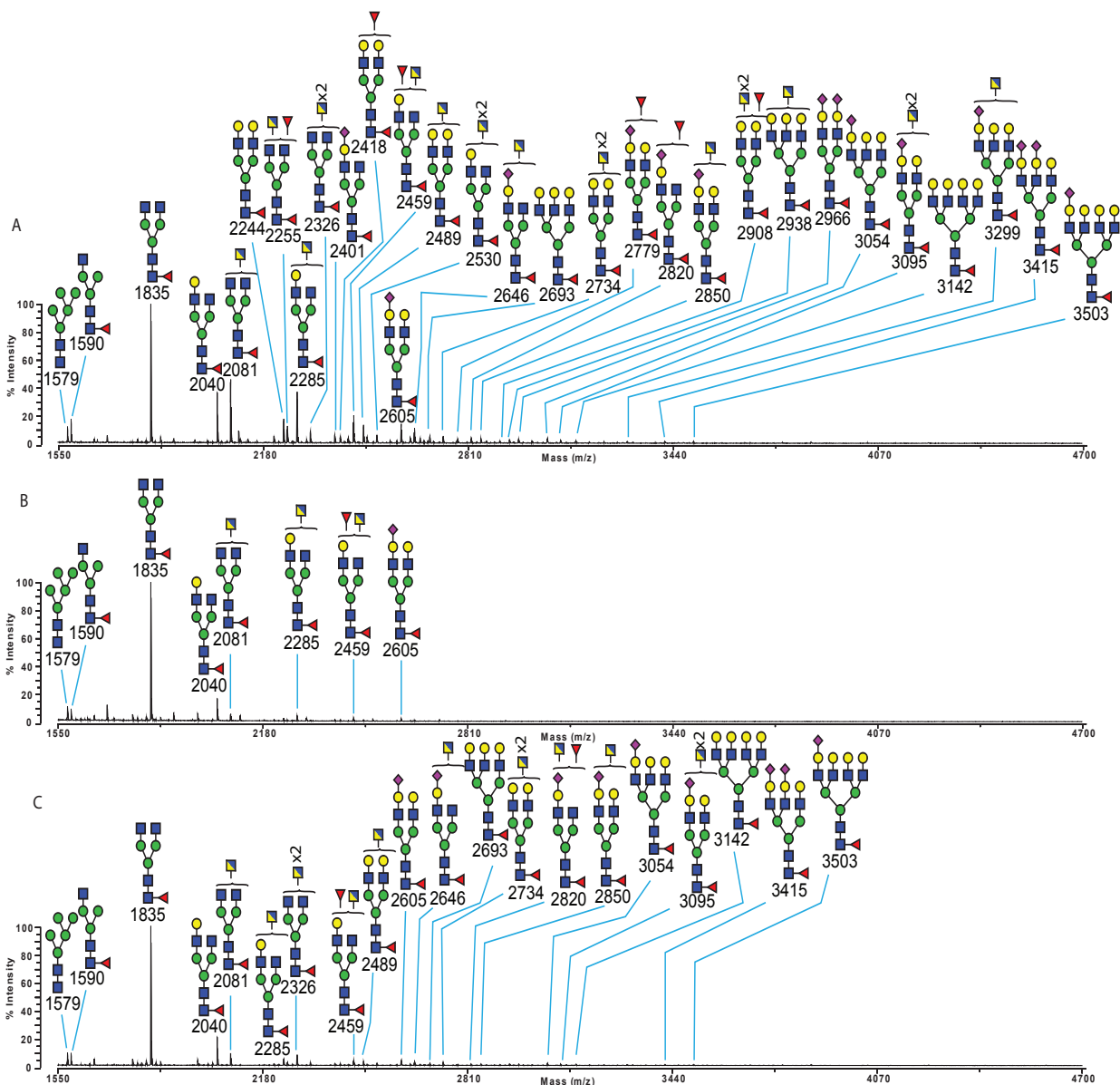

Figure 5.14 MALDI-TOF spectra of N-glycans of IgG1-Fc mutants with multiple N-glycosylation sites and C309 absence.
The recombinant constructs are D221N/C309L/C575A (A), D221N/C309L/N563A/C575A and D221N/N297A/C309L/C575A. See the legend of Figure 3.1 for assignments.

5.2.3.3 Impacts of cysteines interplay on N-glycosylation of IgG1-Fc mutants from HEK293 cell system

The impacts of the presence of cysteine residues on the mutants (Figure 5.15 and 5.16) was assessed as that on the mutants from the CHO cell system (**Section 5.2.2.3**). The N-glycan profiles of Hexa-Fc, mutant C575A and C309L are highly similar in terms of structures and relative abundance. In these three materials, the pool of complex glycans are mainly constituted of bi- and tri-antennary structures with trace amounts of sialylation that is detectable, such as *m/z* 2605. Truncated structures, *m/z* 1835
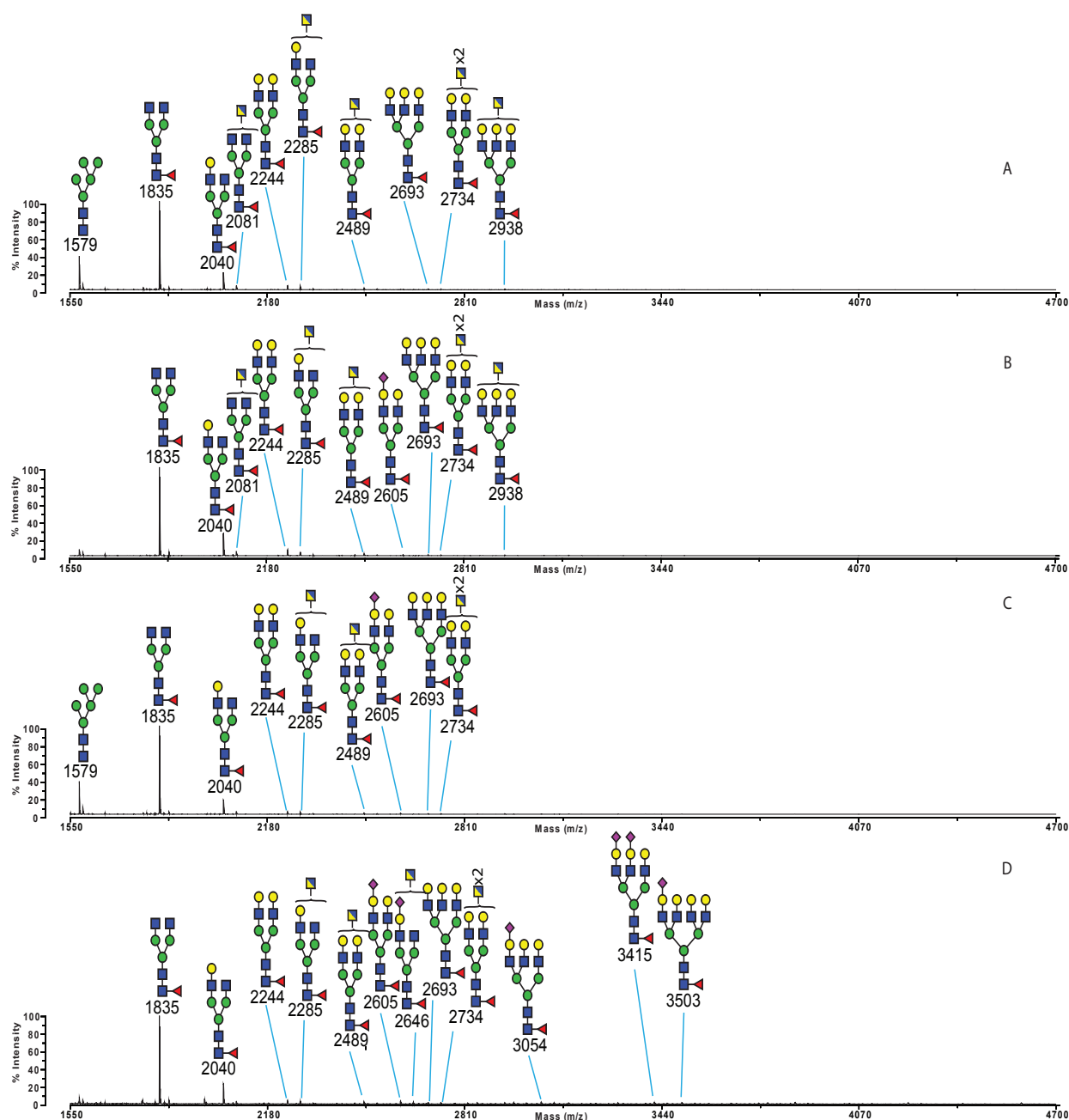
Figure 5.15 MALDI-TOF spectra of N-glycans IgG-Fc mutants.
The recombinant constructs are Hexa-Fc (A), C575A (B), C309L (C) and C309L/C575A (D). See the legend of Figure 3.1 for assignments.
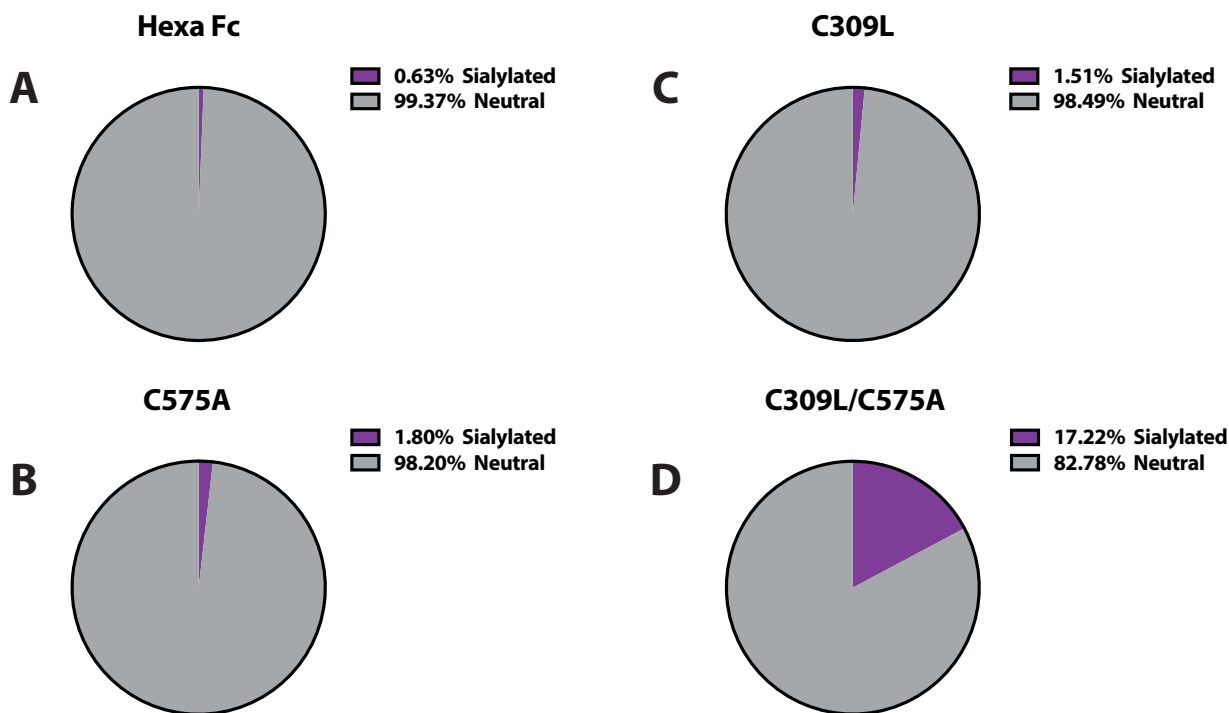
**Figure 5.16 Estimation of sialylation level of IgG1-Fc mutants with varied disulfide bonding status.**
The recombinant constructs are indicated above these pie chart. See the legends of Figure 5.2 &5.8 for detailed interpretation.

and 2040, are dominant in spectra; the content of high mannose structures varied in between these three materials to reach up to intermediate level of relative abundance. N-glycosylation of the fourth mutant, C309L/C575A, is still dominated by m/z 1835 and 2040, whereas signals for tri- and tetra-antennary structure carrying sialic acid residues are detected in this mutant to show its distinction (Figure 5.15D).

Unlike the same mutants expressed in CHO cell system, there is no precise difference between mutants C575A and C309L (Figure 5.16B and C) for their sialylation, but they have a slightly higher ratio of sialylated structures than that of Hexa-Fc (0.63%, Figure 5.16A) as 1.8% and 1.51%, respectively. Even higher ratio (17.22%) of sialylation is observed on mutant C309L/C575A without both cysteine residues (Figure 5.16D).

5.3 Discussion

In this study, we characterised the N-glycosylation of recombinant Fc fragments with varied numbers of N-glycosylation sites and cysteines in two popular cell line systems, CHO and HEK293, structurally and semi-quantitatively. N-glycosylation on these Fc fusion mutants generally retain the bi-antennary complex structures as the major species and acquired cell-specific structural features from these two recruited cell lines (North et al., 2010, Lo et al., 2013). N-glycosylation of each mutant is mainly

different in the ratios of branching and antennary modifications. In particular, the levels of sialylation on N-glycans varied with the changes of N-glycosylation sites and cysteine residues between mutants.

In the CHO cell system, the previously described patterns of sialylation on Hexa-Fc mutants (Blundell et al., 2017) are generally consistent in this study on matching mutants with knockouts of one or both cysteine residues forming relatively reduced multimers or monomers. N-glycosylation sites from insertion (N221) or tailpiece addition (N563) can be highly glycosylated. Their co-presence on a Fc sequence can contribute to a maintenance of the level of sialylation regardless of the status of multimerisation via disulfide bonding (**Section 5.2.2.2**). However, once the N297 site is presented on the sequence, a relatively lower ratio of sialylation is detected, and this might indicate that the presence of N297 could have a negative correlation with the overall sialylation. Glycoproteomic experiments should be conducted to assess the exact location of each glycans on these glycosites to clarify the glycoform on each site in future with more materials. This would be very helpful to determine the actual contribution of the profiled structural features such as sialylation or branching or antennary modifications from each N-glycosylation site. This information would also help to understand the mechanism of this observed enhancement of certain structural features on a site-specific basis. In addition, the influence of cysteine residues on the levels of sialyation was compared. An interesting phenomenon is observed between the groups with/without the presence of N297. Mutants without N297 have higher sialic acid contents in their multimeric forms (C575A) than the monomeric ones with an exception of a pair of mutants D221N/N297A/C575A and D221N/N297A/C309L/C575A which have nearly equivalent levels of sialylation. Nonetheless, the mutants with N297 have slightly higher contents of sialic acid in their monomeric forms (C309L/C575A) than the multimeric ones with the exception for one pair of mutants D221N/C575A and D221N/C309L/575A. This observation might indicate that multimerization of Fc mutants might form preferred quaternary structures for the sialylation of N221 and N563. Instead, the increased sialylation on mutants with N297 might correlate to an improved exposure of their N-glycans for relevant components in the monomeric forms.

The HEK cell system has been employed in the production of IgGs to overcome the limitation of the CHO cell system which cannot assemble α2,6 linked sialic acids onto glycans. Although the HEK cell system is more capable of producing human glycan epitopes such as α2,6 linked sialic acid, our results agree with previous studies to suggest that the total level of sialylation on Fc mutants produced in the HEK cell system is dramatically lower than that in the CHO cell system (Gugliotta et al., 2017, Croset et al., 2012). The lowered sialylation may be mainly caused by the poor galactosylation of glycans on mutants expressed in the HEK cell system because terminal sialylation is dependent on the availability of Gal, but the truncated bi-antennary structure, *m/z* 1835, is always the base peak in all examined mutants. This downregulation of sialylation may also correlate to the capability of HEK cells to express more diverse terminal moieties. For instance, mutant D221N/N297A/C575A with mutation of N297A does not lead to a dramatic increase in the abundance of sialylated structures in comparison

with D221N/C575A. Instead, it leads to a significant increase of potentially truncated, branch fuco-sylated, or LacdiNAc forming structures (Figure 5.12C). This would imply a competition of these various modifications on N-glycans antennae with sialylation in the HEK cell system. The effect of N221 and N563 to enhance total sialylation is limited to a few mutants in this system, and the suggested influence of cysteine residues are also not generally consistent with that observed in the CHO cell lines. Interestingly, regardless of the cell system employed or the multimerization status, there is a pair of mutant D221N/N297A/C575A and D221N/N297A/C309L/C575A that always maintain a relatively high level of sialylation. This observation would indicate that there might be preferences for sites and/or protein conformations to accommodate terminal modifications such as sialylation.

Multiple studies have compared the glycosylation of CHO and HEK cell systems for secretory protein production (Croset et al., 2012, Heffner et al., 2018, Bohm et al., 2015). The CHO cell system is more capable of producing well branched, highly sialylated and even LacNAc elongated N-glycans. The HEK cell system can assemble more neutral, less branched and sialylated structures with various human glycan moieties. The impacts of these featured glycans on glycan-lectin interactions was investigated by our collaborator who examined all the Fc mutants expressed in CHO cell system against a range of lectin receptors (Table 5.3). The results indicate that the sole presence of N563 on sequence such as N297A/C575A or N297A/C309L/C575A would confer a consistent moderate binding affinity of these mutants to Siglec-1. This indicates the accessibility of the enhanced sialylation on this N-glycosylation site to Siglec-1 lectin in either polymeric or monomeric status. Besides, the sole presence of N221 would confer a broad range of lectin reactivities on these two mutants as D221N/N297A/N563A/C575A and D221N/N297A/C309L/N563A/C575A to Siglec 1-4 and C-type lectins (*e.g.* CD23 and Dectin-2). Expanded binding reactivities are observed on the monomeric mutant, and its interactions with C-type lectin might not only depend on the presence of glycan ligands but also their accesability. Because neither of the monomeric C309L/N297A/C575A nor C309L/N563A/C575A with different available N-glycosylation site had binding affinity to the examined C-type lectins, even though they might have a higher content of high mannose structures than the mutant D221N/N297A/C309L/N563A/C575A which could be a ligand for these types of lectins. Thereby, the glycan and nearby sequence of the N221 N-glycosylation site could be a research target to identify a favourable binding moiety for C-type lectins on this recombinant Fc domain. In the presence of multiple N-glycosylation sites, the reactivity of mutants to Siglec receptors was generally enabled, with exception of mutant C309L which was only reactive to tested CD23. The monomeric mutants were generally conferred with a higher binding affinity and/or a broader binding reactivity to C-type lectins. For instance, the multimeric mutant D221N/C575A had binding affinity to Siglec 1-4 and occasionally to CD23, whereas its monomeric counterpart enabled binding to all tested C-type lectin with a high binding and enhanced binding strength to Siglec 1-4, simultaneously. There is only one exception that a monomeric mutant was less reactive to glycan receptors as C309L/N563A/C575A which could only occasionally bind to Siglec 4, but its multimeric counterpart was bound by Siglec 1-2 as well as Dectin-1, -2 and DC-SIGN

Table 5.3 The interactions between mutants and selected Lectins

	Siglec-1	Siglec-2	Siglec-3	Siglec-4	CD23	Dec-tin-1	Dec-tin-2	DC-SIGN	MBL	DEC-205
IgG1-Fc	–	–	–	–	–	–	–	–	–	ND
N297A/C575A	+	–	–	–	–/+	–	–	–	–	ND
C309L/N297A/C575A	++	–	–	–/+	–	–	–	–	–	ND
D221N/N297A/C575A	++	–	–	–	–/+	–	–	–	–	ND
D221N/C309L/N297A/C575A	++	–/+	–	+++	–/+	–	–	–	–	ND
D221N/C575A	++	+	–	+	–/+	–	–	–	–	ND
D221N/C309L/C575A	++++	++++	++++	++++	++++	+++	+++	+	+++	+++
N563A/C575A	++	–/+	–	–	–	–/+	–/+	–/+	–	ND
C309L/N563A/C575A	–	–	–	–/+	–	–	–	–	–	ND
D221N/N297A/N563A/C575A	++	–/+	+	+	+	–	–/+	–	–	ND
D221N/C309L/N297A/N563A/C575A	++	+++	++	+++	+	+	++	–/+	–	++++
D221N/N563A/C575A	+	–	–	–/+	–	+	–/+	–	–	ND
D221N/C309L/N563A/C575A	++	+++	+	+++	+	+	+	–/+	–	ND
Hexa-Fc	++	++	–	–/+	+	–	–	+	–	+
C309L	–	–	–	–	+	–	–	–	–	–
C575A	+	–	–	–	–/+	–	–	–	–	ND
C309L/C575A	+++	+++	+++	+++	+++	++	++	+	+	+++

ND, not detected; –, no binding; –/+, binding in some experiments; +, binding; ++, moderate binding; +++, high binding; ++++, very high binding. Mutants are coloured for the same N-glycosylation site possession to group together. IgG1-Fc is included as a negative control with available N-glycosylation site N297 for lectin binding examinations. The table is adapted from (Blundell et al., 2019)

occasionally.

Functionally, the binding of mutants to Siglec-4 would be beneficial to control neuropathy (Wong et al., 2016) and Siglec-2 binding can be involved in attenuating BCR signalling (Seite et al., 2010). However, due to Siglec-2 does not bind to α2,3 linked sialic acid which is normally produced in CHO cells, it is likely that the observed Siglec-2 binding might follow a glycan-independent approach but the enhanced glycosylation might still be involved in promoting the binding of glycan-independent ligands to Siglec-2 receptors (Alborzian Deh Sheikh et al., 2018).

Besides, dectin-1 was shown to be able to associate with FcγRIIB via linking by IgG1 to block complement signalling via tyrosine phosphorylation by syk family kinase to attenuate C5a receptor functions (Karsten et al., 2012). As the binding IgG1 was required to be well galactosylated to achieve such binding, relatively high level of galactosylation on the N-glycans of mutants as D221N/C309L/C575A and C309L/C575A, but less abundant sialylation, might provide the desired ligands to bind this lectin and engage in the suggested biological events. Monomeric mutants such as D221N/C309L/C575A and C309L/C575A could both bind to DC-SIGN, this binding would be beneficial for immune cells to express inhibitory FcRs and consequently suppress immune responses (Anthony et al., 2011). In addition, branching and even antennary elongation observed on mutants such as D221N/N297A/C309L/N563A/C575A or N297A/C575A could be a ligand candidate for galectin family lectins. Due to the CHO system only being able to produce α2,3 linked terminal sialic acids, it is very likely that galectin family proteins such as galectin-1 and -3 can be reactive to the branched and/or elongated N-glycans assembled on these mutants. Thereby, it could be interesting to examine the reactivities of these mutants with galectin family proteins, and potential functional implications such as to engage in or interfere with certain protein cross-linking by galectins.

Mutants with enhanced glycosylation also obtained reactivities against other receptors. For instance, D221N/C309L/N297A/N563A/C575A was shown to have a broad range of interactions with FcRs (FcγRI, FcγRIIb AND FcγRIIIA) and even influenza hemagglutinins (Blundell et al., 2019). This would indicate a pronounced contribution of the N-glycans assembled on the Fc mutants by the CHO cell system. This also highlights the importance of understanding the functions of the structural features obtained from the CHO cell system, as branching, sialylation and antennary elongation all impact on lectin dependent/independent interactions.

However, to date there is no comparable information available for these mutants expressed in the HEK cell system, which present more varied terminal moieties such as LacdiNAc and Lewis antigens. Lewis-x antigen has been shown to be able to induce potent immune responses (Hsu et al., 2007), whereas the expression of lewis antigens in cancers has also been correlated to the cancer progression (Blanas et al., 2018). LacdiNAc was also proposed as a ligand, expressed on helminths parasites, which

could be bound by galectin-3 from macrophages to induce immunorecognition, and the fucosylated version of this moiety could also be expressed on parasites and be recognised by immune system (van den Berg et al., 2004). Besides, glycodelin expressed mainly in the uterus system can also express LacdiNAc with sialylation or fucosylation (Dell et al., 1995), and the expression of LacdiNAc epitopes has been reported positively or negatively related to the progression of different cancers (Hirano et al., 2014). Thus, it is possible that LacdiNAc might be engaged in relevant events to regulate tumour progression or foetus developments. Subsequently, without a precise improvement of sialylation of Fc mutants expressed in the HEK cell system, the effects of these various epitopes should be critically evaluated to ensure their compatibility and effectiveness. Alternatively, a genetic modification for the CHO cell system has been successfully conducted in which the relevant glycosyltransferases to produce α2,6 linked sialic acids onto IgG1 could be introduced (Raymond et al., 2015). Thereby, a further modification on this described system to knockout α2,3 sialyltransferases combines with the N-glycosylation site manipulation technique applied in this work could be a considerable avenue for achieving the biomedical purpose to generate consistent and efficient Fc-fusion protein with α2,6 linked sialic acid to promote strong immunomodulation effects for countering autoimmune diseases.

Due to the limitations of time the antennary organisation and modification were unable to be fully defined. However, these undefined structural features can be continued by further MS/MS experiments and GC-MS linkage analysis. Moreover, glycoproteomics experiments are also needed to verify the exact glycoform on each N-glycosylation site to evaluate the applicability of these glycans in lectin-glycan interactions as well as to understand the manner of site-specific glycosylation on these N-glycosylation sites.

Chapter VI. Cryoglobulin

Chapter 6 Cryoglobulin

6.1 Introduction

Cryoglobulin is a subgroup of pathogenic immunoglobulins which features as reversible precipitation under low temperature *in vitro*. Cryoglobulins can principally cause a type of vasculitis (Braun et al., 2007), named cryoglobulinaemic vasculitis, in small to medium sized vessels in vivo.

Historically, Wintrobe and Buell first described cold-precipitating proteins identified from the serum of a multiple myeloma patient in 1933 (Wintrobe. and Buell., 1933). The name, cryoglobulin, was then given by Lerner and Waston to these temperature-dependent precipitating proteins (Lerner and Watson, 1947). In 1974, Brouet proposed a widely accepted classification of cryoglobulins based on the observation of clonality and constitutions of immunoglobulins from clinical practices (Brouet et al., 1974). In this classification, cryoglobulins are classified into three categories as type I, II and III (Figure 6.1).

Type I cryoglobulins refer to cryoglobulins with a single constituent of monoclonal immunoglobulins, mainly IgG. The monoclonality of these cryoglobulins results from the expansion of a single clone, possibly in malignancy or indolence or secondary to other lymphoproliferative disorders (Auscher and Guinand, 1964, Somer, 1966, Meltzer et al., 1966, Slavin et al., 1971, Klein et al., 1972).

Type II and III cryoglobulins are both mixed cryoglobulins which are featured as an antigen-antibody association of IgM and IgG (Lospalluto et al., 1962, Auscher and Guinand, 1964, Costanzi et al.,

Type I — Monoclonal IgG

Type II — Polyclonal IgG+ Monoclonal IgM

Type III — Polyclonal IgG+ Polyclonal IgM
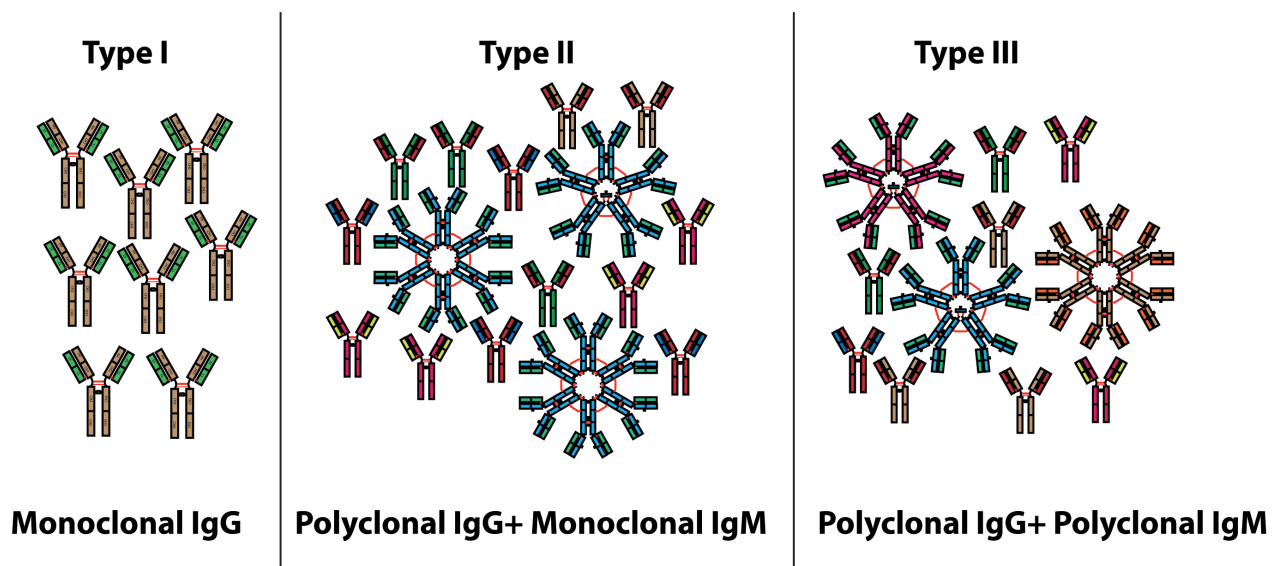
Figure 6.1 Schematic paradigms of the classification of cryoglobulins

1965). The type II cryoglobulins commonly consist of polyclonal IgGs and monoclonal IgM which has rheumatoid factor (RF) activity. The RF activity enables monoclonal IgM to interact with the Fc of polyclonal IgGs. Type III cryoglobulins are immune complexes consisting of polyclonal IgM and IgG. These two types of mixed cryoglobulins are often relevant to chronic infectious diseases, including viral (hepatitis B and C), bacterial or parasitic infections, and auto-immune diseases as systematic lupus erythematosus or rheumatoid arthritis (Lospalluto et al., 1962, Auscher and Guinand, 1964, Costanzi et al., 1965).

Cryoglobulins are not significantly different from normal immunoglobulins in general structure, and they have similar electrophoretic mobility to serum mono- or polyclonal immunoglobulins (Meltzer and Franklin, 1966, Shihabi, 1996, Musset et al., 1992). Their heterogeneous origins complicates the determination of identical sequences that could correlate to their cryogenic properties, and thereby no consistent protein sequence or characters have been suggested to be responsible for the cryogenic activities.

Even if there is no identical sequence moiety that could be correlated to the cryogenic properties of cryoglobulin, the impacts of glycosylation on the formation of cryoprecipitate have been progressively studied on a particular subtype of IgGs—murine IgG3. Human and murine IgG3s can self-associate via spontaneous Fc-Fc interaction as type I cryoglobulins (Grey et al., 1968, Capra and Kunkel, 1970). This complex formation could lead to a temperature dependent insolubility of IgG3 molecules, whereas formation of IgG3 cryoprecipitate required RF activities of IgG3 to cause relevant biological consequences (Berney et al., 1992). Panka and colleagues firstly reported that a direct mutagenesis on a N-glycosylation site (N471S) of the CH3 domain of IgG3 RF could reduce its ability of self-association (Panka, 1997). A study with hybrid IgG which had identical heavy chain sequences expressed in different cells found that the pathogenic activity of IgG3 cryoglobulins was decreased while there was an increase in the level of galactosylation (Mizuochi et al., 2001). Another study investigated cryogenic activities of IgG3 RF antibodies with the same heavy chain sequence plus a knockout of N471 (N471T) on the CH3 domain. They concluded that the CH3 domain N-glycosylation site was not involved in the cryo-precipitation and suggested that the level of galactosylation was negatively correlated to the cryogenic activities of IgG3 RF cryoglobulins (Kuroki et al., 2002). The same group then amended their suggestion in the following study to that the level of sialylation was responsible for the cryogenic activities of IgG3 cryoglobulins as they examined more cryogenic variants of IgG3 (Kuroda et al., 2005). They found that the galactosylation level of these variants was not identical, and the cryogenic IgG3s, which remained in solution after the initial cryo-precipitation *in vitro*, was then unable to form cryoprecipitate and had a relatively higher content of sialic acids (Kuroda et al., 2005). This evidence might imply that the cryogenic activities of cryoglobulins could solely be determined by the status of their glycosylation. The protective effect of sialylation on IgG3 cryoglobulin was further assessed by a group using mutagenesis approaches to enhance the level of sialylation on IgG3 sequences. They found

that the nephritogenic activity of IgG3 cryoglobulin was dependent on the level of sialylation which also negatively regulated the cryogenic activity (Otani et al., 2012).

As glycosylation was shown to have possible regulatory impact on the cryogenic activity of IgG3 cryoglobulins, it was hypothesised that the observed role of glycosylation on murine IgG3 cryoglobulin could be translated to that of mixed type cryoglobulins. Thereby, a system that employed an established cryoglobulin mouse model which overexpresses thymic stromal lymphopoietin (TSLP) to produce type III mixed cryoglobulins was assessed (Taneda et al., 2001). TSLP is a cytokine which was originally found to facilitate the development of IgM[+] B cells (Friend et al., 1994, Sims et al., 2000). Overexpressing this cytokine in mouse models could result in a systematic increase of polyclonal IgM and IgG production and the deposition of these antibodies in kidney (Taneda et al., 2001, Astrakhan et al., 2007). The cryogenic activities of these deposited polyclonal antibodies resulted in the development of cryoglobulinaemic glomerulonephritis and type III cryoglobulins in model animals (Taneda et al., 2001, Astrakhan et al., 2007). This recruited mouse model is identified to mimic two important features of human cryoglobulinaemic glomerulonephritis: 1) macrophage influx and 2) occlusion of glomerular capillaries (Taneda et al., 2001). However, *in vivo* pathogenic studies of this mouse model indicated that the pathogenesis of their cryoglobulins were independent from Fcγ receptors and complement cascade (Muhlfeld et al., 2004, Guo et al., 2009).

In this chapter, we aim to characterise the N-glycan repertoires and verify the protein compositions of IgG and cryoprecipitate from the TSLP transgenic mouse model. By combing information from both aspects, we aim to understand the role of glycosylation in the cryogenic activity of type III cryoglobulins.

6.2 Results

6.2.1 Research Strategies

The study was undertaken in collaboration was with Professor Alan Salama of Royal Free Hospital, University College London, London. They maintained the model animal resources and provided IgG materials and cryo-precipitates for analysis. We received two sets of samples to analyse: 1) murine IgG materials which were purified from TSLP and WT murine sera by a protein G column; 2) cryo-precipitates harvested from TSLP murine serum had been incubated at 4 °C for 24 hours for cryoprecipitation. The glycomics and proteomics strategies for this project are summarised in Figure 6.2

TSLP &WT IgGs and TSLP cryoprecipitates

↓

Reduction and carboxymethylation

↓

Tryptic digestion

↓

PNGaseF digeston → Glycan Fraction → Permethylation

Peptide Fraction

↓ Online LC ES-MS/MS (Synapt)

↓ MALDI-TOF profiling and MALDI-TOF/TOF sequencing
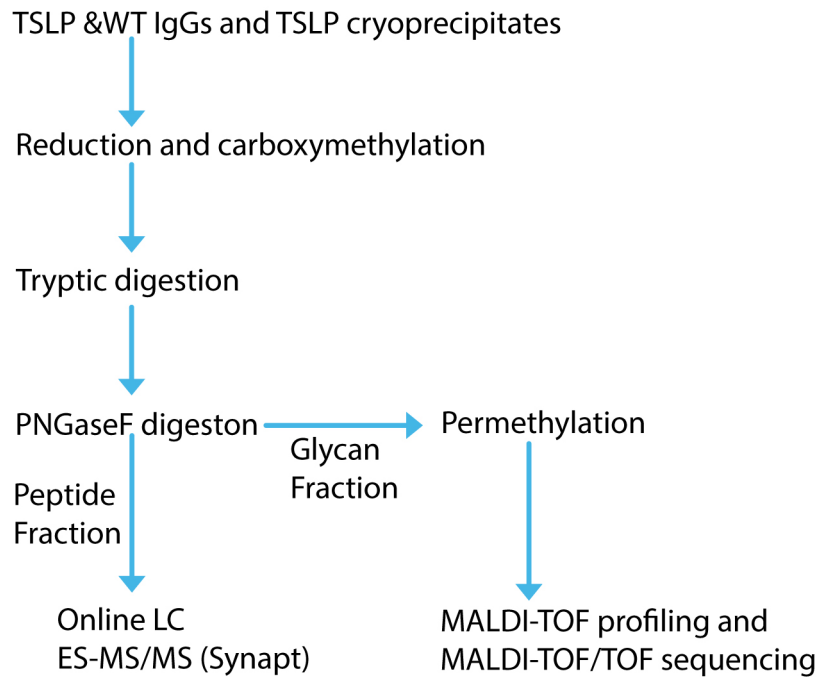
Figure 6.2 Schematic paradigm of research strategies applied in this project

6.2.2 N-glycosylation of murine IgG from transgenic TSLP and WT mice

To profile the N-glycan repertoires of IgG materials from both mouse groups, MALDI-TOF MS based glycomic analysis was performed. Figure 6.3 shows a comparison of IgG N-glycan profiles between TSLP (Figure 6.3A) and WT mice (Figure 6.3B). The N-glycan profile of WT IgGs N-glycans is generally consistent with previous publications (Blomme et al., 2011). Bi-antennary structures are exclusively observed in this spectrum, and three core-fucosylated bi-antennary structures (*m/z* 1835, 2040 and 2244) are predominant in abundance with varied galactosylation levels (n=0-2). Sialylation on the antennary LacNAc is also observed (*m/z* 241, 2635, 2852 and 3026). Instead of observing a NeuAc sialic acid capped terminal, NeuGc is found on all sialylated terminals of antennae. A di-sialylated bi-antennary structure at *m/z* 2852 is the only one without core-fucosylation which could indicate possible origin from other serum glycoproteins as it was the base peak in the spectrum of previous analysis of total murine serum glycome (Blomme et al., 2011). In addition, a core-fucosylated bi-antennary structure, *m/z* 2839, is identified near the baseline carrying a sialyl-LacNAc and a putative Gal-α-Gal.

IgG materials from TSLP mice exhibit a comparable glycoform to that of WT with an exception for two other high-mannose structures in the spectrum as Man5 (*m/z* 1579) and Man6 (*m/z* 1783). Besides, semi-quantitatively, a decrease of sialylation in TSLP is discernible as the relative abundance of mono-sialylated bi-antennary structure *m/z* 2635 becomes lower in comparison with its asialylated precursor *m/z* 2244. The ratio of relative abundance of *m/z* 2635 to 2244 becomes 63% in the TSLP from 134% in the WT control.
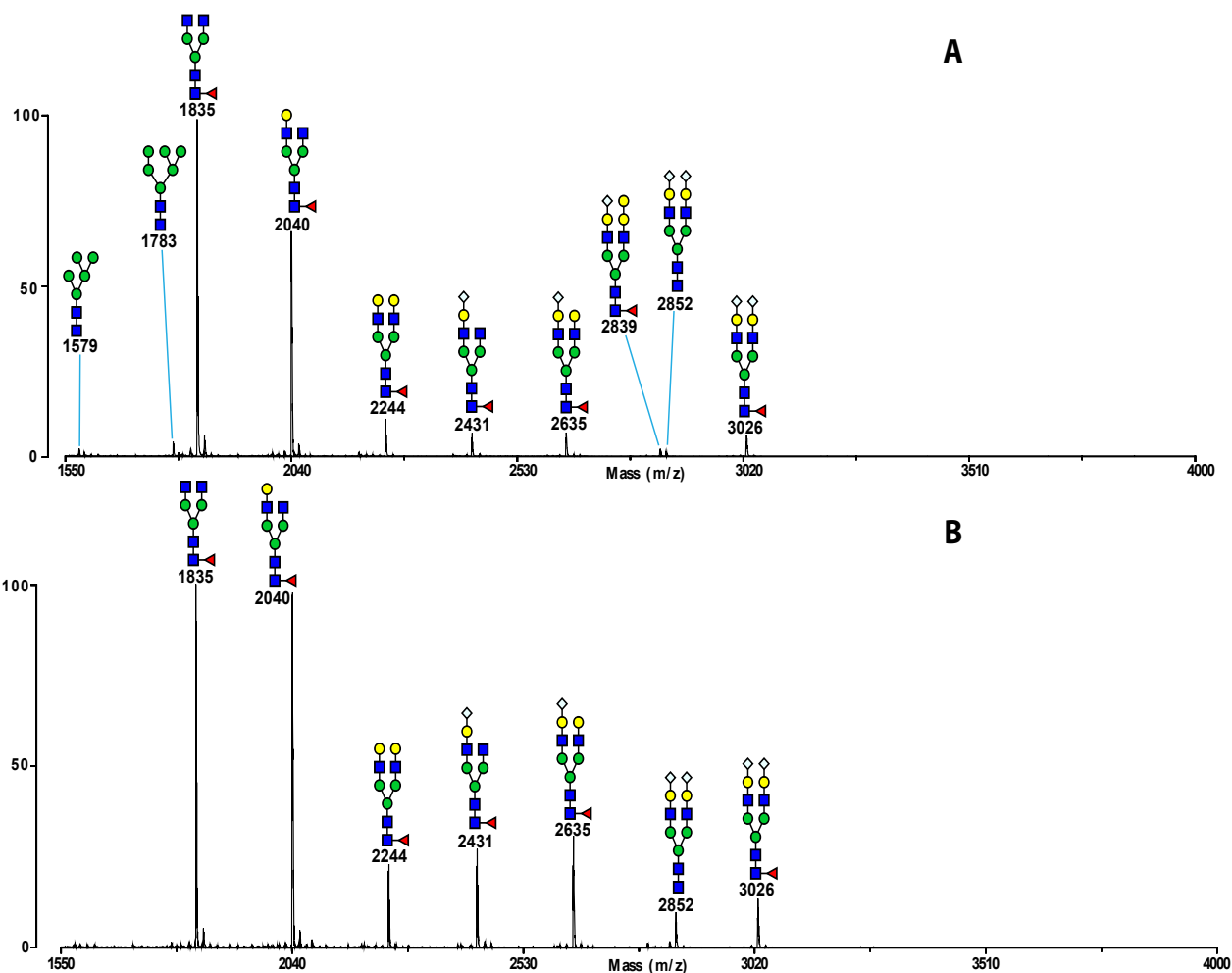
Figure 6.3 MALDI-TOF spectra of murine IgGs from TSLP and WT serum.
N-glycoprofiles of IgG N-glycans are shown for TSLP(A) and WT(B). A protein G purification was performed after sample collection. N-glycans were extracted from the purified IgGs. N-glycans were then permethylated and 50% Acetonitrile/Water fractions were employed for data acquisition. All the ions are observed as [M+Na]$^+$, and all the assignments are based on knowledge of biosynthetic pathway as well as compositional information.

6.2.3 Fragmentation analysis for m/z 2839 from TSLP IgG N-glycans

To verify the presence of NeuGc and Gal-α-Gal moieties on this murine N-glycans from the IgG of TSLP murine serum, fragmentation analysis was carried out on a MALDI-TOF/TOF instrument with CID fragmentation. Figure 6.4 presents the spectrum of interested ion *m/z* 2839. Sialyl-LacNAc with NeuGc on this structure is validated by the presence of a pair of b- and y-ions as *m/z* 877 and 1984. The linear arrangement of this sialyl-LacNAc antenna is verified by observing a series of y ions as *m/z* 2434, 2230, 1984 and 1780. The antenna with Gal-α-Gal is also unambiguously validated as discerning a sequential fragmentation from its terminal Gal to base Man as y ions *m/z* 2621, 2417, 2172 and 1967. Core-fucosylation on reducing end GlcNAc is verified as observing *m/z* 2388 which corresponds to a loss of reducing end Fuc-GlcNAc di-saccharide.
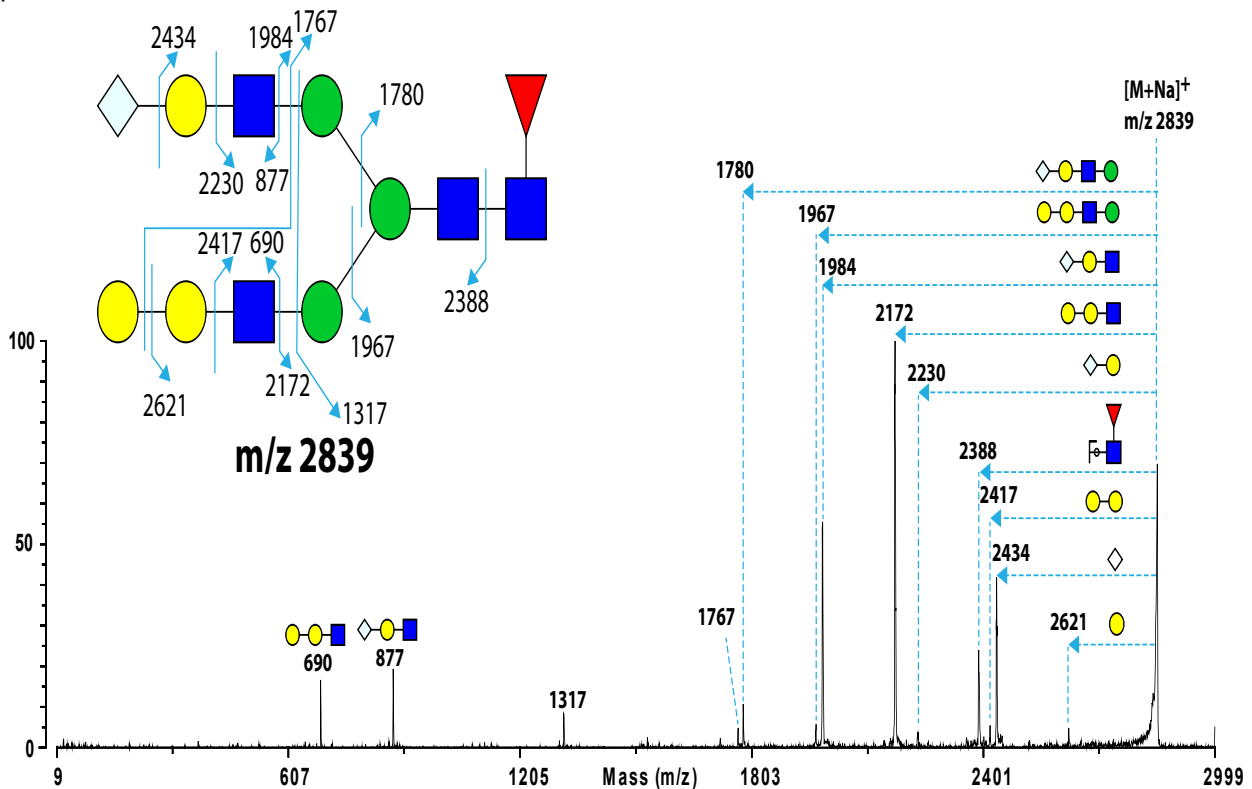
Figure 6.4 MALDI-TOF/TOF spectrum of m/z 2839 selected from TSLP IgG glycans.
The ion is selected from the screened N-glycan repertoire (**Section 6.2.2**) and to verify the suggested structural feature. All ions were observed as [M+Na]⁺ and all the assignments are based on knowledge of the fragmentation pathway for permethylated glycans.

6.2.4 Proteomics analysis for protein composition of TSLP and WT murine IgGs

To simultaneously verify the purity of IgG preparations, deglycosylated peptides of both materials were subject to a standard proteomics analysis by an LC-MS/MS on the Waters Synapt system. A Mascot search was conducted for acquired MS and MS/MS data automatically to verify protein identities in both samples.

The Mascot Search automatically assigned different hits across species for each material, so that these hits were manually filtered to the experimental species, mus musculus (mouse). 45 hits remain in the filtered list of TSLP material (Table 6.1), major hits are immunoglobulin kappa (IgK) chains from Fab region. Three hits for the constant region of Immunoglobulin γ (IgG) chain are found as IgG1 (secreted/membrane bound) and IgG2B. Interestingly, hits for immunoglobulin mu (IgM) and joining (J-chain) chain are also found in this list which could indicate the presence of IgM pentamers in this material (Yoo et al., 1999). The ratio of IgG to IgM is preliminarily evaluated by their protein product intensity sum (MS/MS) as approximately above 10:1 (summed intensity of IgG secreted form and IgG2 γ chain: IgM=5734061:555021).

Table 6.1 Filtered and Simplified Proteomics hit list for TSLP IgG material

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
P01672	Ig kappa chain V-III region PC 7940 OS=Mus musculus PE=1 SV=1	6480.232	204263
P01665	Ig kappa chain V-III region PC 7043 OS=Mus musculus PE=1 SV=1	5981.903	205406
P01654	Ig kappa chain V-III region PC 2880/PC 1229 OS=Mus musculus PE=1 SV=1	5832.497	219113
P03977	Ig kappa chain V-III region 50S10.1 OS=Mus musculus PE=1 SV=1	5717.088	209702
P01660	Ig kappa chain V-III region PC 3741/TEPC 111 OS=Mus musculus PE=1 SV=1	4892.04	193410
P01673	Ig kappa chain V-III region PC 2485/PC 4039 OS=Mus musculus PE=1 SV=1	4831.107	179074
P01662	Ig kappa chain V-III region ABPC 22/PC 9245 OS=Mus musculus PE=1 SV=1	3526.082	50228
P01837	Ig kappa chain C region OS=Mus musculus PE=1 SV=1	3167.019	1440960
P01869	Ig gamma-1 chain C region membrane-bound form OS=Mus musculus GN=Ighg1 PE=1 SV=2	3126.71	4942667
P01868	Ig gamma-1 chain C region secreted form OS=Mus musculus GN=Ighg1 PE=1 SV=1	3126.71	5016984
P01872	Ig mu chain C region OS=Mus musculus GN=Ighm PE=1 SV=2	1961.831	555021
P18528	Ig heavy chain V region 6.96 OS=Mus musculus PE=4 SV=1	1943.353	124560
P01661	Ig kappa chain V-III region MOPC 63 OS=Mus musculus PE=1 SV=1	1649.125	25189
P01756	Ig heavy chain V region MOPC 104E OS=Mus musculus PE=1 SV=1	1478.154	311298
P01867	Ig gamma-2B chain C region OS=Mus musculus GN=Igh-3 PE=1 SV=3	1443.758	791394
P04945	Ig kappa chain V-VI region NQ2-6.1 OS=Mus musculus PE=2 SV=1	1090.817	30048
P01674	Ig kappa chain V-III region PC 2154 OS=Mus musculus PE=1 SV=1	783.616	22590
P01642	Ig kappa chain V-V region L7 (Fragment) OS=Mus musculus GN=Gm10881 PE=1 SV=1	764.854	37693
P01639	Ig kappa chain V-V region MOPC 41 OS=Mus musculus GN=Gm5571 PE=1 SV=1	750.658	40862
P01631	Ig kappa chain V-II region 26-10 OS=Mus musculus PE=1 SV=1	683.509	156173
P01644	Ig kappa chain V-V region HP R16.7 OS=Mus musculus PE=1 SV=1	606.928	163375
P01638	Ig kappa chain V-V region L6 (Fragment) OS=Mus musculus PE=4 SV=1	499.366	225936
P01843	Ig lambda-1 chain C region OS=Mus musculus PE=1 SV=1	405.870	56196
P01592	Immunoglobulin J chain OS=Mus musculus GN=Jchain PE=1 SV=4	388.214	31100
P01646	Ig kappa chain V-V region HP 123E6 OS=Mus musculus PE=1 SV=1	317.935	28218
P01643	Ig kappa chain V-V region MOPC 173 OS=Mus musculus PE=1 SV=1	288.993	135157
P01680	Ig kappa chain V-IV region S107B OS=Mus musculus PE=4 SV=1	179.877	14894
P01637	Ig kappa chain V-V region T1 OS=Mus musculus PE=4 SV=1	146.962	213324

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list

Table 6.2 Filtered and simplified Proteomics hit list for WT IgG material

Protein Acession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
P01654	Ig kappa chain V-III region PC 2880/PC 1229 OS=Mus musculus PE=1 SV=1	5238.799	126085
P01665	Ig kappa chain V-III region PC 7043 OS=Mus musculus PE=1 SV=1	4003.874	103237
P01672	Ig kappa chain V-III region PC 7940 OS=Mus musculus PE=1 SV=1	3921.558	100958
P01673	Ig kappa chain V-III region PC 2485/PC 4039 OS=Mus musculus PE=1 SV=1	3729.992	92274
P01837	Ig kappa chain C region OS=Mus musculus PE=1 SV=1	2622.865	1305180
P01867	Ig gamma-2B chain C region OS=Mus musculus GN=Igh-3 PE=1 SV=3	2508.231	2160084
P01868	Ig gamma-1 chain C region secreted form OS=Mus musculus GN=Ighg1 PE=1 SV=1	1994.344	1406196
P01869	Ig gamma-1 chain C region membrane-bound form OS=Mus musculus GN=Ighg1 PE=1 SV=2	1984.746	1372542
P07724	Serum albumin OS=Mus musculus GN=Alb PE=1 SV=3	1434.491	845898
P01756	Ig heavy chain V region MOPC 104E OS=Mus musculus PE=1 SV=1	1227.859	81975
P01631	Ig kappa chain V-II region 26-10 OS=Mus musculus PE=1 SV=1	1115.306	196090
P01642	Ig kappa chain V-V region L7 (Fragment) OS=Mus musculus GN=Gm10881 PE=1 SV=1	1003.723	46086
P04945	Ig kappa chain V-VI region NQ2-6.1 OS=Mus musculus PE=2 SV=1	699.294	13872
P03987	Ig gamma-3 chain C region OS=Mus musculus PE=1 SV=2	522.987	75549
P01662	Ig kappa chain V-III region ABPC 22/PC 9245 OS=Mus musculus PE=1 SV=1	459.591	16668
P01635	Ig kappa chain V-V region K2 (Fragment) OS=Mus musculus PE=1 SV=1	409.43	52569
P18528	Ig heavy chain V region 6.96 OS=Mus musculus PE=4 SV=1	332.195	31622
P02089	Hemoglobin subunit beta-2 OS=Mus musculus GN=Hbb-b2 PE=1 SV=2	294.789	5177
Q8BGD8	Cytochrome c oxidase assembly factor 6 homolog OS=Mus musculus GN=Coa6 PE=1 SV=1	289.167	2325
P01661	Ig kappa chain V-III region MOPC 63 OS=Mus musculus PE=1 SV=1	191.566	8684
Q8BGZ7	Keratin_ type II cytoskeletal 75 OS=Mus musculus GN=Krt75 PE=1 SV=1	124.092	41447
Q922U2	Keratin_ type II cytoskeletal 5 OS=Mus musculus GN=Krt5 PE=1 SV=1	111.748	8241
Q9Z331	Keratin_ type II cytoskeletal 6B OS=Mus musculus GN=Krt6b PE=1 SV=3	105.857	7402
Q6NXH9	Keratin_ type II cytoskeletal 73 OS=Mus musculus GN=Krt73 PE=1 SV=1	73.879	15172

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list

Unlike the TLSP material, there are 48 hits in the filtered list of WT material (Table 6.2) without any hit correspond to J-chain or IgM. Instead, other serum proteins and common protein contaminants were detected in this material including hemoglobulin, serum albumin and keratin at relatively trace amounts. Besides, hits of IgK are still high-scored in the list, and four hits for IgG chains have been identified including a new IgG3 hit. The listed contaminants are not N-glycosylated (Ku et al., 2010, Anguizola et al., 2013), therefore the characterised glycome for WT IgG material should be representative for its N-glycosylation.

6.2.5 N-glycosylation of cryo-precipitate from TSLP transgenic murine sera

The glycomics analyses for cryo-precipitates were performed on three independent batches of cryo-precipitates harvested from multiple TSLP model mice. As described in **Section 6.1**, the employed mouse strain was characterised as producing type III cryo-globulins which consisted of polyclonal IgMs and IgGs. In addition, to gain a maximum quantity of cryoprecipitates, protein G columns were not applied to these materials for purification in our collaborator's laboratory.
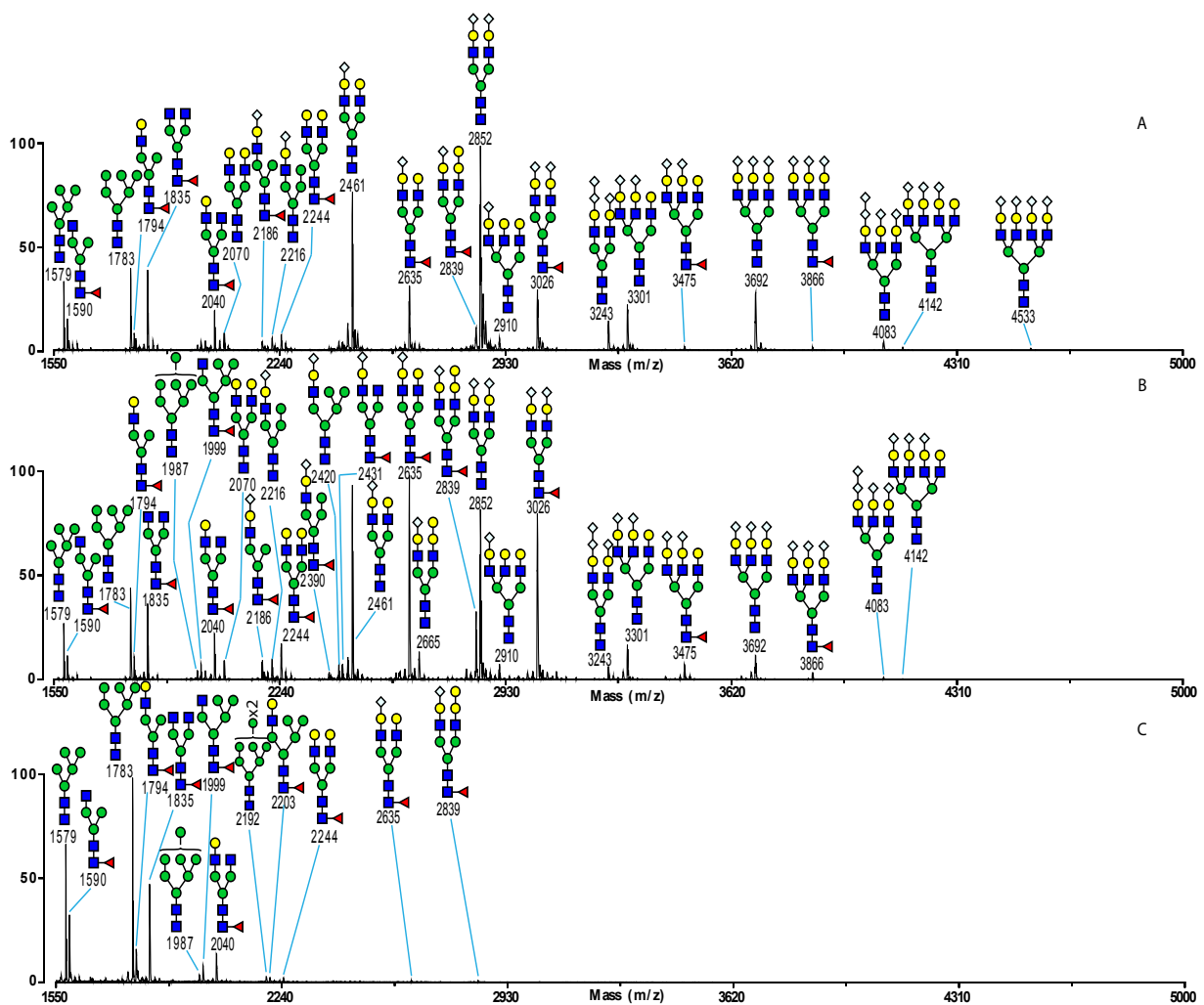
Figure 6.5 MALDI-TOF spectra of three independent batches of TSLP cryopreicipitates
These three batches are first batch (A), second batch (B) and third batch (C). They were all directly analysed without a prior protein G purification. See the legend of Figure 6.3 for assignments.

Figure 6.5 shows the spectra of N-glycans from all batches of tested cryoprecipitate materials. The first and second batches of cryoprecipitate generally share simmilar glycoforms (Figure 6.5A and B). Their glycoforms consist of a full range of N-glycans from high-mannose to complex type. Sialylation is a common modification on these multi-antennary hybrid and complex glycans, such as *m/z* 2216 and 3026. The overall glycomes are dominated by sialylated bi-antennary structures with and without core-fucosylation (*e.g. m/z* 2461, 2635, 2852 and 3026). However, the previously described three core-fucosylated bi-antennary structures (*m/z* 1835, 2040 and 2244) become less dominant in abundance than in the characterised TSLP IgG (**Section 6.2.1**). Structures at *m/z* 3243 and 4083 are found in both spectra which have an extra NeuGc residue in composition with all LacNAc antennary terminals already fully capped with sialic acids. Thus, this observation raised a speculation that both batches of materials contain transferrin. This is because transferrin has been characterised as a serum glycoprotein specifically carrying N-glycans with di-sialylated LacNAc antennae (Freeze, 2006, Rymen et al., 2012).

In contrast to the above two cryoprecipitate materials, there is no detection of tri- and tetra-antennary structures in the third batch of cryoprecipitates (Figure 6.5C). Fully galactosylated bi-antennary structures such as *m/z* 2244, 2635 and 2839 are of relatively low abundance. High mannose glycans are predominant in this batch of N-glycans at *m/z* 1579 and 1783. Two core-fucosylated bi-antennary structures at *m/z* 1835 and 2040 as well as truncated mono-antennary structure *m/z* 1590 are also dominant in the spectrum.

As the third batch of materials has a strikingly less complex N-glycoform, remaining materials was subject to further verification for of both glycan structures and protein identities.

6.2.6 Fragmentation analysis for m/z 3243 of N-glycans from TSLP cryo-precipitate

To verify the proposed di-sialylated LacNAc, *m/z* 3243 was subject to fragmentation analysis as shown in Figure 6.6. A pair of b- and y-ions are identified in the spectrum at *m/z* 1268 and 1998 which correspond to the cleavage of a di-sialylated LacNAc sequence. Notably, a β elimination of NeuGc-Gal di-saccharide is also observed as a pair of cz ions at *m/z* 650 and 2616 which reflects the presence of a β1,3 linkage Type I arrangement of terminal sialylated LacNAc moieties. NeuGc Sialyl-LacNAc was also verified by the presence of a by ion pair as *m/z* 877 and 2389.
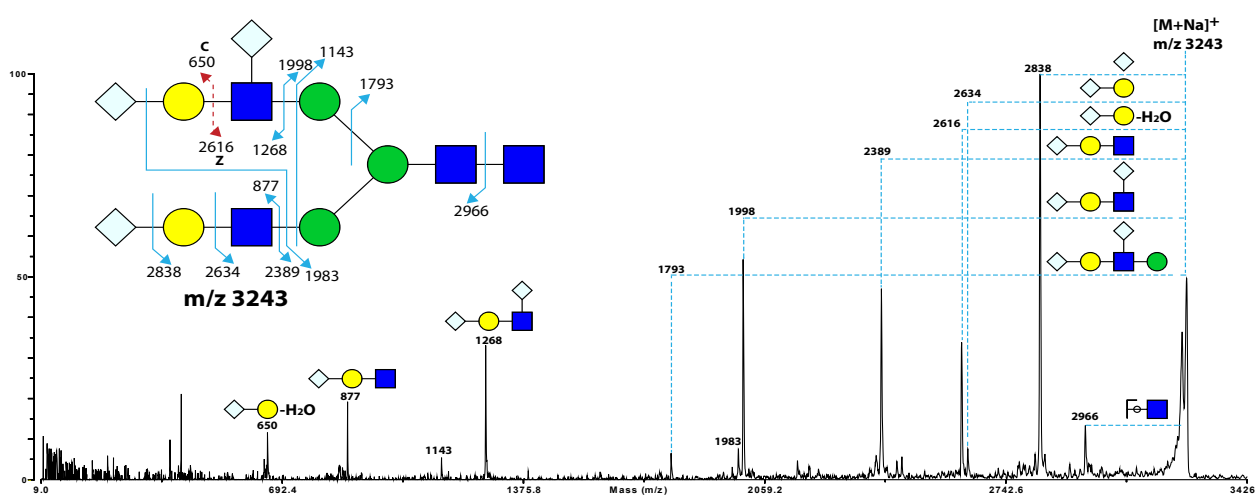
Figure 6.6 MALDI-TOF/TOF spectrum of m/z 3243 selected from the first batch of TSLP cryoprecipitates. This ion wask peaked for validation of the presence of di-sialylated antenna. See the legend of Figure 6.4 for the assignments.

6.2.7 Proteomics analysis for cryo-precipiate from transgenic TSLP mice

There are 79, 17 and 89 hits identified after species filtration as described in **Section 6.2.3** for the first, second and third batches of cryoprecipitate materials (**Section.6.2.5**, Table 6.3, 6.4 and 6.5), respectively. Hits for IgG1 and IgM are consistently present in all cryoprecipitate materials which agrees with the previous suggestion for mixed-cryoglobulins yielded from this animal model (Taneda et al., 2001). However, other serum proteins also abundantly appear in these materials including serotransferrin, hemoglobulin, serum albumin and apolipoprotein; the serum components and their relative abundance are varied between different batches. The complexity of cryoprecipitate components could be the result of random association or steric trapping with these serum components during the process of cryoglobulin precipitation. Besides, the ratio of IgM to serotransferrin to IgG is generally evaluated as in **Section 6.2.4** as approximately 7.7:4.8:4.5 in the first batch, 2.1: 0.6: 0.11 in the second batch and 4.5:0.6:0.19 in the third batch. The dominant presence of serotransferrins and IgM in these cryo-precipitates could imply their contributions to the profiled N-glycan repertoire of each batch of

materials. Moreover, other Ig proteins such as immunoglobulin epsilon (IgE), alpha (IgA) and haptoglobulin are found in some batch of these materials. Hence the profiled overall N-glycan repertoire for each cryo-precipitate material would rather be composed of a complex mixture dominated by IgM and serotransferrins other than that of IgGs.

Table 6.3 Filtered and simplified proteomics hit list for the first batch of TSLP cryoprecipitates

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
Q00623	Apolipoprotein A-I OS=Mus musculus GN=Apoa1 PE=1 SV=2	9129.411	4534233
P01872	Ig mu chain C region OS=Mus musculus GN=Ighm PE=1 SV=2	6428.866	7784539
P07724	Serum albumin OS=Mus musculus GN=Alb PE=1 SV=3	5536.757	4.77E+7
P07759	Serine protease inhibitor A3K OS=Mus musculus GN=Serpina3k PE=1 SV=2	4883.753	1243170
P06330	Ig heavy chain V region AC38 205.12 OS=Mus musculus PE=1 SV=1	3644.426	821486
P01757	Ig heavy chain V region J558 OS=Mus musculus PE=1 SV=1	3374.99	550841
Q92lI1	Serotransferrin OS=Mus musculus GN=Tf PE=1 SV=1	3191.249	4885248
Q61646	Haptoglobin OS=Mus musculus GN=Hp PE=1 SV=1	2592.372	1935282
P09813	Apolipoprotein A-II OS=Mus musculus GN=Apoa2 PE=1 SV=2	2447.953	332790
Q00898	Alpha-1-antitrypsin 1-5 OS=Mus musculus GN=Serpina1e PE=1 SV=1	2403.225	2444066
P01864	Ig gamma-2A chain C region secreted form OS=Mus musculus PE=1 SV=1	2176.89	2869060
P01837	Ig kappa chain C region OS=Mus musculus PE=1 SV=1	1839.351	856369
Q61838	Alpha-2-macroglobulin OS=Mus musculus GN=A2m PE=1 SV=3	1715.35	4830746
Q91X72	Hemopexin OS=Mus musculus GN=Hpx PE=1 SV=2	1633.988	2431502
P01666	Ig kappa chain V-III region PC 7183 OS=Mus musculus PE=1 SV=1	1603.055	144261
P01655	Ig kappa chain V-III region PC 7132 OS=Mus musculus PE=1 SV=1	1430.772	116570
P02088	Hemoglobin subunit beta-1 OS=Mus musculus GN=Hbb-b1 PE=1 SV=2	1416.045	312807
P01654	Ig kappa chain V-III region PC 2880/PC 1229 OS=Mus musculus PE=1 SV=1	1406.32	106034
P01670	Ig kappa chain V-III region PC 6684 OS=Mus musculus PE=1 SV=1	1396.951	102473
P01645	Ig kappa chain V-V region HP 93G7 OS=Mus musculus PE=1 SV=1	1360.305	246223
P01867	Ig gamma-2B chain C region OS=Mus musculus GN=Igh-3 PE=1 SV=3	1351.854	1188242
P01942	Hemoglobin subunit alpha OS=Mus musculus GN=Hba PE=1 SV=2	1324.426	256690
P22599	Alpha-1-antitrypsin 1-2 OS=Mus musculus GN=Serpina1b PE=1 SV=2	1287.915	2232720

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other glycoprotein contaminants are highlighted in red fonts

Table 6.3 Filtered and simplified proteomics hit list for the first batch of TSLP cryoprecipitates (continued 1)

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
Q00897	Alpha-1-antitrypsin 1-4 OS=Mus musculus GN=Serpina1d PE=2 SV=1	1259.595	2175716
P18528	Ig heavy chain V region 6.96 OS=Mus musculus PE=4 SV=1	1094.943	115546
Q00896	Alpha-1-antitrypsin 1-3 OS=Mus musculus GN=Serpina1c PE=1 SV=2	913.569	2199785
P07758	Alpha-1-antitrypsin 1-1 OS=Mus musculus GN=Serpina1a PE=1 SV=4	903.898	2198355
P02104	Hemoglobin subunit epsilon-Y2 OS=Mus musculus GN=Hbb-y PE=1 SV=2	828.664	150267
P01868	Ig gamma-1 chain C region secreted form OS=Mus musculus GN=Ighg1 PE=1 SV=1	817.666	528157
P01869	Ig gamma-1 chain C region membrane-bound form OS=Mus musculus GN=Ighg1 PE=1 SV=2	817.666	528157
P07309	Transthyretin OS=Mus musculus GN=Ttr PE=1 SV=1	814.047	231638
Q60590	Alpha-1-acid glycoprotein 1 OS=Mus musculus GN=Orm1 PE=1 SV=1	808.477	326885
P01646	Ig kappa chain V-V region HP 123E6 OS=Mus musculus PE=1 SV=1	791.680	142582
P01027	Complement C3 OS=Mus musculus GN=C3 PE=1 SV=3	763.400	4309022
P07361	Alpha-1-acid glycoprotein 2 OS=Mus musculus GN=Orm2 PE=1 SV=1	739.834	250844
P06728	Apolipoprotein A-IV OS=Mus musculus GN=Apoa4 PE=1 SV=3	642.745	310476
P01865	Ig gamma-2A chain C region membrane-bound form OS=Mus musculus GN=Igh-1a PE=1 SV=3	601.191	1241075
P01863	Ig gamma-2A chain C region A allele OS=Mus musculus GN=Ighg PE=1 SV=1	601.191	1241075
P01647	Ig kappa chain V-V region HP 124E1 OS=Mus musculus PE=1 SV=1	568.625	103641
Q03734	Serine protease inhibitor A3M OS=Mus musculus GN=Serpina3m PE=1 SV=2	558.584	473916
P08226	Apolipoprotein E OS=Mus musculus GN=Apoe PE=1 SV=2	519.021	318716
P01878	Ig alpha chain C region OS=Mus musculus PE=1 SV=1	518.168	173706
P06336	Ig epsilon chain C region OS=Mus musculus PE=4 SV=2	467.883	410617
Q3TB92	Allergin-1 OS=Mus musculus GN=Milr1 PE=1 SV=2	437.248	233752
P01843	Ig lambda-1 chain C region OS=Mus musculus PE=1 SV=1	380.966	342691
P01635	Ig kappa chain V-V region K2 (Fragment) OS=Mus musculus PE=1 SV=1	336.805	95395

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other glycoprotein contaminants are highlighted in red fonts

Table 6.3 Filtered and simplified proteomics hit list for the first batch of TSLP cryoprecipitates (continued 2)

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
Q91WP6	Serine protease inhibitor A3N OS=Mus musculus GN=Serpina3n PE=1 SV=1	299.910	369685
Q80X76	Serine protease inhibitor A3F OS=Mus musculus GN=Serpina3f PE=1 SV=3	274.598	280393
P23953	Carboxylesterase 1C OS=Mus musculus GN=Ces1c PE=1 SV=4	267.993	171093
Q5I2A0	Serine protease inhibitor A3G OS=Mus musculus GN=Serpina3g PE=2 SV=2	263.456	320970
P29621	Serine protease inhibitor A3C OS=Mus musculus GN=Serpina3c PE=2 SV=1	262.196	314358
Q06890	Clusterin OS=Mus musculus GN=Clu PE=1 SV=1	260.121	130580
P28665	Murinoglobulin-1 OS=Mus musculus GN=Mug1 PE=1 SV=3	258.3	825870
O08677	Kininogen-1 OS=Mus musculus GN=Kng1 PE=1 SV=1	203.364	239774
Q9JL35	High mobility group nucleosome-binding domain-containing protein 5 OS=Mus musculus GN=Hmgn5 PE=1 SV=2	200.840	203632
P28666	Murinoglobulin-2 OS=Mus musculus GN=Mug2 PE=2 SV=2	151.511	463741
P01898	H-2 class I histocompatibility antigen_ Q10 alpha chain OS=Mus musculus GN=H2-Q10 PE=1 SV=3	149.179	68983
Q8VCT4	Carboxylesterase 1D OS=Mus musculus GN=Ces1d PE=1 SV=1	147.382	105092
Q9QXC1	Fetuin-B OS=Mus musculus GN=Fetub PE=1 SV=1	140.098	62587
A6X935	Inter alpha-trypsin inhibitor_ heavy chain 4 OS=Mus musculus GN=Itih4 PE=1 SV=2	136.573	200929
Q7TMY4	THO complex subunit 7 homolog OS=Mus musculus GN=Thoc7 PE=1 SV=2	136.109	94309

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other glycoprotein contaminants are highlighted in red fonts

Table 6.4 Filtered and simplified proteomics hit list for the second batch of TSLP cryoprecipitates

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
P01872	Ig mu chain C region OS=Mus musculus GN=Ighm PE=1 SV=2	4116.517	2102348
P07724	Serum albumin OS=Mus musculus GN=Alb PE=1 SV=3	2835.722	6960735
Q92I1I	Serotransferrin OS=Mus musculus GN=Tf PE=1 SV=1	1142.25	661035
Q00623	Apolipoprotein A-I OS=Mus musculus GN=Apoa1 PE=1 SV=2	649.577	223787
P07759	Serine protease inhibitor A3K OS=Mus musculus GN=Serpina3k PE=1 SV=2	637.577	76491
P02088	Hemoglobin subunit beta-1 OS=Mus musculus GN=Hbb-b1 PE=1 SV=2	591.460	215960
P06336	Ig epsilon chain C region OS=Mus musculus PE=4 SV=2	303.482	201196
P01867	Ig gamma-2B chain C region OS=Mus musculus GN=Igh-3 PE=1 SV=3	266.783	72863
Q91X72	Hemopexin OS=Mus musculus GN=Hpx PE=1 SV=2	263.504	345236
Q61838	Alpha-2-macroglobulin OS=Mus musculus GN=A2m PE=1 SV=3	251.292	403719
P01027	Complement C3 OS=Mus musculus GN=C3 PE=1 SV=3	188.557	395643
P06728	Apolipoprotein A-IV OS=Mus musculus GN=Apoa4 PE=1 SV=3	148.086	46524
P01942	Hemoglobin subunit alpha OS=Mus musculus GN=Hba PE=1 SV=2	143.107	35749
P01868	Ig gamma-1 chain C region secreted form OS=Mus musculus GN=Ighg1 PE=1 SV=1	121.568	43668
P01869	Ig gamma-1 chain C region membrane-bound form OS=Mus musculus GN=Ighg1 PE=1 SV=2	121.568	43668
Q5FW57	Glycine N-acyltransferase-like protein OS=Mus musculus GN=Gm4952 PE=1 SV=3	115.899	139368

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other glycoprotein contaminants are highlighted in red fonts

Table 6.5 Filtered and simplified proteomics hit list for the third batch of TSLP cryoprecipitates

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
P02088	Hemoglobin subunit beta-1 OS=Mus musculus GN=Hbb-b1 PE=1 SV=2	16033.95	2964519
P01942	Hemoglobin subunit alpha OS=Mus musculus GN=Hba PE=1 SV=2	8678.739	3402022
P02089	Hemoglobin subunit beta-2 OS=Mus musculus GN=Hbb-b2 PE=1 SV=2	8134.476	1883105
P07724	Serum albumin OS=Mus musculus GN=Alb PE=1 SV=3	6566.179	1.34E+7
P01872	Ig mu chain C region OS=Mus musculus GN=Ighm PE=1 SV=2	6535.993	4518571
P02104	Hemoglobin subunit epsilon-Y2 OS=Mus musculus GN=Hbb-y PE=1 SV=2	4282.938	1288346
Q00623	Apolipoprotein A-I OS=Mus musculus GN=Apoa1 PE=1 SV=2	3473.553	600370
P01837	Ig kappa chain C region OS=Mus musculus PE=1 SV=1	2460.063	482082
P01592	Immunoglobulin J chain OS=Mus musculus GN=Jchain PE=1 SV=4	2137.985	164238
Q921I1	Serotransferrin OS=Mus musculus GN=Tf PE=1 SV=1	1311.009	650212
P01665	Ig kappa chain V-III region PC 7043 OS=Mus musculus PE=1 SV=1	1076.698	31507
P01638	Ig kappa chain V-V region L6 (Fragment) OS=Mus musculus PE=4 SV=1	1072.916	73379
P06330	Ig heavy chain V region AC38 205.12 OS=Mus musculus PE=1 SV=1	988.492	81031
Q61646	Haptoglobin OS=Mus musculus GN=Hp PE=1 SV=1	985.019	210914
P00761	Trypsin OS=Sus scrofa PE=1 SV=1	873.658	42699
P01756	Ig heavy chain V region MOPC 104E OS=Mus musculus PE=1 SV=1	832.258	50424
P01654	Ig kappa chain V-III region PC 2880/PC 1229 OS=Mus musculus PE=1 SV=1	786.432	29668
P01672	Ig kappa chain V-III region PC 7940 OS=Mus musculus PE=1 SV=1	716.687	23213
Q91X72	Hemopexin OS=Mus musculus GN=Hpx PE=1 SV=2	684.060	329182
Q61838	Pregnancy zone protein OS=Mus musculus GN=Pzp PE=1 SV=3	596.133	774841
P01843	Ig lambda-1 chain C region OS=Mus musculus PE=1 SV=1	587.677	36525
P01798	Ig heavy chain V-III region E109 OS=Mus musculus PE=1 SV=1	508.701	17926

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other glyco-protein contaminants are highlighted in red fonts

Table 6.5 Filtered and simplified proteomics hit list for the third batch of TSLP cryoprecipitates (continued 1)

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
Q60590	Alpha-1-acid glycoprotein 1 OS=Mus musculus GN=Orm1 PE=1 SV=1	488.477	50771
P18528	Ig heavy chain V region 6.96 OS=Mus musculus PE=4 SV=1	402.424	10734
Q00897	Alpha-1-antitrypsin 1-4 OS=Mus musculus GN=Serpina1d PE=1 SV=1	391.239	135084
P01631	Ig kappa chain V-II region 26-10 OS=Mus musculus PE=1 SV=1	383.932	52121
P01727	Ig lambda-1 chain V region S43 OS=Mus musculus PE=1 SV=1	380.859	40502
P01864	Ig gamma-2A chain C region secreted form OS=Mus musculus PE=1 SV=1	375.523	73999
P01869	Ig gamma-1 chain C region membrane-bound form OS=Mus musculus GN=Ighg1 PE=1 SV=2	368.37	97974
P01868	Ig gamma-1 chain C region secreted form OS=Mus musculus GN=Ighg1 PE=1 SV=1	368.37	99347
Q00896	Alpha-1-antitrypsin 1-3 OS=Mus musculus GN=Serpina1c PE=1 SV=2	365.462	146994
P08226	Apolipoprotein E OS=Mus musculus GN=Apoe PE=1 SV=2	359.403	63678
P22599	Alpha-1-antitrypsin 1-2 OS=Mus musculus GN=Serpina1b PE=1 SV=2	349.082	151456
P60710	Actin_ cytoplasmic 1 OS=Mus musculus GN=Actb PE=1 SV=1	300.339	44705
P23953	Carboxylesterase 1C OS=Mus musculus GN=Ces1c PE=1 SV=4	287.091	95209
P01027	Complement C3 OS=Mus musculus GN=C3 PE=1 SV=3	286.144	1022329
P07361	Alpha-1-acid glycoprotein 2 OS=Mus musculus GN=Orm2 PE=1 SV=1	240.320	3951
Q00898	Alpha-1-antitrypsin 1-5 OS=Mus musculus GN=Serpina1e PE=1 SV=1	237.749	69864
O08677	Kininogen-1 OS=Mus musculus GN=Kng1 PE=1 SV=1	218.196	79809
P07759	Serine protease inhibitor A3K OS=Mus musculus GN=Serpina3k PE=1 SV=2	224.915	39901
P26645	Myristoylated alanine-rich C-kinase substrate OS=Mus musculus GN=Marcks PE=1 SV=2	188.427	19091
P21614	Vitamin D-binding protein OS=Mus musculus GN=Gc PE=1 SV=2	177.624	2421536
Q06890	Clusterin OS=Mus musculus GN=Clu PE=1 SV=1	156.101	24410
P01865	Ig gamma-2A chain C region membrane-bound form OS=Mus musculus GN=Igh-1a PE=1 SV=3	137.262	30339

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other Ig contaminants are highlighted in red fonts

Table 6.5 Filtered and simplified Proteomics hit list for the third batch of TSLP cryoprecipitates (continued 2)

Protein Accession	Protein Description	Protein Score	Protein Matched Peptide Intensity Sum
P01863	Ig gamma-2A chain C region_ A allele OS=Mus musculus GN=Ighg PE=1 SV=1	137.262	37346
Q03734	Serine protease inhibitor A3M OS=Mus musculus GN=Serpina3m PE=1 SV=2	119.420	20877
P01729	Ig lambda-2 chain V region MOPC 315 OS=Mus musculus PE=1 SV=1	116.784	28705
P01878	Ig alpha chain C region OS=Mus musculus PE=1 SV=1	111.750	19019
Q91WP6	Serine protease inhibitor A3N OS=Mus musculus GN=Serpina3n PE=1 SV=1	109.296	16055
Q7TMY4	THO complex subunit 7 homolog OS=Mus musculus GN=Thoc7 PE=1 SV=2	85.623	17060
P01867	Ig gamma-2B chain C region OS=Mus musculus GN=Igh-3 PE=1 SV=3	84.376	19113
Q8BID6	Zinc finger and BTB domain-containing protein 46 OS=Mus musculus GN=Zbtb46 PE=1 SV=2	54.791	6139
Q61147	Ceruloplasmin OS=Mus musculus GN=Cp PE=1 SV=2	52.224	40271
P68033	Actin_ alpha cardiac muscle 1 OS=Mus musculus GN=Actc1 PE=1 SV=1	47.394	8914

Hits for IgM and IgG are underlined. Only one hit for isoforms of the other sequences with the same score is preserved in list. Transferrin and other glycoprotein contaminants are highlighted in red fonts

6.3 Discussion

In this chapter we characterised the N-glycan repertoire of IgG and cryoprecipitate materials from the TSLP transgenic mouse model (**Section 6.2.2** and **6.2.5**) and verified their protein compositions by proteomics (**Section 6.2.4** and **6.2.7**). These results exhibit the complexity of this type of cryoprecipitation.

The glycomics study of murine IgGs illustrated a relatively similar glycome between the materials from the TSLP overexpressing mouse model and wide-type control, with a minor semi-quantitative difference of sialylation. Both glycomes are dominated by bi-antennary core-fucosylated N-glycans with varied galactosylation and sialyaltion. Proteomic studies reveal that impurities exist in both IgG materials. The WT IgG material has a relatively less severe contamination as its major contaminants were not reported to have N-glycosylation, therefore there is a minor impact of the characterised glycome for WT IgG. Instead, TSLP IgG material has a minor content of IgM and its joint J chain contaminants. These two Ig contaminants can be N-glycosylated with heterogeneous structures ranging from high mannose to complex type. Their actual contributions to the abundance of structures presented in the glycome of TSLP IgG are difficult to evaluate. Thereby, it is challenging to make an accurate semi-quantitative comparison of glycan features such as sialylation between the actual TSLP and WT IgGs. In addition, even the TSLP mouse model was suggested to circulate a high load of cryoglobulins in its body, it would still have an extent of immunoglobulins which are not cryo-precipitative (Taneda et al., 2001). The recruited approach to harvest IgGs from murine sera is not specifically designed to separate cryogenic and non-cryogenic antibodies. Thus, it is likely that the glycoforms of both groups of antibodies were mixed in the TSLP IgG and then the significance of targeted alteration on glycans might be diluted.

The study of cryoprecipitates exhibits more diverse results. As the glycomics analysis was directly performed on cryoprecipitates without pre-purification, the results of glycomic profiling show a collection of polyclonal IgM and IgG glycans from type III mixed cryoglobulins with possibly some degree of contamination from serum proteins. Inconsistent glyco-profiles were recovered from three independent batches of TSLP cryoprecipitates. Sialylation is dominant in 2 of 3 batches, whereas only trace level of sialylation was found in the third batch of cryoprecipitates. Moreover, fragmentation analysis for selected glycans from the highly sialylated batches of cryoprecipitates verified the presence of di-sialylated LacNAc antennae which are a typical glycan moiety of serotransferrin N-glycans. Subsequent proteomics analysis confirmed the presence of serotransferrin in all cryoprecipitate materials with varied abundance which generally corresponds to the observed levels of sialylation in each batch of cryoprecipitates. IgM which is highly ranked in the proteomics list might mainly contribute to the dominant high mannose, hybrid and non-sialylated complex type of glycans as observed in the third batch of cryoprecipitate with the relatively lowest levels of serotransferrin and IgG. There are few

glycomic studies of murine IgM glycosylation via various technologies (Anderson et al., 1985, Leibiger et al., 1998, Wang et al., 2003) and they have a consistent observation for a complicated N-glycan repertoire ranging from high mannose to complex type. But they had various suggestions for the exact glycoforms of murine IgM due to the difference of employed cell lines for IgM production. There was a report to suggest that heterogeneity of glycosylation did exist between murine IgM's and even another report suggested that glycosylation of a human IgM produced in a hybridoma cell was different from that at its parental cells (Fukuta et al., 2000, Wright et al., 1990). Hence, it is necessary to validate the major N-glycosylation patterns of TSLP IgM, and this might be helpful to clarify the contribution of IgM glycans to the observed glycomes of cryo-precipitates. In addition, the contribution of IgG for the glyco-profile might be indistinguishable in the analysed cryoprecipitates as it has the lowest protein abundance in the compared glycoproteins and a glycoform that is likely to partly overlapped with the combined glycomes of serotransferrin and IgM. There is still an approach to clarify the glycomes of IgG in these cryo-precipiates. A glycoproteomic study would allow site-specific identification of the glycan structures and linking peptide backbones, simultaneously. Due to the complexity of the cryo-precipiate composition, an enrichment of IgM or IgG should be undertaken before the glycoproteomics experiments by methods such as electrophoresis and affinity chromatography.

Based on our mass spectrometric data, we would suggest that the analysed cryoprecipiates of type III cryoglobulins produce a complicated N-glycan repertoire from a mixture of serum components. The complex itself contains the suggested IgM and IgG components, with IgM in predominance in protein composition. Other serum components such as serotransferrin and other immunoglobulins possibly associated with the suggested IgM and IgG complex in a relatively random manner. This random association could be the result of a physiochemical restriction or RF activities of various Ig proteins during the precipitation process. Therefore, the complex N-glycan repertoire might not directly reflect its biological relevance to the proposed modulatory functions of IgG glycans for the cryogenic activities of cryoglobulins. However, the research could still be continued on this avenue, because this study was mainly restricted to the challenges of recovering optimal Ig materials. Procedures to purify the cryoprecipitates with affinity chromatography and then re-perform the cryo-precipitation on the purified materials for obtaining relatively high degree of cryogenic antibodies would be informative. However, obtaining enough experimental material from the TSLP transgenic mice is challenging, and heterogeneity might be introduced by a large pool of glycans from various mice. Therefore, an optimisation of the sampling strategy would be in a great need to improve the recovery rate and specific separation of cryogenic antibodies economically.

After incorporating the glycomics and proteomics results of TSLP cryoglobulins and IgGs, it is hard to suggest a structural feature or change of N-glycosylation that correlates to the cryogenic activities of the mixed cryoglobulins from TSLP mice. This inconclusive suggestion might reflect the differences of this work to the referred IgG3 studies. The TSLP cryoglobulins are mainly consist of IgM, IgG1 and IgG2. They are unable to self-associate as IgG3 does, so that their Ig-association and even cryogenic

activities could mainly be based on their RF activities to form immune complexes (Kolopp-Sarda and Miossec, 2018, Abdelmoula et al., 1989). Murine IgG3 exclusively has two extra N-glycosylation sites in its sequence, and one site located at the CH3 domain of IgG3 was studied for its involvement in cryogenic activities (Plummer et al., 1984, Panka, 1997, Kuroki et al., 2002). There have been contradictory suggestions for the involvement of this site (Panka, 1997, Kuroki et al., 2002), but this controversy cannot exclude the possibility, because these two studies did not perform a consistent mutation on the murine IgG3 sequence. The influence of different amino acid choice for mutation could have cause the controversial observations. Moreover, as previous studies shown that additional N-glycosylation sites on protein sequence could influence the overall N-glycosylation of glycoproteins (Blundell et al., 2017, Picanco-Castro and Swiech, 2018) resulting in hyper-glycosylation. Thus, it is likely that the impacts of the two extra N-glycosylation sites of IgG3 on its N-glycosylation and cryogenic activity might be underevaluated in previous mutagenesis studies (Panka, 1997; Kuroki et al., 2002). However, this potential influence of glycosylation is very unlikely to have on TSLP cryoglobulins. Besides, as a monoclonal cryoglobulin, the sequence of the studied IgG3 has been well characterised (Plummer et al., 1984), and even some positively charged residues on its heavy chain variable regions were determined as a secondary factor to influence its cryogenic activities (Panka et al., 1995, Kuroki et al., 2002). On the contrary, several early studies suggested that monoclonal cryoglobulins were more negatively charged (Erickson et al., 1982, Gerber-Jenson et al., 1981, Middaugh and Litman, 1978) and there have been no relevant studies for the polyclonal cryoglobulins from the TSLP mouse model. Thus, detailed sequencing studies are necessary to clarify the contribution of amino acid sequence to the cryogenic activities of polyclonal cryoglobulin.

Concluding Remarks

Concluding Remark

In this thesis, the focus of study is on the mass spectrometry-based structural elucidation of glycans from three different systems: PSG1, IgG (Fc) and canine melanoma, which are all immunologically important. They modulate or facilitate immune systems to induce or suppress immune responses for particular objectives .The obtained structural information of their glycans is assisting investigations of their roles in immunity. Their correlation with immune system mechanisms could be invaluable to guide medical design to facilitate or inhibit certain functionalities to cure patients from challenging diseases by patient's own immunity. In addition, the observed macro- and/or micro-heterogeneity of these three glycosylation systems also highlights the importance of mass spectrometric strategies for studying such heterogenous glycans and even larger glycoconjugates.

In the PSG1 project (chapter 3), I carried out glycomics and glycoproteomics studies for the structural definition of glycans as well as their site-occupancy on the sequence of PSG1. The results of native PSG1 N-glycosylation illustrated a quantitatively well-regulated core-fucosylation as well as a site-specific branching on N-glycan structures. The characterised N-glycome indicated potential reactivities between PSG1 and immunomodulatory lectins such as galectins and siglecs. These potential reactivities might be associated with the unsolved interacting mechanism of PSG1 to a range of immune cells, including macrophages, to trigger cytokine-dependent immunomodulation. In addition, the distribution of N-glycans on this protein sequence has been revealed by the glycoproteomics study. This information could be helpful for future studies to refine quaternary structures of this protein to predict the possible interaction of its molecular motifs to potential receptors. It is also important for future studies to investigate the mechanism of site-specific N-glycosylation as well as related functional studies. This is because the specific site-occupancy on PSG1 N-domain by N-glycans are shown by our results and our collaborator's functional studies to be able to interfere with protein-protein interactions between PSG1 and integrins. Therefore, this aspect of macro-heterogeneity in glycosylation could be an interesting direction to investigate for understanding the mechanism to manipulate the functionality of a protein by site-specific glycosylation. The importance of combining glycomics and glycoproteomics studies in order to identify the structural determinants of glycan related interactions is also highlighted by this work.

In the canine melanoma project (Chapter 4), we characterised the N-, O- and glycolipid glycomes of a stage III melanoma that was surgically removed from a dog participating in a cancer vaccine clinical trial at the University of Missouri. The glycans of canine melanoma generally share the typical structural features such as core-fucosylation, sialylation and linear poly-LacNAc extension with its

human counterparts. This indicates that it could be an excellent model for pre-clinical studies of melanoma from a glycobiological viewpoint. The consistent expression of oncofoetal antigens such as linear poly-LacNAc shows suggested connections between the reproductive system and cancer by the Hu-FEDs hypothesis. Based on knowledge of human and murine immune systems, the presence of shared gangliosides would benefit cancer progression through manipulating the cytotoxic activities of NK cells via a ligation of ganglioside and inhibitory siglec-7 receptor (Yamaji et al., 2002). More recently, the signalling of murine siglec-G was shown to be able to impair the cross-presentation of the MHC I complex on dendritic cells (Ding et al., 2016). The subsequent phosphorylation of NOX2 by SHP-1 triggered by siglec-G receptor could result in an excessive hydrolysis of exogenous antigens in DCs phagosomes. Consequently, the formation of MHC I complexes was restricted, and the severity of this effect was associated with the expression level of siglec-G on DCs. Thus, it is likely that tumour cells might recruit this approach to minimise the induction of cross-presentation of their antigens to the immune system and successive CTL responses. However, whether these mechanisms could still be viable in the canine immune system is still largely unknown, because the functional study of this animal model has lagged far behind. Future studies of lectin expression and distribution, in particular galectins and siglecs, in the tumour microenvironment of the canine system would help to get a deeper understanding of glycan interactions with immune cells as well as melanoma cells. Moreover, thorough studies of cell signalling triggered by the lectin-glycan interactions on these cells in the canine system would give invaluable information to critically evaluate the role of glycosylation on the development of melanoma as well as the responding mechanism of immune system to them in this canine system. More importantly, the comparability of canine melanoma to human melanoma and correpsonding immune systems could then be fully understood from a glycobiological viewpoint; and the developing glyco-immune strategy would subsequently be optimised to treat human melanoma in future.

In the Fc fusion protein project (chapter 5), we analysed the N-glycosylation of engineered Fc fusion proteins from CHO and HEK293 cell systems. Manipulation of N-glycosylation site and polymerisation status on the Fc fusion proteins was shown to be able to improve the overall sialylation level and lectin interactions on glycans for potential treatments of autoimmune disease. In the CHO cell system, hyper-sialylation could be observed on Fc fusion proteins and monomeric fusion proteins had a superior binding affinity to lectins, while their polymeric counterparts were preferable for Fc receptors. Branching and elongation of N-glycan structures could also be observed on these Fc proteins, and this might implicate their interactions with galectins to engage in a wider range of immunological events. However, in the HEK293 cell system, a wider range of truncated or LacdiNAc related epitopes were found on the N-glycans. This reflects the differences of these two expression systems in terms of glycosylation machinery, and the characterised difference could help to choose the optimal cell lines for Fc fusion protein production to pursue certain structural features. This engineering approach could be a new avenue to study and optimise the N-glycosylation on Fc fusion proteins to regulate the

relevant biological functions, especially its immunomodulatory effects. It is also an ideal alternative to the original intravenous immunologulin treatment for suppressing auto-immune responses because of lower cost in production and more reliable quality control for glycosylation and subsequent clinical effects.

In the cryoglobulin project (Chapter 6), I aimed to rigorously characterise the N-glycosylation of mixed type of polyclonal cryoglobulins. However, practical difficulties in sample preparation led to a challenge to extract valid information from the obtained results. I carried out direct glycomics and proteomics studies for the cryoprecipitates from murine serum. The diverse N-glycan profiles of cryoprecipitates indicated a possible heterogeneity between specimens, and proteomics sequencing verified the complexity of composition in these cryoprecipitates as accommodating a substantial population of other serum (glyco)proteins. Hence, the actual extent and features of N-glycosylation of cryoglobulins become difficult to be evaluated. Besides, the protein-level structural information of this type of cryoglobulins has not been well-defined. It is also necessary to comprehensively characterise the biochemical properties of this type of cryoglobulins from protein to post-translational levels .

The mass spectrometry strategy is invaluable in these studies of glycans and glycoconjugates. The MALDI-TOF/TOF based characterisation is powerful and sensitive to enable a combined MS and MS/MS analysis of glycans from merely a few micrograms of protein materials. Glycan terminal moieties and antennary arrangement of LacNAc units can generally be characterised, and further glycosidic linkages of monosaccharide can be studied with linkage specific enzymatic digestion and GC-MS. The structural information obtained via the employed glycomics strategies should be able to facilitate unambiguous definition of glycan structures from mammalian glycoproteins and glycolipids. There is still a challenge in identifying the sulphation on glycans via permethylation based mass spectrometry analysis. To preserve the thermally labile sulphate group, a bespoke permethylation in cold environment was tried. But the results might not be able to define the presence of low levels of sulphation on glycans. Alternatively, as the development of hydrophilic interaction liquid chromatography (HILIC), a LC based pre-screening could be employed to verify the presence of sulphation prior to mass spectrometric analysis. This chromatography was shown to be able to separate glycan isomers on peptides (Badgett et al., 2018), and it is logical that the sulphated glycans could be distinctively separated based on their enhanced negative charge and unique composition of sulphate groups to the other glycans. The glycomics studies could be specifically carried out on an enriched population of sulphated glycans via an off-line LC system. Then the subsequent glycomics procedure for sulphated glycans can be confidently optimised and conducted for structural elucidation.

The outstanding sensitivity of the Q-STAR instrument enabled the glycoproteomic study of PSG1, and the results facilitated the identification of site-occupation and site-specific glycoforms. This information is invaluable to understand the glycosylation machinery and the potential functionality of

glycans on specific sites. As a reversed-phase column was incorporated with the instrument, the retention of peptides without glycosylation could lead to a suppression of detection for glycopeptides. In future work, a HILIC column could be incorporated for enrichment of glycopeptides, possibly gaining a more targeting mass spectrometric analysis for glycans and even for the sulphated glycan species.

Overall, in this thesis, I characterised three glycosylation systems. Information arising from my work is being used by our collaborators in their ongoing functional studies. Hopefully some of these studies could be transferred into clinical application in the near future. Besides, the mass spectrometric strategies could even be further developed with more advanced instrumentation as well as improved chromatographic separation technologies.

References

References

Abdelmoula, M., Spertini, F., Shibata, T., Gyotoku, Y., Luzuy, S., Lambert, P. H. & Izui, S. 1989. Igg3 Is The Major Source Of Cryoglobulins In Mice. J Immunol, 143, 526-32.

Aigner, S., Sthoeger, Z. M., Fogel, M., Weber, E., Zarn, J., Ruppert, M., Zeller, Y., Vestweber, D., Stahel, R., Sammar, M. & Altevogt, P. 1997. Cd24, A Mucin-Type Glycoprotein, Is A Ligand For P-Selectin On Human Tumor Cells. Blood, 89, 3385-95.

Albertolle, M. E., Hassis, M. E., Ng, C. J., Cuison, S., Williams, K., Prakobphol, A., Dykstra, A. B., Hall, S. C., Niles, R. K., Ewa Witkowska, H. & Fisher, S. J. 2015. Mass Spectrometry-Based Analyses Showing The Effects Of Secretor And Blood Group Status On Salivary N-Glycosylation. Clin Proteomics, 12, 29.

Alborzian Deh Sheikh, A., Akatsu, C., Imamura, A., Abdu-Allah, H. H. M., Takematsu, H., Ando, H., Ishida, H. & Tsubata, T. 2018. Proximity Labeling Of Cis-Ligands Of Cd22/Siglec-2 Reveals Stepwise Alpha2,6 Sialic Acid-Dependent And -Independent Interactions. Biochem Biophys Res Commun, 495, 854-859.

Alexander, A. N., Huelsmeyer, M. K., Mitzey, A., Dubielzig, R. R., Kurzman, I. D., Macewen, E. G. & Vail, D. M. 2006. Development Of An Allogeneic Whole-Cell Tumor Vaccine Expressing Xenogeneic Gp100 And Its Implementation In A Phase Ii Clinical Trial In Canine Patients With Malignant Melanoma. Cancer Immunol Immunother, 55, 433-42.

Amano, M., Galvan, M., He, J. & Baum, L. G. 2003. The St6gal I Sialyltransferase Selectively Modifies N-Glycans On Cd45 To Negatively Regulate Galectin-1-Induced Cd45 Clustering, Phosphatase Modulation, And T Cell Death. J Biol Chem, 278, 7469-75.

Anderson, D. R., Atkinson, P. H. & Grimes, W. J. 1985. Major Carbohydrate Structures At Five Glycosylation Sites On Murine Igm Determined By High Resolution 1H-Nmr Spectroscopy. Arch Biochem Biophys, 243, 605-18.

Anguizola, J., Matsuda, R., Barnaby, O. S., Hoy, K. S., Wa, C., Debolt, E., Koke, M. & Hage, D. S. 2013. Review: Glycation Of Human Serum Albumin. Clin Chim Acta, 425, 64-76.

Anthony, R. M., Kobayashi, T., Wermeling, F. & Ravetch, J. V. 2011. Intravenous Gammaglobulin Suppresses Inflammation Through A Novel T(H)2 Pathway. Nature, 475, 110-3.

Anthony, R. M., Wermeling, F. & Ravetch, J. V. 2012. Novel Roles For The Igg Fc Glycan. Ann N Y Acad Sci, 1253, 170-80.

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. Biochim Biophys Acta, 1473, 4-8.

Arnold, J. N., Wormald, M. R., Suter, D. M., Radcliffe, C. M., Harvey, D. J., Dwek, R. A., Rudd, P. M. & Sim, R. B. 2005. Human Serum Igm Glycosylation: Identification Of Glycoforms That Can Bind To Mannan-Binding

Lectin. J Biol Chem, 280, 29080-7.

Aruffo, A., Kolanus, W., Walz, G., Fredman, P. & Seed, B. 1991. Cd62/P-Selectin Recognition Of Myeloid And Tumor Cell Sulfatides. Cell, 67, 35-44.

Astrakhan, A., Omori, M., Nguyen, T., Becker-Herman, S., Iseki, M., Aye, T., Hudkins, K., Dooley, J., Farr, A., Alpers, C. E., Ziegler, S. F. & Rawlings, D. J. 2007. Local Increase In Thymic Stromal Lymphopoietin Induces Systemic Alterations In B Cell Development. Nature Immunology, 8, 522.

Auscher, C. & Guinand, S. 1964. [Investigation Of A Cryo-Precipitable Beta2a-Globulin]. Clin Chim Acta, 9, 40-8.

Avril, T., Floyd, H., Lopez, F., Vivier, E. & Crocker, P. R. 2004. The Membrane-Proximal Immunoreceptor Tyrosine-Based Inhibitory Motif Is Critical For The Inhibitory Signaling Mediated By Siglecs-7 And -9, Cd33-Related Siglecs Expressed On Human Monocytes And Nk Cells. The Journal Of Immunology, 173, 6841.

Badgett, M. J., Mize, E., Fletcher, T., Boyes, B. & Orlando, R. 2018. Predicting The Hilic Retention Behavior Of The N-Linked Glycopeptides Produced By Trypsin Digestion Of Immunoglobulin Gs (Iggs). Journal Of Biomolecular Techniques : Jbt, 29, 98-104.

Baldus, S. E., Zirbes, T. K., Monig, S. P., Engel, S., Monaca, E., Rafiqpoor, K., Hanisch, F. G., Hanski, C., Thiele, J., Pichlmaier, H. & Dienes, H. P. 1998. Histopathological Subtypes And Prognosis Of Gastric Cancer Are Correlated With The Expression Of Mucin-Associated Sialylated Antigens: Sialosyl-Lewis(A), Sialosyl-Lewis(X) And Sialosyl-Tn. Tumour Biol, 19, 445-53.

Ballesteros, A., Mentink-Kane, M. M., Warren, J., Kaplan, G. G. & Dveksler, G. S. 2015. Induction And Activation Of Latent Transforming Growth Factor-Beta1 Are Carried Out By Two Distinct Domains Of Pregnancy-Specific Glycoprotein 1 (Psg1). J Biol Chem, 290, 4422-31.

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.

Bast Jr, R. C., Feeney, M., Lazarus, H., Nadler, L. M., Colvin, R. B. & Knapp, R. C. 1981. Reactivity Of A Monoclonal Antibody With Human Ovarian Carcinoma. Journal Of Clinical Investigation, 68, 1331-1337.

Baum, L. G., Derbin, K., Perillo, N. L., Wu, T., Pang, M. & Uittenbogaart, C. 1996. Characterization Of Terminal Sialic Acid Linkages On Human Thymocytes. Correlation Between Lectin-Binding Phenotype And Sialyltransferase Expression. J Biol Chem, 271, 10793-9.

Baum, L. G., Pang, M., Perillo, N. L., Wu, T., Delegeane, A., Uittenbogaart, C. H., Fukuda, M. & Seilhamer, J. J. 1995. Human Thymic Epithelial Cells Express An Endogenous Lectin, Galectin-1, Which Binds To Core 2 O-Glycans On Thymocytes And T Lymphoblastoid Cells. J Exp Med, 181, 877-87.

Baumheter, S., Singer, M. S., Henzel, W., Hemmerich, S., Renz, M., Rosen, S. D. & Lasky, L. A. 1993. Binding Of L-Selectin To The Vascular Sialomucin Cd34. Science, 262, 436-8.

Bause, E., Hettkamp, H. & Legler, G. 1982. Conformational Aspects Of N-Glycosylation Of Proteins. Studies With Linear And Cyclic Peptides As Probes. Biochem J, 203, 761-8.

Belisle, J. A., Horibata, S., Jennifer, G. A. A., Petrie, S., Kapur, A., André, S., Gabius, H.-J., Rancourt, C., Connor, J., Paulson, J. C. & Patankar, M. S. 2010. Identification Of Siglec-9 As The Receptor For Muc16 On Human Nk Cells, B Cells, And Monocytes. Molecular Cancer, 9, 118.

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.

Berg, E. L., Mcevoy, L. M., Berlin, C., Bargatze, R. F. & Butcher, E. C. 1993. L-Selectin-Mediated Lymphocyte Rolling On Madcam-1. Nature, 366, 695-8.

Berg, E. L., Robinson, M. K., Mansson, O., Butcher, E. C. & Magnani, J. L. 1991. A Carbohydrate Domain Common To Both Sialyl Le(A) And Sialyl Le(X) Is Recognized By The Endothelial Cell Leukocyte Adhesion Molecule Elam-1. J Biol Chem, 266, 14869-72.

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.

Berney, T., Fulpius, T., Shibata, T., Reininger, L., Van Snick, J., Shan, H., Weigert, M., Marshak-Rothstein, A. & Izui, S. 1992. Selective Pathogenicity Of Murine Rheumatoid Factors Of The Cryoprecipitable Igg3 Subclass. Int Immunol, 4, 93-9.

Billingham, R. E., Brent, L. & Medawar, P. B. 1956. The Antigenic Stimulus In Transplantation Immunity. Nature, 178, 514-519.

Blanas, A., Sahasrabudhe, N. M., Rodriguez, E., Van Kooyk, Y. & Van Vliet, S. J. 2018. Corrigendum: Fucosylated Antigens In Cancer: An Alliance Toward Tumor Progression, Metastasis, And Resistance To Chemotherapy. Front Oncol, 8, 150.

Blixt, O., Collins, B. E., Van Den Nieuwenhof, I. M., Crocker, P. R. & Paulson, J. C. 2003. Sialoside Specificity Of The Siglec Family Assessed Using Novel Multivalent Probes - Identification Of Potent Inhibitors Of Myelin-Associated Glycoprotein. Journal Of Biological Chemistry, 278, 31007-31019.

Blois, S. M. & Barrientos, G. 2014. Galectin Signature In Normal Pregnancy And Preeclampsia. J Reprod Immunol, 101-102, 127-134.

Blois, S. M., Sulkowski, G., Tirado-Gonzalez, I., Warren, J., Freitag, N., Klapp, B. F., Rifkin, D., Fuss, I., Strober, W. & Dveksler, G. S. 2014. Pregnancy-Specific Glycoprotein 1 (Psg1) Activates Tgf-Beta And Prevents Dextran Sodium Sulfate (Dss)-Induced Colitis In Mice. Mucosal Immunol, 7, 348-58.

Blomme, B., Van Steenkiste, C., Grassi, P., Haslam, S. M., Dell, A., Callewaert, N. & Van Vlierberghe, H. 2011. Alterations Of Serum Protein N-Glycosylation In Two Mouse Models Of Chronic Liver Disease Are Hepatocyte And Not B Cell Driven. Am J Physiol Gastrointest Liver Physiol, 300, G833-42.

Blow, N. 2009. Glycobiology: A Spoonful Of Sugar. Nature, 457, 617-20.

Blundell, P. A., Le, N. P. L., Allen, J., Watanabe, Y. & Pleass, R. J. 2017. Engineering The Fragment Crystallizable (Fc) Region Of Human Igg1 Multimers And Monomers To Fine-Tune Interactions With Sialic Acid-Dependent Receptors. J Biol Chem, 292, 12994-13007.

Blundell, P. A., Lu, D., Wilkinson, M., Dell, A., Haslam, S. & Pleass, R. J. 2019. Insertion Of N-Terminal Hinge Glycosylation Enhances Interactions Of The Fc Region Of Human Igg1 Monomers With Glycan-Dependent Receptors And Blocks Hemagglutination By The Influenza Virus. J Immunol, 202, 1595-1611.

Bogoevska, V., Nollau, P., Lucka, L., Grunow, D., Klampe, B., Uotila, L. M., Samsen, A., Gahmberg, C. G. &

Wagener, C. 2007. Dc-Sign Binds Icam-3 Isolated From Peripheral Human Leukocytes Through Lewis X Residues. Glycobiology, 17, 324-33.

Bohm, E., Seyfried, B. K., Dockal, M., Graninger, M., Hasslacher, M., Neurath, M., Konetschny, C., Matthiessen, P., Mitterer, A. & Scheiflinger, F. 2015. Differences In N-Glycosylation Of Recombinant Human Coagulation Factor Vii Derived From Bhk, Cho, And Hek293 Cells. Bmc Biotechnol, 15, 87.

Bohn, H., Kraus, W. & Winckler, W. 1983. Purification And Characterization Of Two New Soluble Placental Tissue Proteins (Pp13 And Pp17). Oncodev Biol Med, 4, 343-50.

Bojic-Trbojevic, Z., Jovanovic Krivokuca, M., Stefanoska, I., Kolundzic, N., Vilotic, A., Kadoya, T. & Vicovac, L. 2018. Integrin Beta1 Is Bound To Galectin-1 In Human Trophoblast. J Biochem, 163, 39-50.

Bolton, A. E., Clough, K. J., Stoker, R. J., Pockley, A. G., Mowles, E. A., Westwood, O. M. R. & Chapman, M. G. 1987. Identification Of Placental Protein 14 As An Immunosuppressive Factor In Human Reproduction. The Lancet, 329, 593-595.

Bondt, A., Rombouts, Y., Selman, M. H., Hensbergen, P. J., Reiding, K. R., Hazes, J. M., Dolhain, R. J. & Wuhrer, M. 2014. Immunoglobulin G (Igg) Fab Glycosylation Analysis Using A New Mass Spectrometric High-Throughput Profiling Method Reveals Pregnancy-Associated Changes. Mol Cell Proteomics, 13, 3029-39.

Bondt, A., Wuhrer, M., Kuijper, T. M., Hazes, J. M. & Dolhain, R. J. 2016. Fab Glycosylation Of Immunoglobulin G Does Not Associate With Improvement Of Rheumatoid Arthritis During Pregnancy. Arthritis Res Ther, 18, 274.

Braun, G. S., Horster, S., Wagner, K. S., Ihrler, S. & Schmid, H. 2007. Cryoglobulinaemic Vasculitis: Classification And Clinical And Therapeutic Aspects. Postgrad Med J, 83, 87-94.

Brockhausen, I. 1999. Pathways Of O-Glycan Biosynthesis In Cancer Cells. Biochim Biophys Acta, 1473, 67-95.

Brockhausen, I. & Stanley, P. 2015. O-Galnac Glycans. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Brouet, J. C., Clauvel, J. P., Danon, F., Klein, M. & Seligmann, M. 1974. Biologic And Clinical Significance Of Cryoglobulins. A Report Of 86 Cases. Am J Med, 57, 775-88.

Bruhns, P., Iannascoli, B., England, P., Mancardi, D. A., Fernandez, N., Jorieux, S. & Daeron, M. 2009. Specificity And Affinity Of Human Fcgamma Receptors And Their Polymorphic Variants For Human Igg Subclasses. Blood, 113, 3716-25.

Burda, P. & Aebi, M. 1999. The Dolichol Pathway Of N-Linked Glycosylation. Biochim Biophys Acta, 1426, 239-57.

Burton, G. J. & Jauniaux, E. 2004. Placental Oxidative Stress: From Miscarriage To Preeclampsia. Journal Of The Society For Gynecologic Investigation, 11, 342-352.

Camolotto, S., Racca, A., Rena, V., Nores, R., Patrito, L. C., Genti-Raimondi, S. & Panzetta-Dutari, G. M. 2010. Expression And Transcriptional Regulation Of Individual Pregnancy-Specific Glycoprotein Genes In Differentiating Trophoblast Cells. Placenta, 31, 312-9.

Cancer Genome Atlas, N. 2015. Genomic Classification Of Cutaneous Melanoma. Cell, 161, 1681-96.

Canis, K., Anzengruber, J., Garenaux, E., Feichtinger, M., Benamara, K., Scheiflinger, F., Savoy, L. A., Reipert, B. M. & Malisauskas, M. 2018. In-Depth Comparison Of N-Glycosylation Of Human Plasma-Derived Factor Viii And Different Recombinant Products: From Structure To Clinical Implications. J Thromb Haemost.

Capon, D. J., Chamow, S. M., Mordenti, J., Marsters, S. A., Gregory, T., Mitsuya, H., Byrn, R. A., Lucas, C., Wurm, F. M., Groopman, J. E. & Et Al. 1989. Designing Cd4 Immunoadhesins For Aids Therapy. Nature, 337, 525-31.

Capra, J. D. & Kunkel, H. G. 1970. Aggregation Of Gamma-G3 Proteins: Relevance To The Hyperviscosity Syndrome. J Clin Invest, 49, 610-21.

Cariappa, A., Takematsu, H., Liu, H., Diaz, S., Haider, K., Boboila, C., Kalloo, G., Connole, M., Shi, H. N., Varki, N., Varki, A. & Pillai, S. 2009. B Cell Antigen Receptor Signal Strength And Peripheral B Cell Development Are Regulated By A 9-O-Acetyl Sialic Acid Esterase. J Exp Med, 206, 125-38.

Carrillo-Conde, B., Song, E. H., Chavez-Santoscoy, A., Phanse, Y., Ramer-Tait, A. E., Pohl, N. L., Wannemuehler, M. J., Bellaire, B. H. & Narasimhan, B. 2011. Mannose-Functionalized "Pathogen-Like" Polyanhydride Nanoparticles Target C-Type Lectin Receptors On Dendritic Cells. Mol Pharm, 8, 1877-86.

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.

Chan, A. L., Morris, H. R., Panico, M., Etienne, A. T., Rogers, M. E., Gaffney, P., Creighton-Kempsford, L. & Dell, A. 1991. A Novel Sialylated N-Acetylgalactosamine-Containing Oligosaccharide Is The Major Complex-Type Structure Present In Bowes Melanoma Tissue Plasminogen Activator. Glycobiology, 1, 173-85.

Chang, W. W., Yu, C. Y., Lin, T. W., Wang, P. H. & Tsai, Y. C. 2006. Soyasaponin I Decreases The Expression Of Alpha2,3-Linked Sialic Acid On The Cell Surface And Suppresses The Metastatic Potential Of B16f10 Melanoma Cells. Biochem Biophys Res Commun, 341, 614-9.

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. 2016. Evidence For Differential Glycosylation Of Trophoblast Cell Types. Mol Cell Proteomics, 15, 1857-66.

Chen, Y., Hu, D., Yabe, R., Tateno, H., Qin, S. Y., Matsumoto, N., Hirabayashi, J. & Yamamoto, K. 2011. Role Of Malectin In Glc(2)Man(9)Glcnac(2)-Dependent Quality Control Of Alpha1-Antitrypsin. Mol Biol Cell, 22, 3559-70.

Chester, M. A. 1998. Iupac-Iub Joint Commission On Biochemical Nomenclature (Jcbn). Nomenclature Of Glycolipids--Recommendations 1997. Eur J Biochem, 257, 293-8.

Choo, M., Tan, H. L., Ding, V., Castangia, R., Belgacem, O., Liau, B., Hartley-Tassell, L., Haslam, S. M., Dell, A. & Choo, A. 2017. Characterization Of H Type 1 And Type 1 N-Acetyllactosamine Glycan Epitopes On Ovarian Cancer Specifically Recognized By The Anti-Glycan Monoclonal Antibody Mab-A4. J Biol Chem, 292, 6163-6176.

Chu, S. Y., Vostiar, I., Karki, S., Moore, G. L., Lazar, G. A., Pong, E., Joyce, P. F., Szymkowski, D. E. & Desjarlais, J. R. 2008. Inhibition Of B Cell Receptor-Mediated Activation Of Primary Human B Cells By Coengagement Of Cd19 And Fcgammariib With Fc-Engineered Antibodies. Mol Immunol, 45, 3926-33.

Ciolczyk-Wierzbicka, D., Bodzioch, M., Gil, D., Zmudzinska, D., Dembinska-Kiec, A. & Laidler, P. 2007. Expression Of Fucosyltransferases Contributes To Melanoma Invasive Phenotype. Med Chem, 3, 418-24.

Clark, G. 2017. Cancer Immune-Based Therapy. United States Patent Application 20170246275.

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., Patankar, M. S., Hinsch, K. D. & Oehninger, S. 1995. New Concepts In Human Sperm-Zona Pellucida Interaction. Hum Reprod, 10 Suppl 1, 31-7.

Clark, G. F. & Schust, D. J. 2013. Manifestations Of Immune Tolerance In The Human Female Reproductive Tract. Front Immunol, 4, 26.

Cole, R. B. 2010. Electrospray And Maldi Mass Spectrometry : Fundamentals, Instrumentation, Practicalities, And Biological Applications, Hoboken, N.j., Wiley.

Collins, T., Read, M. A., Neish, A. S., Whitley, M. Z., Thanos, D. & Maniatis, T. 1995. Transcriptional Regulation Of Endothelial Cell Adhesion Molecules: Nf-Kappa B And Cytokine-Inducible Enhancers. Faseb J, 9, 899-909.

Coloma, M. J., Trinh, R. K., Martinez, A. R. & Morrison, S. L. 1999. Position Effects Of Variable Region Carbohydrate On The Affinity And In Vivo Behavior Of An Anti-(1-->6) Dextran Antibody. J Immunol, 162, 2162-70.

Costanzi, J. J., Coltman, C. A., Jr., Clark, D. A., Tennenbaum, J. I. & Criscuolo, D. 1965. Cryoblobulinemia Associated With A Macroglobulin; Studies Of A 17.5S Cryoprecipitating Factor. Am J Med, 39, 163-72.

Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. 2003. An Evolving Hierarchical Family Classification For Glycosyltransferases. J Mol Biol, 328, 307-17.

Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., Van Den Berg, T. K., Van Der Merwe, P. A., Watt, S. M. & Varki, A. 1998. Siglecs: A Family Of Sialic-Acid Binding Lectins. Glycobiology, 8, V.

Crocker, P. R., Paulson, J. C. & Varki, A. 2007. Siglecs And Their Roles In The Immune System. Nature Reviews Immunology, 7, 255-266.

Croset, A., Delafosse, L., Gaudry, J. P., Arod, C., Glez, L., Losberger, C., Begue, D., Krstanovic, A., Robert, F., Vilbois, F., Chevalet, L. & Antonsson, B. 2012. Differences In The Glycosylation Of Recombinant Proteins Expressed In Hek And Cho Cells. J Biotechnol, 161, 336-48.

Cummings, R. D. & Mcever, R. P. 2015. C-Type Lectins. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Cummings, R. D. & Pierce, J. M. 2014. The Challenge And Promise Of Glycomics. Chem Biol, 21, 1-15.

Curtis, B. M., Scharnowske, S. & Watson, A. J. 1992. Sequence And Expression Of A Membrane-Associated C-Type Lectin That Exhibits Cd4-Independent Binding Of Human Immunodeficiency Virus Envelope Glycoprotein Gp120. Proc Natl Acad Sci U S A, 89, 8356-60.

Czajkowsky, D. M., Andersen, J. T., Fuchs, A., Wilson, T. J., Mekhaiel, D., Colonna, M., He, J., Shao, Z., Mitchell, D. A., Wu, G., Dell, A., Haslam, S., Lloyd, K. A., Moore, S. C., Sandlie, I., Blundell, P. A. & Pleass, R. J. 2015. Developing The Ivig Biomimetic, Hexa-Fc, For Drug And Vaccine Applications. Sci Rep, 5, 9526.

Czajkowsky, D. M., Hu, J., Shao, Z. & Pleass, R. J. 2012. Fc-Fusion Proteins: New Developments And Future Perspectives. Embo Mol Med, 4, 1015-28.

Dalton, C. F., Laird, S. M., Serle, E., Saravelos, H., Warren, M. A., Li, T. C. & Bolton, A. E. 1995. The Measurement Of Ca 125 And Placental Protein 14 In Uterine Flushings In Women With Recurrent Miscarriage; Relation To Endometrial Morphology. Human Reproduction, 10, 2680-2684.

Dalziel, M., Whitehouse, C., Mcfarlane, I., Brockhausen, I., Gschmeissner, S., Schwientek, T., Clausen, H., Burchell, J. M. & Taylor-Papadimitriou, J. 2001. The Relative Activities Of The C2gnt1 And St3gal-I Glycosyltransferases Determine O-Glycan Structure And Expression Of A Tumor-Associated Epitope On Muc1. J Biol Chem, 276, 11007-15.

Damsky, W. E., Jr & Bosenberg, M. 2010. Mouse Melanoma Models And Cell Lines. Pigment Cell Melanoma Res, 23, 853-9.

De Lange, J., Ly, L. V., Lodder, K., Verlaan-De Vries, M., Teunisse, A. F., Jager, M. J. & Jochemsen, A. G. 2012. Synergistic Growth Inhibition Based On Small-Molecule P53 Activation As Treatment For Intraocular Melanoma. Oncogene, 31, 1105-16.

Degroote, S., Wolthoorn, J. & Van Meer, G. 2004. The Cell Biology Of Glycosphingolipids. Semin Cell Dev Biol, 15, 375-87.

Dekkers, G., Plomp, R., Koeleman, C. A., Visser, R., Von Horsten, H. H., Sandig, V., Rispens, T., Wuhrer, M. & Vidarsson, G. 2016. Multi-Level Glyco-Engineering Techniques To Generate Igg With Defined Fc-Glycans. Sci Rep, 6, 36964.

Dell, A., Lu, D., Haslam, S. M. & Clark, G. F. 2018. Towards A Novel Cancer Vaccine: Characterisation Of The Glycome Of Canine Melanoma Cells. Oxford Univ Press Inc.

Dell, A., Morris, H. R., Easton, R. L., Panico, M., Patankar, M., Oehniger, S., Koistinen, R., Koistinen, H., Seppala, M. & Clark, G. F. 1995. Structural Analysis Of The Oligosaccharides Derived From Glycodelin, A Human Glycoprotein With Potent Immunosuppressive And Contraceptive Activities. J Biol Chem, 270, 24116-26.

Dell, A., Morris, H. R., Egge, H., Von Nicolai, H. & Strecker, G. 1983. Fast-Atom-Bombardment Mass-Spectrometry For Carbohydrate-Structure Determination. Carbohydrate Research, 115, 41-52.

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.

Demetriou, M., Nabi, I. R., Coppolino, M., Dedhar, S. & Dennis, J. W. 1995. Reduced Contact-Inhibition And Substratum Adhesion In Epithelial Cells Expressing Glcnac-Transferase V. J Cell Biol, 130, 383-92.
Dempster, A. J. 1918. A New Method Of Positive Ray Analysis. Physical Review, 11, 316-325.

Dennis, J. W., Laferte, S., Waghorne, C., Breitman, M. L. & Kerbel, R. S. 1987. Beta 1-6 Branching Of Asn-Linked Oligosaccharides Is Directly Associated With Metastasis. Science, 236, 582-5.

Diamond, M. S., Staunton, D. E., Marlin, S. D. & Springer, T. A. 1991. Binding Of The Integrin Mac-1 (Cd11b/

Cd18) To The Third Immunoglobulin-Like Domain Of Icam-1 (Cd54) And Its Regulation By Glycosylation. Cell, 65, 961-71.

Ding, Y., Guo, Z., Liu, Y., Li, X., Zhang, Q., Xu, X., Gu, Y., Zhang, Y., Zhao, D. & Cao, X. 2016. The Lectin Siglec-G Inhibits Dendritic Cell Cross-Presentation By Impairing Mhc Class I–Peptide Complex Formation. Nature Immunology, 17, 1167.

Diskin, S., Cao, Z., Leffler, H. & Panjwani, N. 2009. The Role Of Integrin Glycosylation In Galectin-8-Mediated Trabecular Meshwork Cell Adhesion And Spreading. Glycobiology, 19, 29-37.

Dithmar, S., Albert, D. M. & Grossniklaus, H. E. 2000. Animal Models Of Uveal Melanoma. Melanoma Res, 10, 195-211.

Dodd, R. B. & Drickamer, K. 2001. Lectin-Like Proteins In Model Organisms: Implications For Evolution Of Carbohydrate-Binding Activity. Glycobiology, 11, 71R-9R.

Domon, B. & E Costello, C. 1988. A Systematic Nomenclature For Carbohydrate Fragmentations In Fab-Ms/Ms Spectra Of Glycoconjugates.

Drickamer, K. 1992. Engineering Galactose-Binding Activity Into A C-Type Mannose-Binding Protein. Nature, 360, 183-6.

Drickamer, K. & Dodd, R. B. 1999. C-Type Lectin-Like Domains In Caenorhabditis Elegans: Predictions From The Complete Genome Sequence. Glycobiology, 9, 1357-69.

Drickamer, K. & Taylor, M. E. 1993. Biology Of Animal Lectins. Annu Rev Cell Biol, 9, 237-64.

Duksin, D. & Bornstein, P. 1977. Changes In Surface Properties Of Normal And Transformed Cells Caused By Tunicamycin, An Inhibitor Of Protein Glycosylation. Proc Natl Acad Sci U S A, 74, 3433-7.

Dyatlovitskaya, E. V., Kryukova, E. V., Suskova, V. S., Emez, V. I. & Bergelson, L. D. 1990. Immunomodulatory Effects Of Human Placenta Gangliosides. Biomedical Science, 1, 397-400.

Earl, L. A., Bi, S. & Baum, L. G. 2010. N- And O-Glycans Modulate Galectin-1 Binding, Cd45 Signaling, And T Cell Death. J Biol Chem, 285, 2232-44.

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, 881-90.

Eppihimer, M. J., Russell, J., Anderson, D. C., Wolitzky, B. A. & Granger, D. N. 1997. Endothelial Cell Adhesion Molecule Expression In Gene-Targeted Mice. Am J Physiol, 273, H1903-8.

Erickson, B. W., Gerber-Jenson, B., Wang, A. C. & Litman, G. W. 1982. Molecular Basis For The Temperature-Dependent Insolubility Of Cryoglobulins--Xi. Sequence Comparison Of The Heavy-Chain Variable Regions Of The Human Cryoimmunoglobulins Mce And Hil By Metric Analysis. Mol Immunol, 19, 357-65.

Evans, J., Salamonsen, L. A., Menkhorst, E. & Dimitriadis, E. 2015. Dynamic Changes In Hyperglycosylated Human Chorionic Gonadotrophin Throughout The First Trimester Of Pregnancy And Its Role In Early Placentation. Hum Reprod, 30, 1029-38.

Falconer, R. A., Errington, R. J., Shnyder, S. D., Smith, P. J. & Patterson, L. H. 2012. Polysialyltransferase: A New Target In Metastatic Cancer. Curr Cancer Drug Targets, 12, 925-39.

Ferrara, C., Grau, S., Jager, C., Sondermann, P., Brunker, P., Waldhauer, I., Hennig, M., Ruf, A., Rufer, A. C., Stihle, M., Umana, P. & Benz, J. 2011. Unique Carbohydrate-Carbohydrate Interactions Are Required For High Affinity Binding Between Fcgammariii And Antibodies Lacking Core Fucose. Proc Natl Acad Sci U S A, 108, 12669-74.

Ferrara, N. & Davis-Smyth, T. 1997. The Biology Of Vascular Endothelial Growth Factor. Endocr Rev, 18, 4-25.

Frank, M., Walker, R. C., Lanzilotta, W. N., Prestegard, J. H. & Barb, A. W. 2014. Immunoglobulin G1 Fc Domain Motions: Implications For Fc Engineering. J Mol Biol, 426, 1799-811.

Freeze, H. H. 2006. Genetic Defects In The Human Glycome. Nat Rev Genet, 7, 537-51.

Frickel, E. M., Riek, R., Jelesarov, I., Helenius, A., Wuthrich, K. & Ellgaard, L. 2002. Trosy-Nmr Reveals Interaction Between Erp57 And The Tip Of The Calreticulin P-Domain. Proc Natl Acad Sci U S A, 99, 1954-9.

Friend, S. L., Hosier, S., Nelson, A., Foxworthe, D., Williams, D. E. & Farr, A. 1994. A Thymic Stromal Cell Line Supports In Vitro Development Of Surface Igm+ B Cells And Produces A Novel Growth Factor Affecting B And T Lineage Cells. Exp Hematol, 22, 321-8.

Fu, C., Zhao, H., Wang, Y., Cai, H., Xiao, Y., Zeng, Y. & Chen, H. 2016. Tumor-Associated Antigens: Tn Antigen, Stn Antigen, And T Antigen. Hla, 88, 275-286.

Fukuda, M. N., Sasaki, H., Lopez, L. & Fukuda, M. 1989. Survival Of Recombinant Erythropoietin In The Circulation: The Role Of Carbohydrates. Blood, 73, 84-9.

Fukumori, T., Takenaka, Y., Yoshii, T., Kim, H. R., Hogan, V., Inohara, H., Kagawa, S. & Raz, A. 2003. Cd29 And Cd7 Mediate Galectin-3-Induced Type Ii T-Cell Apoptosis. Cancer Res, 63, 8302-11.

Fukuta, K., Abe, R., Yokomatsu, T., Kono, N., Nagatomi, Y., Asanagi, M., Shimazaki, Y. & Makino, T. 2000. Comparative Study Of The N-Glycans Of Human Monoclonal Immunoglobulins M Produced By Hybridoma And Parental Cells. Arch Biochem Biophys, 378, 142-50.

Galli, C., Bernasconi, R., Solda, T., Calanca, V. & Molinari, M. 2011. Malectin Participates In A Backup Glycoprotein Quality Control Pathway In The Mammalian Er. Plos One, 6, E16304.

Garcia-Vallejo, J. J., Van Liempt, E., Da Costa Martins, P., Beckers, C., Van Het Hof, B., Gringhuis, S. I., Zwaginga, J. J., Van Dijk, W., Geijtenbeek, T. B., Van Kooyk, Y. & Van Die, I. 2008. Dc-Sign Mediates Adhesion And Rolling Of Dendritic Cells On Primary Human Umbilical Vein Endothelial Cells Through Lewisy Antigen Expressed On Icam-2. Mol Immunol, 45, 2359-69.

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.

Geijtenbeek, T. B., Torensma, R., Van Vliet, S. J., Van Duijnhoven, G. C., Adema, G. J., Van Kooyk, Y. & Figdor, C. G. 2000. Identification Of Dc-Sign, A Novel Dendritic Cell-Specific Icam-3 Receptor That Supports Primary Immune Responses. Cell, 100, 575-85.

Gencler, B. & Gonul, M. 2016. Cutaneous Side Effects Of Braf Inhibitors In Advanced Melanoma: Review Of The Literature. Dermatol Res Pract, 2016, 5361569.

Gerber-Jenson, B., Kazin, A., Kehoe, J. M., Scheffel, C., Erickson, B. W. & Litman, G. W. 1981. Molecular

Basis For The Temperature-Dependent Insolubility Of Cryoglobulins. X. The Amino Acid Sequence Of The Heavy Chain Variable Region Of Mce. J Immunol, 126, 1212-6.

Gill, D. J., Chia, J., Senewiratne, J. & Bard, F. 2010. Regulation Of O-Glycosylation Through Golgi-To-Er Relocation Of Initiation Enzymes. J Cell Biol, 189, 843-58.

Glinskii, O. V., Sud, S., Mossine, V. V., Mawhinney, T. P., Anthony, D. C., Glinsky, G. V., Pienta, K. J. & Glinsky, V. V. 2012. Inhibition Of Prostate Cancer Bone Metastasis By Synthetic Tf Antigen Mimic/Galectin-3 Inhibitor Lactulose-L-Leucine. Neoplasia, 14, 65-73.

Green, P. J., Tamatani, T., Watanabe, T., Miyasaka, M., Hasegawa, A., Kiso, M., Yuen, C. T., Stoll, M. S. & Feizi, T. 1992. High Affinity Binding Of The Leucocyte Adhesion Molecule L-Selectin To 3'-Sulphated-Le(A) And -Le(X) Oligosaccharides And The Predominance Of Sulphate In This Interaction Demonstrated By Binding Studies With A Series Of Lipid-Linked Oligosaccharides. Biochem Biophys Res Commun, 188, 244-51.

Grey, H. M., Kohler, P. F., Terry, W. D. & Franklin, E. C. 1968. Human Monoclonal Gamma G-Cryoglobulins With Anti-Gamma-Globulin Activity. J Clin Invest, 47, 1875-84.

Gringhuis, S. I., Den Dunnen, J., Litjens, M., Van Der Vlist, M. & Geijtenbeek, T. B. H. 2009. Carbohydrate-Specific Signaling Through The Dc-Sign Signalosome Tailors Immunity To Mycobacterium Tuberculosis, Hiv-1 And Helicobacter Pylori. Nature Immunology, 10, 1081.

Grossniklaus, H. E., Barron, B. C. & Wilson, M. W. 1995. Murine Model Of Anterior And Posterior Ocular Melanoma. Curr Eye Res, 14, 399-404.

Gugliotta, A., Ceaglio, N., Raud, B., Forno, G., Mauro, L., Kratje, R. & Oggero, M. 2017. Glycosylation And Antiproliferative Activity Of Hyperglycosylated Ifn-Alpha2 Potentiate Hek293 Cells As Biofactories. Eur J Pharm Biopharm, 112, 119-131.

Guo, S., Muhlfeld, A. S., Wietecha, T. A., Peutz-Kootstra, C. J., Kowalewska, J., Yi, K., Spencer, M., Pichaiwong, W., Nimmerjahn, F., Hudkins, K. L. & Alpers, C. E. 2009. Deletion Of Activating Fcgamma Receptors Does Not Confer Protection In Murine Cryoglobulinemia-Associated Membranoproliferative Glomerulonephritis. Am J Pathol, 175, 107-18.

Ha, C. T., Waterhouse, R., Warren, J., Zimmermann, W. & Dveksler, G. S. 2008. N-Glycosylation Is Required For Binding Of Murine Pregnancy-Specific Glycoproteins 17 And 19 To The Receptor Cd9. Am J Reprod Immunol, 59, 251-8.

Ha, C. T., Wu, J. A., Irmak, S., Lisboa, F. A., Dizon, A. M., Warren, J. W., Ergun, S. & Dveksler, G. S. 2010. Human Pregnancy Specific Beta-1-Glycoprotein 1 (Psg1) Has A Potential Role In Placental Vascular Morphogenesis. Biol Reprod, 83, 27-35.

Hakomori, S. 1983. Tumor-Associated Glycolipid Antigens Defined By Monoclonal Antibodies. Bull Cancer, 70, 118-26.

Hakomori, S. 1984. Tumor-Associated Carbohydrate Antigens. Annu Rev Immunol, 2, 103-26.

Hamerman, J. A. & Lanier, L. L. 2006. Inhibition Of Immune Responses By Itam-Bearing Receptors. Sci Stke, 2006, Re1.

Handschuh, K., Guibourdenche, J., Tsatsaris, V., Guesnon, M., Laurendeau, I., Evain-Brion, D. & Fournier, T. 2007. Human Chorionic Gonadotropin Produced By The Invasive Trophoblast But Not The Villous Trophoblast Promotes Cell Invasion And Is Down-Regulated By Peroxisome Proliferator-Activated Receptor-Gam-

ma. Endocrinology, 148, 5011-5019.

Hao, P., Adav, S. S., Gallart-Palau, X. & Sze, S. K. 2017. Recent Advances In Mass Spectrometric Analysis Of Protein Deamidation. Mass Spectrom Rev, 36, 677-692.

Harris, L. K., Jones, C. J. & Aplin, J. D. 2009. Adhesion Molecules In Human Trophoblast - A Review. Ii. Extravillous Trophoblast. Placenta, 30, 299-304.

Hart, G. W. & Copeland, R. J. 2010. Glycomics Hits The Big Time. Cell, 143, 672-6.

Hasegawa, M., Orita, T., Kojima, T., Tomonoh, K., Hirata, Y. & Ochi, N. 1992. Improvement In The Heterogeneous N-Termini And The Defective N-Glycosylation Of Human Interleukin-6 By Genetic Engineering. Eur J Biochem, 210, 9-12.

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.

Hayward, N. K., Wilmott, J. S., Waddell, N., Johansson, P. A., Field, M. A., Nones, K., Patch, A. M., Kakavand, H., Alexandrov, L. B., Burke, H., Jakrot, V., Kazakoff, S., Holmes, O., Leonard, C., Sabarinathan, R., Mularoni, L., Wood, S., Xu, Q., Waddell, N., Tembe, V., Pupo, G. M., De Paoli-Iseppi, R., Vilain, R. E., Shang, P., Lau, L. M. S., Dagg, R. A., Schramm, S. J., Pritchard, A., Dutton-Regester, K., Newell, F., Fitzgerald, A., Shang, C. A., Grimmond, S. M., Pickett, H. A., Yang, J. Y., Stretch, J. R., Behren, A., Kefford, R. F., Hersey, P., Long, G. V., Cebon, J., Shackleton, M., Spillane, A. J., Saw, R. P. M., Lopez-Bigas, N., Pearson, J. V., Thompson, J. F., Scolyer, R. A. & Mann, G. J. 2017. Whole-Genome Landscapes Of Major Melanoma Subtypes. Nature, 545, 175-180.

Head, J. R. & Billingham, R. E. 1986. Concerning The Immunology Of The Uterus. American Journal Of Reproductive Immunology And Microbiology : Ajrim, 10, 76-81.

Heffner, K. M., Wang, Q., Hizal, D. B., Can, O. & Betenbaugh, M. J. 2018. Glycoengineering Of Mammalian Expression Systems On A Cellular Level. Adv Biochem Eng Biotechnol.

Heikinheimo, M., Jalanko, H., Renlund, M., Rapola, J. & Wahlstrom, T. 1987. Studies On The Elevated Amniotic Fluid Sp1 In Meckel's Syndrome: Modified Glycosylation Of Sp1. Placenta, 8, 427-32.

Hemmerich, S., Bertozzi, C. R., Leffler, H. & Rosen, S. D. 1994. Identification Of The Sulfated Monosaccharides Of Glycam-1, An Endothelial-Derived Ligand For L-Selectin. Biochemistry, 33, 4820-9.

Hemmerich, S., Leffler, H. & Rosen, S. D. 1995. Structure Of The O-Glycans In Glycam-1, An Endothelial-Derived Ligand For L-Selectin. J Biol Chem, 270, 12035-47.

Hemmerich, S. & Rosen, S. D. 1994. 6'-Sulfated Sialyl Lewis X Is A Major Capping Group Of Glycam-1. Biochemistry, 33, 4830-5.

Hernandez, B., Adissu, H. A., Wei, B. R., Michael, H. T., Merlino, G. & Simpson, R. M. 2018. Naturally Occurring Canine Melanoma As A Predictive Comparative Oncology Model For Human Mucosal And Other Triple Wild-Type Melanomas. Int J Mol Sci, 19.

Hernandez, J. D., Nguyen, J. T., He, J., Wang, W., Ardman, B., Green, J. M., Fukuda, M. & Baum, L. G. 2006. Galectin-1 Binds Different Cd43 Glycoforms To Cluster Cd43 And Regulate T Cell Death. J Immunol, 177, 5328-36.

Hibbs, M. L., Tarlinton, D. M., Armes, J., Grail, D., Hodgson, G., Maglitto, R., Stacker, S. A. & Dunn, A. R.

1995. Multiple Defects In The Immune System Of Lyn-Deficient Mice, Culminating In Autoimmune Disease. Cell, 83, 301-11.

Hirabayashi, J. & Kasai, K. 1984. Human Placenta Beta-Galactoside-Binding Lectin. Purification And Some Properties. Biochem Biophys Res Commun, 122, 938-44.

Hirano, K., Matsuda, A., Shirai, T. & Furukawa, K. 2014. Expression Of Lacdinac Groups On N-Glycans Among Human Tumors Is Complex. Biomed Res Int, 2014, 981627.

Hirschberg, K., Rodger, J. & Futerman, A. H. 1993. The Long-Chain Sphingoid Base Of Sphingolipids Is Acylated At The Cytosolic Surface Of The Endoplasmic Reticulum In Rat Liver. Biochem J, 290 (Pt 3), 751-7.

Hodis, E., Watson, I. R., Kryukov, G. V., Arold, S. T., Imielinski, M., Theurillat, J. P., Nickerson, E., Auclair, D., Li, L., Place, C., Dicara, D., Ramos, A. H., Lawrence, M. S., Cibulskis, K., Sivachenko, A., Voet, D., Saksena, G., Stransky, N., Onofrio, R. C., Winckler, W., Ardlie, K., Wagle, N., Wargo, J., Chong, K., Morton, D. L., Stemke-Hale, K., Chen, G., Noble, M., Meyerson, M., Ladbury, J. E., Davies, M. A., Gershenwald, J. E., Wagner, S. N., Hoon, D. S., Schadendorf, D., Lander, E. S., Gabriel, S. B., Getz, G., Garraway, L. A. & Chin, L. 2012. A Landscape Of Driver Mutations In Melanoma. Cell, 150, 251-63.

Hoffmann, E. & Stroobant, V. 2007. Mass Spectrometry: Principles And Applications.

Holmen Larsson, J. M., Thomsson, K. A., Rodriguez-Pineiro, A. M., Karlsson, H. & Hansson, G. C. 2013. Studies Of Mucus In Mouse Stomach, Small Intestine, And Colon. Iii. Gastrointestinal Muc5ac And Muc2 Mucin O-Glycan Patterns Reveal A Regiospecific Distribution. Am J Physiol Gastrointest Liver Physiol, 305, G357-63.

Holmes, E. H., Ostrander, G. K., Clausen, H. & Graem, N. 1987. Oncofetal Expression Of Lex Carbohydrate Antigens In Human Colonic Adenocarcinomas. Regulation Through Type 2 Core Chain Synthesis Rather Than Fucosylation. J Biol Chem, 262, 11331-8.

Hossler, P., Khattak, S. F. & Li, Z. J. 2009. Optimal And Consistent Protein Glycosylation In Mammalian Cell Culture. Glycobiology, 19, 936-49.

Hoti, N., Yang, S., Hu, Y., Shah, P., Haffner, M. C. & Zhang, H. 2018. Overexpression Of Alpha (1,6) Fucosyltransferase In The Development Of Castration-Resistant Prostate Cancer Cells. Prostate Cancer Prostatic Dis, 21, 137-146.

Hsiao, C. T., Cheng, H. W., Huang, C. M., Li, H. R., Ou, M. H., Huang, J. R., Khoo, K. H., Yu, H. W., Chen, Y. Q., Wang, Y. K., Chiou, A. & Kuo, J. C. 2017. Fibronectin In Cell Adhesion And Migration Via N-Glycosylation. Oncotarget, 8, 70653-70668.

Hsu, S. C., Tsai, T. H., Kawasaki, H., Chen, C. H., Plunkett, B., Lee, R. T., Lee, Y. C. & Huang, S. K. 2007. Antigen Coupled With Lewis-X Trisaccharides Elicits Potent Immune Responses In Mice. J Allergy Clin Immunol, 119, 1522-8.

Hunt, J. S., Andrews, G. K. & Wood, G. W. 1987. Normal Trophoblasts Resist Induction Of Class I Hla. J Immunol, 138, 2481-7.

Ikehara, Y., Ikehara, S. K. & Paulson, J. C. 2004. Negative Regulation Of T Cell Receptor Signaling By Siglec-7 (P70/Airm) And Siglec-9. J Biol Chem, 279, 43117-25.

Imai, Y., Lasky, L. A. & Rosen, S. D. 1993. Sulphation Requirement For Glycam-1, An Endothelial Ligand For L-Selectin. Nature, 361, 555-7.

Imai, Y., Singer, M. S., Fennie, C., Lasky, L. A. & Rosen, S. D. 1991. Identification Of A Carbohydrate-Based Endothelial Ligand For A Lymphocyte Homing Receptor. J Cell Biol, 113, 1213-21.

Iwamori, M., Shimomura, J., Tsuyuhara, S. & Nagai, Y. 1984. Gangliosides Of Various Rat Tissues: Distribution Of Ganglio-N-Tetraose-Containing Gangliosides And Tissue-Characteristic Composition Of Gangliosides. J Biochem, 95, 761-70.

Jandus, C., Boligan, K. F., Chijioke, O., Liu, H., Dahlhaus, M., Demoulins, T., Schneider, C., Wehrli, M., Hunger, R. E., Baerlocher, G. M., Simon, H. U., Romero, P., Munz, C. & Von Gunten, S. 2014. Interactions Between Siglec-7/9 Receptors And Ligands Influence Nk Cell-Dependent Tumor Immunosurveillance. J Clin Invest, 124, 1810-20.

Jefferis, R. 2009. Glycosylation As A Strategy To Improve Antibody-Based Therapeutics. Nat Rev Drug Discov, 8, 226-34.

Jeschke, U., Wang, X., Briese, V., Friese, K. & Stahn, R. 2003. Glycodelin And Amniotic Fluid Transferrin As Inhibitors Of E-Selectin-Mediated Cell Adhesion. Histochemistry And Cell Biology, 119, 345-354.

Jiang, Z., Zhang, H., Liu, C., Yin, J., Tong, S., Lv, J., Wei, S. & Wu, S. 2018. Beta3gnt8 Promotes Colorectal Cancer Cells Invasion Via Cd147/Mmp2/Galectin3 Axis. Front Physiol, 9, 588.

Jimbow, K., Roth, S. I., Fitzpatrick, T. B. & Szabo, G. 1975. Mitotic Activity In Non-Neoplastic Melanocytes In Vivo As Determined By Histochemical, Autoradiographic, And Electron Microscope Studies. J Cell Biol, 66, 663-70.

Jin, L., Mclean, P. A., Neel, B. G. & Wortis, H. H. 2002. Sialic Acid Binding Domains Of Cd22 Are Required For Negative Regulation Of B Cell Receptor Signaling. J Exp Med, 195, 1199-205.

Jones, C. J. P., Carter, A. M., Aplin, J. D. & Enders, A. C. 2007. Glycosylation At The Fetomaternal Interface In Hemomonochorial Placentae From Five Widely Separated Species Of Mammal: Is There Evidence For Convergent Evolution? Cells Tissues Organs, 185, 269-284.

Jonuleit, H., Schmitt, E., Schuler, G., Knop, J. & Enk, A. H. 2000. Induction Of Interleukin 10-Producing, Nonproliferating Cd4(+) T Cells With Regulatory Properties By Repetitive Stimulation With Allogeneic Immature Human Dendritic Cells. J Exp Med, 192, 1213-22.

Julkunen, M., Koistinen, R., Sjöberg, J., Rutanen, E. M., Wahlström, T. & Seppälä, M. 1986. Secretory Endometrium Synthesizes Placental Protein 14. Endocrinology, 118, 1782-1786.

Julkunen, M., Rutanen, E. M., Koskimies, A., Ranta, T., Bohn, H. & Seppälä, M. 1985. Distribution Of Placental Protein 14 In Tissues And Body Fluids During Pregnancy. Bjog: An International Journal Of Obstetrics & Gynaecology, 92, 1145-1151.

Julkunen, M., Seppälä, M. & Jänne, O. A. 1991. Molecular Cloning Of Complementary Dnas For Two Human Endometrial Proteins And Cellular Localization Of Their Messenger Rnas. Annals Of The New York Academy Of Sciences.

Julkunen, M., Wahlström, T., Seppälä, M., Koistinen, R., Koskimies, A., Stenman, U. H. & Bohn, H. 1984. Detection And Localization Of Placental Protein 14-Like Protein In Human Seminal Plasma And In The Male Genital Tract. Archives Of Andrology, 12 Suppl, 59-67.

Kailemia, M. J., Ruhaak, L. R., Lebrilla, C. B. & Amster, I. J. 2014. Oligosaccharide Analysis By Mass Spec-

trometry: A Review Of Recent Developments. Anal Chem, 86, 196-212.

Kanitakis, J., Al-Rifai, I., Faure, M. & Claudy, A. 1998. Differential Expression Of The Cancer Associated Antigens T (Thomsen-Friedenreich) And Tn To The Skin In Primary And Metastatic Carcinomas. J Clin Pathol, 51, 588-92.

Kannagi, R., Cai, B. H., Huang, H. C., Chao, C. C. & Sakuma, K. 2018. Gangliosides And Tumors. Methods Mol Biol, 1804, 143-171.

Kannagi, R. & Hakomori, S.-I. 1983. Glycosphingolipids As Tumor-Associated And Differentiation Markers56. Jnci: Journal Of The National Cancer Institute, 71, 231-251.

Kannagi, R., Yin, J., Miyazaki, K. & Izawa, M. 2008. Current Relevance Of Incomplete Synthesis And Neo-Synthesis For Cancer-Associated Alteration Of Carbohydrate Determinants--Hakomori's Concepts Revisited. Biochim Biophys Acta, 1780, 525-31.

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.

Karas, M. & Hillenkamp, F. 1988. Laser Desorption Ionization Of Proteins With Molecular Masses Exceeding 10,000 Daltons. Analytical Chemistry, 60, 2299-2301.

Karlsson, K. A. 1970. Sphingolipid Long Chain Bases. Lipids, 5, 878-91.

Karsten, C. M., Pandey, M. K., Figge, J., Kilchenstein, R., Taylor, P. R., Rosas, M., Mcdonald, J. U., Orr, S. J., Berger, M., Petzold, D., Blanchard, V., Winkler, A., Hess, C., Reid, D. M., Majoul, I. V., Strait, R. T., Harris, N. L., Kohl, G., Wex, E., Ludwig, R., Zillikens, D., Nimmerjahn, F., Finkelman, F. D., Brown, G. D., Ehlers, M. & Kohl, J. 2012. Anti-Inflammatory Activity Of Igg1 Mediated By Fc Galactosylation And Association Of FcgammaRIIb And Dectin-1. Nat Med, 18, 1401-6.

Katayama, Y., Hidalgo, A., Chang, J., Peired, A. & Frenette, P. S. 2005. Cd44 Is A Physiological E-Selectin Ligand On Neutrophils. J Exp Med, 201, 1183-9.

Kawano, T., Takasaki, S., Tao, T. W. & Kobata, A. 1991. N-Linked Sugar Chains Of Mouse B16 Melanoma Cells And Their Low-Metastasizing Variant Selected By Wheat Germ Agglutinin. Glycobiology, 1, 375-85.

Kelleher, D. J. & Gilmore, R. 2006. An Evolving View Of The Eukaryotic Oligosaccharyltransferase. Glycobiology, 16, 47R-62R.

Kelleher, D. J., Karaoglu, D., Mandon, E. C. & Gilmore, R. 2003. Oligosaccharyltransferase Isoforms That Contain Different Catalytic Stt3 Subunits Have Distinct Enzymatic Properties. Mol Cell, 12, 101-11.

Kelm, S., Gerlach, J., Brossmer, R., Danzer, C. P. & Nitschke, L. 2002. The Ligand-Binding Domain Of Cd22 Is Needed For Inhibition Of The B Cell Receptor Signal, As Demonstrated By A Novel Human Cd22-Specific Inhibitor Compound. J Exp Med, 195, 1207-13.

Kemna, M. J., Plomp, R., Van Paassen, P., Koeleman, C. A. M., Jansen, B. C., Damoiseaux, J. G. M. C., Cohen Tervaert, J. W. & Wuhrer, M. 2017. Galactosylation And Sialylation Levels Of Igg Predict Relapse In Patients With Pr3-Anca Associated Vasculitis. Ebiomedicine, 17, 108-118.

Khanna, C., Lindblad-Toh, K., Vail, D., London, C., Bergman, P., Barber, L., Breen, M., Kitchell, B., Mcneil, E., Modiano, J. F., Niemi, S., Comstock, K. E., Ostrander, E., Westmoreland, S. & Withrow, S. 2006. The Dog As A Cancer Model. Nature Biotechnology, 24, 1065.

Kim, Y. J. & Varki, A. 1997. Perspectives On The Significance Of Altered Glycosylation Of Glycoproteins In Cancer. Glycoconj J, 14, 569-76.

King, A., Gardner, L. & Loke, Y. W. 1996. Human Decidual Leukocytes Do Not Proliferate In Response To Either Extravillous Trophoblast Or Allogeneic Peripheral Blood Lymphocytes. J Reprod Immunol, 30, 67-74.

Kinoshita, M., Mitsui, Y., Kakoi, N., Yamada, K., Hayakawa, T. & Kakehi, K. 2014. Common Glycoproteins Expressing Polylactosamine-Type Glycans On Matched Patient Primary And Metastatic Melanoma Cells Show Different Glycan Profiles. J Proteome Res, 13, 1021-33.

Kiyoshi, M., Tsumoto, K., Ishii-Watabe, A. & Caaveiro, J. M. M. 2017. Glycosylation Of Igg-Fc: A Molecular Perspective. Int Immunol, 29, 311-317.

Klein, M., Danon, F., Brouet, J. C., Signoret, Y. & Seligmann, M. 1972. [Immunochemical Study Of 130 Human Cryoglobulins]. Rev Eur Etud Clin Biol, 17, 948-57.

Knochenmuss, R. 2006. Ion Formation Mechanisms In Uv-Maldi. Analyst, 131, 966-86.

Koistinen, H., Koistinen, R., Dell, A., Morris, H. R., Easton, R. L., Patankar, M. S., Oehninger, S., Clark, G. F. & Seppälä, M. 1996. Glycodelin From Seminal Plasma Is A Differentially Glycosylated Form Of Contraceptive Glycodelin-A. Molecular Human Reproduction, 2, 759-765.

Koistinen, R., Heikinheimo, M., Rutanen, E. M., Stenman, U. H., Lee, J. N. & Seppala, M. 1981. Concanavalin A Binding Of Pregnancy-Specific Beta-1-Glycoprotein In Normal Pregnancy And Trophoblastic Disease. Oncodev Biol Med, 2, 179-82.

Kolasinska, E., Przybylo, M., Janik, M. & Litynska, A. 2016. Towards Understanding The Role Of Sialylation In Melanoma Progression. Acta Biochim Pol, 63, 533-41.

Kolopp-Sarda, M. N. & Miossec, P. 2018. Cryoglobulins: An Update On Detection, Mechanisms And Clinical Contribution. Autoimmun Rev, 17, 457-464.

Kolter, T. 2012. Ganglioside Biochemistry. Isrn Biochem, 2012, 506160.

Komori, H. & Ito, M. 1995. Conversion Of Short-Chain Ceramides To Short-Chain Ceramide Gm3 In B16 Melanoma Cells. Febs Lett, 374, 299-302.

Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. 2013. Unraveling The Mechanism Of Electrospray Ionization. Anal Chem, 85, 2-9.

Kontermann, R. E. 2011. Strategies For Extended Serum Half-Life Of Protein Therapeutics. Curr Opin Biotechnol, 22, 868-76.

Koopman, L. A., Kopcow, H. D., Rybalov, B., Boyson, J. E., Orange, J. S., Schatz, F., Masch, R., Lockwood, C. J., Schachter, A. D., Park, P. J. & Strominger, J. L. 2003. Human Decidual Natural Killer Cells Are A Unique Nk Cell Subset With Immunomodulatory Potential. J Exp Med, 198, 1201-12.

Kozlov, G., Bastos-Aristizabal, S., Maattanen, P., Rosenauer, A., Zheng, F., Killikelly, A., Trempe, J. F., Thomas, D. Y. & Gehring, K. 2010. Structural Basis Of Cyclophilin B Binding By The Calnexin/Calreticulin P-Domain. J Biol Chem, 285, 35551-7.

Ku, N. O., Toivola, D. M., Strnad, P. & Omary, M. B. 2010. Cytoskeletal Keratin Glycosylation Protects Epithelial Tissue From Injury. Nat Cell Biol, 12, 876-85.

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. Journal Of Biological Chemistry, 278, 28619-28634.

Kuroda, Y., Kuroki, A., Kikuchi, S., Funase, T., Nakata, M. & Izui, S. 2005. A Critical Role For Sialylation In Cryoglobulin Activity Of Murine Igg3 Monoclonal Antibodies. J Immunol, 175, 1056-61.

Kuroki, A., Kuroda, Y., Kikuchi, S., Lajaunias, F., Fulpius, T., Pastore, Y., Fossati-Jimack, L., Reininger, L., Toda, T., Nakata, M., Kojima, N., Mizuochi, T. & Izui, S. 2002. Level Of Galactosylation Determines Cryoglobulin Activity Of Murine Igg3 Monoclonal Rheumatoid Factor. Blood, 99, 2922-8.

Labriola, C., Cazzulo, J. J. & Parodi, A. J. 1995. Retention Of Glucose Units Added By The Udp-Glc:glycoprotein Glucosyltransferase Delays Exit Of Glycoproteins From The Endoplasmic Reticulum. J Cell Biol, 130, 771-9.

Ladenson, R. P., Schwartz, S. O. & Ivy, A. C. 1949. Incidence Of The Blood Groups And The Secretor Factor In Patients With Pernicious Anemia And Stomach Carcinoma. Am J Med Sci, 217, 194-7.

Lanteri, M., Giordanengo, V., Vidal, F., Gaudray, P. & Lefebvre, J. C. 2002. A Complete Alpha1,3-Galactosyltransferase Gene Is Present In The Human Genome And Partially Transcribed. Glycobiology, 12, 785-92.

Lau, E., Feng, Y., Claps, G., Fukuda, M. N., Perlina, A., Donn, D., Jilaveanu, L., Kluger, H., Freeze, H. H. & Ronai, Z. A. 2015. The Transcription Factor Atf2 Promotes Melanoma Metastasis By Suppressing Protein Fucosylation. Sci Signal, 8, Ra124.

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. G., Oh, M. J., Park, S. Y., An, H. J. & Kim, J. H. 2018. Inhibition Of Poly-Lacnac Biosynthesis With Release Of Cmp-Neu5ac Feedback Inhibition Increases The Sialylation Of Recombinant Epo Produced In Cho Cells. Sci Rep, 8, 7273.

Lee, C. L., Chiu, P. C. N., Pang, P. C., Chu, I. K., Lee, K. F., Koistinen, R., Koistinen, H., Seppälä, M., Morris, H. R., Tissot, B., Panico, M., Dell, A. & Yeung, W. S. B. 2011. Glycosylation Failure Extends To Glycoproteins In Gestational Diabetes Mellitus: Evidence From Reduced A2-6 Sialylation And Impaired Immunomodulatory Activities Of Pregnancy-Related Glycodelin-A. Diabetes, 60, 909-917.

Lee, C. L., Pang, P. C., Yeung, W. S. B., Tissot, B., Panico, M., Lao, T. T. H., Chu, I. K., Lee, K. F., Chung, M. K., Lam, K. K. W., Koistinen, R., Koistinen, H., Seppälä, M., Morris, H. R., Dell, A. & Chiu, P. C. N. 2009A. Effects Of Differential Glycosylation Of Glycodelins On Lymphocyte Survival. Journal Of Biological Chemistry, 284, 15084-15096.

Lee, J. N., Grudzinskas, J. G. & Chard, T. 1979. Circulating Levels Of Pregnancy Proteins In Early And Late Pregnancy In Relation To Placental Tissue Concentration. Br J Obstet Gynaecol, 86, 888-90.

Lee, M. J., Heo, Y. M., Hong, S. H., Kim, K. & Park, S. 2009B. The Binding Properties Of Glycosylated And Non-Glycosylated Tim-3 Molecules On Cd4cd25 T Cells. Immune Netw, 9, 58-63.

Leffler, H. & Barondes, S. H. 1986. Specificity Of Binding Of Three Soluble Rat Lung Lectins To Substituted And Unsubstituted Mammalian Beta-Galactosides. J Biol Chem, 261, 10119-26.

Leibiger, H., Kersten, B., Albersheim, P. & Darvill, A. 1998. Structural Characterization Of The Oligosac-

charides Of A Human Monoclonal Anti-Lipopolysaccharide Immunoglobulin M. Glycobiology, 8, 497-507.

Leibiger, H., Wüstner, D., Stigler, R. D. & Marx, U. 1999. Variable Domain-Linked Oligosaccharides Of A Human Monoclonal Igg: Structure And Influence On Antigen Binding. The Biochemical Journal, 338 (Pt 2), 529-538.

Lenter, M., Levinovitz, A., Isenmann, S. & Vestweber, D. 1994. Monospecific And Common Glycoprotein Ligands For E- And P-Selectin On Myeloid Cells. J Cell Biol, 125, 471-81.

Lerner, A. B. & Watson, C. J. 1947. Studies Of Cryoglobulins; Unusual Purpura Associated With The Presence Of A High Concentration Of Cryoglobulin (Cold Precipitable Serum Globulin). Am J Med Sci, 214, 410-5.

Levinovitz, A., Muhlhoff, J., Isenmann, S. & Vestweber, D. 1993. Identification Of A Glycoprotein Ligand For E-Selectin On Mouse Myeloid Cells. J Cell Biol, 121, 449-59.

Lewis, A. M., Croughan, W. D., Aranibar, N., Lee, A. G., Warrack, B., Abu-Absi, N. R., Patel, R., Drew, B., Borys, M. C., Reily, M. D. & Li, Z. J. 2016. Understanding And Controlling Sialylation In A Cho Fc-Fusion Process. Plos One, 11, E0157111.

Li, X., Jiang, J., Zhao, X., Wang, J., Han, H., Zhao, Y., Peng, B., Zhong, R., Ying, W. & Qian, X. 2013. N-Glycoproteome Analysis Of The Secretome Of Human Metastatic Hepatocellular Carcinoma Cell Lines Combining Hydrazide Chemistry, Hilic Enrichment And Mass Spectrometry. Plos One, 8, E81921.

Li, Z. H., Wang, L. L., Liu, H., Muyayalo, K. P., Huang, X. B., Mor, G. & Liao, A. H. 2018. Galectin-9 Alleviates Lps-Induced Preeclampsia-Like Impairment In Rats Via Switching Decidual Macrophage Polarization To M2 Subtype. Front Immunol, 9, 3142.

Lin, T. M., Halbert, S. P. & Spellacy, W. N. 1974. Measurement Of Pregnancy-Associated Plasma Proteins During Human Gestation. J Clin Invest, 54, 576-82.

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.

Lisboa, F. A., Warren, J., Sulkowski, G., Aparicio, M., David, G., Zudaire, E. & Dveksler, G. S. 2011. Pregnancy-Specific Glycoprotein 1 Induces Endothelial Tubulogenesis Through Interaction With Cell Surface Proteoglycans. J Biol Chem, 286, 7577-86.

Litynska, A., Przybylo, M., Pochec, E., Hoja-Lukowicz, D., Ciolczyk, D., Laidler, P. & Gil, D. 2001. Comparison Of The Lectin-Binding Pattern In Different Human Melanoma Cell Lines. Melanoma Res, 11, 205-12.

Liu, F. T. 2005. Regulatory Roles Of Galectins In The Immune Response. Int Arch Allergy Immunol, 136, 385-400.

Liu, W., Ramachandran, V., Kang, J., Kishimoto, T. K., Cummings, R. D. & Mcever, R. P. 1998. Identification Of N-Terminal Residues On P-Selectin Glycoprotein Ligand-1 Required For Binding To P-Selectin. J Biol Chem, 273, 7078-87.

Lo, C. Y., Antonopoulos, A., Dell, A., Haslam, S. M., Lee, T. & Neelamegham, S. 2013. The Use Of Surface Immobilization Of P-Selectin Glycoprotein Ligand-1 On Mesenchymal Stem Cells To Facilitate Selectin Mediated Cell Tethering And Rolling. Biomaterials, 34, 8213-22.

Lospalluto, J., Dorward, B., Miller, W., Jr. & Ziff, M. 1962. Cryoglobulinemia Based On Interaction Between A Gamma Macroglobulin And 7S Gamma Globulin. Am J Med, 32, 142-7.

Lowe, J. B. 2002. Glycosylation In The Control Of Selectin Counter-Receptor Structure And Function. Immunol Rev, 186, 19-36.

Lu, J. & Gu, J. 2015. Significance Of Beta-Galactoside Alpha2,6 Sialyltranferase 1 In Cancers. Molecules, 20, 7509-27.

Macauley, M. S., Crocker, P. R. & Paulson, J. C. 2014. Siglec-Mediated Regulation Of Immune Cell Function In Disease. Nat Rev Immunol, 14, 653-66.

Magalhaes, A., Duarte, H. O. & Reis, C. A. 2017. Aberrant Glycosylation In Cancer: A Novel Molecular Mechanism Controlling Metastasis. Cancer Cell, 31, 733-735.

Mallick, P., Schirle, M., Chen, S. S., Flory, M. R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B. & Aebersold, R. 2007. Computational Prediction Of Proteotypic Peptides For Quantitative Proteomics. Nat Biotechnol, 25, 125-31.

Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. & Zagulin, V. A. 1973. The Mass-Reflectron A New Nonmagnetic Time-Of-Flight High Resolution Mass-Spectrometer. Zhurnal Ehksperimental'noj I Teoreticheskoj Fiziki, 64, 82-89.

Mandal, D. K., Bhattacharyya, L., Koenig, S. H., Brown, R. D., 3Rd, Oscarson, S. & Brewer, C. F. 1994. Studies Of The Binding Specificity Of Concanavalin A. Nature Of The Extended Binding Site For Asparagine-Linked Carbohydrates. Biochemistry, 33, 1157-62.

Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., Almeida, R., Silva, F., Morais, V., Costa, J., Kihlberg, J., Clausen, H. & Reis, C. A. 2004. Role Of The Human St6galnac-I And St6galnac-Ii In The Synthesis Of The Cancer-Associated Sialyl-Tn Antigen. Cancer Res, 64, 7050-7.

Martinez, F. F., Knubel, C. P., Sanchez, M. C., Cervi, L. & Motran, C. C. 2012. Pregnancy-Specific Glycoprotein 1A Activates Dendritic Cells To Provide Signals For Th17-, Th2-, And Treg-Cell Polarization. Eur J Immunol, 42, 1573-84.

Matsumoto, T., Takahashi, N., Kojima, T., Yoshioka, Y., Ishikawa, J., Furukawa, K., Ono, K., Sawada, M., Ishiguro, N. & Yamamoto, A. 2016. Soluble Siglec-9 Suppresses Arthritis In A Collagen-Induced Arthritis Mouse Model And Inhibits M1 Activation Of Raw264.7 Macrophages. Arthritis Res Ther, 18, 133.

Matsumura, K., Higashida, K., Ishida, H., Hata, Y., Yamamoto, K., Shigeta, M., Mizuno-Horikawa, Y., Wang, X., Miyoshi, E., Gu, J. & Taniguchi, N. 2007. Carbohydrate Binding Specificity Of A Fucose-Specific Lectin From Aspergillus Oryzae: A Novel Probe For Core Fucose. J Biol Chem, 282, 15700-8.

Mcever, R. P., Moore, K. L. & Cummings, R. D. 1995. Leukocyte Trafficking Mediated By Selectin-Carbohydrate Interactions. J Biol Chem, 270, 11025-8.

Mechref, Y. 2012. Use Of Cid/Etd Mass Spectrometry To Analyze Glycopeptides. Curr Protoc Protein Sci, Chapter 12, Unit 12 11 1-11.

Medawar, P. B. 1953. Some Immunological And Endocrinological Problems Raised By The Evolution Of Viviparity In Vertebrates. Symp Soc Exp Biol, 7, 320-337.

Mellquist, J. L., Kasturi, L., Spitalnik, S. L. & Shakin-Eshleman, S. H. 1998. The Amino Acid Following An Asn-X-Ser/Thr Sequon Is An Important Determinant Of N-Linked Core Glycosylation Efficiency. Biochemistry, 37, 6833-6837.

Meltzer, M. & Franklin, E. C. 1966. Cryoglobulinemia--A Study Of Twenty-Nine Patients. I. Igg And Igm Cryoglobulins And Factors Affecting Cryoprecipitability. Am J Med, 40, 828-36.

Meltzer, M., Franklin, E. C., Elias, K., Mccluskey, R. T. & Cooper, N. 1966. Cryoglobulinemia--A Clinical And Laboratory Study. Ii. Cryoglobulins With Rheumatoid Factor Activity. Am J Med, 40, 837-56.

Middaugh, C. R. & Litman, G. W. 1978. Investigations Of The Molecular Basis For The Temperature-Dependent Insolubility Of Cryoglobulins. Vi. Quenching By Acrylamide Of The Intrinsic Tryptophan Fluorescence Of Cryoglobulin And Non-Cryoglobulin Igm Proteins. Biochim Biophys Acta, 535, 33-43.

Miller, N., Vile, R. G. & Hart, I. R. 1996. Selectin Ligands On Human Melanoma Cells. Glycoconj J, 13, 33-43.

Mimura, Y., Ashton, P. R., Takahashi, N., Harvey, D. J. & Jefferis, R. 2007. Contrasting Glycosylation Profiles Between Fab And Fc Of A Human Igg Protein Studied By Electrospray Ionization Mass Spectrometry. J Immunol Methods, 326, 116-26.

Miyoshi, E., Moriwaki, K. & Nakagawa, T. 2008. Biological Function Of Fucosylation In Cancer Biology. J Biochem, 143, 725-9.

Mizuochi, T., Pastore, Y., Shikata, K., Kuroki, A., Kikuchi, S., Fulpius, T., Nakata, M., Fossati-Jimack, L., Reininger, L., Matsushita, M., Fujita, T. & Izui, S. 2001. Role Of Galactosylation In The Renal Pathogenicity Of Murine Immunoglobulin G3 Monoclonal Cryoglobulins. Blood, 97, 3537-43.

Moiseeva, E. P., Williams, B., Goodall, A. H. & Samani, N. J. 2003. Galectin-1 Interacts With Beta-1 Subunit Of Integrin. Biochem Biophys Res Commun, 310, 1010-6.

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.

Moore, K. L., Stults, N. L., Diaz, S., Smith, D. F., Cummings, R. D., Varki, A. & Mcever, R. P. 1992. Identification Of A Specific Glycoprotein Ligand For P-Selectin (Cd62) On Myeloid Cells. J Cell Biol, 118, 445-56.

Moore, T. & Dveksler, G. S. 2014. Pregnancy-Specific Glycoproteins: Complex Gene Families Regulating Maternal-Fetal Interactions. Int J Dev Biol, 58, 273-80.

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., 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.

Morrow, D. M., Xiong, N., Getty, R. R., Ratajczak, M. Z., Morgan, D., Seppala, M., Riittinen, L., Gewirtz, A. M. & Tykocinski, M. L. 1994. Hematopoietic Placental Protein 14: An Immunosuppressive Factor In Cells Of The Megakaryocytic Lineage. American Journal Of Pathology, 145, 1485-1495.

Muhlfeld, A. S., Segerer, S., Hudkins, K., Farr, A. G., Bao, L., Kraus, D., Holers, V. M., Quigg, R. J. & Alpers, C. E. 2004. Overexpression Of Complement Inhibitor Crry Does Not Prevent Cryoglobulin-Associated Membranoproliferative Glomerulonephritis. Kidney Int, 65, 1214-23.

Musset, L., Diemert, M. C., Taibi, F., Thi Huong Du, L., Cacoub, P., Leger, J. M., Boissy, G., Gaillard, O. & Galli, J. 1992. Characterization Of Cryoglobulins By Immunoblotting. Clin Chem, 38, 798-802.

Nagalla, S. R., Snyder, C. K., Michaels, J. E., Laughlin, M. J., Roberts, C. T., Balaji, M., Balaji, V., Seshiah, V. & Rao, P. V. 2015. Maternal Serum Biomarkers For Risk Assessment In Gestational Diabetes. A Potential Universal Screening Test To Predict Gdm Status. Indian J Endocrinol Metab, 19, 155-9.

Nakamori, S., Kameyama, M., Imaoka, S., Furukawa, H., Ishikawa, O., Sasaki, Y., Kabuto, T., Iwanaga, T., Matsushita, Y. & Irimura, T. 1993. Increased Expression Of Sialyl Lewisx Antigen Correlates With Poor Survival In Patients With Colorectal Carcinoma: Clinicopathological And Immunohistochemical Study. Cancer Res, 53, 3632-7.

Nemansky, M., Thotakura, N. R., Lyons, C. D., Ye, S., Reinhold, B. B., Reinhold, V. N. & Blithe, D. L. 1998. Developmental Changes In The Glycosylation Of Glycoprotein Hormone Free Alpha Subunit During Pregnancy. J Biol Chem, 273, 12068-76.

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.

Nimmerjahn, F. & Ravetch, J. V. 2008. Fcgamma Receptors As Regulators Of Immune Responses. Nat Rev Immunol, 8, 34-47.

Nishiya, A. T., Massoco, C. O., Felizzola, C. R., Perlmann, E., Batschinski, K., Tedardi, M. V., Garcia, J. S., Mendonca, P. P., Teixeira, T. F. & Zaidan Dagli, M. L. 2016. Comparative Aspects Of Canine Melanoma. Vet Sci, 3.

Nishizumi, H., Taniuchi, I., Yamanashi, Y., Kitamura, D., Ilic, D., Mori, S., Watanabe, T. & Yamamoto, T. 1995. Impaired Proliferation Of Peripheral B Cells And Indication Of Autoimmune Disease In Lyn-Deficient Mice. Immunity, 3, 549-60.

Nisselbaum, J. S. & Bernfeld, P. 1956. The Properties Of Two Glycoproteins Isolated From The Plasma Of Normal And Tumor Bearing Mice. Journal Of The American Chemical Society, 78, 687-689.

Nitschke, L., Carsetti, R., Ocker, B., Kohler, G. & Lamers, M. C. 1997. Cd22 Is A Negative Regulator Of B-Cell Receptor Signalling. Curr Biol, 7, 133-43.

Norgard, K. E., Moore, K. L., Diaz, S., Stults, N. L., Ushiyama, S., Mcever, R. P., Cummings, R. D. & Varki, A. 1993. Characterization Of A Specific Ligand For P-Selectin On Myeloid Cells. A Minor Glycoprotein With Sialylated O-Linked Oligosaccharides. J Biol Chem, 268, 12764-74.

North, S. J., Huang, H. H., Sundaram, S., Jang-Lee, J., Etienne, A. T., Trollope, A., Chalabi, S., Dell, A., Stanley, P. & Haslam, S. M. 2010. Glycomics Profiling Of Chinese Hamster Ovary Cell Glycosylation Mutants Reveals N-Glycans Of A Novel Size And Complexity. J Biol Chem, 285, 5759-75.

O'brien, T. J., Beard, J. B., Underwood, L. J. & Shigemasa, K. 2002. The Ca 125 Gene: A Newly Discovered Extension Of The Glycosylated N-Terminal Domain Doubles The Size Of This Extracellular Superstructure. Tumor Biology, 23, 154-169.

O'keefe, T. L., Williams, G. T., Batista, F. D. & Neuberger, M. S. 1999. Deficiency In Cd22, A B Cell-Specific Inhibitory Receptor, Is Sufficient To Predispose To Development Of High Affinity Autoantibodies. J Exp Med, 189, 1307-13.

O'keefe, T. L., Williams, G. T., Davies, S. L. & Neuberger, M. S. 1996. Hyperresponsive B Cells In Cd22-Deficient Mice. Science, 274, 798-801.

O'riordan, R. T. 2014. Pregnancy-Specific Glycoprotein Function, Conservation And Receptor Investigation. University College Cork.

Ochwat, D., Hoja-Lukowicz, D. & Litynska, A. 2004. N-Glycoproteins Bearing Beta1-6 Branched Oligosaccharides From The A375 Human Melanoma Cell Line Analysed By Tandem Mass Spectrometry. Melanoma Res, 14, 479-85.

Okamoto, N., Uchida, A., Takakura, K., Kariya, Y., Kanzaki, H., Riittinen, L., Koistinen, R., Seppälä, M. & Mori, T. 1991. Suppression By Human Placental Protein 14 Of Natural Killer Cell Activity. American Journal Of Reproductive Immunology, 26, 137-142.

Olivari, S. & Molinari, M. 2007. Glycoprotein Folding And The Role Of Edem1, Edem2 And Edem3 In Degradation Of Folding-Defective Glycoproteins. Febs Lett, 581, 3658-64.

Osborne, J. C., Jr., Rosen, S. W., Nilsson, B., Calvert, I. & Bohn, H. 1982. Physicochemical Studies Of Pregnancy-Specific Beta 1-Glycoprotein: Unusual Ultracentrifugal And Circular Dichroic Properties. Biochemistry, 21, 5523-8.

Otani, M., Kuroki, A., Kikuchi, S., Kihara, M., Nakata, J., Ito, K., Furukawa, J., Shinohara, Y. & Izui, S. 2012. Sialylation Determines The Nephritogenicity Of Igg3 Cryoglobulins. J Am Soc Nephrol, 23, 1869-78.

Overwijk, W. W. & Restifo, N. P. 2001. B16 As A Mouse Model For Human Melanoma. Current Protocols In Immunology, Chapter 20, Unit-20.1.

Oz, O. K., Campbell, A. & Tao, T. W. 1989. Reduced Cell-Adhesion To Fibronectin And Laminin Is Associated With Altered Glycosylation Of Beta-1 Integrins In A Weakly Metastatic Glycosylation Mutant. International Journal Of Cancer, 44, 343-347.

Paclik, D., Danese, S., Berndt, U., Wiedenmann, B., Dignass, A. & Sturm, A. 2008. Galectin-4 Controls Intestinal Inflammation By Selective Regulation Of Peripheral And Mucosal T Cell Apoptosis And Cell Cycle. Plos One, 3, E2629.

Pang, P. C., Haslam, S. M., Dell, A. & Clark, G. F. 2016. The Human Fetoembryonic Defense System Hypothesis: Twenty Years On. Mol Aspects Med, 51, 71-88.

Panka, D. J. 1997. Glycosylation Is Influential In Murine Igg3 Self-Association. Mol Immunol, 34, 593-8.

Panka, D. J., Salant, D. J., Jacobson, B. A., Minto, A. W. & Marshak-Rothstein, A. 1995. The Effect Of Vh Residues 6 And 23 On Igg3 Cryoprecipitation And Glomerular Deposition. Eur J Immunol, 25, 279-84.

Parekh, R. B., Dwek, R. A., Sutton, B. J., Fernandes, D. L., Leung, A., Stanworth, D., Rademacher, T. W., Mizuochi, T., Taniguchi, T., Matsuta, K. & Et Al. 1985. Association Of Rheumatoid Arthritis And Primary Osteoarthritis With Changes In The Glycosylation Pattern Of Total Serum Igg. Nature, 316, 452-7.

Park, J. S., Withers, S. S., Modiano, J. F., Kent, M. S., Chen, M., Luna, J. I., Culp, W. T. N., Sparger, E. E., Rebhun, R. B., Monjazeb, A. M., Murphy, W. J. & Canter, R. J. 2016. Canine Cancer Immunotherapy Studies:

Linking Mouse And Human. J Immunother Cancer, 4, 97.

Pasek, M., Duk, M., Podbielska, M., Sokolik, R., Szechinski, J., Lisowska, E. & Krotkiewski, H. 2006. Galactosylation Of Igg From Rheumatoid Arthritis (Ra) Patients--Changes During Therapy. Glycoconj J, 23, 463-71.

Patankar, M., Oehninger, S. G. & Clark, G. F. 1993. The Role Of Glycococonjugates In Human Sperm–Zona Pellucida Binding. Glycobiology, 3, 521.

Patankar, M. S., Jing, Y., Morrison, J. C., Belisle, J. A., Lattanzio, F. A., Deng, Y., Wong, N. K., Morris, H. R., Dell, A. & Clark, G. F. 2005. Potent Suppression Of Natural Killer Cell Response Mediated By The Ovarian Tumor Marker Ca125. Gynecologic Oncology, 99, 704-713.

Paul, W. S., H 1960. Apparatus For Separating Charged Particles Of Different Specific Charges. United States Patent Application 2939952.

Perillo, N. L., Marcus, M. E. & Baum, L. G. 1998. Galectins: Versatile Modulators Of Cell Adhesion, Cell Proliferation, And Cell Death. J Mol Med (Berl), 76, 402-12.

Perillo, N. L., Pace, K. E., Seilhamer, J. J. & Baum, L. G. 1995. Apoptosis Of T Cells Mediated By Galectin-1. Nature, 378, 736-9.

Perillo, N. L., Uittenbogaart, C. H., Nguyen, J. T. & Baum, L. G. 1997. Galectin-1, An Endogenous Lectin Produced By Thymic Epithelial Cells, Induces Apoptosis Of Human Thymocytes. J Exp Med, 185, 1851-8.

Picanco-Castro, V. & Swiech, K. 2018. Recombinant Glycoprotein Production : Methods And Protocols.
Pillai, S., Netravali, I. A., Cariappa, A. & Mattoo, H. 2012. Siglecs And Immune Regulation. Annu Rev Immunol, 30, 357-92.

Pinho, S. S. & Reis, C. A. 2015. Glycosylation In Cancer: Mechanisms And Clinical Implications. Nat Rev Cancer, 15, 540-55.

Plummer, T. H., Jr., Elder, J. H., Alexander, S., Phelan, A. W. & Tarentino, A. L. 1984. Demonstration Of Peptide:n-Glycosidase F Activity In Endo-Beta-N-Acetylglucosaminidase F Preparations. J Biol Chem, 259, 10700-4.

Pochec, E., Rydlewska, M., Przybylo, M. & Litynska, A. 2015. Diverse Expression Of N-Acetylglucosaminyltransferase V And Complex-Type Beta1,6-Branched N-Glycans In Uveal And Cutaneous Melanoma Cells. Acta Biochim Pol, 62, 323-8.

Pockley, A. G. & Bolton, A. E. 1989. Placental Protein 14 (Pp14) Inhibits The Synthesis Of Interleukin-2 And The Release Of Soluble Interleukin-2 Receptors From Phytohaemagglutinin-Stimulated Lymphocytes. Clinical And Experimental Immunology, 77, 252-256.

Pockley, A. G. & Bolton, A. E. 1990. The Effect Of Human Placental Protein 14 (Pp14) On The Production Of Interleukin-1 From Mitogenically Stimulated Mononuclear Cell Cultures. Immunology, 69, 277-281.

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.

Przybylo, M., Pochec, E., Link-Lenczowski, P. & Litynska, A. 2008. Beta1-6 Branching Of Cell Surface Glycoproteins May Contribute To Uveal Melanoma Progression By Up-Regulating Cell Motility. Mol Vis, 14, 625-36.

Quast, I., Keller, C. W., Maurer, M. A., Giddens, J. P., Tackenberg, B., Wang, L. X., Munz, C., Nimmerjahn, F., Dalakas, M. C. & Lunemann, J. D. 2015. Sialylation Of Igg Fc Domain Impairs Complement-Dependent Cytotoxicity. J Clin Invest, 125, 4160-70.

Rabinovich, G. A., Liu, F. T., Hirashima, M. & Anderson, A. 2007. An Emerging Role For Galectins In Tuning The Immune Response: Lessons From Experimental Models Of Inflammatory Disease, Autoimmunity And Cancer. Scand J Immunol, 66, 143-58.

Rabinovich, G. A. & Toscano, M. A. 2009. Turning 'Sweet' On Immunity: Galectin-Glycan Interactions In Immune Tolerance And Inflammation. Nat Rev Immunol, 9, 338-52.

Rapoport, E., Mikhalyov, I., Zhang, J., Crocker, P. & Bovin, N. 2003. Ganglioside Binding Pattern Of Cd33-Related Siglecs. Bioorganic And Medicinal Chemistry Letters, 13, 675-678.

Ravetch, J. V. & Lanier, L. L. 2000. Immune Inhibitory Receptors. Science, 290, 84-9.

Raymond, C., Robotham, A., Spearman, M., Butler, M., Kelly, J. & Durocher, Y. 2015. Production Of Alpha2,6-Sialylated Igg1 In Cho Cells. Mabs, 7, 571-83.

Roberts, A. A., Amano, M., Felten, C., Galvan, M., Sulur, G., Pinter-Brown, L., Dobbeling, U., Burg, G., Said, J. & Baum, L. G. 2003. Galectin-1-Mediated Apoptosis In Mycosis Fungoides: The Roles Of Cd7 And Cell Surface Glycosylation. Mod Pathol, 16, 543-51.

Roepstorff, P. & Fohlman, J. 1984. Proposal For A Common Nomenclature For Sequence Ions In Mass Spectra Of Peptides. Biomed Mass Spectrom, 11, 601.

Roopenian, D. C. & Akilesh, S. 2007. Fcrn: The Neonatal Fc Receptor Comes Of Age. Nat Rev Immunol, 7, 715-25.

Rosen, S., Singer, M., Yednock, T. & Stoolman, L. 1985. Involvement Of Sialic Acid On Endothelial Cells In Organ-Specific Lymphocyte Recirculation. Science, 228, 1005-1007.

Rosen, S. D., Chi, S. I., True, D. D., Singer, M. S. & Yednock, T. A. 1989. Intravenously Injected Sialidase Inactivates Attachment Sites For Lymphocytes On High Endothelial Venules. J Immunol, 142, 1895-902.

Rout, U. K., Wang, J., Paria, B. C. & Armant, D. R. 2004. Alpha5beta1, Alphavbeta3 And The Platelet-Associated Integrin Alphaiibbeta3 Coordinately Regulate Adhesion And Migration Of Differentiating Mouse Trophoblast Cells. Dev Biol, 268, 135-51.

Ruan, S. & Lloyd, K. O. 1992. Glycosylation Pathways In The Biosynthesis Of Gangliosides In Melanoma And Neuroblastoma Cells: Relative Glycosyltransferase Levels Determine Ganglioside Patterns. Cancer Res, 52, 5725-31.

Ruiz-Canada, C., Kelleher, D. J. & Gilmore, R. 2009. Cotranslational And Posttranslational N-Glycosylation Of Polypeptides By Distinct Mammalian Ost Isoforms. Cell, 136, 272-83.

Rutherfurd, K. J., Chou, J. Y. & Mansfield, B. C. 1995. A Motif In Psg11s Mediates Binding To A Receptor On The Surface Of The Promonocyte Cell Line Thp-1. Mol Endocrinol, 9, 1297-305.

Rye, P. D., Fodstad, O., Emilsen, E. & Bryne, M. 1998. Invasion Potential And N-Acetylgalactosamine Expression In A Human Melanoma Model. Int J Cancer, 75, 609-14.

Rymen, D., Keldermans, L., Race, V., Regal, L., Deconinck, N., Dionisi-Vici, C., Fung, C. W., Sturiale, L., Rosnoblet, C., Foulquier, F., Matthijs, G. & Jaeken, J. 2012. Cog5-Cdg: Expanding The Clinical Spectrum. Orphanet J Rare Dis, 7, 94.

Sako, D., Chang, X. J., Barone, K. M., Vachino, G., White, H. M., Shaw, G., Veldman, G. M., Bean, K. M., Ahern, T. J., Furie, B. & Et Al. 1993. Expression Cloning Of A Functional Glycoprotein Ligand For P-Selectin. Cell, 75, 1179-86.

Sancho, D. & Reis E Sousa, C. 2012. Signaling By Myeloid C-Type Lectin Receptors In Immunity And Homeostasis. Annu Rev Immunol, 30, 491-529.

Scallon, B. J., Tam, S. H., Mccarthy, S. G., Cai, A. N. & Raju, T. S. 2007. Higher Levels Of Sialylated Fc Glycans In Immunoglobulin G Molecules Can Adversely Impact Functionality. Mol Immunol, 44, 1524-34.

Scarborough, R. M., Rose, J. W., Hsu, M. A., Phillips, D. R., Fried, V. A., Campbell, A. M., Nannizzi, L. & Charo, I. F. 1991. Barbourin. A Gpiib-Iiia-Specific Integrin Antagonist From The Venom Of Sistrurus M. Barbouri. J Biol Chem, 266, 9359-62.

Schadendorf, D., Van Akkooi, A. C. J., Berking, C., Griewank, K. G., Gutzmer, R., Hauschild, A., Stang, A., Roesch, A. & Ugurel, S. 2018. Melanoma. Lancet, 392, 971-984.

Schallus, T., Jaeckh, C., Feher, K., Palma, A. S., Liu, Y., Simpson, J. C., Mackeen, M., Stier, G., Gibson, T. J., Feizi, T., Pieler, T. & Muhle-Goll, C. 2008. Malectin: A Novel Carbohydrate-Binding Protein Of The Endoplasmic Reticulum And A Candidate Player In The Early Steps Of Protein N-Glycosylation. Mol Biol Cell, 19, 3404-14.

Schnaar, R. L. & Kinoshita, T. 2015. Glycosphingolipids. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Scudder, P., Uemura, K., Dolby, J., Fukuda, M. N. & Feizi, T. 1983. Isolation And Characterization Of An Endo-Beta-Galactosidase From Bacteroides Fragilis. Biochem J, 213, 485-94.

Seeberger, P. H. 2015. Monosaccharide Diversity. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Seite, J. F., Cornec, D., Renaudineau, Y., Youinou, P., Mageed, R. A. & Hillion, S. 2010. Ivig Modulates Bcr Signaling Through Cd22 And Promotes Apoptosis In Mature Human B Lymphocytes. Blood, 116, 1698-704.

Sgroi, D., Varki, A., Braesch-Andersen, S. & Stamenkovic, I. 1993. Cd22, A B Cell-Specific Immunoglobulin Superfamily Member, Is A Sialic Acid-Binding Lectin. J Biol Chem, 268, 7011-8.

Shain, A. H., Yeh, I., Kovalyshyn, I., Sriharan, A., Talevich, E., Gagnon, A., Dummer, R., North, J., Pincus, L., Ruben, B., Rickaby, W., D'arrigo, C., Robson, A. & Bastian, B. C. 2015. The Genetic Evolution Of Melanoma From Precursor Lesions. N Engl J Med, 373, 1926-36.

Shanley, D. K., Kiely, P. A., Golla, K., Allen, S., Martin, K., O'riordan, R. T., Ball, M., Aplin, J. D., Singer, B. B., Caplice, N., Moran, N. & Moore, T. 2013. Pregnancy-Specific Glycoproteins Bind Integrin Alphaiibbeta3 And Inhibit The Platelet-Fibrinogen Interaction. Plos One, 8, E57491.

Shen, L., Luo, Z., Wu, J., Qiu, L., Luo, M., Ke, Q. & Dong, X. 2017. Enhanced Expression Of Alpha2,3-Linked Sialic Acids Promotes Gastric Cancer Cell Metastasis And Correlates With Poor Prognosis. Int J Oncol, 50,

1201-1210.

Shihabi, Z. K. 1996. Analysis And General Classification Of Serum Cryoglobulins By Capillary Zone Electrophoresis. Electrophoresis, 17, 1607-12.

Sims, J. E., Williams, D. E., Morrissey, P. J., Garka, K., Foxworthe, D., Price, V., Friend, S. L., Farr, A., Bedell, M. A., Jenkins, N. A., Copeland, N. G., Grabstein, K. & Paxton, R. J. 2000. Molecular Cloning And Biological Characterization Of A Novel Murine Lymphoid Growth Factor. J Exp Med, 192, 671-80.

Sindhura, B. R., Hegde, P., Chachadi, V. B., Inamdar, S. R. & Swamy, B. M. 2017. High Mannose N-Glycan Binding Lectin From Remusatia Vivipara (Rvl) Limits Cell Growth, Motility And Invasiveness Of Human Breast Cancer Cells. Biomed Pharmacother, 93, 654-665.

Sjoberg, E. R., Powell, L. D., Klein, A. & Varki, A. 1994. Natural Ligands Of The B Cell Adhesion Molecule Cd22 Beta Can Be Masked By 9-O-Acetylation Of Sialic Acids. J Cell Biol, 126, 549-62.

Slavin, R. G., Suriano, J. R. & Dreesman, G. 1971. Studies On Cryogelglobulinemia Associated With Iga Myeloma Proteins. Int Arch Allergy Appl Immunol, 40, 739-48.

Snyder, S. K., Wessells, J. L., Waterhouse, R. M., Dveksler, G. S., Wessner, D. H., Wahl, L. M. & Zimmermann, W. 2001. Pregnancy-Specific Glycoproteins Function As Immunomodulators By Inducing Secretion Of Il-10, Il-6 And Tgf-B1 By Human Monocytes. 45, 205-216.

Somer, T. 1966. The Viscosity Of Blood, Plasma And Serum In Dys- And Paraproteinemias. Acta Med Scand Suppl, 456, 1-97.

Sousa, M. & Parodi, A. J. 1995. The Molecular Basis For The Recognition Of Misfolded Glycoproteins By The Udp-Glc:glycoprotein Glucosyltransferase. Embo J, 14, 4196-203.

Sperandio, M., Gleissner, C. A. & Ley, K. 2009. Glycosylation In Immune Cell Trafficking. Immunol Rev, 230, 97-113.

Srinivasan, N., Bane, S. M., Ahire, S. D., Ingle, A. D. & Kalraiya, R. D. 2009. Poly N-Acetyllactosamine Substitutions On N- And Not O-Oligosaccharides Or Thomsen-Friedenreich Antigen Facilitate Lung Specific Metastasis Of Melanoma Cells Via Galectin-3. Glycoconj J, 26, 445-56.

Stamper, H. B., Jr. & Woodruff, J. J. 1977. An In Vitro Model Of Lymphocyte Homing. I. Characterization Of The Interaction Between Thoracic Duct Lymphocytes And Specialized High-Endothelial Venules Of Lymph Nodes. J Immunol, 119, 772-80.

Stanley, P., Taniguchi, N. & Aebi, M. 2015. N-Glycans. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Stei, M. M., Loeffler, K. U., Holz, F. G. & Herwig, M. C. 2016. Animal Models Of Uveal Melanoma: Methods, Applicability, And Limitations. Biomed Res Int, 2016, 4521807.

Steinman, R. M. 2000. Dc-Sign: A Guide To Some Mysteries Of Dendritic Cells. Cell, 100, 491-4.

Stephens, W. E. 1946. A Pulsed Mass Spectrometer With Time Dispersion. Bull. Am. Phys. Soc.

Stillman, B. N., Hsu, D. K., Pang, M., Brewer, C. F., Johnson, P., Liu, F. T. & Baum, L. G. 2006. Galectin-3 And Galectin-1 Bind Distinct Cell Surface Glycoprotein Receptors To Induce T Cell Death. J Immunol, 176,

778-89.

Stoolman, L. M. & Rosen, S. D. 1983. Possible Role For Cell-Surface Carbohydrate-Binding Molecules In Lymphocyte Recirculation. J Cell Biol, 96, 722-9.

Stoolman, L. M., Tenforde, T. S. & Rosen, S. D. 1984. Phosphomannosyl Receptors May Participate In The Adhesive Interaction Between Lymphocytes And High Endothelial Venules. J Cell Biol, 99, 1535-40.

Strohl, W. R. 2009. Optimization Of Fc-Mediated Effector Functions Of Monoclonal Antibodies. Curr Opin Biotechnol, 20, 685-91.

Sturm, A., Lensch, M., Andre, S., Kaltner, H., Wiedenmann, B., Rosewicz, S., Dignass, A. U. & Gabius, H. J. 2004. Human Galectin-2: Novel Inducer Of T Cell Apoptosis With Distinct Profile Of Caspase Activation. J Immunol, 173, 3825-37.

Subedi, G. P., Hanson, Q. M. & Barb, A. W. 2014. Restricted Motion Of The Conserved Immunoglobulin G1 N-Glycan Is Essential For Efficient Fcgammariiia Binding. Structure, 22, 1478-88.

Svennerholm, L. 1963. Chromatographic Separation Of Human Brain Gangliosides. J Neurochem, 10, 613-23.

Swanson, B. J., Mcdermott, K. M., Singh, P. K., Eggers, J. P., Crocker, P. R. & Hollingsworth, M. A. 2007. Muc1 Is A Counter-Receptor For Myelin-Associated Glycoprotein (Siglec-4A) And Their Interaction Contributes To Adhesion In Pancreatic Cancer Perineural Invasion. Cancer Res, 67, 10222-9.

Sweeney, J. G., Liang, J., Antonopoulos, A., Giovannone, N., Kang, S., Mondala, T. S., Head, S. R., King, S. L., Tani, Y., Brackett, D., Dell, A., Murphy, G. F., Haslam, S. M., Widlund, H. R. & Dimitroff, C. J. 2018. Loss Of Gcnt2/I-Branched Glycans Enhances Melanoma Growth And Survival. Nat Commun, 9, 3368.

Taguchi, T., Kitajima, K., Muto, Y., Yokoyama, S., Inoue, S. & Inoue, Y. 1995. Proton Nmr Study Of The Trimannosyl Unit In A Pentaantennary N-Linked Decasaccharide Structure. Complete Assignment Of The Proton Resonances And Conformational Characterization. Eur J Biochem, 228, 822-9.

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.

Takematsu, H., Diaz, S., Stoddart, A., Zhang, Y. & Varki, A. 1999. Lysosomal And Cytosolic Sialic Acid 9-O-Acetylesterase Activities Can Be Encoded By One Gene Via Differential Usage Of A Signal Peptide-Encoding Exon At The N Terminus. J Biol Chem, 274, 25623-31.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. & Matsuo, 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.

Taneda, S., Segerer, S., Hudkins, K. L., Cui, Y., Wen, M., Segerer, M., Wener, M. H., Khairallah, C. G., Farr, A. G. & Alpers, C. E. 2001. Cryoglobulinemic Glomerulonephritis In Thymic Stromal Lymphopoietin Transgenic Mice. Am J Pathol, 159, 2355-69.

Tejwani, V., Andersen, M. R., Nam, J. H. & Sharfstein, S. T. 2018. Glycoengineering In Cho Cells: Advances In Systems Biology. Biotechnol J, 13, E1700234.

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.

Testa, U., Castelli, G. & Pelosi, E. 2017. Melanoma: Genetic Abnormalities, Tumor Progression, Clonal Evolution And Tumor Initiating Cells. Med Sci (Basel), 5.

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 Lett, 583, 1728-35.

Todeschini, A. R., Dos Santos, J. N., Handa, K. & Hakomori, S. I. 2007. Ganglioside Gm2-Tetraspanin Cd82 Complex Inhibits Met And Its Cross-Talk With Integrins, Providing A Basis For Control Of Cell Motility Through Glycosynapse. J Biol Chem, 282, 8123-33.

Tribulatti, M. V., Cattaneo, V., Hellman, U., Mucci, J. & Campetella, O. 2009. Galectin-8 Provides Costimulatory And Proliferative Signals To T Lymphocytes. J Leukoc Biol, 86, 371-80.

Tribulatti, M. V., Mucci, J., Cattaneo, V., Aguero, F., Gilmartin, T., Head, S. R. & Campetella, O. 2007. Galectin-8 Induces Apoptosis In The Cd4(High)Cd8(High) Thymocyte Subpopulation. Glycobiology, 17, 1404-12.

Umana, P., Jean-Mairet, J., Moudry, R., Amstutz, H. & Bailey, J. E. 1999. Engineered Glycoforms Of An Antineuroblastoma Igg1 With Optimized Antibody-Dependent Cellular Cytotoxic Activity. Nat Biotechnol, 17, 176-80.

Valmu, L., Alfthan, H., Hotakainen, K., Birken, S. & Stenman, U. H. 2006. Site-Specific Glycan Analysis Of Human Chorionic Gonadotropin Beta-Subunit From Malignancies And Pregnancy By Liquid Chromatography--Electrospray Mass Spectrometry. Glycobiology, 16, 1207-18.

Van De Geijn, F. E., Wuhrer, M., Selman, M. H., Willemsen, S. P., De Man, Y. A., Deelder, A. M., Hazes, J. M. & Dolhain, R. J. 2009. Immunoglobulin G Galactosylation And Sialylation Are Associated With Pregnancy-Induced Improvement Of Rheumatoid Arthritis And The Postpartum Flare: Results From A Large Prospective Cohort Study. Arthritis Res Ther, 11, R193.

Van Den Berg, T. K., Honing, H., Franke, N., Van Remoortere, A., Schiphorst, W. E., Liu, F. T., Deelder, A. M., Cummings, R. D., Hokke, C. H. & Van Die, I. 2004. Lacdinac-Glycans Constitute A Parasite Pattern For Galectin-3-Mediated Immune Recognition. J Immunol, 173, 1902-7.

Van Gisbergen, K. P., Geijtenbeek, T. B. & Van Kooyk, Y. 2005. Close Encounters Of Neutrophils And Dcs. Trends Immunol, 26, 626-31.

Van Kooyk, Y. & Rabinovich, G. A. 2008. Protein-Glycan Interactions In The Control Of Innate And Adaptive Immune Responses. Nat Immunol, 9, 593-601.

Varki, A. 2007. Glycan-Based Interactions Involving Vertebrate Sialic-Acid-Recognizing Proteins. Nature, 446, 1023-9.

Varki, A. & Kornfeld, S. 2015. Historical Background And Overview. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Varki, A., Schnaar, R. L. & Crocker, P. R. 2015. I-Type Lectins. In: Rd, Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. (Eds.) Essentials Of Glycobiology. Cold Spring Harbor (Ny).

Von Lampe, B., Stallmach, A. & Riecken, E. O. 1993. Altered Glycosylation Of Integrin Adhesion Molecules In Colorectal Cancer Cells And Decreased Adhesion To The Extracellular Matrix. Gut, 34, 829-36.

Walker, G. J., Soyer, H. P., Terzian, T. & Box, N. F. 2011. Modelling Melanoma In Mice. Pigment Cell Melanoma Res, 24, 1158-76.

Walther, T., Karamanska, R., Chan, R. W., Chan, M. C., Jia, N., Air, G., Hopton, C., Wong, M. P., Dell, A., Malik Peiris, J. S., Haslam, S. M. & Nicholls, J. M. 2013. Glycomic Analysis Of Human Respiratory Tract Tissues And Correlation With Influenza Virus Infection. Plos Pathog, 9, E1003223.

Walz, G., Aruffo, A., Kolanus, W., Bevilacqua, M. & Seed, B. 1990. Recognition By Elam-1 Of The Sialyl-Lex Determinant On Myeloid And Tumor Cells. Science, 250, 1132-5.

Wang, F., Nakouzi, A., Angeletti, R. H. & Casadevall, A. 2003. Site-Specific Characterization Of The N-Linked Oligosaccharides Of A Murine Immunoglobulin M By High-Performance Liquid Chromatography/Electrospray Mass Spectrometry. Anal Biochem, 314, 266-80.

Weis, W. I. & Drickamer, K. 1996. Structural Basis Of Lectin-Carbohydrate Recognition. Annu Rev Biochem, 65, 441-73.

Weis, W. I., Taylor, M. E. & Drickamer, K. 1998. The C-Type Lectin Superfamily In The Immune System. Immunol Rev, 163, 19-34.

Wiley, W. C. & Mclaren, I. H. 1955. Time-Of-Flight Mass Spectrometer With Improved Resolution. Review Of Scientific Instruments, 26, 1150-1157.

Wilkins, P. P., Moore, K. L., Mcever, R. P. & Cummings, R. D. 1995. Tyrosine Sulfation Of P-Selectin Glycoprotein Ligand-1 Is Required For High Affinity Binding To P-Selectin. J Biol Chem, 270, 22677-80.

Wilson, C. M., Roebuck, Q. & High, S. 2008. Ribophorin I Regulates Substrate Delivery To The Oligosaccharyltransferase Core. Proc Natl Acad Sci U S A, 105, 9534-9.

Wilson, T. J., Fuchs, A. & Colonna, M. 2012. Cutting Edge: Human Fcrl4 And Fcrl5 Are Receptors For Iga And Igg. J Immunol, 188, 4741-5.

Wintrobe., M. & Buell., M. 1933. Hyperproteinemia Associated With Multiple Myeloma:with Report Of A Case In Which An Extraordinary Hyperproteinemia Was Associated With Thrombosis Of The Retinal Veins And Symptoms Suggesting Raynaud's Disease. Bulletin Of The Johns Hopkins Hospital, 156-165.

Wong, A. H., Fukami, Y., Sudo, M., Kokubun, N., Hamada, S. & Yuki, N. 2016. Sialylated Igg-Fc: A Novel Biomarker Of Chronic Inflammatory Demyelinating Polyneuropathy. J Neurol Neurosurg Psychiatry, 87, 275-9.

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.

Wright, J. F., Shulman, M. J., Isenman, D. E. & Painter, R. H. 1990. C1 Binding By Mouse Igm. The Effect Of Abnormal Glycosylation At Position 402 Resulting From A Serine To Asparagine Exchange At Residue 406 Of The Mu-Chain. J Biol Chem, 265, 10506-13.

Wu, H., Shi, X. L., Zhang, H. J., Song, Q. J., Yang, X. B., Hu, W. D., Mei, G. L., Chen, X., Mao, Q. S. & Chen, Z. 2016A. Overexpression Of St3gal-I Promotes Migration And Invasion Of Hcclm3 In Vitro And Poor Prognosis In Human Hepatocellular Carcinoma. Onco Targets Ther, 9, 2227-36.

Wu, Y., Lan, C., Ren, D. & Chen, G. Y. 2016B. Induction Of Siglec-1 By Endotoxin Tolerance Suppresses The

Innate Immune Response By Promoting Tgf-Beta1 Production. J Biol Chem, 291, 12370-82.

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.

Yamashiro, S., Ruan, S., Furukawa, K., Tai, T., Lloyd, K. O., Shiku, H. & Furukawa, K. 1993. Genetic And Enzymatic Basis For The Differential Expression Of Gm2 And Gd2 Gangliosides In Human Cancer Cell Lines. Cancer Res, 53, 5395-400.

Yamashita, K., Kamerling, J. P. & Kobata, A. 1982. Structural Study Of The Carbohydrate Moiety Of Hen Ovomucoid. Occurrence Of A Series Of Pentaantennary Complex-Type Asparagine-Linked Sugar Chains. J Biol Chem, 257, 12809-14.

Yamashita, M. & Fenn, J. B. 1984. Electrospray Ion Source. Another Variation On The Free-Jet Theme. The Journal Of Physical Chemistry, 88, 4451-4459.

Yaniv, E., Borovsky, Z., Mishan-Eisenberg, G. & Rachmilewitz, J. 2003. Placental Protein 14 Regulates Selective B Cell Responses. Cellular Immunology, 222, 156-163.

Yeh, J. C., Hiraoka, N., Petryniak, B., Nakayama, J., Ellies, L. G., Rabuka, D., Hindsgaul, O., Marth, J. D., Lowe, J. B. & Fukuda, M. 2001. Novel Sulfated Lymphocyte Homing Receptors And Their Control By A Core1 Extension Beta 1,3-N-Acetylglucosaminyltransferase. Cell, 105, 957-69.

Yoo, E. M., Coloma, M. J., Trinh, K. R., Nguyen, T. Q., Vuong, L. U., Morrison, S. L. & Chintalacharuvu, K. R. 1999. Structural Requirements For Polymeric Immunoglobulin Assembly And Association With J Chain. J Biol Chem, 274, 33771-7.

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.

Yu, S. Y., Wu, S. W., Hsiao, H. H. & Khoo, K. H. 2009. Enabling Techniques And Strategic Workflow For Sulfoglycomics Based On Mass Spectrometry Mapping And Sequencing Of Permethylated Sulfated Glycans. Glycobiology, 19, 1136-49.

Yu, X., Baruah, K., Harvey, D. J., Vasiljevic, S., Alonzi, D. S., Song, B. D., Higgins, M. K., Bowden, T. A., Scanlan, C. N. & Crispin, M. 2013. Engineering Hydrophobic Protein-Carbohydrate Interactions To Fine-Tune Monoclonal Antibodies. J Am Chem Soc, 135, 9723-32.

Zamze, S., Wing, D. R., Wormald, M. R., Hunter, A. P., Dwek, R. A. & Harvey, D. J. 2001. A Family Of Novel, Acidic N-Glycans In Bowes Melanoma Tissue Plasminogen Activator Have L2/Hnk-1-Bearing Antennae, Many With Sulfation Of The Fucosylated Chitobiose Core. Eur J Biochem, 268, 4063-78.

Zelensky, A. N. & Gready, J. E. 2005. The C-Type Lectin-Like Domain Superfamily. Febs J, 272, 6179-217.

Zhang, J. Q., Nicoll, G., Jones, C. & Crocker, P. R. 2000. Siglec-9, A Novel Sialic Acid Binding Member Of The Immunoglobulin Superfamily Expressed Broadly On Human Blood Leukocytes. J Biol Chem, 275, 22121-6.

Zhong, X., Ma, W., Meade, C. L., Tam, A. S., Llewellyn, E., Cornell, R., Cote, K., Scarcelli, J. J., Marshall, J. K., Tzvetkova, B., Figueroa, B., Dinino, D., Sievers, A., Lee, C., Guo, J., Mahan, E., Francis, C., Lam, K., D'antona, A. M., Zollner, R., Zhu, H. L., Kriz, R., Somers, W. & Lin, L. 2018. Transient Cho Expression Platform For Robust Antibody Production And Its Enhanced N-Glycan Sialylation On Therapeutic Glycoproteins. Bio-

technol Prog.

Zhu, C., Anderson, A. C., Schubart, A., Xiong, H., Imitola, J., Khoury, S. J., Zheng, X. X., Strom, T. B. & Kuchroo, V. K. 2005. The Tim-3 Ligand Galectin-9 Negatively Regulates T Helper Type 1 Immunity. Nat Immunol, 6, 1245-52.

Zhu, D., Mccarthy, H., Ottensmeier, C. H., Johnson, P., Hamblin, T. J. & Stevenson, F. K. 2002. Acquisition Of Potential N-Glycosylation Sites In The Immunoglobulin Variable Region By Somatic Mutation Is A Distinctive Feature Of Follicular Lymphoma. Blood, 99, 2562-8.

Zhuo, Y. & Bellis, S. L. 2011. Emerging Role Of Alpha2,6-Sialic Acid As A Negative Regulator Of Galectin Binding And Function. J Biol Chem, 286, 5935-41.

Zöllner, O., Lenter, M. C., Blanks, J. E., Borges, E., Steegmaier, M., Zerwes, H. G. & Vestweber, D. 1997. L-Selectin From Human, But Not From Mouse Neutrophils Binds Directly To E-Selectin. The Journal Of Cell Biology, 136, 707-716.

Zufferey, R., Knauer, R., Burda, P., Stagljar, I., Te Heesen, S., Lehle, L. & Aebi, M. 1995. Stt3, A Highly Conserved Protein Required For Yeast Oligosaccharyl Transferase Activity In Vivo. Embo J, 14, 4949-60.

Appendix

Aminio acids and their residual molecular weight

Amino Acid	Three letter code	One letter code	Redisue mass(Da)
Alanine	Ala	A	71.037
Arginine	Arg	R	156.101
Asparagine	Asn	N	114.043
Aspartic acid	Asp	D	115.027
Cysteine	Cys	C	103.009
Glutamic acid	Glu	E	129.043
Glutamine	Gln	Q	128.059
Glycine	Gly	G	57.021
Histidine	His	H	137.059
Isoleucine	Ile	I	113.084
Leucine	Leu	L	113.084
Lysine	Lys	K	128.095
Methionine	Met	M	131.040
Phenylalanine	Phe	F	147.068
Proline	Pro	P	97.053
Serine	Ser	S	87.032
Threonine	Thr	T	101.048
Tryptophan	Trp	W	186.079
Tyrosine	Tyr	Y	163.063
Valine	Val	V	99.068