Glycoprotein Interactions Studied Using Electrospray Ionization Mass Spectrometry

by

Yilin Wang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry
University of Alberta

© Yilin Wang, 2019

# Abstract

This thesis focuses on the development and application of electrospray ionization mass spectrometry (ESI-MS) based methods for characterizing glycoproteins (GPs) and investigating their interactions with glycan binding proteins (GBPs) and drugs.

A multipronged approach was developed to identify and quantify water soluble GBP-GP interactions. First, the *catch-and-release* (CaR)−ESI−MS assay, implemented with ion mobility separation prior to GBP "release" (i.e., CaR$^{IMS}$−ESI−MS), was employed to rapidly identify GBP-GP binding in solution. The apparent affinity ($K_{a,app}$) of the GBP for the GP was then determined using the competitive *proxy ligand*-ESI−MS binding assay. Finally, screening of *N*-glycan libraries enzymatically released from the GPs against the GBP revealed the binding partners. Measurements performed at multiple GBP concentrations allow for affinity ranking of the released *N*-glycans (grouped as compositional isomers). This approach was demonstrated using the known interactions between a C-terminal domain fragment of human galectin-3 (hGal-3C) and three human serum GPs, α-1-acid glycoprotein (AGP), haptoglobin phenotype 1−1 (Hp1−1) and α-2-macroglobulin (α2M).

As an extension of this work, high-resolution MS was employed to characterize the microheterogeneity of AGP, following treatment with glycosidases, a sialidase, an α1,3-4 fucosidase and PNGase F. Using estimated glycoform concentrations, the affinities of warfarin, an anticoagulant drug, were measured by ESI-MS titration experiments. All major unfucosylated glycoforms of asialo-AGP showed similar binding affinities ($\sim 10^5$

$M^{-1}$) to warfarin, which is consistent with isothermal titration calorimetry (ITC) measurements. Binding properties of warfarin to PNGase F treated AGP, where *N*-glycans were partially or completely removed, were determined and compared. Unglycosylated AGP showed no binding to warfarin while binding increased with increasing level of AGP glycosylation, indicating the significant roles of *N*-glycans in promoting AGP recognition of warfarin.

# Preface

The research work presented in Chapter 2 of this thesis has been published as: Wang, Y.; Park, H.; Lin, H.; Kitova, E.N; Klassen, J.S. Multipronged ESI–MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins. *Anal. Chem*. **2019**, *91*(3), 2140-2147. I was responsible for data collection, analysis and manuscript preparation. Hong Lin performed preliminary experiments and assisted with optimizing experimental conditions and the composition of the manuscript. Heajin Park contributed to the extraction and analysis of *N*-glycans from glycoproteins. Elena N. Kitova performed parts of the *proxy-ligand* binding experiments and assisted with the manuscript edits. John S. Klassen was the supervisory author.

Chapter 3 includes preliminary data of an on-going project. I was responsible for performing experiments, processing data and writing. Elena N. Kitova assisted with data analysis.

# Acknowledgement

It is a genuine pleasure to express my sincere gratitude to my advisor Prof. John S. Klassen for the continuous support of my graduate research, for his patience, inspiration and guidance. I am also grateful to my supervisor and defense committee members, Prof. Todd L. Lowary, Prof. Liang Li and Prof. Matthew Macauley, for their support and insightful comments into my research.

I would like to extend my gratitude to Dr. Elena N. Kitova, Dr. Heajin Park, Dr. Pavel I. Kitov and Hong Lin, for their constructive advice, extensive knowledge and technical support. I would also like to thank Prof. Christopher Cairo for providing us with the human galectin-3 and Blake Zheng for his assistance in protein purification and analysis.

I also had a great pleasure of working with previous and present excellent colleagues in the Klassen group, Dr. Ling Han, Dr. Jun Li, Dr. Mickey Richards, Dr. Amr El-hawiet, Jianing Li, Yajie Chen, Zhixiong Li, Erick Baez, Linh Nguyen, Md Reazul Islam and Km Shams-Ud-Doha. I treasure all the stimulating discussions and the fun we have had in the last three years.

I would like to acknowledge the financial support from the Department of Chemistry, University of Alberta.

Finally, special thanks to my parents and my friend, Xu Hong Chen, for their endless encouragement and spiritual support.

# Table of Contents

**Interactions with Glycoproteins**

**Chapter 3 Influence of *N*-glycosylation on Glycoprotein Interactions Evaluated by Electrospray Ionization Mass Spectrometry**

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| *Ab* | Abundance of gas-phase ions |
| AC | Alternating current |
| ADC | Analogue-to-digital converter |
| AGP | α-1-acid glycoprotein |
| AT | Ion-mobility separation arrival time |
| BCA | Bovine carbonic anhydrase |
| BPH | Benign prostatic hyperplasia |
| CaR | *Catch-and-release* |
| CE | Capillary electrophoresis |
| CE | Capillary electrophoresis |
| CEM | Chain injection model |
| CID | Collision-induced dissociation |
| CIEF | Capillary isoelectrofocusing |
| CRM | Charged residue model |
| CZE | Capillary zone electrophoresis |
| DC | Direct current |
| DMSO | Dimethyl sulfoxide |
| DTT | Dithiothreitol |
| *Du* | Duty cycle |
| ECD | Electron-capture dissociation |
| ECM | Extracellular matrix |
| EDTA | Ethylenediaminetetraacetic acid |
| EOF | Electroosmotic flow |
| ER | Endoplasmic reticulum |
| ESI | Electrospray ionization |

| | |
|---|---|
| ETD | Electron-transfer dissociation |
| FAC-FD | Frontal affinity chromatography using fluorescence detection |
| fPSA | Free prostate specific antigen |
| Fuc | Fucose |
| GBP | Glycan binding protein |
| Glc | Glucose |
| GlcNAc | *N*-acetylglucosamine |
| GP | Glycoprotein |
| GPI | Glycosylphosphatidylinositol |
| HCD | Higher energy collisional dissociation |
| Hex | Hexose |
| HexNAc | *N*-acetylhexosamine |
| hGal-3C | C terminal domain of human galectin-3 |
| HILIC | Hydrophilic interaction chromatography |
| hK3 | Human   kallikrein 3 |
| Hp1-1 | haptoglobin phenotype 1-1 |
| HPAEC-PAD | High-performance anion-exchange chromatography with pulsed amperometric detection |
| HPLC | High-performance liquid chromatography |
| IAA | Iodoacetamide |
| IEM | Ion evaporation model |
| IMS | Ion mobility separation |
| IRMPD | Infrared multiphoton dissociation |
| ITC | Isothermal titration calorimetry |
| ITC | Isothermal titration calorimetry |
| $K_a$ | Association constant |
| $K_{a,app}$ | Apparent association constant |

| | |
|---|---|
| L | Ligand |
| LacNAc | *N*-acetyllactosamine |
| LC | Liquid Chromatography |
| LIF | Laser-induced fluorescence |
| LNnO | Lacto-N-neooctaose |
| $L_{proxy}$ | Proxy ligand |
| m/z | Mass-to-charge ratio |
| MAA | *Maackia amurensis* agglutinin |
| Man | Mannose |
| MD | Molecular dynamics |
| MEKC | Micellar electrokinetic capillary chromatography |
| M-LAC | Multi-lectin affinity chromatography |
| MS | Mass spectrometry |
| MW | Molecular weight |
| ND | Not detected |
| NEU 2 | Human neuraminidase 2 |
| *N*-glycan | NG |
| NHS | N-hydroxysuccinimide |
| P | Protein |
| PGC | Porous graphitized carbon chromatography |
| PHI | Prostate Health Index |
| PhoSL | *Pholiota squarrosa* lectin |
| $P_{ref}$ | Reference protein |
| PSA | Prostate specific antigen |
| PTMs | Post-translation modifications |
| *R* | Abundance ratio |
| $R_{eq}$ | Sensor response at saturation |

| | |
|---|---|
| *Res* | Resolution |
| RF | Radio frequency |
| *RF* | Response factor |
| RFU | relative fluorescence units |
| RP | Reversed phase |
| $R_{sat}$ | Sensor response at saturation |
| scFv | Single chain variable fragment |
| SEC | Size exclusion chromatography |
| Sia | Sialic acid |
| SNA | *Sambucus nigra* agglutinin |
| SPR | Surface plasmon resonance |
| SRIG | Stacked ring ion guides |
| TEAA | Triethylammonium acetate |
| TFA | Trifluoroacetic acid |
| TJA-II | *Trichosanthes japonica* agglutinin II |
| TOF | Time-of-flight |
| tPSA | Total prostate specific antigen |
| T-wave | Travelling-wave |
| UEA-I | *Ulex europaeus* agglutinin I |
| UHMR | Ultra-high mass range |
| α2M | α-2-macroglobulin |

# Chapter 1

# Characterization of Noncovalent Glycoprotein Interactions Using Electrospray Ionization Mass spectrometry

## 1.1 Introduction

Carbohydrates, together with proteins, lipids, and nucleic acids, are one of the four major building blocks in living systems. They are widely distributed in almost all organisms in nature, exhibiting remarkable complexity and diversity.[1] The heterogeneity of carbohydrates originates from the structure and sequence of the building blocks, monosaccharides (ten widely used in mammalian systems),[2] the types and the stereochemical configuration of glycosidic bonds formed between various types and number of monosaccharides, and the branching pattern.[3] Carbohydrates play critical biological roles in modulating structural and functional properties of attached protein and lipids, mediating intrinsic and extrinsic recognition (including signaling, clearance, intracellular and pathogen adhesion) and are involved in many pathophysiological processes.[1,4] Carbohydrates, naturally existing as monosaccharides, oligosaccharides, polysaccharides and glycoconjuates (i.e., glycoproteins, glycolipids, proteoglycans and glycosylphosphatidylinositol (GPI)), are found in different organelles (i.e., nuclei, Golgi apparatus, endoplasmic reticulum (ER), lysosome, cytosol), on the cell surface and extracellular matrix (ECM).[5]

Almost 50% of proteins in nature are glycosylated.[6] Unlike the synthesis of proteins, which is coded in the genome, glycosylation is not driven by a template but a process catalyzed by the actions of numerous glycotransferases; over 200 glycotransferases genes have been identified.[7] Attachment of glycans on the surface of proteins can account for large molecular weight of the total masses and thus can significantly alter the solubility, stability and conformation of proteins.[8] Glycans attached to proteins directly participate in numerous physiological and pathological processes through recognition of

glycan-binding proteins (GBPs), which mediate various functions including cell-cell communication, host-pathogen and ligand-receptor interactions, signaling transduction, macromolecules transportation, and immune response.[5,7,9,10] Alterations in protein glycosylation are associated with many diseases and can serve as a biomarker and as a contributor to cancer.[8] Additionally, stability, efficacy and immunogenicity of therapeutic proteins (i.e. antibodies, Fc-fusion proteins, growth factors, cytokines, hormones, and therapeutic enzymes) are known to be affected by the glycosylation status, which impacts their pharmacodynamics and pharmacokinetics.[11,12] Due to the significant biological roles of glycans conjugated to proteins, a number of analytical methods have been developed to characterize glycoprotein interactions and elucidate the influences of glycans on their binding specificity and affinity.

Isothermal titration calorimetry (ITC) is one of the most widely used techniques to probe protein-carbohydrate and protein-(glyco)protein interactions by measuring the heat released or absorbed during a binding event, providing a complete thermodynamic profile of protein interactions.[13,14] The association constant, $K_a$, binding enthalpy, $\Delta H$, and the binding stoichiometry can be accurately determined from a single experiment. The change of free energy of binding, $\Delta G$, can be calculated from $K_a$, and the entropy of binding $\Delta S$ can be subsequently obtained. While it is relatively easy to perform, conventional ITC generally requires large amounts of sample (~1 mg of protein) and is low throughput (2-3 hours per titration).[15] The advent of modern ITC, such as Nano ITC (TA instruments) and $ITC_{200}$ (Microcal), has greatly reduced the sample consumption by three times and improved throughput (30-40 min per titration), but higher protein concentration (two folds) is needed, which can lead to protein aggregation.[15,16]

Surface plasmon resonance (SPR) is another widespread approach for characterizing biomolecular interactions.[17,18] This technique relies on monitoring the variation of the angle of reflectivity of the incident light induced by binding events in close proximity to the metal layer (typically gold or silver). In this assay, a ligand immobilized on the matrix

interacts with analyte (protein) solution for a period of time after measuring the response of running buffer (baseline), and then the ligand/analyte complexes gradually dissociate as the analyte solution is exchanged by running buffer. $K_a$, and the maximum sensor response achieved at saturation, $R_{sat}$, can be determined by non-linear fitting to of Langmuir binding isotherm to a plot of responses at the equilibrium sensor response ($R_{eq}$) versus analyte concentrations.[17] Thermodynamic parameters, such as $\Delta G$, can be directly calculated from $K_a$, while $\Delta H$ and $\Delta S$ can be obtained by performing experiments at different temperatures. Although SPR offers high sensitivity and relatively low sample consumption ($\sim$ng), immobilization of one of the interacting species through coupling (i.e., amine coupling, thiol coupling, aldehyde coupling) might potentially alter the nature of the interactions and ligand affinity.[18] It is also limited by the diffusion of the analyte from the bulk solution to the surface (mass transport), especially for binding with fast kinetics ($k_{on} > 10^6$ $M^{-1}$ $s^{-1}$ and $k_{off} > 10^{-1}$ $M^{-1}$ $s^{-1}$).[18,19] Moreover, affinities are also difficult to obtained from multimeric proteins.

Frontal affinity chromatography using fluorescence detection (FAC-FD) is a unique tool for analysis of protein-glycan interactions with small sample consumption (0.5 pmol of glycan/assay).[20] It is based on the measurements of elution volumes of fluorescence-labelled glycans flowing through a lectin-immobilized column. Retardation of the elution volumes ($V-V_0$) are observed for glycans that are associated with the lectins compared to a non-interacting control glycan.[20] $K_a$ values can be obtained from the amount of glycan bound to the lectins ($[A]_0(V-V_0)$, where $[A]_0$ is the initial concentration of glycan), and $[A]_0$ is not necessarily required in cases where it is negligible relative to $K_a$ values.[20,21] However, labelling of glycans through monoamine coupling (e.g., pyridylamination) gives rise to open ring structures and binding properties of lectins may be modified by immobilization.[22] Besides, total amount of the immobilized lectins must be obtained by a concentration dependence analysis prior to the determination of $K_a$ values of glycans of interest, and thus the total analysis can take several days.[20] Another

drawback of this technique is that it cannot be applied to the analysis of samples containing a mixture of glycans,[20] which limits its use in *N*-glycan libraries extracted from glycoproteins.

Capillary electrophoresis (CE) in combination with laser-induced fluorescence (LIF) detection is an ultrasensitive tool for the detection of fluorescence-labelled glycans and the evaluation of GBP interactions with fluorescently labelled glycans extracted from glycoproteins.[21] By monitoring the migration time of negatively charged fluorescence-labelled glycans in an electrolyte in the presence/absence of a GBP, binding specificities of glycans in a library can be simultaneously determined and binding affinities can be obtained by linear or non-linear regression analysis of changes in migration of glycans upon the addition of a GBP and the corresponding concentrations of the GBP. [21,23] However, stable electroosmotic flow (EOF) needs be controlled or corrected to ensure the accurate measurement of electrophoretic mobility and to avoid the capillary-wall adsorption of glycans, and coating of the capillary inner wall may be used in some cases.[23] Similar to FAC-FD, fluorescent labelling of glycans is required in CE.

Semi-quantitative methods such as glycan microarray can also be used to probe glycan binding specificity. Typically, a buffer containing different concentrations of labelled GBPs (e.g., fluorescence-labelled, His-tagged, biotinylated, Fc fusion) interact with a library of defined glycans from glycoproteins immobilized onto an *N*-hydroxysuccinimide (NHS)-derivatized glass slides, and the binding specificity of the GBP is determined by ranking based on the relative fluorescence units (RFU) detected either directly or indirectly by immunochemical fluorescence detection.[24] Notably, only glycans with ranking values that exhibit concentration dependence of GBP are included in the evaluation of GBP specificity.[24] Similarly, glycoprotein microarray, where GBP interacts with glycoprotein immobilized on slides (e.g., expoxide-coated surface, nitrocellulose slides), can be exploited for the analysis of glycosylation alterations of natural glycoprtoeins.[25,26] However, weak glycan-GBP or GP-GBP interactions might not

be detected because of the washing steps, resulting in false negatives. Besides, the density of glycans printed on the surface, different immobilization and detection methods may affect binding affinities, and thus resulting in unreliable total fluorescence readout.[27–29] Electrospray ionization mass spectrometry (ESI-MS) represents a powerful tool for identification and quantification of protein-ligand, protein-carbohydrate and other non-covalent interactions in vitro.[29–33] As a "soft" analytical method, it relies on detecting noncovalent interactions transferred from solution to the gas phase through the ESI process without labelling or immobilization.[34] Binding stoichiometry can be established and multiple binding equilibria as well as $K_a$ values in the range of $10^2$-$10^7$ $M^{-1}$ can be determined simultaneously.[34,35] Typically, less than a minute is needed for acquiring a mass spectrum with less than fmol of sample consumption per analysis. Moreover, this assay can be applied to the screening of carbohydrate libraries[36,37] against proteins and characterization of protein-glycoprotein interactions qualitatively and quantitatively.[38] Most importantly, binding affinities measured by ESI-MS are in good agreement with values obtained by other binding methods.[30,39,40] However, as reliable detection and measurements are based on the underlying assumption that protein-complex ions detected in the gas phase represent their solution states, limitations also exist in ESI-MS assays. Before presenting the work in this thesis, an overview of basic principles of ESI, followed by the MS instrumentation, ESI-MS based assays and potential limitations of ESI, are given below.

## 1.2 Electrospray Ionization (ESI) Mass Spectrometry

### 1.2.1 Electrospray ionization

ESI is a prevalent soft (with little or no fragmentation) ionization technique where ions are transferred from solution to the gas phase at atmospheric pressure. It enables the detection of non-volatile biomolecules (such as peptides, proteins and nucleotides) through the assistance of a volatile solvent with low sample consumption (typically 1-100 µM) and expands the mass range by the formation of multiply charged ions, which also enhances the dissociation efficiency in tandem MS.

In a conventional ESI device, an analyte solution is loaded into an electrically conducting capillary held at a high electric potential around $\pm$ 2 - $\pm$ 6 kV.[41] Under an electric field of ~$10^6$ V m$^{-1}$, charges are separated in the analyte solution and results in a cone (Taylor cone) extended from the capillary tip, carrying excess positively charged analytes (in positive mode)[42] As soon as the electrostatic force surmounts the surface tension of the solution, a fine jet is ejected from the distal end of the Taylor cone and disintegrates into charged droplets, forming an electrospray plume. Shrinkage of the droplets initially of micrometer in diameter occurs as a result of solvent evaporation. As the surface area/volume ratio of the droplets increases, Rayleigh limit $z_R$ (shown in Eq.1.1), where electrostatic repulsion is balanced by the surface tension, is reached.[43] At or around 10-20% below the Rayleigh limit, surface tension is overcome by Coulombic repulsion, and these droplets become parent droplets and eject numerous smaller droplets (offspring droplets) carrying off 2-5% of their mass and 5-20% of their charge during each cycle of this process (termed droplet jet fission), until radii are decreased to a few nanometers.[41]

$$z_R = \frac{8\pi}{e}\sqrt{\varepsilon_0 \gamma R_d^{\,3}} \tag{1.1}$$

where $e$ is the elementary charge, $\varepsilon_0$ is the vacuum permittivity, $\gamma$ is the surface tension of solvent, and $R_d$ is the droplet radius.
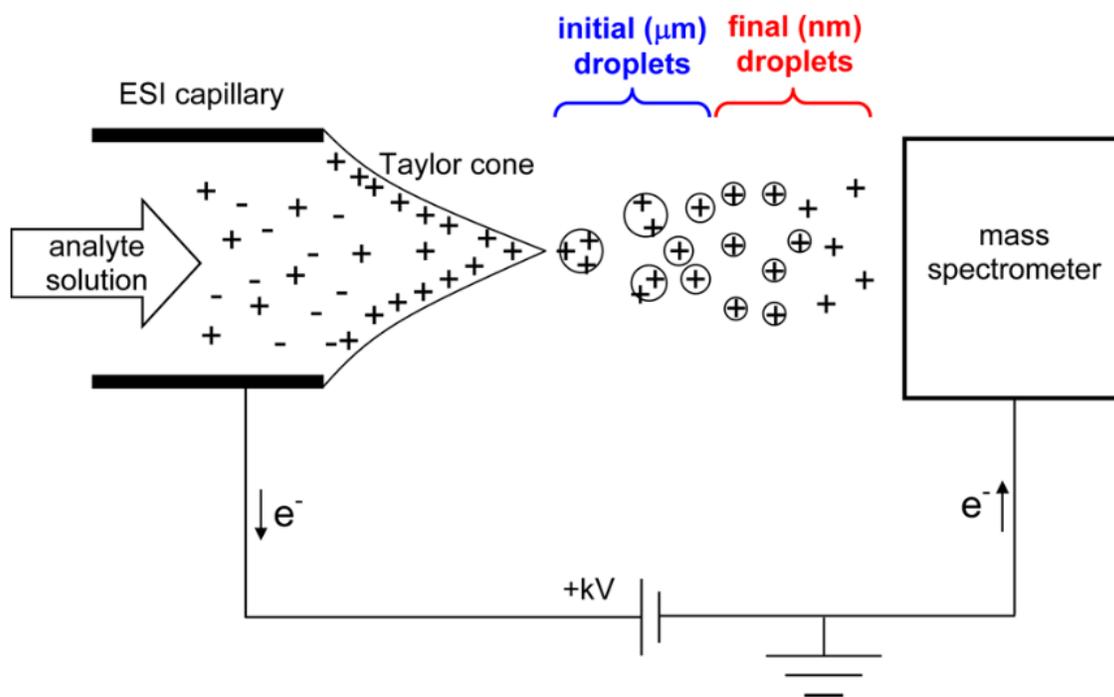
**Figure 1.1** Schematic representation of the process of ESI in positive ion mode, adapted from Ref 44.

Nano-ESI is a downscaling of conventional ESI, where 1-5 μL of analyte solution is infused into a borosilicate glass capillary with an orifice of 1-4 μM diameter.[42] Voltages around 500 - 1000 V are applied to the capillary through either a coated conductive material (e.g. gold) or an inserted metal wire. The size of initial droplets produced from nano-ESI is around 10 times reduced (hundreds of nanometers), leading to higher tolerance of buffer salts compared to conventional ESI. Non-surface active analytes such as oligosaccharides and glycoconjugates benefit from increased surface/volume ratio of parent ions in nano-ESI, resulting in higher sensitivity.[45] Furthermore, high spray stability is realized for analytes containing solvents with high surface tensions at lower capillary voltages.[45]
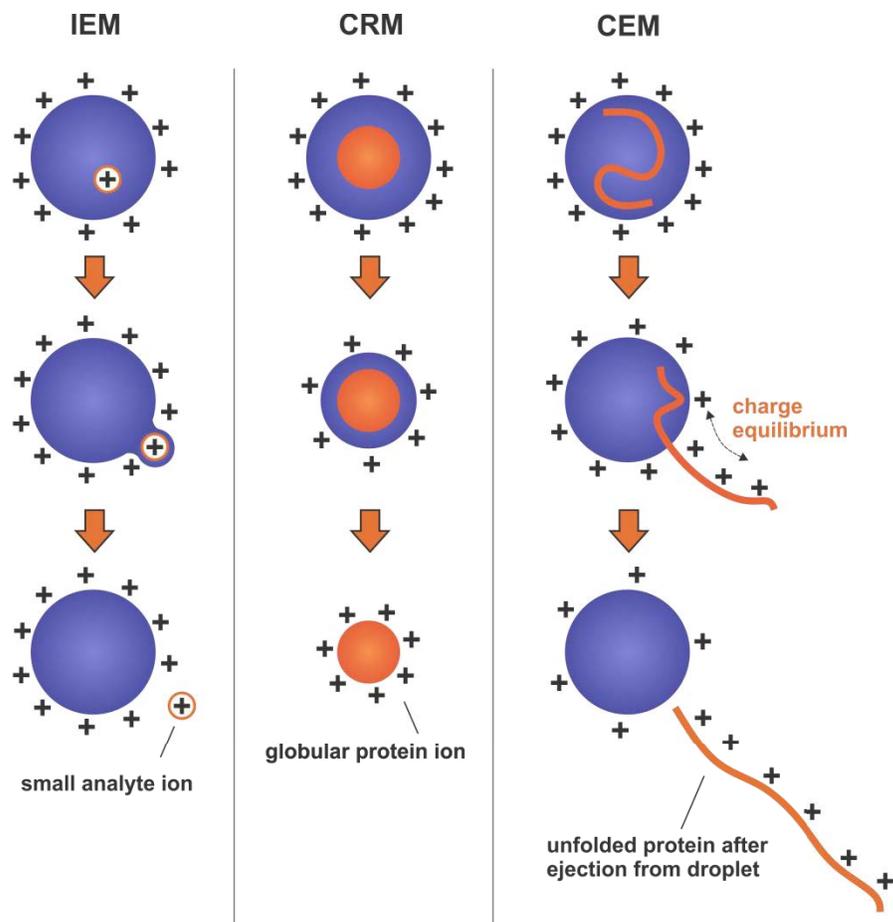
**Figure 1.2** Mechanisms of ion formation in ESI, adapted from Ref 44.

Three mechanisms (as shown in Figure 1.2) have been proposed to illustrate the mechanisms of analyte ionization in ESI. In the ion evaporation model (IEM): as droplets shrink to around 10 nm in diameter by solvent evaporation, a strong electric field compensates for the ionic solvation energies and solvated ions are ejected directly from the surface of the droplets.[46] It is believed that species with low MW (<~3000 Da) prefer IEM, as large analytes with globular shape (such folded proteins) possess more charge than would be expected from field evaporation.[41] The charged residue model (CRM), however, can well explain the multiple charges carried by large molecules during the ESI process. As solvent evaporates, highly charged nanodroplets undergo successive jet fissions until they only contain one analyte molecule, and charge of the last droplet

8

containing solvent is then transferred to the analyte forming a gas ion by thermal declustering.[44] Protonation state of analyte in this process, independent of the physicochemical properties of the ion, is determined by the Rayleigh stability limit of droplets which is related to their surface area (the nature of the solvent). That experimental charge states of globular proteins observed in ESI were in proximity to $[m+z_RH]^{z_R+}$ (where $z_R$ and m is the Rayleigh charge and the mass of a protein) further supports the CRM theory for large globular species.[47,48] However, unfolded proteins, resulted from denaturation (e.g. elevated temperature, pH changes, disulfide disruption), mutation or absence of cofactor, typically exhibit higher charge distribution than natively folded proteins.[49] Chain ejection model (CEM) has been proposed to account for this behavior, which was demonstrated by molecular dynamics (MD) simulations.[44,49] In this model, the interior hydrophobic residues of proteins, exposed to the hydrophilic solvent due to protein unfolding, are gradually expelled and experience proton equilibration with the highly charged droplets, and are finally ejected as highly charged unfolded gas phase ions.[44,49]

**1.2.2 MS instrumentation**

**1.2.2.1 Hybrid Quadrupole Time of Flight mass spectrometer**

A Synapt G2-S quadrupole ion-mobility separation time-of-flight (Q-IMS-TOF) mass spectrometer (Waters UK Ltd., Manchester, UK), equipped with a nanoESI source was used in Chapter 2. A schematic illustration of the instrument is shown in Figure 1.3. By applying an electric potential on the platinum wire inserted in a borosilicate glass capillary, analytes in buffered solutions are ionized by nanoESI at atmospheric pressure and enter the ionization source through a Z spray™. The ion beam expands and is captured by the entrance of a two-stage ion funnel, StepWave™. Ions are efficiently radially confined and guided by radio frequency (RF) voltages with the same amplitude but opposite phase applied to a stack of ring electrodes with decreasing radii of the central aperture. Ions in the first stage are readily directed to the non-coaxial second stage

while neutral molecules are extracted into an exhaust pipe. The focused ions then pass through the quadrupole, where ions of interest are selectively filtered based on *m/z*. Ions can be subjected to collision-induced dissociation (CID) TriWave region filled with helium (He) gas. And transmitted ions are then detected by the TOF mass analyzer. A brief description of the quadrupole, Travelling-wave (T-wave) and TOF parts are given below.
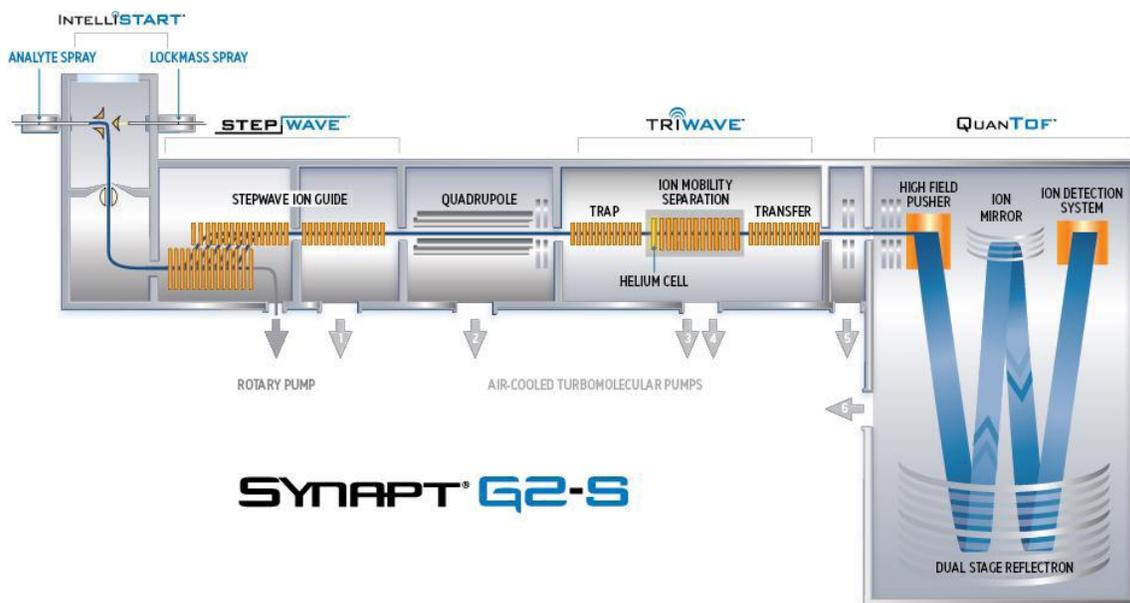


**Figure 1.3** A schematic diagram of the Synapt G2-S Q-IMS-TOF mass spectrometer, adapted from Waters user's manual.

### 1.2.2.1.1 Quadrupole

Quadrupole mass analyzer comprises four parallel rod electrodes with hyperbolic or circular cross-sections extending in the z direction. Ions focused at the entrance of the quadrupole experience a time-depend (direct current (DC)) and a time-independent (alternating current (AC)) potential applied on the rods and are attracted or repulsed periodically. The electric potential of a hyperbolic quadrupole can be described as

$$\phi(x,y) = [U + V\cos(\omega t)] \frac{x^2 - y^2}{2r_0} \tag{1.2}$$

where $\phi$ is the potential distribution, $x$ and y are coordinates along the x and y axis, $U$ and $V$ are the magnitudes of a DC and a radiofrequency (RF), $\omega$ is the angular frequency, and $r_o$ is the distance from the center (z axis) to the inner surface of an electrode. Electrodes along the X axis are held at potentials of the same amplitude but opposite in sign as the electrodes along the Y axis.
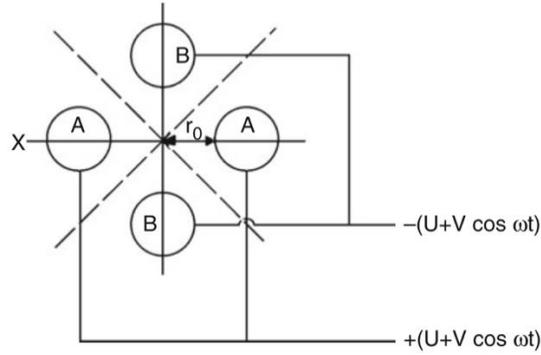


**Figure 1.4** Schematic diagrams of cross-sections of a cylindrical quadrupole, adapted from Ref 42.

According to Newton's law of $F=m_i a$, trajectory of any ion can be described as:

$$\frac{d^2x}{dt^2} + \frac{ez}{m_i r_0{}^2}[U+V\cos(\omega t)]x=0 \tag{1.3}$$

$$\frac{d^2y}{dt^2} - \frac{ez}{m_i r_0{}^2}[U+V\cos(\omega t)]y=0 \tag{1.4}$$

$$\frac{d^2z}{dt^2}=0 \tag{1.5}$$

where $e$ is the electron charge, $z$ is the charge number, $m_i$ is the mass of the ion in kg. Mathieu's equation can be derived as

$$\frac{d^2u}{d\xi^2}+\left[a_u+2q_u\cos2\xi\right]u=0 \tag{1.6}$$

$$a_u = \frac{8ezU}{m_i r_0^2 \omega^2} \qquad (1.7)$$

$$q_u = \frac{4ezV}{m_i r_0^2 \omega^2} \qquad (1.8)$$

where $\xi = \omega t/2$, $u$ is either $x$ or $y$. Solutions to Eq. 1.6 can be obtained by plotting $a$ and $q$, also known as the stability diagram (as shown in Figure 1.5), which can be used to determine the stability of ions in a quadrupole at difference $m/z$. Practically, the ratio of $a$ to $q$, or $2U/V$, is always kept at constant so that only certain ions with restricted $m/z$ can traverse the quadrupole, while the others hit the rods, being neutralized and pumped away.[50] By decreasing the slope ($2U/V$) of a straight line with an intercept at zero (also known as the mass scan line), the band pass region (represented by the width, $\Delta q$) become wider, allowing a increasing $m/z$ range pass through (increased sensitivity), while the resolution is reduced. Therefore, the quadrupole acts as a mass filter by variation of the magnitude of $U$ and $V$ while keeping $2U/V$ at constant.
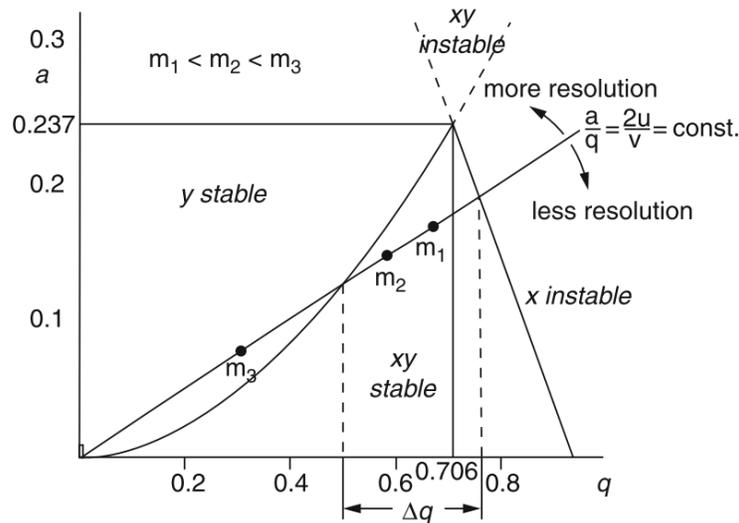


**Figure 1.5** The stability diagram of a quadrupole analyzer, adapted from Ref 42.

The RF-only operation mode of the quadrupole can be achieved by setting $U$ to zero (corresponding to the scan line equivalent to the $q$ axis), and in this case, ions at a wide

range of m/z are transmitted. The lowest *m/z* of ions with stable trajectories is determined by the right-hand *q*-intercept.[51]

## 1.2.2.1.2 Travelling-wave

The T-wave in the Synapt G2-S consists of three consecutive T-wave stacked ring ion guides (SRIG): trap, IMS and transfer T-wave. The ion guide is composed of a stack of ring electrodes arranged orthogonally to the ion pathway with opposite phases of radially confining RF voltages applied to adjacent electrodes (as shown in Figure 1.6).[52] Ions in the electrodes experience axial "traps" resulting from the alternating phases of RF voltages, slowing down the axial motion of ions.[52] Accordingly, a pulsed DC voltage is superimposed on the RF applied ring electrodes at constant time intervals, propelling the axial ion transmission, and consequently, reducing the transit time.[53]

Ions exiting from the quadrupole enter the trap T-wave, where low (below 5V) or no travelling wave is applied and the gas is kept at around $10^{-2}$ mbar. Ions are gated into the IMS cell by only applying a modulated DC voltage (typically $\pm$ 5V) to the last electrode.[53]

IMS is a technique that separates ions under a weak electric field in a gaseous atmosphere on the basis of their mobility, which is affected by their size, shape and charge. In principle, at pressure above 0.2 mbar (typically 2-5 mbar $N_2$, which provides higher resolution compared to He)[54], ions with lower mobility are unable to keep up with the travelling wave in the T-wave IMS cell, resulting in longer transit time as they roll over the wave, while more mobile ions are less overtaken by the wave and traverse more quickly.[52] Therefore, ions are separated based on their mobility and SRIG can serve as an IMS T-wave. In the Synapt IMS T-wave, a helium cell is placed prior to the IMS cell to overcome the potential gas flow from the high-pressure nitrogen IMS cell.[54]
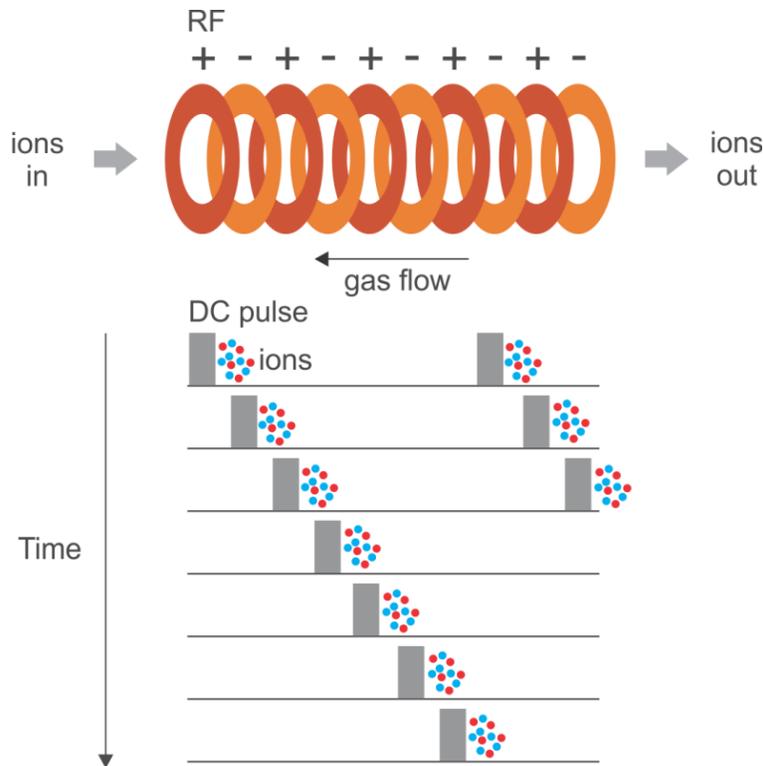
**Figure 1.6** Illustration of the operation of a T-wave SRIG, adapted from Ref 53.

Ions separated in the IMS cell are then transmitted to the gas-filled transfer T-wave, which is typically operated at $10^{-2}$ mbar gas pressure, 1-2 V and 300 m/s of travelling wave to maintain the mobility separation. Ions are then transported and detected by the TOF mass analyzer.

Both transfer and trap T-wave can serve as collision cells by applying additional DC voltages to each ring electrode. Dissociated ions produced in the trap T-wave can be further separated in the IMS T-wave, while the transfer T-wave adds another dimension of fragmentation after separation. CID occurs when precursor ions, selected by the quadrupole, collide with neutral gas (Argon in Synapt instrument) and a faction of their kinetic energy is converted into internal energy.[55] Deposition of the internal energy can lead to breakage of the non-covalent and covalent bonds and dissociation of the ions. Optimized transportation, fragmentation and separation of ions can be achieved by

adjusting the velocity of the travelling wave, magnitude of the DC voltage and gas pressure.[52]

**1.2.2.1.3 TOF mass analyzer**

TOF mass analyzers, in which ions of different $m/z$ are detected according to their flight time in a field-free drift tube, provide high sensitivity (high ion transmission and good duty cycle ($Du$)), a high acquisition rate (up to 30 scans per second in Synapt G2-S), and a wide mass range (up to 100,000 m/z in the Synapt G2-S). Ions transmitted from transfer T-wave are focused into an ion beam by transfer lens and are accelerated orthogonally into a flight tube by a high-field pusher (typically at 5-10 kV). The velocity of an ion in a field-free region after acceleration can be calculated as:

$$v = \sqrt{\frac{2ezU}{m_i}}$$

(1.9)

where $m_i$ is the mass of an ion, $U$ is the acceleration voltage, $e$ is the electron charge, $z$ is the charge number. If the ion travels a distance $s$, the travelling time $t$ is:

$$t = \frac{s}{\sqrt{2eU}} \sqrt{\frac{m_i}{z}}$$

(1.10)

Ions of different kinetic energies then enter a dual-stage reflectron (shown in Figure 1.7), which consists of multiple ring electrodes at a two-stage increasing negative electric potentials. Ions are decelerated by the retarding electric field (typically set to around 5–10% higher than $U$) until their velocities reach 0 and are reflected back to the detector.[42] More energetic ions penetrate deeper and thus stay longer in the reflectron than the ions with less kinetic energy. Reflectrons result in significant correction of spatial and velocity spreads of ions of same $m/z$ and the resolving power is also improved due to the extended flight paths.[56] Arrival times of ions of different $m/z$ and the intensity of the signal are accurately recorded and converted to mass spectra by an analogue-to-digital converter (ADC).
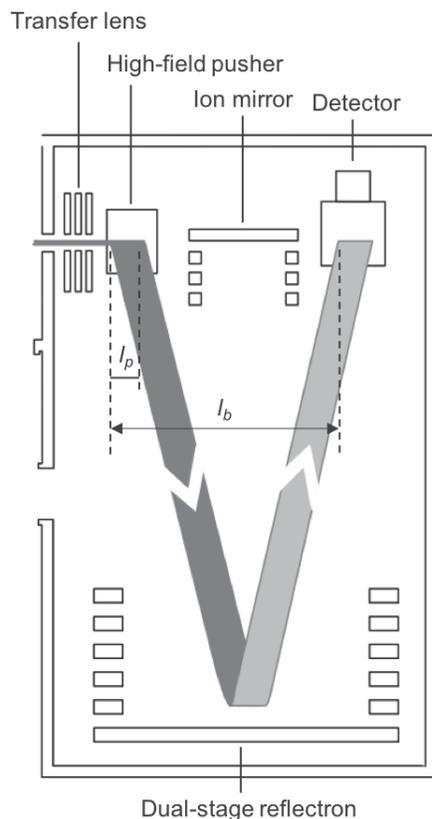
**Figure 1.7** TOF analyzer of Synapt G2-S in the sensitivity mode, adapted from Waters user's manual.

The efficiency of the TOF mass analyzer, *Du*, is determined by the ratio of the length of the ion package before acceleration, $l_p$, to the x-axial distance ions travel from the pusher to the detector, $l_b$. For the high-mass ion at m/z$_{max}$, *Du* is defined as $l_p$ / $l_b$, while for the lightest ion at m/z$_{low}$, *Du* is given as:

$$Du = \frac{l_p}{l_b} \sqrt{\frac{m_i/z_{low}}{m_i/z_{max}}} \tag{1.11}$$

### 1.2.2.2 Orbitrap mass spectrometer

A Q Exactive Ultra-high mass range (UHMR) hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, US), equipped with a nano ESI source

was used in Chapter 3. It is a Fourier transform based mass spectrometer offering high resolution ($Res$ = 20,000 – 200,000) and high mass accuracy (0.5 – 3 ppm).[42]
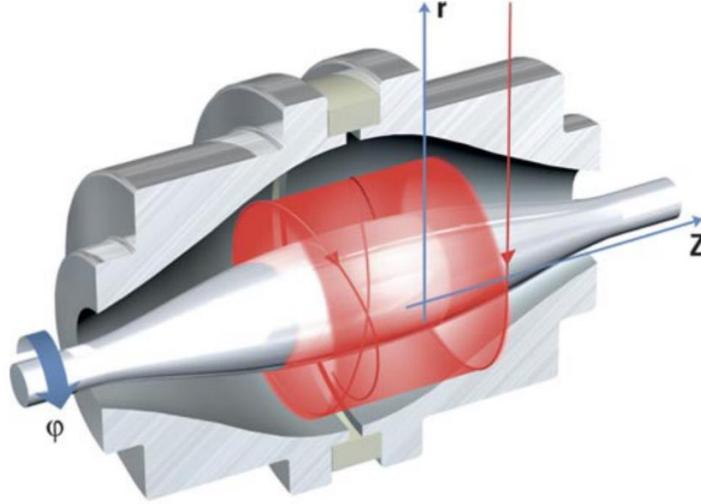


**Figure 1.8** Orbitrap mass analyzer, adapted from Ref 57.

The Orbitrap mass analyzer is composed of two electrodes (a spindle-like central electrode at high voltage and a grounded outer barrel-like electrode separated into a half by a dielectric ceramic ring) as shown in Figure 1.8. An accurate quadro-logarithmic field $U(r,z)$ (as shown in Eq. 1.12) is formed as a result of the shape of the electrodes.

$$U(\text{r,z}) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}(R_m)^2 \ln\left(\frac{r}{R_m}\right) + \text{C} \qquad (1.12)$$

where $r$ and $z$ are cylindrical coordinates, $k$ is the field curvature determined by the shape of the electrodes and the potential applied, $R_m$ is the characteristic radius (the maximum orbital radii of trapped ions) and $C$ is a constant.[58]

An ion package is injected as a focused beam tangentially into the Orbitrap analyzer off the equatorial plane of symmetry ($z$ =0) from an RF-only nitrogen filled multipole termed the C-trap and starts to rotate around the central electrode with decreasing radius under increasing electrode potential (termed "electrodynamic squeezing").[58] Simultaneously,

ions directed by the axial electric field $E_z$ oscillate along the z axis at frequency $\omega_z$ (Eq. 1.13), while rotation frequency $\omega_\varphi$ corresponding to angular motion (Eq. 1.14) and radial oscillation $\omega_r$ corresponding to radial motion (Eq. 1.15) under the radial field $E_r$ can be derived.[59]

$$\omega_z = \sqrt{k \cdot \frac{ze}{m_i}} \qquad (1.13)$$

$$\omega_\varphi = \frac{\omega_z}{\sqrt{2}} \sqrt{\left(\frac{R_m}{R_c}\right)^2 - 1} \qquad (1.14)$$

$$\omega_r = \omega_z \sqrt{\left(\frac{R_m}{R_c}\right)^2 - 2} \qquad (1.15)$$

where $R_c$ is circular radius of ions. Only the harmonic axial oscillations frequency $\omega_z$ is independent of initial positions and kinetic energies of ions, which can be detected in time domain by the image current, generated from the pair of outer electrodes after amplification and then converted to mass to charge spectrum by Fourier transform. The mass resolving power *Res* can be calculated as Eq. 1.16.

$$Res = \frac{m_i}{\Delta m_i} = \frac{1}{2\Delta\omega_z} \sqrt{k \cdot \frac{ze}{m_i}} \qquad (1.16)$$

However, the actual resolution of Orbitrap is limited by the time of the detectable signal. Therefore, ultra-high vacuum (typically $\sim 2 \times 10^{-10}$ mbar) is required to reduce the collisions with residual gas, which can result in loss of the coherence of ion packages and potential ion fragmentation. Besides, another prerequisite for the successful image current detection is the proper ion injection. A curved RF-only gas filled multipole, C trap, serves as an external ion storage device prior to the injection, where ions are collected and thermalized through collisions with nitrogen gas. Ions are ejected from the C-trap through

the gap between the inner electrodes by applying fast high-voltage electric pulses (~ hundreds of nanoseconds), and subsequently converge on the entrance of the Orbitrap.[42] Shown in Figure 1.9 is a schematic diagram of the Thermo Fisher Q Exactive UHMR mass spectrometer used in the present work. Gaseous ions generated from nanoESI at atmospheric pressure are drawn into the ion transfer capillary by decreasing pressure and transported to the RF-lens where they are captured and focused into an ion beam. Ions are then transmitted to the low pressure ion optics regions and enter the injection flatapole, where they are trapped, focused and desolvated by maintaining a negative potential on the injection flatapole lens and a positive potential on inter-flatapole lens. Subsequently, ejected ions are focused and guided through the bent flatapole by an axial DC field and a focusing RF field. Neutral particles and solvent droplets are removed due to the $90^{o}$ arc. A hyperbolic segmented quadrupole (HyperQuad[TM]) mass filter is employed to selectively transfer ions with improved ion transmission and optimized isolation window. Exiting from the gas-free multipole, ions release their kinetic energy through collisions with gas and pass through the C trap in an ion beam and detected in the Orbitrap mass analyzer. Ion fragmentation can be induced by employing the HCD (Higher energy collisional dissociation) cell, a straight multipole connected to the C trap.
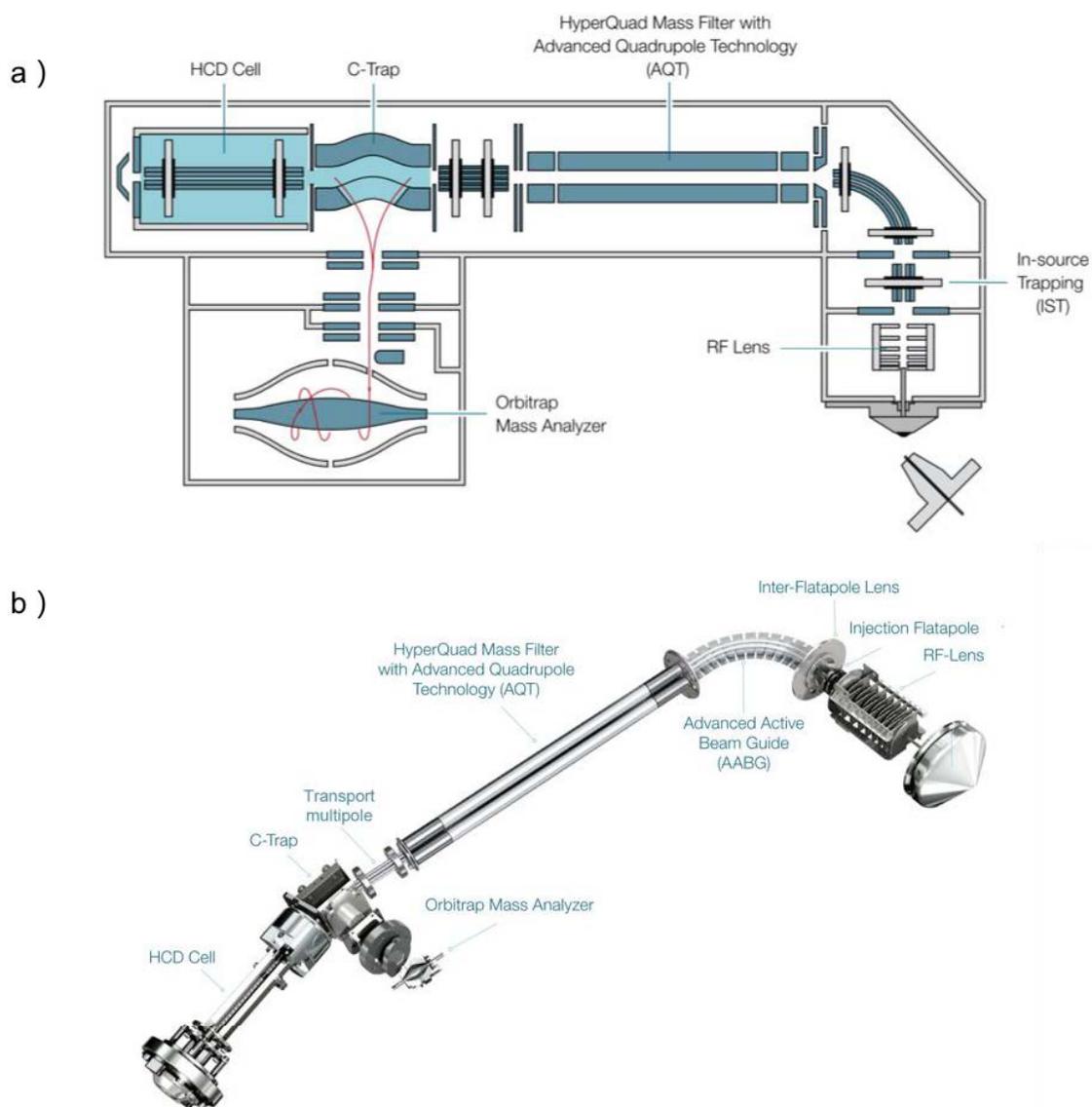
**Figure 1.9** a), b) Schematic diagrams of the Thermo Fisher Q Exactive UHMR mass spectrometer, adapted from Thermo Fisher user's manual.

## 1.3 ESI-MS Based Assays

### 1.3.1 Direct ESI-MS assay

Interactions between proteins and their ligands (including proteins, carbohydrates, lipids, nucleic acids, and small molecules) in solution can be quantified using the direct ESI-MS assay.[34] For a protein (P) with a single binding site, a complex (PL) is formed with the addition of a ligand (L), and equilibrium is established in solution as:

$$P + L \rightleftharpoons PL \tag{1.17}$$

Theoretically, the concentration ratio ($R$) of the ligand-bound (PL) and free protein (P) at equilibrium can be represented by the total abundance ($Ab$) ratio the corresponding ions measured by ESI-MS (as shown in Eq.1.18).

$$R = \frac{\sum Ab(\text{PL})}{\sum Ab(\text{P})} = \frac{[\text{PL}]}{[\text{P}]} \tag{1.18}$$

As the signal of image detection in FT-ICR-MS is proportional to the abundance and charge states (n) of ions,[60] $R$ is determined as:[61]

$$R = \frac{\sum_n \left( Ab(\text{PL})^{n+}/n \right)}{\sum_n \left( Ab(\text{P})^{n+}/n \right)} = \frac{[\text{PL}]}{[\text{P}]} \tag{1.19}$$

With known initial concentrations of protein ($[\text{P}]_0$) and ligand ($[\text{L}]_0$), $K_a$ can be calculated as:

$$K_a = \frac{[\text{PL}]}{[\text{P}][\text{L}]} = \frac{R}{[\text{L}]_0 - \frac{R}{1+R}[\text{P}]_0} \tag{1.20}$$

Normally, a reliable $K_a$ is determined by multiple measurements performed at a fixed analyte concentration (typically P) and a series of concentrations of the other one (typically L). And nonlinear regression analysis of the experimentally determined concentration- dependence of the fraction of ligand-bound protein, i.e., $R/(R + 1)$ is employed to extract the $K_a$ (as shown in Eq.1.21).[34]

$$\frac{R}{R+1} = \frac{1 + K_a[\text{L}]_0 + K_a[\text{P}]_0 - \sqrt{(1 - K_a[\text{L}]_0 + K_a[\text{P}]_0)^2 + 4K_a[\text{L}]_0}}{2K_a[\text{P}]_0} \tag{1.21}$$

Practically, for protein and ligand concentrations ranging from 0.1 to 1000 μM, measured $R$ values vary from 0.05 to 20, corresponding to $K_a$ in a range of $10^3$ to $10^7$ M$^{-1}$.[34] Competitive binding in conjunction with direct ESI-MS assay can be applied in interactions with much higher affinity.[34]

**1.3.2 *Catch-and-release* (CaR) ESI-MS assay**

In cases where molecular weight (MW) of protein-ligand complexes cannot be precisely determined by ESI-MS as a result of the heterogeneity (i.e., glycoproteins, due to the presence of multiple glycosylation forms (glycoforms)) or size of the protein, *catch-and-release* (CaR) ESI-MS assay is implemented to identify the ligands.[62] In this assay, ions corresponding to complexes formed by protein 'catching' ligands in solution are isolated by quadrupole mass filter, followed by the 'release' of bound ligands by CID and the detection by a mass spectrometer.[63] The identity of ligands can be determined by either accurate mass analysis or in conjunction with IMS or MS/MS in cases of isomeric ligands. CaR ESI-MS has been successfully employed in the identification protein ligands, including drugs, carbohydrates, glycolipids, peptides and proteins.[38,62–68]

**1.3.3 *Proxy ligand* ESI-MS assay**

However, CaR ESI-MS assay only provides qualitative identification of the bound ligands, and the application of direct ESI-MS assay for quantification of binding affinity is limited in the case of a protein binding to a ligand incorporated in a system (such as lipids contained in model membranes), where complexes are detected due to the gas phase dissociation, and thus are affected by the different response factors for bound and unbound protein ions during the ESI process.[69] Therefore, a competitive binding assay, *proxy ligand* ESI-MS, where a proxy ligand ($L_{proxy}$) binding to the target protein (P) with a known affinity ($K_{a,proxy}$) competes with the ligand (L), is employed. An equilibrium is established as:

$$PL_{proxy} \rightleftharpoons L_{proxy} + P + L \rightleftharpoons PL \tag{1.24}$$

And in this system, the concentrations of P ($[P]_0$), $L_{proxy}$ ($[L_{proxy}]_0$) and L ($[L]_0$) at equilibrium are related to their initial concentrations by mass balance:

$$[P]_0 = [P] + [PL_{proxy}] + [PL] \tag{1.25}$$

$$[L_{proxy}]_0 = [L_{proxy}] + [PL_{proxy}] \tag{1.26}$$

$$[L]_0 = [L] + [PL] \tag{1.27}$$

The concentration of unbound P decreases as P binds to L, and thus the relative abundance of $PL_{proxy}$ to free P ($R_{proxy}$, as shown in Eq.1.28) increases. Accordingly, interactions between P and L can be quantified by monitoring $R_{proxy}$ by ESI-MS.[69]

$$R_{proxy} = \frac{\sum Ab(PL_{proxy})}{\sum Ab(P)} = \frac{[PL_{proxy}]}{[P]} \tag{1.28}$$

$K_a$ can be calculated from Eq.1.29

$$
K_a = \frac{R}{\left[ [L]_0 - \dfrac{R}{1+R+R_{proxy}} \right][P]_0} = \frac{1}{\dfrac{[L]_0}{R} - \dfrac{[P]_0}{1+R_{proxy}+R}}
$$

$$
= \frac{1}{\left( [L_{proxy}]_0 - \dfrac{R_{proxy}}{K_{a,proxy}} \right) \left( \dfrac{[L]_0}{R_{proxy}[P]_0 - (R_{proxy}+1)\left([L_{proxy}]_0 - \dfrac{R_{proxy}}{K_{a,proxy}}\right)} - \dfrac{1}{R_{proxy}} \right)} \tag{1.29}
$$

$[L]_0$ can be expressed as Eq. 1.30 by rearranging Eq. 1.29.

$$
[L]_0 = \left( \frac{R_{proxy}[P]_0}{[L_{proxy}]_0 - \dfrac{R_{proxy}}{K_{a,proxy}}} - R_{proxy} - 1 \right) \left( \frac{1}{K_a} + \frac{[L_{proxy}]_0}{R_{proxy}} - \frac{1}{K_{a,proxy}} \right) \tag{1.30}
$$

By titrating a range of L concentrations into P and $L_{proxy}$ at fixed concentrations, $K_a$ can be determined from fitting Eq 1.30 over the titration curve of $[L]_0$ versus experimentally measured $R_{proxy}$.[70]

## 1.4 Potential pitfalls of ESI-MS assays

Although detecting protein-ligand complexes by ESI-MS assays is relatively easy in terms of simplicity (without labelling or immobilization) and speed ( typically less than a minute per spectrum), reliability of the binding data needs to be examined. One of the prerequisites is that equilibrium abundances of species in solution must be maintained equally as those of corresponding ions during the ESI process and in gas phase. These abundances can be affected by physical and chemical processes, which can ultimately lead to inaccurate $K_a$ values and binding stoichiometry. Three common sources of errors in ESI-MS measurements and available strategies to tackle these problems are discussed below.[34]

### 1.4.1 In-source dissociation

Collision-induced dissociation of protein-ligand complexes during ESI-MS reduces the abundance of PL relative to P, resulting in decreased $K_a$ values or even false negatives in binding measurements. The occurrence of in-source dissociation is influenced by many factors, such as choice of ion sources, instrumental parameters, accumulation time, and the MW and gas phase stability of the complex.[71] While low affinity interactions generally exhibit low gas stabilities, some complexes formed by strong ionic bonding are less stable in the gas phase than those stabilized by hydrophobic interactions.[72–77] Identification of in-source dissociation can be achieved by monitoring $R$ (relative abundance of bound to unbound proteins) at different ion source parameters.

Although "gentle" sampling conditions ((i.e., low temperatures of drying gas and sampling capillary, low potentials applied on lens and short accumulation time) attenuate the degree of in-source dissociation, protein ion signal is usually sacrificed.[77] Therefore, source parameters need to be carefully adjusted to preserve the interactions in the gas phase and obtain adequate signal-to-noise ratio at the same time. In addition, stabilizing additives such as free amino acids, dimethyl sulfoxide (DMSO), and imidazole have been shown to protect labile protein-ligand interactions against gas phase dissociation due to

the reduced Coulombic repulsion from decreased charge states and enhanced cooling effect of adduct dissociation.[74–76,78,79] A competitive binding assay, the reference ligand ESI-MS approach, which involves a reference ligand of P with a known binding affinity and high kinetics stability, provides an alternative way to quantify protein-ligand interactions that are susceptible to in-source dissociation.[73,75]

## 1.4.2 Nonspecific binding

False positives of the binding data can be produced by nonspecific binding of free ligands to protein and protein-ligand complexes during the ESI process.[80,81] Detection of multiple ligands-bound protein ions exhibiting a Poisson-like distribution and measurements of ligand concentration-dependent $K_a$ potentially result from the formation of nonspecific protein-ligand complexes. The occurrence of nonspecific ligand binding can be explained according to the CRM model (described in Section 1.2.1, shown in Figure 1.10). In the cases where more than one analyte ions are contained in a nanodroplet produced from parent droplets, solvent evaporation can produce gaseous ions with nonspecifically interacting protein and ligands. Increasing ligand concentration contributes to a higher possibility of the formation of nonspecific protein-ligand complexes, which can be minimized by simply reducing the concentration.[81] However, this strategy cannot be applied to relatively low affinity interactions ($K_a < 10^4$ M$^{-1}$) as detectable complexes ions are typically obtained at high ligand concentrations (>0.05 mM).[34]
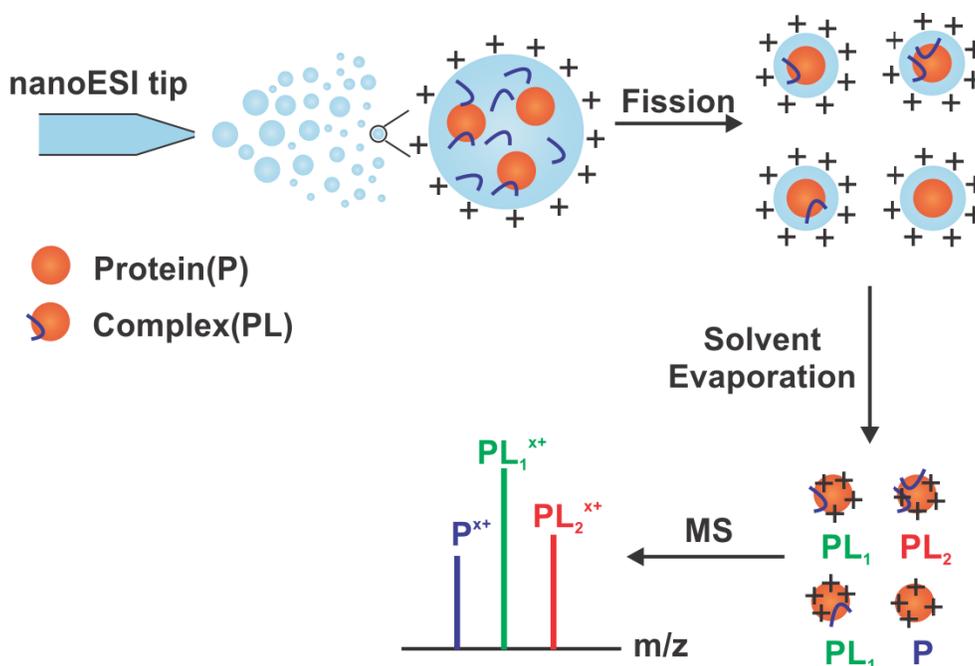
**Figure 1.10** Schematic diagram of occurrence of nonspecific binding during the ESI process under the positive ion mode, adapted from Ref 34.

A number of methods have been developed to correct nonspecific interactions in ESI.[71,82–88] The reference protein method, involving the introduction of a non-interacting reference protein ($P_{ref}$) into the solution, is the most straightforward approach to quantitatively correct for the contribution of nonspecific binding with the underlying assumption that in-source dissociation is absent and that the distribution of nonspecific protein-ligand complexes is not influenced by the nature of proteins.[82,83] The measured (apparent) abundance ($Ab_{app}$) of complexes in the presence of nonspecific binding can be described as:

$$Ab_0(\text{PL}_N) = \left[ Ab_{app}(\text{PL}_N) - f_{1,P_{ref}} Ab(\text{PL}_{N-1}) - \ldots f_{N,P_{ref}} Ab(\text{P}) \right] / f_{0,P_{ref}} \qquad (1.31)$$

where $f_{N,Pref}$ is the fraction of $P_{ref}$ that undergoes nonspecific binding of N molecules to L.[82] Notably, this method for correcting nonspecific binding has been successfully applied to protein interactions with proteins, carbohydrate, divalent metal ions, amino acids and peptides.[82–84,89]

### 1.4.3 Non-uniform response factors

Abundances of P and PL detected by ESI-MS are correlated to their corresponding solution concentrations via response factors ($RF$s), which can be discriminated due to the ionization process, mass analyzer, detection efficiencies and in-source dissociation in the cases of labile complexes.[90] The ratio between the solution concentrations is given by:

$$\frac{[\text{PL}]}{[\text{P}]}=\frac{RF_{\text{P}}Ab(\text{PL})}{RF_{\text{PL}}Ab(\text{P})}=RF_{\text{P/PL}}\frac{Ab(PL)}{Ab(\text{P})} \tag{1.32}$$

where $RF_{\text{P}}$ and $RF_{\text{PL}}$ are response factors of P and PL. Determination of $K_a$ values by direct ESI-MS assay and *proxy ligand* ESI-MS assay is under the assumption that $RF$ values are uniform (i.e. $RF_{P/PL} \approx 1$), which is typically valid where P and PL have similar MW (i.e. $MW_{PL}/MW_{P} \leq 110\%$).[91] However, in some cases where size and surface properties of P and PL are different (i.e. comparable size of L to P, presence of high MW solute), non-uniform response factors are expected.[92,93] Due to the comparable MWs of P (GBP in Chapter 2, GP in Chapter 3) and PL (GBP-$L_{proxy}$ in Chapter 2, GP-warfarin in Chapter 3), non-uniform response factors are negligible in the present work.

## 1.5 The Present Work

Chapter 2 focuses on developing a multipronged ESI-MS approach to identify and quantify glycan-mediated interactions between water soluble GBPs and GPs. First, GBP-GP binding in solution is be rapidly detected by CaR-ESI-MS assay in combination with IMS. The $K_a$ of the GBP for the GP is then measured using the competitive *proxy ligand*-ESI-MS binding assay. Finally, glycans that are recognized by the GBP is identified by screening of enzymatically released N-glycans from the GP against the GBP using ESI-MS. Affinities of released N-glycans (grouped as compositional isomers) is ranked by measurements performed at multiple GBP concentrations.

High-resolution native mass spectrometry has emerged as a promising analytical tool to study protein glycosylation and its influence on glycoprotein interactions. In Chapter 3, high-resolution mass spectrometry is employed to examine the conclusions of a previous

study on the impact of *N*-glycan branching and fucosylation on human alpha-1-acid glycoprotein (AGP) interactions with warfarin, an anticoagulant drug binder. And as an extension, a comprehensive investigation on the influence of glycosylation on AGP-warfarin binding was carried out by enzymatical removal of fucose residues and *N*-glycans from AGP. First, a Q-Exactive UHMR Orbitrap mass spectrometry is employed to elucidate the microheterogeneity of AGP, and following treatments with sialidase and fucosidase. Then, *N*-glycans are completely removed from AGP by PNGase F. Finally, interactions of glycoforms of asialo-AGP and warfarin are quantified by direct ESI-MS and compared to the previously reported values and ITC measurements. The impact of *N*-glycosylation on AGP-warfarin interaction is further investigated on the PNGase F treated AGP.

Chapter 4 comprises a summary of this thesis, some future work and preliminary data on probing biomarkers in prostate cancer.

## 1.6 References

(1) Ghazarian, H.; Idoni, B.; Oppenheimer, S. B. A Glycobiology Review: Carbohydrates, Lectins and Implications in Cancer Therapeutics. *Acta Histochemica*. **2011**, 236–247.

(2) Herget, S.; Toukach, P. V; Ranzinger, R.; Hull, W. E.; Knirel, Y. A.; Von Der Lieth, C. W. Statistical Analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): Characteristics and Diversity of Bacterial Carbohydrates in Comparison with Mammalian Glycans. *BMC Struct. Biol.* **2008**, *8*.

(3) Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. *Nat. Chem. Biol.* **2010**, *6* (10), 713–723.

(4) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49.

(5) Zhao, Y. Y.; Takahashi, M.; Gu, J. G.; Miyoshi, E.; Matsumoto, A.; Kitazume, S.; Taniguchi, N. Functional Roles of N-Glycans in Cell Signaling and Cell Adhesion in Cancer. *Cancer Science*. **2008**, 1304–1310.

(6) Apweiler, R.; Hermjakob, H.; Sharon, N. On the Frequency of Protein Glycosylation, as Deduced from Analysis of the SWISS-PROT Database. *Biochim. Biophys. Acta* **1999**, *1473* (1), 4–8.

(7) Cummings, R. D.; Pierce, J. M. The Challenge and Promise of Glycomics. *Chem. Biol.* **2014**, *21* (1), 1–15.

(8) Stowell, S. R.; Ju, T.; Cummings, R. D. Protein Glycosylation in Cancer. *Annu. Rev. Pathol.* **2015**, *10*, 473–510.

(9) Zeng, X.; Andrade, C. A. S.; Oliveira, M. D. L.; Sun, X. L. Carbohydrate-Protein Interactions and Their Biosensing Applications. *Analytical and Bioanalytical Chemistry*. **2012**, 3161–3176.

(10) Paulson, J. C.; Blixt, O.; Collins, B. E. Sweet Spots in Functional Glycomics. *Nature Chemical Biology*. **2006**, 238–248.

(11)  Zhong, X.; Somers, W. Recent Advances in Glycosylation Modifications in the Context of Therapeutic Glycoproteins. In *Integrative Proteomics*; InTech, **2012**; Vol. i, 13.

(12)  Costa, A. R.; Rodrigues, M. E.; Henriques, M.; Oliveira, R.; Azeredo, J. Glycosylation: Impact, Control and Improvement during Therapeutic Protein Production. *Critical Reviews in Biotechnology*. Informa Healthcare USA, Inc **2014**, 281–299.

(13)  Liang, Y. Applications of Isothermal Titration Calorimetry in Protein Science. *Acta Biochimica et Biophysica Sinica*. **2008**, 565–576.

(14)  Dam, T. K.; Brewer, C. F. Thermodynamic Studies of Lectin−Carbohydrate Interactions by Isothermal Titration Calorimetry. *Chem. Rev.* **2002**, *102* (2), 387–430.

(15)  Torres, F. E.; Recht, M. I.; Coyle, J. E.; Bruce, R. H.; Williams, G. Higher Throughput Calorimetry: Opportunities, Approaches and Challenges. *Current Opinion in Structural Biology*. **2010**, 598–605.

(16)  Rajarathnam, K.; Rösgen, J. Isothermal Titration Calorimetry of Membrane Proteins - Progress and Challenges. *Biochimica et Biophysica Acta - Biomembranes*. **2014**, 69–77.

(17)  Huber, W.; Mueller, F. Biomolecular Interaction Analysis in Drug Discovery Using Surface Plasmon Resonance Technology. *Curr. Pharm. Des.* **2007**, *12* (31), 3999–4021.

(18)  De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. Kinetic Characterization by Surface Plasmon Resonance-Based Biosensors: Principle and Emerging Trends. *Cell. Mol. Bioeng.* **2008**, *1* (4), 204–215.

(19)  Tudos, A. J.; Schasfoort, R. B. M. *Handbook of Surface Plasmon Resonance*; Schasfoort, R. B. M., Ed.; Royal Society of Chemistry: Cambridge, **2017**.

(20)  Tateno,    H.;    Nakamura-Tsuruta,    S.;    Hirabayashi,    J.    Frontal    Affinity

Chromatography: Sugar-Protein Interactions. *Nat. Protoc.* **2007**, *2* (10), 2529–2537.

(21) Kasai, K. *Lectins*; Hirabayashi, J., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, **2014**; Vol. 1200.

(22) Iwaki, J.; Hirabayashi, J. Carbohydrate-Binding Specificity of Human Galectins: An Overview by Frontal Affinity Chromatography. *Trends Glycosci. Glycotechnol.* **2018**, *30* (172), 137–153.

(23) Chen, Z.; Weber, S. G. Determination of Binding Constants by Affinity Capillary Electrophoresis, Electrospray Ionization Mass Spectrometry and Phase-Distribution Methods. *TrAC - Trends Anal. Chem.* **2008**, *27* (9), 738–748.

(24) Smith, D. F.; Song, X.; Cummings, R. D. Use of Glycan Microarrays to Explore Specificity of Glycan-Binding Proteins. In *Methods in Enzymology*; **2010**; Vol. 480, 417–444.

(25) Patwa, T. H.; Zhao, J.; Andersen, M. A.; Simeone, D. M.; Lubman, D. M. Screening of Glycosylation Patterns in Serum Using Natural Glycoprotein Microarrays and Multi-Lectin Fluorescence Detection. *Anal. Chem.* **2006**, *78* (18), 6411–6421.

(26) Patwa, T.; Li, C.; Simeone, D. M.; Lubman, D. M. Glycoprotein Analysis Using Protein Microarrays and Mass Spectrometry. *Mass Spectrom. Rev.* **2010**, *29* (5), 830–844.

(27) Stowell, S. R.; Cummings, R. D. *Galectins*; Stowell, S. R., Cummings, R. D., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, **2015**; Vol. 1207.

(28) Kilcoyne, M.; Gerlach, J. Q.; Kane, M.; Joshi, L. Surface Chemistry and Linker Effects on Lectin-Carbohydrate Recognition for Glycan Microarrays. *Anal. Methods* **2012**, *4* (9), 2721–2728.

(29) Han, L.; Shams-Ud-Doha, K.; Kitova, E. N.; Klassen, J. S. Screening

Oligosaccharide Libraries against Lectins Using the Proxy Protein Electrospray Ionization Mass Spectrometry Assay. *Anal. Chem.* **2016**, *88* (16), 8224–8231.

(30) Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; St-Pierre, Y.; Klassen, J. S. Human Milk Oligosaccharide Specificities of Human Galectins. Comparison of Electrospray Ionization Mass Spectrometry and Glycan Microarray Screening Results. *Anal. Chem.* **2017**, *89* (9), 4914–4921.

(31) El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Carbohydrate-Protein Interactions by Electrospray Ionization Mass Spectrometry Analysis. *Biochemistry* **2012**, *51* (21), 4244–4253.

(32) Liu, L.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Fatty Acid Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (2), 310–318.

(33) Cubrilovic, D.; Biela, A.; Sielaff, F.; Steinmetzer, T.; Klebe, G.; Zenobi, R. Quantifying Protein-Ligand Binding Constants Using Electrospray Ionization Mass Spectrometry: A Systematic Binding Affinity Study of a Series of Hydrophobically Modified Trypsin Inhibitors. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (10), 1768–1777.

(34) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441.

(35) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. Affinities of Recombinant Norovirus P Dimers for Human Blood Group Antigens. *Glycobiology* **2013**, *23* (3), 276–285.

(36) El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; St-Pierre, Y.; Klassen, J. S. High-Throughput Label- and Immobilization-Free Screening of Human Milk Oligosaccharides Against Lectins. *Anal. Chem.* **2017**, *89* (17), 8713–8722.

(37) El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; Bode, L.;

Hage, N.; Falcone, F. H.; Klassen, J. S. Screening Natural Libraries of Human Milk Oligosaccharides against Lectins Using CaR-ESI-MS. *Analyst* **2018**, *143* (2), 536–548.

(38) Wang, Y.; Park, H.; Lin, H.; Kitova, E. N.; Klassen, J. S. Multipronged ESI-MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins. *Anal. Chem.* **2019**, *91* (3), 2140–2147.

(39) Yao, Y.; Shams-Ud-Doha, K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Carbohydrate Interactions Using Liquid Sample Desorption Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2014**, *26* (1), 98–106.

(40) Rademacher, C.; Shoemaker, G. K.; Kim, H. S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T.; et al. Ligand Specificity of CS-35, a Monoclonal Antibody That Recognizes Mycobacterial Lipoarabinomannan: A Model System for Oligofuranoside-Protein Recognition. *J. Am. Chem. Soc.* **2007**, *129* (34), 10489–10502.

(41) Liuni, P.; Wilson, D. J. Understanding and Optimizing Electrospray Ionization Techniques for Proteomic Analysis. *Expert Rev. Proteomics* **2011**, *8* (2), 197–209.

(42) Gross, J. H. *Mass Spectrometry*; Springer International Publishing: Cham, **2017**; Vol. 56.

(43) Rayleigh, Lord. XX. On the Equilibrium of Liquid Conducting Masses Charged with Electricity. *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **1882**, *14* (87), 184–186.

(44) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9.

(45) Karas, M.; Bahr, U.; Dülcks, T. Nano-Electrospray Ionization Mass Spectrometry: Addressing Analytical Problems beyond Routine. *Fresenius. J. Anal. Chem.* **2000**, *366* (6–7), 669–676.

(46)   Iribarne, J. V. On the Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.* **1976**, *64* (6), 2287.

(47)   Kaltashov, I. A.; Mohimen, A. Estimates of Protein Surface Areas in Solution by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2005**, *77* (16), 5370–5379.

(48)   McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. Release of Native-like Gaseous Proteins from Electrospray Droplets via the Charged Residue Mechanism: Insights from Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2015**, *137* (39), 12667–12676.

(49)   Metwally, H.; Duez, Q.; Konermann, L. Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry. *Anal. Chem.* **2018**, *90* (16), 10069–10077.

(50)   Miller, P. E.; Denton, M. B. The Quadrupole Mass Filter: Basic Operating Concepts. *J. Chem. Educ.* **1986**, *63* (7), 617.

(51)   Kero, F. A.; Pedder, R. E.; Yost, R. A. Quadrupole Mass Analyzers: Theoretical and Practical Considerations. In *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*; **2009**.

(52)   Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a Travelling Wave-Based Radio-Frequency-Only Stacked Ring Ion Guide. *Rapid Commun. Mass Spectrom.* **2004**, *18* (20), 2401–2414.

(53)   Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. An Investigation of the Mobility Separation of Some Peptide and Protein Ions Using a New Hybrid Quadrupole/Travelling Wave IMS/Oa-ToF Instrument. *Int. J. Mass Spectrom.* **2007**, *261* (1), 1–12.

(54)   Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in Travelling Wave Ion Mobility Resolution. *Rapid Commun. Mass Spectrom.* **2011**, *25* (11), 1559–1566.

(55) Shukla, A. K.; Futrell, J. H. Tandem Mass Spectrometry: Dissociation of Ions by Collisional Activation. *J. Mass Spectrom.* **2000**, *35* (9), 1069–1090.

(56) Guilhaus, M. Special Feature: Tutorial. Principles and Instrumentation in Time-of-Flight Mass Spectrometry. Physical and Instrumental Concepts. *J. Mass Spectrom.* **1995**, *30* (11), 1519–1532.

(57) Marshall, A. G. Fourier Transform Mass Spectrometry. *Spectrosc. Biomed. Sci.* **2018**, 87–105.

(58) Perry, R. H.; Cooks, R. G.; Noll, R. J. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. *Mass Spectrom. Rev.* **2008**, *27* (6), 661–699.

(59) Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **2000**, *72* (6), 1156–1162.

(60) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.

(61) Wang, W.; Kitova, E. N.; Klassen, J. S. Influence of Solution and Gas Phase Processes on Protein-Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2003**, *75* (19), 4945–4955.

(62) El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Protein Interactions with Isomeric Carbohydrate Ligands Using a Catch and Release Electrospray Ionization-Mass Spectrometry Assay. *Anal. Chem.* **2013**, *85* (16), 7637–7644.

(63) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Applications of a Catch and Release Electrospray Ionization Mass Spectrometry Assay for Carbohydrate Library Screening. *Anal. Chem.* **2012**, *84* (1), 50–58.

(64) Wigger, M.; Eyler, J. R.; Benner, S. A.; Li, W.; Marshall, A. G. Fourier

Transform-Ion Cyclotron Resonance Mass Spectrometric Resolution, Identification, and Screening of Non-Covalent Complexes of Hck Src Homology 2 Domain Receptor and Ligands from a 324-Member Peptide Combinatorial Library. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (10), 1162–1169.

(65)  Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C.; Jia, F.; Cairo, C. W.; Klassen, J. S. Protein-Glycosphingolipid Interactions Revealed Using Catch-and-Release Mass Spectrometry. *Anal. Chem.* **2012**, *84* (18), 7618–7621.

(66)  Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Klassen, J. S. Identifying Carbohydrate Ligands of a Norovirus P Particle Using a Catch and Release Electrospray Ionization Mass Spectrometry Assay. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (1), 111–119.

(67)  Rezaei Darestani, R.; Winter, P.; Kitova, E. N.; Tuszynski, J. A.; Klassen, J. S. Screening Anti-Cancer Drugs against Tubulin Using Catch-and-Release Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (5), 876–885.

(68)  Li, J.; Fan, X.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Eugenio, L.; Ng, K. K. S.; Xiong, Z. J.; Privé, G. G.; Klassen, J. S. Screening Glycolipids Against Proteins in Vitro Using Picodiscs and Catch-and-Release Electrospray Ionization-Mass Spectrometry. *Anal. Chem.* **2016**, *88* (9), 4742–4750.

(69)  Han, L.; Kitova, E. N.; Li, J.; Nikjah, S.; Lin, H.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. Protein-Glycolipid Interactions Studied in Vitro Using ESI-MS and Nanodiscs: Insights into the Mechanisms and Energetics of Binding. *Anal. Chem.* **2015**, *87* (9), 4888–4896.

(70)  Wang, Y.; Park, H.; Lin, H.; Kitova, E. N.; Klassen, J. S. Multipronged ESI-MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins. *Anal. Chem.* **2019**, *91* (3), 2140–2147.

(71)  Lane, L. A.; Ruotolo, B. T.; Robinson, C. V; Favrin, G.; Benesch, J. L. P. A Monte

Carlo Approach for Assessing the Specificity of Protein Oligomers Observed in Nano-Electrospray Mass Spectra. *Int. J. Mass Spectrom.* **2009**, *283* (1–3), 169–177.

(72)  Van Dongen, W. D.; Heck, A. J. R. Binding of Selected Carbohydrates to Apo-Concanavalin A Studied by Electrospray Ionization Mass Spectrometry. *Analyst* **2000**, *125* (4), 583–589.

(73)  El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. Quantifying Labile Protein-Ligand Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (11), 1893–1899.

(74)  Sun, J.; Kitova, E. N.; Klassen, J. S. Method for Stabilizing Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2007**, *79* (2), 416–425.

(75)  Liu, L.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Fatty Acid Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (2), 310–318.

(76)  Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Gas Phase Stabilization of Noncovalent Protein Complexes Formed by Electrospray Ionization. *Anal. Chem.* **2009**, *81* (18), 7801–7806.

(77)  Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. Top-down ESI-ECD-FT-ICR Mass Spectrometry Localizes Noncovalent Protein-Ligand Binding Sites. *J. Am. Chem. Soc.* **2006**, *128* (45), 14432–14433.

(78)  Zhang, H.; Lu, H.; Chingin, K.; Chen, H. Stabilization of Proteins and Noncovalent Protein Complexes during Electrospray Ionization by Amino Acid Additives. *Anal. Chem.* **2015**, *87* (14), 7433–7438.

(79)  Landreh, M.; Alvelius, G.; Johansson, J.; Jörnvall, H. Protective Effects of Dimethyl Sulfoxide on Labile Protein Interactions during Electrospray Ionization. *Anal. Chem.* **2014**, *86* (9), 4135–4139.

(80) Hossain, B. M.; Konermann, L. Pulsed Hydrogen/Deuterium Exchange MS/MS for Studying the Relationship between Noncovalent Protein Complexes in Solution and in the Gas Phase after Electrospray Ionization. *Anal. Chem.* **2006**, *78* (5), 1613–1619.

(81) Wang, W.; Kitova, E. N.; Klassen, J. S. Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Factors Influencing Their Formation and Stability. *Anal. Chem.* **2005**, *77* (10), 3060–3071.

(82) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. Method for Distinguishing Specific from Nonspecific Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2006**, *78* (9), 3010–3018.

(83) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. Nonspecific Interactions Between Proteins and Charged Biomolecules in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (3), 472–481.

(84) Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. Method for Identifying Nonspecific Protein-Protein Interactions in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2007**, *79* (21), 9301–9311.

(85) Daubenfeld, T.; Bouin, A. P.; van der Rest, G. A Deconvolution Method for the Separation of Specific Versus Nonspecific Interactions in Noncovalent Protein-Ligand Complexes Analyzed by ESI-FT-ICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (9), 1239–1248.

(86) Shimon, L.; Sharon, M.; Horovitz, A. A Method for Removing Effects of Nonspecific Binding on the Distribution of Binding Stoichiometries: Application to Mass Spectroscopy Data. *Biophys. J.* **2010**, *99* (5), 1645–1649.

(87) Sun, N.; Sun, J.; Kitova, E. N.; Klassen, J. S. Identifying Nonspecific Ligand Binding in Electrospray Ionization Mass Spectrometry Using the Reporter Molecule Method. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (7), 1242–1250.

(88) Kitova, E. N.; Soya, N.; Klassen, J. S. Identifying Specific Small-Molecule

Interactions Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2011**, *83* (13), 5160–5167.

(89)    Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. Direct Quantification of Protein-Metal Ion Affinities by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2010**, *82* (6), 2170–2174.

(90)    Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. Influence of Response Factors on Determining Equilibrium Association Constants of Non-Covalent Complexes by Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2003**, *38* (5), 491–501.

(91)    Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. Affinities of Shiga Toxins 1 and 2 for Univalent and Oligovalent Pk-Trisaccharide Analogs Measured by Electrospray Ionization Mass Spectrometry. *Glycobiology* **2007**, *17* (10), 1127–1137.

(92)    Chitta, R. K.; Rempel, D. L.; Gross, M. L. Determination of Affinity Constants and Response Factors of the Noncovalent Dimer of Gramicidin by Electrospray Ionization Mass Spectrometry and Mathematical Modeling. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (7), 1031–1038.

(93)    Lin, H.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Ligand Interactions by Direct Electrospray Ionization-Ms Analysis: Evidence of Nonuniform Response Factors Induced by High Molecular Weight Molecules and Complexes. *Anal. Chem.* **2013**, *85* (19), 8919–8922.

## Chapter 2

## Multipronged ESI-MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins

## 2.1 Introduction

Non-covalent interactions between glycans and glycan-binding proteins (GBPs) underpin many important biological processes, including cell recognition and signaling, the immune response, bacterial and viral infections and cancer metastasis.[1] In mammals, glycans are present as free oligosaccharides (e.g., milk oligosaccharides and free *N*-glycans in serum) or associated with lipids (glycolipids), peptides (glycopeptides) and proteins (glycoproteins (GPs) and proteoglycans).[1] Most membrane and secreted proteins produced by mammalian cells are glycosylated with *N*- and *O*-linked glycans.[1]

The detection and characterization of GBP-GP binding is challenging due to the heterogeneity of GPs, which typically exist in many different glycosylation forms (i.e., glycoforms), and the relatively low affinities of most glycan-mediated GBP-GP interactions.[2,3] Water-soluble GBP-GP complexes are commonly identified using affinity methods (co-immunoprecipitation, affinity chromatography).[4,5] However, these methods do not provide any insight into the glycan motifs being recognized by the GBP or the affinities of the interactions. The GP glycan specificities of GBPs can be also probed using glycan microarrays produced from purified/fractionated *O*- and *N*-glycans released chemically or enzymatically (from GPs), or from synthetic oligosaccharides.[6] Microarray screening is rapid and uses relatively small amounts of sample.[7] However, the method is not quantitative, thereby necessitating the use of conventional, low-throughput binding assays, such as isothermal titration calorimetry (ITC),[8] to determine apparent affinities of the interactions. Moreover, chemical modification of the GBP or glycan, which is required for immobilization, may alter their binding properties and false negatives for low affinity interactions are often produced.[9–11] Given the limitations of existing methods,

there is a need for new analytical techniques to aid in the discovery and characterization of GBP-GP interactions.

Electrospray ionization mass spectrometry (ESI-MS) is increasingly used to study glycan interactions with water-soluble GBPs.[12–18] The direct ESI-MS assay, which is based on the detection and quantification of free and ligand-bound protein ions produced from buffered aqueous solutions, can measure the affinity and stoichiometry of GBP-oligosaccharide interactions *in vitro*.[12,13] Importantly, the method is label free and, thereby, avoids artefacts arising from chemical modifications, doesn't require immobilization of one of the binding partners, which necessarily alters the binding thermochemistry, and can directly quantify interactions with affinities ranging from approximately $10^3$ M$^{-1}$ to $10^7$ M$^{-1}$.[12,13] Moreover, because multiple binding equilibria can be monitored simultaneously, the assay is readily applied to mixtures of oligosaccharides.[12,13,19–21] In cases where the relative abundances of the free and ligand-bound GBP cannot be accurately measured by ESI-MS (because of high MW or heterogeneity of the GBP), competitive binding assays can be used to quantify GBP-glycan interactions.[20–22] The *catch-and-release* (CaR)-ESI-MS assay, where ligands 'caught' by the GBP in solution are 'released' as gas-phase ions by collision-induced dissociation (CID) and identified by molecular weight (MW) alone or in combination with ion mobility separation (IMS) or CID fingerprinting, enables high-throughput screening of oligosaccharide libraries.[19,23–27]

Although ESI-MS is now an established method for studying GBP interactions with oligosaccharides and glycolipids, to the best of our knowledge there are no reports describing its application to GBP-GP binding mediated by glycans. Complexes of human serum GP haptoglobin phenotype 1-1 (Hp1-1) with hemo- and myoglobin were identified by ESI-MS and the binding stoichiometry established.[28] However, these interactions do not involve the Hp1-1 glycans. Intact GBP-GP complexes have been successfully transferred from aqueous solution to the gas phase by nanoflow ESI (nanoESI). However,

the resulting complex ions were detected with a gas-phase electrophoretic mobility molecular analyzer and not MS.[29]

Here, we describe a multipronged ESI-MS approach for characterizing glycan-mediated complexes of water-soluble GBPs and GPs. First, CaR-ESI-MS implemented with IMS (CaR[IMS]-ESI-MS) is employed to detect GBP-GP binding in solution. The *proxy ligand*-ESI-MS assay is then used to measure the apparent affinity ($K_{a,app}$) of the GBP for the GP. Finally, to identify glycans that are recognized by the GBP, the *N*-glycans are enzymatically released from the GP and screened, as libraries, against the GBP by ESI-MS. The affinities of released *N*-glycans (grouped as compositional isomers) are ranked based on ESI-MS binding data measured at multiple GBP concentrations.[30] The interactions between a C-terminal domain fragment of human galectin-3 (hGal-3C) to three positive acute phase human serum GPs, Hp1-1, alpha-1-acid glycoprotein (AGP) and alpha-2-macroglobulin ($\alpha$2M), which were previously identified by affinity chromatography and proteomics analysis,[5,31,32] served as model systems to demonstrate the implementation of the method.

## 2.2 Experimental

### 2.2.1 Materials

#### 2.2.1.1 Proteins

The recombinant fragment of the C-terminal carbohydrate recognition domain (residues 107–250) of human galectin-3 (hGal-3C, MW 16,327 Da) was a gift from Prof. C. Cairo (University of Alberta). Human plasma $\alpha$2M and Hp phenotype 1-1 (Hp1-1) were purchased from Athens Research and Technology (Athens, GA). Human plasma AGP and bovine carbonic anhydrase (BCA, MW 27,000 Da) were purchased from Sigma-Aldrich Canada (Oakville, Canada). The single chain fragment (scFv, MW 26,539 Da) of the monoclonal antibody Se155-4 was produced using recombinant technology.[33] The scFv and BCA served as reference proteins ($P_{ref}$) to correct the mass spectra for the formation of nonspecific interactions during the ESI process.[34] All proteins were dialyzed

against an aqueous solution of 200 mM ammonium acetate (pH 6.8) using an Amicon 0.5 mL microconcentrator (EMD Millipore, Billerica, MA) with a MW cut-off of 10 kDa (AGP, BCA and scFv) or 30 kDa (Hp1-1 and α2M) and stored at -20 $^{\circ}$C until needed. The concentrations of the protein stock solutions were estimated by UV absorption (280 nm).

## 2.2.1.2 Oligosaccharides

Lacto-N-neooctaose (LNnO, β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc, MW 1437.36 Da) was purchased from Elicityl SA (Crolles, France) and served as the proxy ligand (L$_{proxy}$). The *N*-glycans of AGP, Hp1-1 and α2M were enzymatically released, as free oligosaccharides, using peptide *N*-glycosidase F (PNGase F) (New England BioLabs, MA, USA),[35] isolated and desalted using a PD MiniTrap G-10 pre-packed column (GE Healthcare, Buckinghamshire, UK) and stored at -20 $^{\circ}$C until needed.

To produce the *N*-glycan libraries, 500 μg of GP (AGP, Hp1-1 or α2M) were dissolved in 500 μL of 8 M urea in 100 mM Tris-HCl (pH 8.0) containing 3 mM EDTA, and incubated at room temperature for 1 h. The denatured GPs were then reduced with 10 μL of 500 mM dithiothreitol (DTT) at room temperature for 1 h followed by alkylation with 23 μL of 500 mM iodoacetamide (IAA) at room temperature for 20 min in the dark. The reaction was quenched by adding 10 μL of 250 mM DTT, and the solution buffer exchanged using a PD MidiTrap G-25 column (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's instructions. The GP solution was subsequently digested with trypsin/chymotrypsin [Substrate/Enzyme (weight/weight) = 50] in 50 mM ammonium bicarbonate (pH 8.0) for 18 h at 37 °C. The reaction was quenched by heat inactivation at 100 °C for 10 min. The resulting glycopeptides were incubated in the presence of 2500 U (5 μL) of PNGase F (New England BioLabs, MA, USA) at 37 °C for 18 h. Released *N*-glycans were first purified using porous graphitized carbon (Hypercarb

cartridges, 100 mg, 1 mL volume, Thermo Fisher Scientific). The porous graphitized carbon cartridge was washed with 1 mL 80% acetonitrile containing 0.1% trifluoroacetic acid (TFA) followed by 2 mL of water. The samples were applied to the cartridges. After washing with 2 mL of water, the *N*-glycans were eluted with 1 mL of 25% acetonitrile containing 0.1% TFA. The collected *N*-glycans were desalted in a second purification step with gel filtration on PD MiniTrap G-10 column (GE Healthcare) according to manufacturer's instructions. The purified *N*-glycans were then lyophilized and stored at −20 °C until used.

## 2.2.2 Mass spectrometry

All MS experiments were carried out using a Synapt G2S quadrupole-ion mobility separation-time-of-flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK) equipped with a nanoflow ESI (nanoESI) source. Mass spectra were measured in both positive and negative ion modes. Cesium iodide (concentration 1 mg mL$^{-1}$) was used for calibration. The nanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.78 mm i.d.), pulled to ~5 μm outer-diameter using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). To perform ESI, a platinum wire was inserted into the solution and a voltage of ~1.0 kV (positive ion mode) or ~0.75 kV (negative ion mode) was applied. Unless otherwise noted, a cone voltage of 20 V was used; the source block temperature was maintained at 60 ºC for all experiments. For the CaR$^{IMS}$-ESI-MS measurements, a Trap voltage of 5 V was used and the Transfer voltage was varied from 10 V to 200 V. IMS was carried out using a wave velocity of 600 m s$^{-1}$ and a Wave Height of 35.0 V; the helium and nitrogen gas flow rates were 180 mL min$^{-1}$ and 90 mL min$^{-1}$, respectively. Argon was used in the Trap and Transfer ion guides at pressures of 2.22 x 10$^{-2}$ mbar and 3.36 x 10$^{-2}$ mbar, respectively. Data acquisition and processing were carried out using MassLynx (v4.1). Spectral deconvolution was performed with the UniDec deconvolution algorithm.[36]

**2.2.3 *Proxy ligand* ESI-MS assay**

The *proxy ligand* ESI-MS assay, which is based on competitive ligand binding and direct ESI-MS measurements, was used to quantify the apparent affinity ($K_{a,app}$) of hGal-3C (P) for a given GP.[21] The assay employs a proxy ligand ($L_{proxy}$), which binds to P with known affinity.[10] The extent of $L_{proxy}$ binding to P, as determined by ESI-MS, provides a quantitative measure of the extent of P binding to the GP.[21] The $K_{a,app}$ was calculated by fitting (Igor Pro 6, WaveMetrics Inc.) Eq. 2.1 to the measured $R_{proxy}$ (the total abundance ratio of the gaseous $PL_{proxy}$ to P ions, Eq. 2.2) over a range of GP concentrations:[21]

$$[GP]_0 = \left( \frac{R_{proxy}[P]_0}{[L_{proxy}]_0 - \dfrac{R_{proxy}}{K_{a,proxy}}} - R_{proxy} - 1 \right) \left( \frac{1}{K_{a,app}} + \frac{[L_{proxy}]_0}{R_{proxy}} - \frac{1}{K_{a,proxy}} \right) \qquad (2.1)$$

where $[P]_0$, $[GP]_0$ and $[L_{proxy}]_0$ are the initial concentrations of P, GP and $L_{proxy}$, respectively. The $[GP]_0$ was estimated from the mass of the sample used to prepare the stock solution and the weighted average MW established from ESI-MS data.

$$R_{proxy} = \frac{\sum Ab(PL_{proxy})}{\sum Ab(P)} = \frac{[PL_{proxy}]}{[P]} \qquad (2.2)$$

**2.2.4 ESI-MS screening of *N*-glycan libraries**

The direct ESI-MS assay was used to screen mixtures of *N*-glycans produced from a given GP against P. Ligands were identified from the measured MWs of intact P-glycan (i.e., PL) complexes observed in the ESI mass spectra. Because the concentrations of the *N*-glycans in a given library were not known, it was not possible to directly evaluate their apparent affinities (of each isomer set) for hGal-3C directly from the mass spectrum. Instead, the percentage change in the $R$ of the gaseous PL to P ions (i.e., $\Delta R\%$) resulting from a change in P concentration (Eq. 2.3) was used to distinguish low affinity ligands (association constant ($K_a$) $\sim 10^3$ $M^{-1}$) from moderate-to-high affinity ($K_a > 10^4$ $M^{-1}$) ligands:[30]

$$\Delta R\% = \frac{R_{initial} - R_{final}}{R_{initial}} \times 100\% \tag{2.3}$$

where $R_{initial}$ and $R_{final}$ are the $R$ ratios (Eq. 2.2) measured at the two P concentrations for a given $N$-glycan ligand (grouped as compositional isomers).

## 2.3 Results and Discussion

### 2.3.1 Detecting hGal-3C interactions with serum GP by CaR[IMS]-ESI-MS

Identification of interactions between water-soluble GBPs and GPs by direct ESI-MS analysis is complicated by the heterogeneity of GPs. Additionally, because of the low affinities typical of GBP-GP interactions, relatively high concentrations are often required to promote the formation of detectable concentrations of complex. This, in turn, can lead to the formation of nonspecific GBP-GP complexes during the ESI process, which may further hinder the identification of specific interactions. As illustrated below, the CaR[IMS]-ESI-MS assay, implemented using a suitable $P_{ref}$, overcomes these challenges and can serve as a sensitive, albeit qualitative, method for establishing the presence of specific GBP-GP interactions *in vitro*.

**AGP.** Human AGP, a monomeric protein found in plasma at concentrations ranging from 0.4 to 1.5 mg mL$^{-1}$, is glycosylated at five different sites with complex sialylated di-,tri and tetra-antennary $N$-glycans, which may be fucosylated.[37,38] Shown in Figures 2.1a and 1b are representative mass spectra acquired in positive ion mode for aqueous 200 mM ammonium acetate solutions (pH 6.8, 25 °C) of AGP (0.23 mg mL$^{-1}$) alone and in the presence of hGal-3C (2.5 µM) and $P_{ref}$ (scFv, 2.5 µM), respectively, using a cone voltage of 20 V. Also shown are the corresponding IMS 2D heat maps (plots of ion $m/z$ versus IMS arrival (drift) time). Inspection of Figure 2.1a reveals three broad features at $m/z$ from approximately 3000 to 4000, which are attributed to the multiple glycoforms of AGP monomer, at charge states +10 to +12 (estimated from differences in $m/z$ of the same putative glycoforms at different charge states). Under the 'gentle' instrumental conditions used to acquire the mass spectrum, which minimize collisional heating of the

gaseous ions, the individual AGP glycoforms are poorly resolved. However, they become distinguishable at higher cone voltages (Figure 2.2).
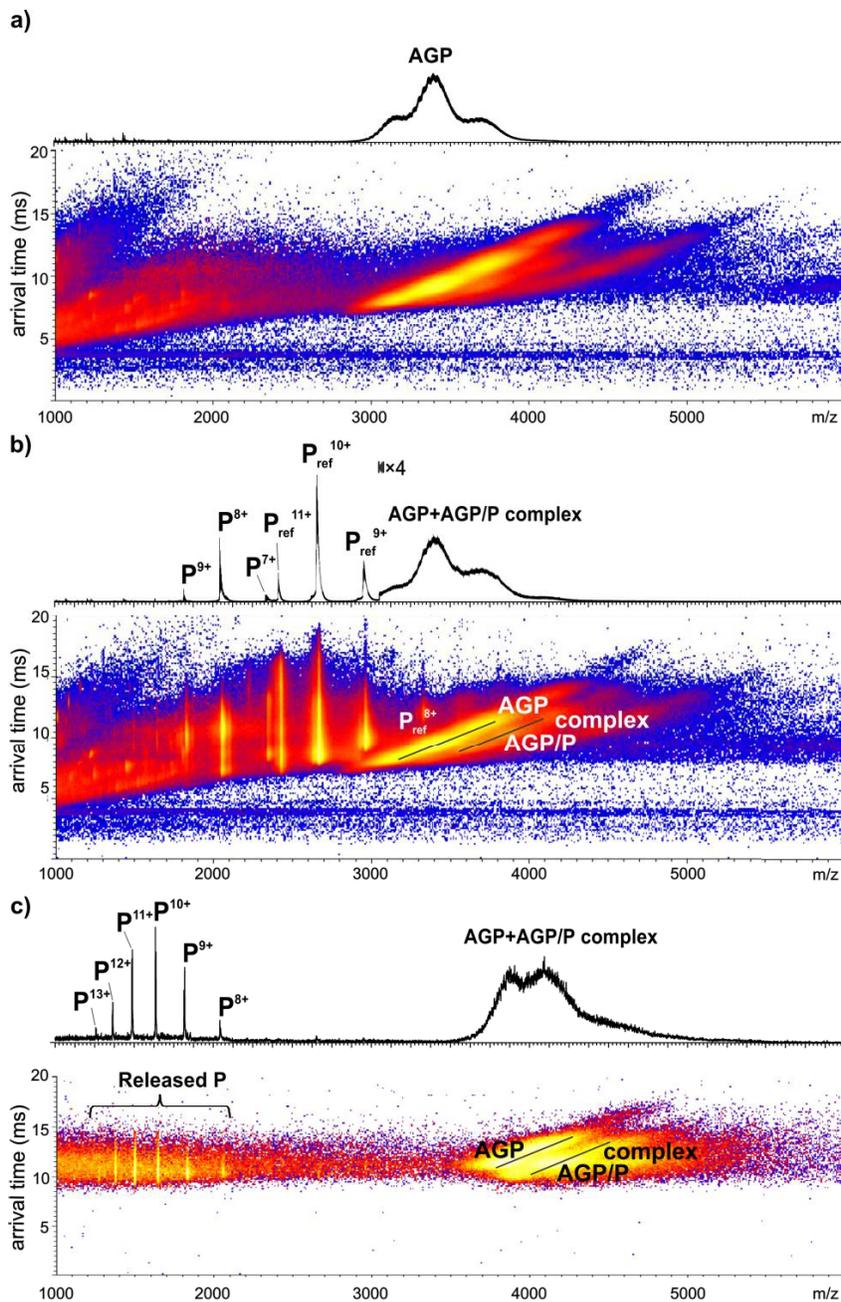


**Figure 2.1.** ESI mass spectrum and corresponding IMS heat map (arrival time versus *m/z*; white-blue-red-yellow: 0~33%~66%~100% counts) acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of AGP (0.23 mg mL$^{-1}$)

(a) alone and (b) with hGal-3C (P, 2.5 μM) and scFv (P$_{ref}$, 2.5 μM). (c) CID mass spectrum and corresponding IMS heat map of ions with *m/z* ≥3800, produced from the solution described in (b), at a Transfer voltage of 90 V.
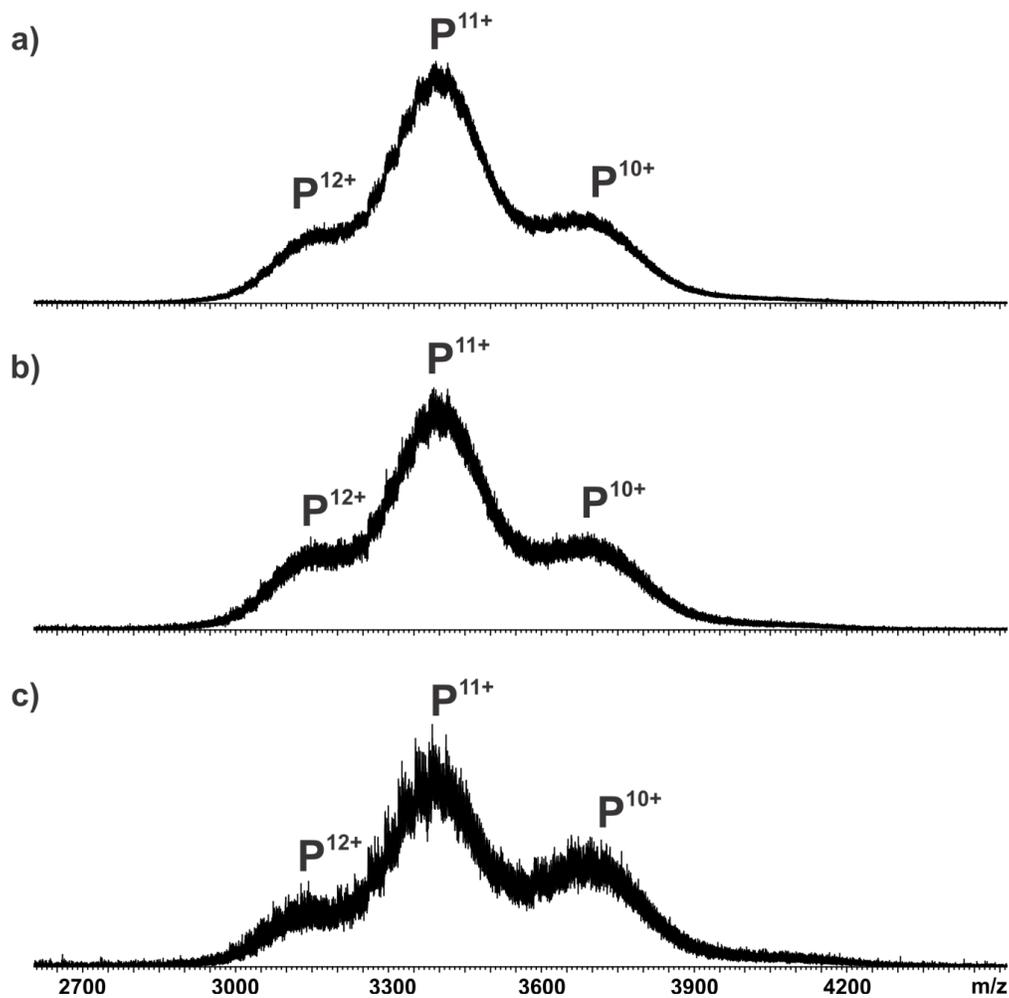


**Figure 2.2.** ESI mass spectra acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of AGP (0.23 mg mL$^{-1}$) at cone voltages (a) 20 V, (b) 60 V and (c) 100 V.

This is presumably a result of shedding buffer (ammonium acetate) adducts due to collisional activation in the source. Analysis of the mass spectra acquired at high cone voltages (e.g. 100 V) allowed for an estimation of the charge states of the AGP ions,

which were then used in the determination of the AGP MW. The weighted average MW of AGP determined from the deconvoluted mass spectrum (using the UniDec deconvolution algorithm)[36] measured at cone voltage 20 V was found to be 37.35 kDa, (Figure 2.3a). Similar values were obtained from mass spectra measured at cone voltage 60 V (37.35 kDa) and 100 V (37.25 kDa) (Figures 2.3b and 2.3c). These values fall within the range of previously reported AGP MWs, 30 kDa to 40 kDa.[39–42]
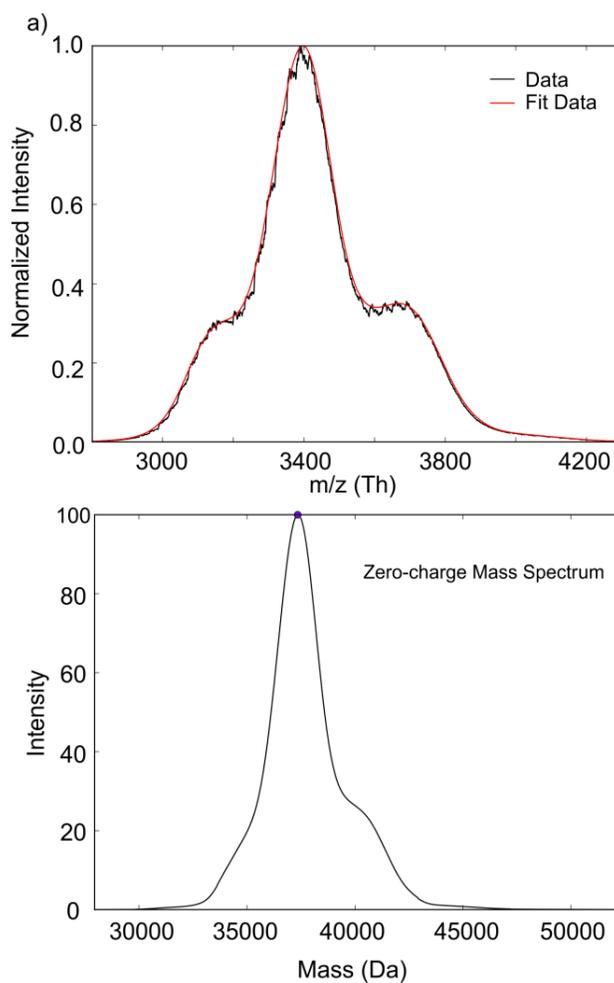


**Figure 2.3.** (a) Analysis of the ESI mass spectrum of AGP acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of AGP (0.23 mg mL$^{-1}$) at cone voltages 20 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 37.35 kDa. Spectral

deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 2800 to 4300; Subtract Minimum - 0.0; Gaussian Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 10 to 12; Mass range 33,000 to 50,000; Sample Mass Every 50.0 Da; Peak FWHM (Th) 80.0; Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.998.
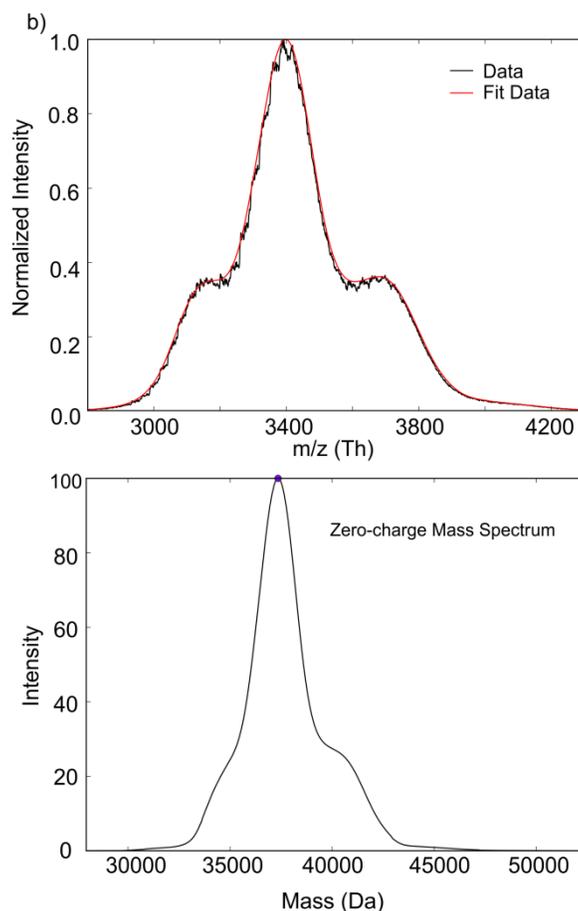


**Figure 2.3.** (b) Analysis of the ESI mass spectrum of AGP acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of AGP (0.23 mg mL$^{-1}$) at cone voltages 60 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 37.35 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 2800 to 4300; Subtract Minimum - 0.0; Gaussian

Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 10 to 12; Mass range 33,000 to 50,000; Sample Mass Every 50.0 Da; Peak FWHM (Th) 80.0; Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting *R* square value is 0.99.
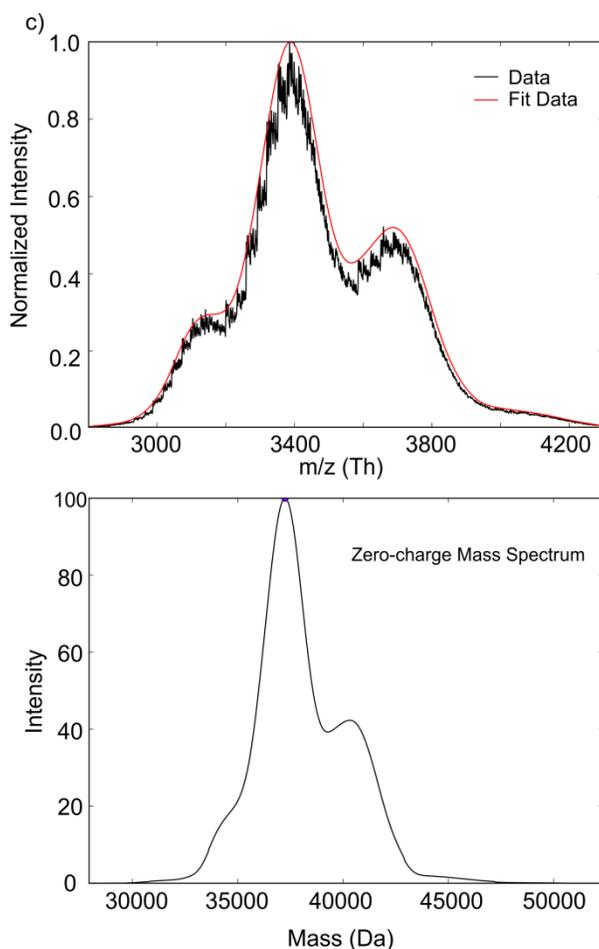


**Figure 2.3.** (c) Analysis of the ESI mass spectrum of AGP acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of AGP (0.23 mg mL$^{-1}$) at cone voltages 100 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 37.25 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 2800 to 4300; Subtract Minimum - 0.0; Gaussian Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 10 to 12;

Mass range 33,000 to 50,000; Sample Mass Every 50.0 Da; Peak FWHM (Th) 80.0; Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.97.



**Figure 2.3.** (d) Analysis of the mass spectrum of the hGal-3C-AGP complex acquired in positive ion mode for 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of hGal-3C (2.5 µM), scFv (2.5 µM) and AGP (0.23 mg mL$^{-1}$) at cone voltage of 20 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 47.70 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 3100 to 4500; Subtract Minimum - 0.0; Gaussian Smoothing - 0; Linear *m/z* (Constant delta *m/z*);

Bin every 1.0; Charge Range - 13 to 15; Mass range 44,000 to 58,000; Sample Mass Every 50.0 Da; Peak FWHM (Th) 80.0; Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.99.



**Figure 2.3.** (e) Analysis of the mass spectrum of the hGal-3C-AGP complex acquired in positive ion mode for 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of hGal-3C (2.5 µM), scFv (2.5 µM) and AGP (0.23 mg mL$^{-1}$) at cone voltage of 20 V. The mass spectrum was extracted from DriftScope in the region of hGal3C-AGP complex corresponding to >55% of the total ion signal. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to a MW of 53.75 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 3500 to 4500; Subtract Minimum - 0.0; Gaussian

Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 13 to 15; Mass range 45,000 to 60,000; Sample Mass Every 50.0 Da; Peak FWHM (Th) 80.0; Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.96.

In Figure 2.1b, ions corresponding to hGal-3C and $P_{ref}$, at charge states +7 to +9 and +9 to +11, respectively, dominated the mass spectrum. However, broad features, presumably related to AGP ions, at *m/z* from approximately 3000 to 4500 were detectable. The presence of more abundant signal at higher *m/z* is suggestive of the presence of hGal-3C-AGP complexes. However, the deconvoluted mass spectrum fails to provide unambiguous evidence of hGal-3C-AGP binding (Figure 2.3d).

Inspection of the heat map in Figure 2.1a reveals that the majority of the AGP ions exhibit IMS arrival times (ATs) ranging from approximately 8 ms to 14 ms. The addition of hGal-3C to the solution resulted in the appearance of a second band, adjacent to that of free AGP but at higher *m/z* (Figure 2.1b). Notably, deconvolution of the mass spectrum associated with the region of the heat map that contains the second band yields a weighted average MW of 53.75 kDa, which is consistent with the expected MW (53.68 kDa) of the 1:1 hGal-3C-AGP complex (Figure 2.3e). To conclusively establish the presence of ions corresponding to hGal-3C-AGP complex, $CaR^{IMS}$-ESI-MS was performed. To implement the assay, ions with *m/z* >3800 (which excludes the free $P_{ref}$ ions, including the +8 charge state, which was visible in the IMS heat map, Figure 2.1b) were selected using the quadrupole mass filter and subjected to collisional activation, post IMS, in the Transfer region at voltages ranging from 10 V to 150 V. At Transfer voltages >70 V, ion signal corresponding to protonated hGal-3C, at charge states +8 to +13, was detected (Figure 2.1c). The broad charge state distribution of the released hGal-3C ions is consistent with the general phenomenon of asymmetric partitioning of charge observed in CID of multiprotein complexes in the gas-phase and suggests that

hGal-3C undergoes partial unfolding during the dissociation process.[43] Importantly, the IMS-ATs (8.8 ms to 14.4 ms) of the hGal-3C ions fall within the same range of the ATs (8.8 ms to 15.0 ms) measured for the putative hGal-3C-AGP complexes, which is further evidence that the hGal-3C ions originated from complexes with AGP (Figure 2.1b). Notably, no ion signal corresponding to the release of the negative control, $P_{ref}$, was detected, suggesting that little, if any, nonspecific binding of hGal-3C to AGP occurred during the ESI process.
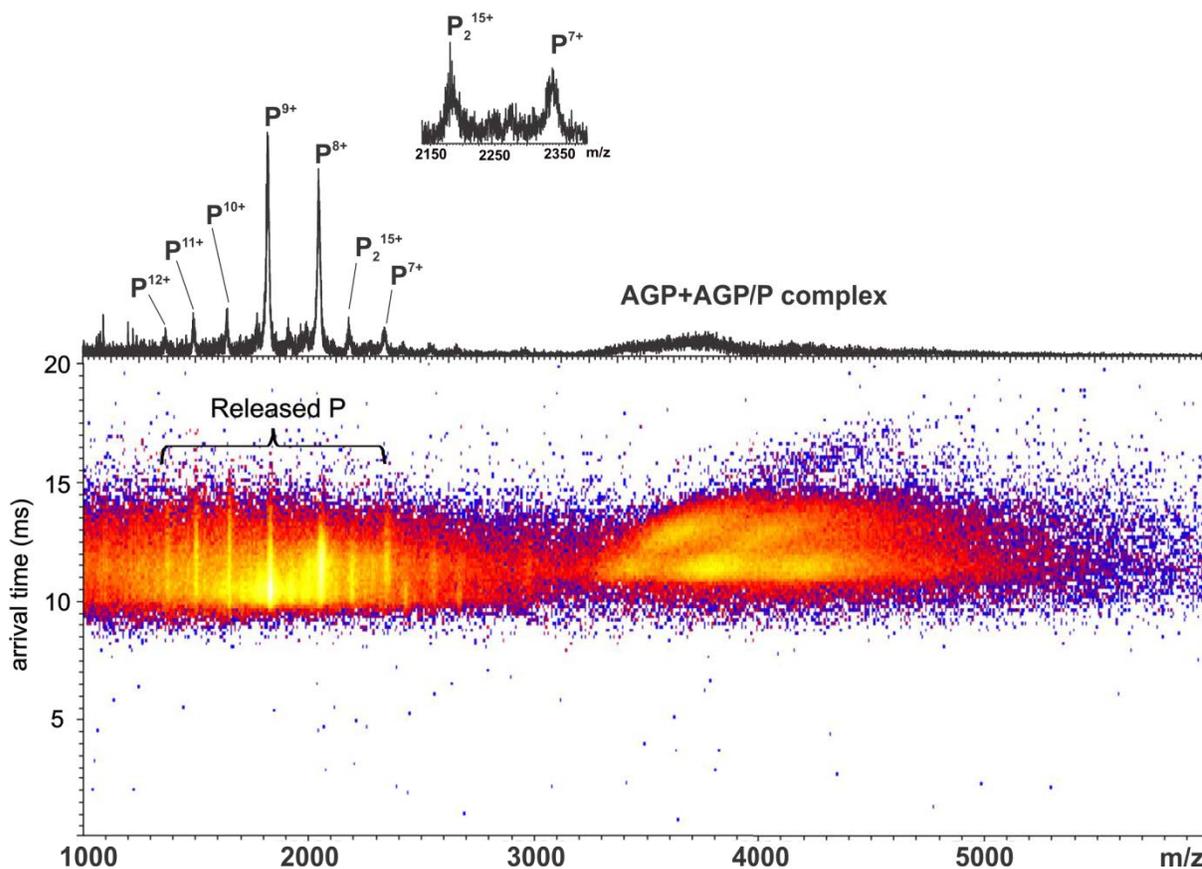


**Figure 2.4.** CID mass spectrum and corresponding IMS heat map (arrival time versus *m/z*; white-blue-red-yellow: 0~33%~66%~100% counts) of ions with *m/z* ≥3800 produced by ESI performed in positive ion mode on a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of hGal-3C (P, 2.5 µM), scFv ($P_{ref}$, 2.5 µM) and AGP (0.23 mg mL$^{-1}$) at a Transfer voltage of 150 V.

Interestingly, at Transfer voltages >120 V, CaR[IMS]-ESI-MS produced ions corresponding to hGal-3C homodimer, albeit at low abundance (Figure 2.4). The observation of dimer suggests the presence of a small amount of the 1:2 complex, i.e., (AGP + 2hGal-3C), and, possibly, higher order complexes in solution. Non-covalent cross linking of hGal-3C bound to AGP, either in solution or after transfer to the gas phase, could, in principle, lead to intact homodimers being ejected upon collisional activation. That a higher collision voltage (energy) is required to observe homodimer ions is consistent with the need to break more intermolecular interactions to release the hGal-3C homodimer, compared to that required for the release of monomer.[44]

Having successfully demonstrated the use of CaR[IMS]-ESI-MS for the detection of interactions between AGP and hGal-3C, the assay was then applied to solutions of hGal-3C with Hp1-1 or α2M in order to establish the general utility of the approach for detecting specific GBP-GP interactions *in vitro*. The results are summarized below.

**Hp1-1.** Human Hp1-1, one of three primary phenotypes of Hp, exists as a non-covalently linked homodimer consisting of a pair of covalently bound α-β subunits.[45] Each β subunit in Hp has four *N*-glycosylation sites comprising mainly bi- and tri-antennary sialylated complex *N*-glycans, that may be mono- or di-fucosylated.[46] Shown in Figures 2.5a and 2.5b are representative mass spectra acquired in positive ion mode for aqueous solutions (200 mM ammonium acetate, pH 6.8 and 25 °C) of Hp1-1 (0.25 mg mL$^{-1}$) alone and in the presence of hGal-3C (2.5 μM) and P$_{ref}$ (scFv, 2.5 μM), respectively, measured at a cone voltage of 20 V. Inspection of the mass spectrum reveals ion signal consistent with the Hp1-1 homodimer at charge states +18 to +21, *vide infra*. Unlike AGP, the charge states of the Hp1-1 ions could not be directly identified from mass spectra acquired at higher cone voltages and were instead determined from UniDec analysis of the mass spectrum. The average MW measured for Hp1-1 determined from the deconvoluted mass spectrum, 93.45 kDa (Figure 2.6), falls within the range of reported MWs, 86 kDa to 96 kDa, of the homodimer.[42,47]
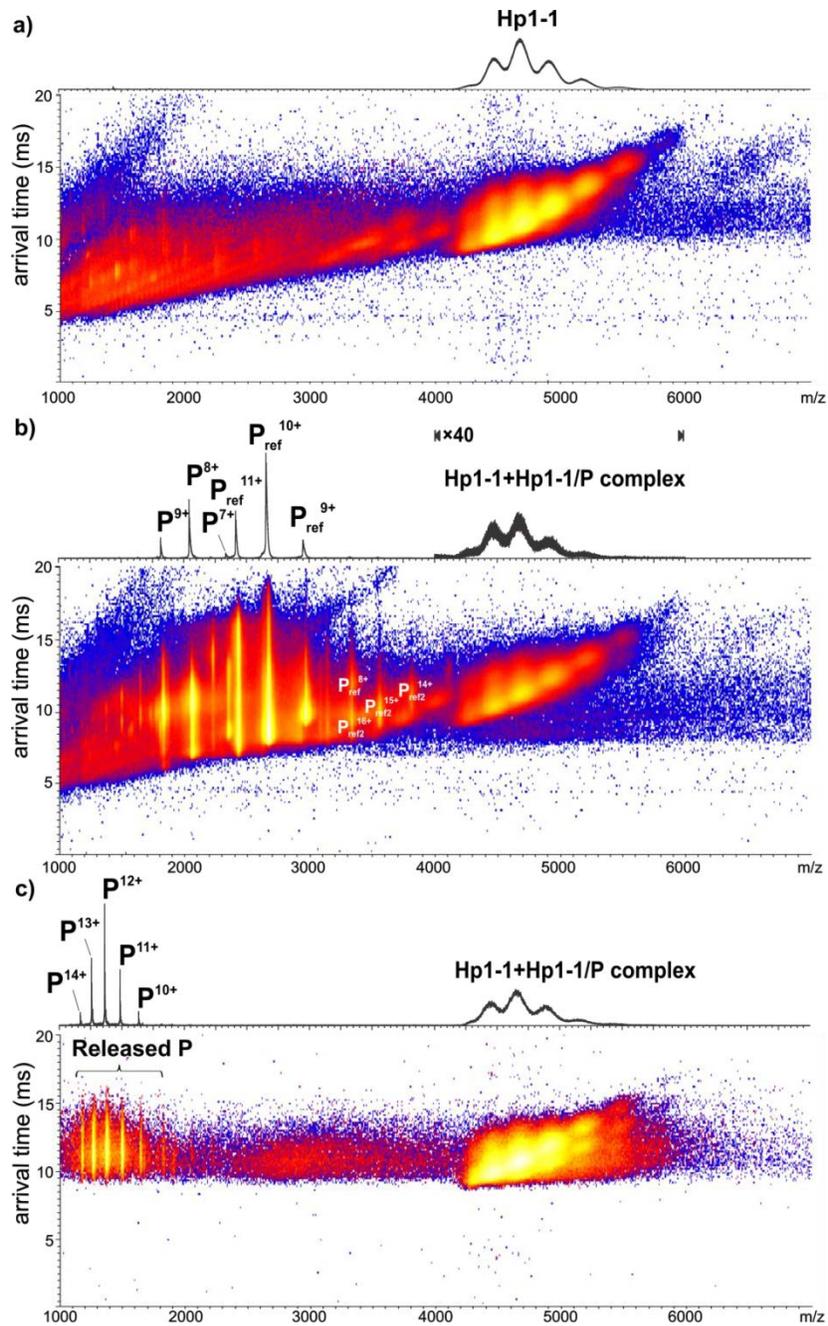
**Figure 2.5.** ESI mass spectrum and corresponding IMS heat map (arrival time versus *m/z*; white-blue-red-yellow: 0~33%~66%~100% counts) acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of Hp1-1 (0.25 mg mL$^{-1}$) (a) alone and (b) with hGal-3C (P, 2.5 µM) and scFv (P$_{ref}$, 2.5 µM). (c) CID mass spectrum and corresponding IMS heat map of ions with *m/z* ≥4200 from solution described in (b), at a Transfer voltage of 90 V.
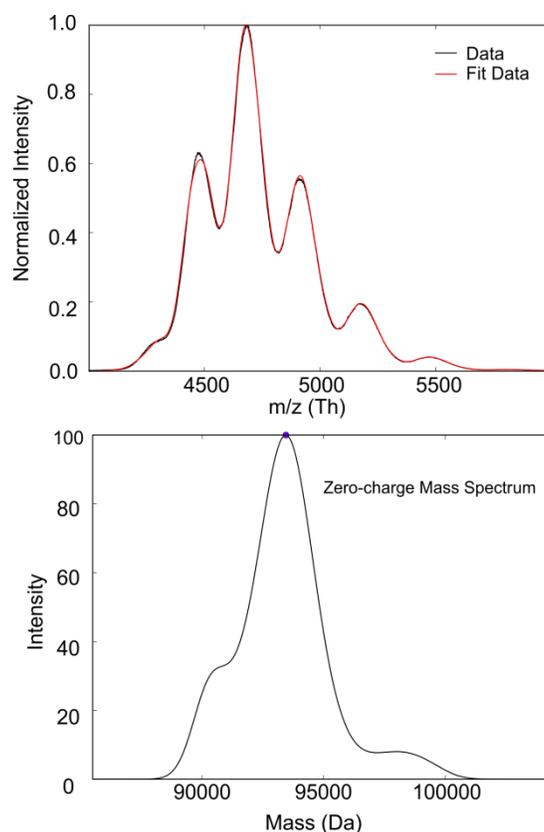
**Figure 2.6.** Analysis of the mass spectrum of Hp1-1 acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of Hp1-1 (0.25 mg mL$^{-1}$) at cone voltage 20 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 93.45 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 4000 to 6000; Subtract Minimum - 0.0; Gaussian Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 16 to 22; Mass range 90,000 to 100,000; Sample Mass Every 50.0 Da; Peak FWHM (Th) 80.0; Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.999. Peaks in Zero-charge Mass Spectrum peaks are labelled as MW values (Da) of detected species.

The addition of hGal-3C to solution did not observably affect the *m/z* or IMS-ATs of the Hp1-1 ions (Figure 2.5b) and analysis of the deconvoluted mass spectrum failed to

provide a MW consistent with the presence of hGal-3C binding to Hp1-1. However, the presence of hGal-3C-bound Hp1-1 ions could be established from CaR$^{IMS}$-ESI-MS performed on ions with $m/z$ >4200. At Transfer voltages >60 V, signal corresponding to protonated hGal-3C, at charge states +10 to +14, was detected; no $P_{ref}$ ions were observed (Figure 2.5c). The IMS-ATs (9.0 ms to 15.6 ms) of the hGal-3C ions fall within the same range of the ATs measured for the Hp1-1 homodimer ions (8.7 ms to 15.6 ms, Figure 2.5b), a finding consistent with the hGal-3C ions originating from hGal-3C-Hp1-1 complexes. In contrast to what was observed for AGP, higher Transfer voltages did not produce signal corresponding to the hGal-3C homodimer (data not shown). The absence of hGal-3C homodimer ions may reflect the lower probability of non-covalent cross-linking of hGal-3C due to the larger size of Hp1-1, compared to AGP.

**α2M.** Human α2M is known to exist as both a covalent homodimer, composed of a pair of identical subunits joined by two disulfide bonds,[48] and a non-covalent dimer of the covalent homodimer (which will be referred to as the α2M homotetramer in this work).[49] Each subunit, which has an estimated MW of 180 kDa,[48] has eight *N*-glycosylation sites comprising both high-mannose and complex type *N*-glycans, which may also be mono- and di-fucosylated.[38] Shown in Figures 2.7a and 2.7b are representative mass spectra acquired in positive mode for aqueous solutions of α2M (1.7 mg mL$^{-1}$) alone and in the presence of hGal-3C (2.5 μM) and $P_{ref}$ (scFv, 2.5 μM), respectively, measured at a cone voltage of 20 V. In Figure 2.7a, two ion distributions, centred at $m/z$ ~10000 and $m/z$ ~13000, with IMS-ATs ranging from 20.0 ms to 31.4 ms and 28.0 ms to 39.0 ms, respectively, were observed. These two distributions are attributed to α2M homodimer, at charge states from +37 to +42, and α2M homotetramer, at charge states from +53 to +61. Based on the relative ion abundances and assuming uniform ESI-MS response factors, approximately 43% and 57% of the α2M subunit exist as dimer and tetramer, respectively. However, given that the homodimer (due to its smaller size) is expected to have a higher ESI-MS response factor than the homotetramer,

it is likely that the homotetramer represents the dominant form in solution. From the deconvoluted mass spectrum, the average MW of the α2M homodimer and homotetramer is estimated to be 358 and 717 kDa, respectively (Figures 2.8 and 2.9).
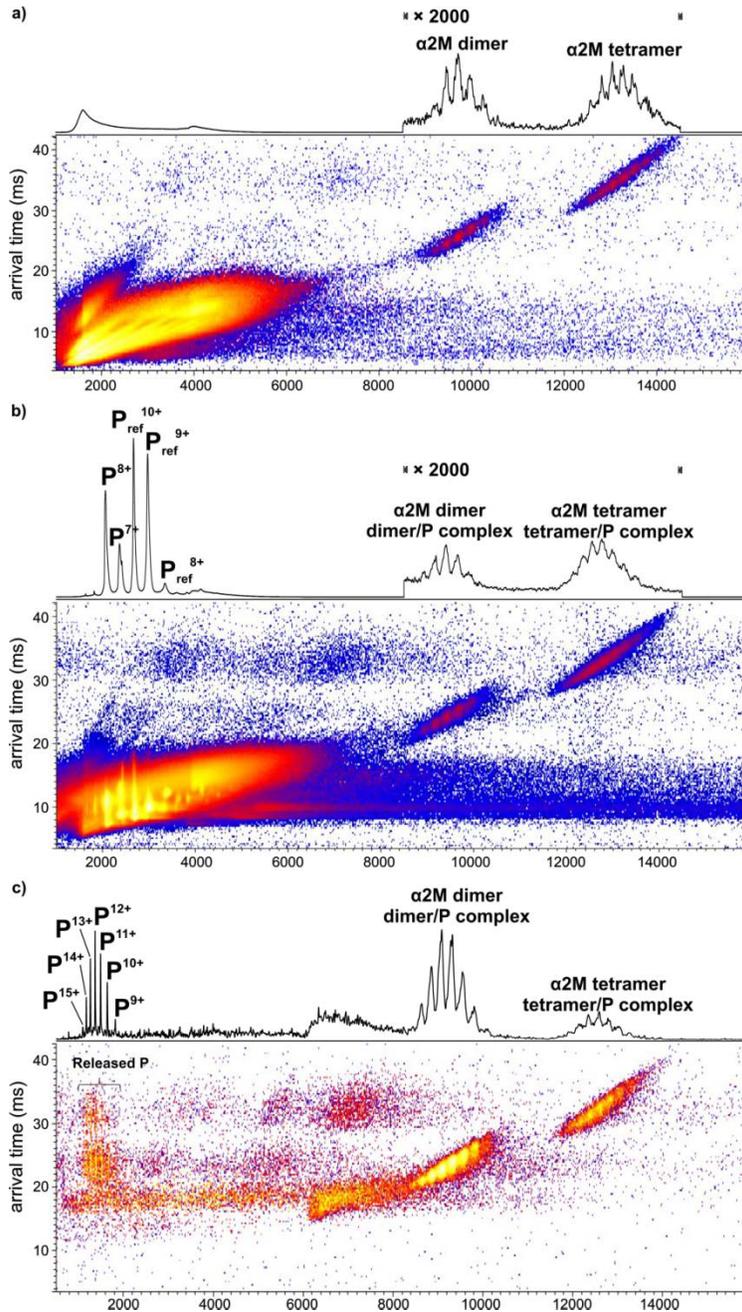


**Figure 2.7.** ESI mass spectrum and corresponding IMS heat map (arrival time versus *m/z*; white-blue-red-yellow: 0~33%~66%~100% counts) acquired in positive ion mode for a

200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of α2M (1.7 mg mL$^{-1}$) (a) alone and (b) with hGal-3C (P, 2.5 µM) and scFv (P$_{ref}$, 2.5 µM). (c) CID mass spectrum and corresponding IMS heat map of ions with *m/z* ≥6400 from solution described in (b), with a Transfer collision energy of 150 V.
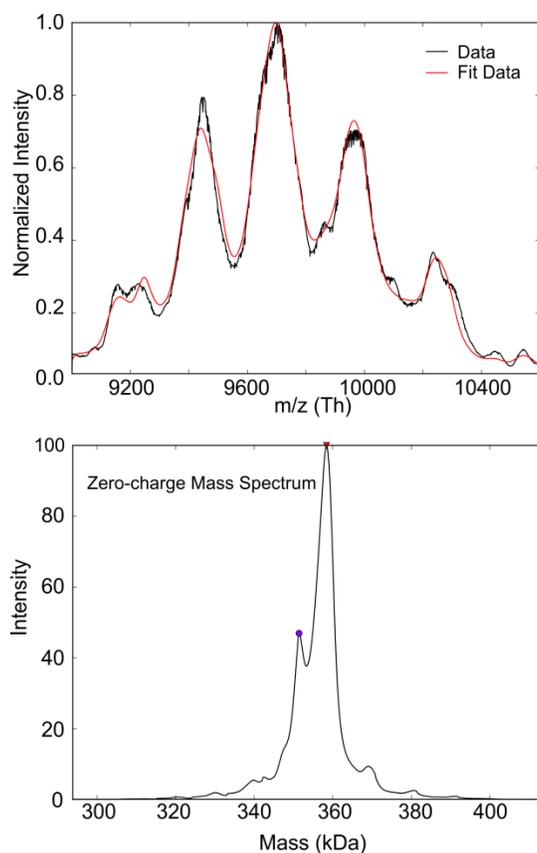


**Figure 2.8.** Analysis of the mass spectrum of the α2M dimer acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of α2M (1.7 mg mL$^{-1}$) at cone voltage 30 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 358.4 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 9000 to 10600; Subtract Minimum - 0.0; Gaussian Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 34 to 39; Mass range 300,000 to 400,000; Sample Mass Every 200.0 Da; Peak FWHM (Th) 60.0;

Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.99.
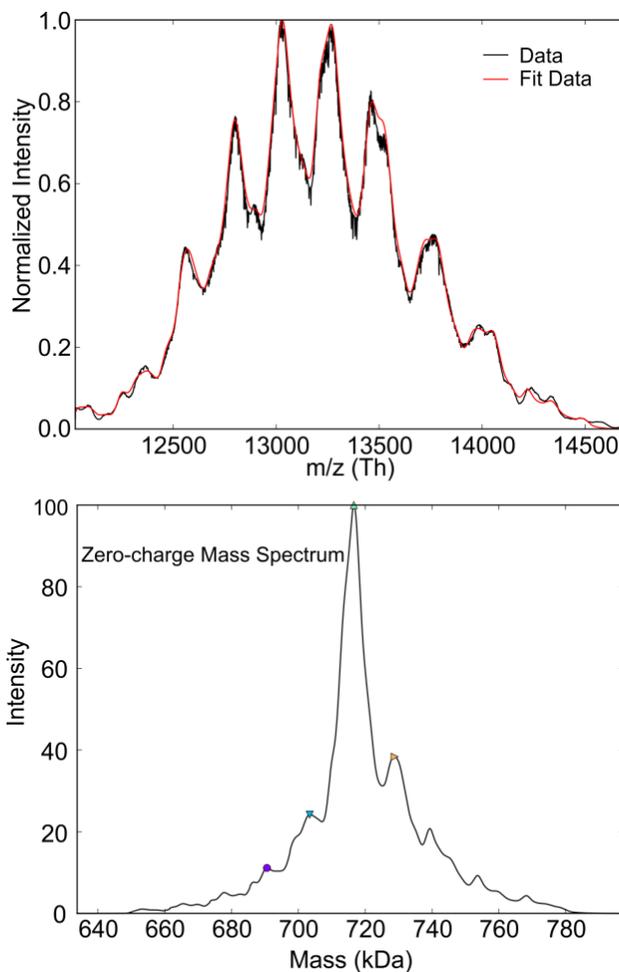


**Figure 2.9.** Analysis of the mass spectrum of the α2M tetramer acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C) of α2M (1.7 mg mL$^{-1}$) at cone voltage 30 V. Fit Data mass spectrum (top panel) and zero charge mass spectrum (bottom panel) correspond to an average MW of 716.6 kDa. Spectral deconvolution was performed with the UniDec deconvolution algorithm using the following parameters: *m/z* range – 10,000 to 15,000; Subtract Minimum - 0.0; Gaussian Smoothing - 0; Linear *m/z* (Constant delta *m/z*); Bin every 1.0; Charge Range - 50 to 59; Mass range 650,000 to 780,000; Sample Mass Every 200.0 Da; Peak FWHM (Th) 50.0;

Peak Shape Function – Gaussian; Spectral raw data were loaded as text files containing intensity and *m/z* values. Fitting R square value is 0.996.

Due to the high MW and heterogeneity of the α2M complexes, it was not possible to identify hGal-3C binding, either from the mass spectrum alone or in conjunction with IMS analysis (Figure 2.7b). However, binding of hGal-3C to both the α2M dimer and tetramer was established by CaR$^{IMS}$-ESI-MS performed on ions with *m/z* >8000. At Transfer voltages >120 V protonated hGal-3C ions, at charge states +9 to +15 and IMS-ATs of 20.0 ms to 26.2 ms, were released from the α2M homodimer. At Transfer voltages >150 V hGal-3C ions, at charge states +10 to +15 and ATs of 28.0 ms to 36.0 ms, were released from the α2M homotetramer (Figure 2.7c). No P$_{ref}$ ions were detected under these conditions. Also, similar to the results obtained for Hp1-1, there was no evidence of higher-order hGal-3C complexes released from the α2M dimers or tetramers. Taken together, the aforementioned results demonstrate that CaR$^{IMS}$-ESI-MS can serve as a general method for establishing binding of soluble GBPs to soluble GPs. However, the assay is qualitative and doesn't provide any insight into the affinities of the interactions. To provide further support for the existence of specific hGal-3C interactions with the serum GPs in solution and to quantify their apparent affinities the *proxy ligand* ESI-MS assay was applied.

**2.3.2 Quantifying hGal-3C interactions with serum GP by *proxy ligand* ESI-MS**

To quantify hGal-3C-GP binding with the *proxy ligand* ESI-MS assay, the GP was titrated into a solution containing hGal-3C and $L_{proxy}$ (lacto-N-neooctaose) and the ratio of the abundances of free and $L_{proxy}$-bound hGal-3C ions ($R_{proxy}$ = $Ab$(hGal-3C+$L_{proxy}$)/$Ab$(hGal-3C)) was monitored by ESI-MS. Shown in Figure 2.10 are representative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, pH 6.8, 25 °C) of hGal-3C (5.3 µM), $L_{proxy}$ (3.3 µM) and $P_{ref}$ (2.5 µM) in the absence (Figure 2.10a) and presence of 1.4 mg mL$^{-1}$ AGP (Figure 2.10b). Ion signal corresponding to both free and $L_{proxy}$-bound hGal-3C was detected in both mass spectra. However, the relative abundance of $L_{proxy}$-bound hGal-3C increased upon introduction of AGP to the solution. This increase is consistent with AGP binding to hGal-3C, which effectively reduces the concentration of hGal-3C available for binding to $L_{proxy}$. Shown in Figure 2.13a is a plot of $R_{proxy}$ versus the molar concentration (calculated based on the average MW, *vide supra*) of AGP. Also shown is the best fit of the 1:1 GBP-GP binding model (eq 1), which corresponds to a $K_{a,app}$ of $(4.0 \pm 0.5) \times 10^5$ M$^{-1}$, to the experimental data. Analogous measurements were performed for hGal-3C binding to Hp1-1 and to α2M (Figures 2.11 and 2.12, respectively) and the corresponding plots of $R_{proxy}$ versus GP concentration are shown in Figures 2.13b and 2.13c, respectively. Similar to what was observed for AGP, the titration data obtained with Hp1-1 and α2M can be well described by the 1:1 binding model, with a $K_{a,app}$ of $(1.9 \pm 0.1) \times 10^5$ M$^{-1}$ and $(2.6 \pm 0.1) \times 10^5$ M$^{-1}$, respectively.

To our knowledge, absolute affinities of hGal-3C for AGP, Hp1-1 and α2M have not been previously reported. However, the $K_{a,app}$ values determined here are consistent with results obtained by affinity chromatography carried out using immobilized hGal-3, which suggested the $K_{a,app}$ of this GBP for serum GPs (unpurified) is on the order of $10^6$ M$^{-1}$.[5] Moreover, in a separate affinity chromatography study, the affinities of a number of mammalian mono-sialylated complex type *N*-glycans for hGal-3 and hGal-3C were

estimated to be $\sim 10^6$ M$^{-1}$.[50]
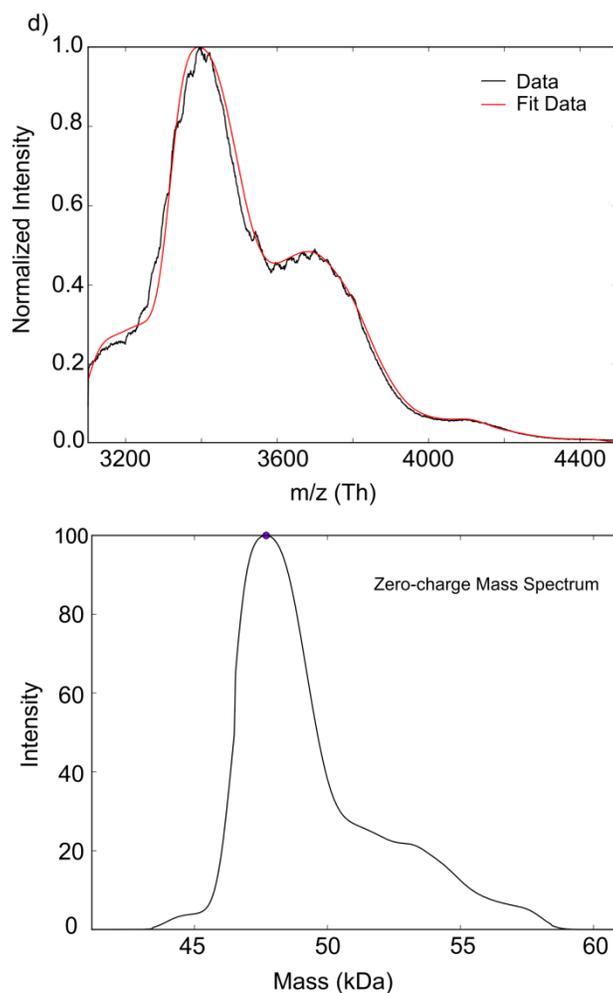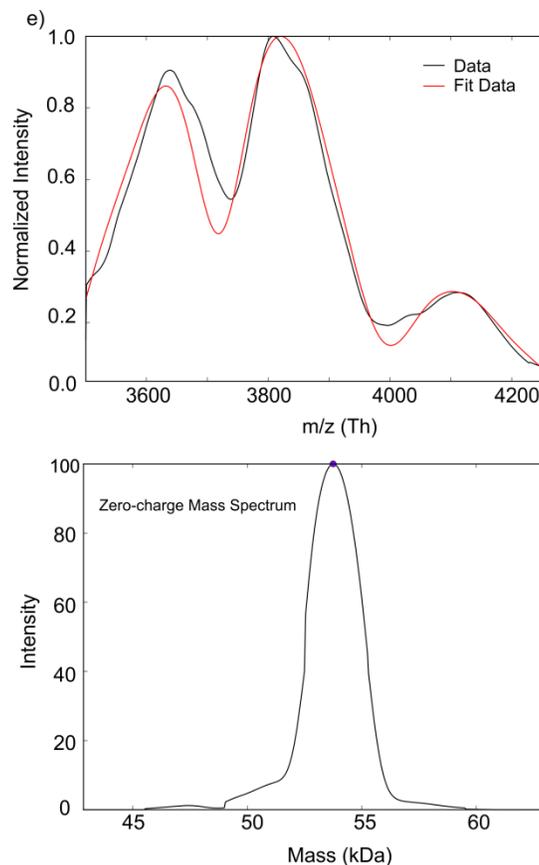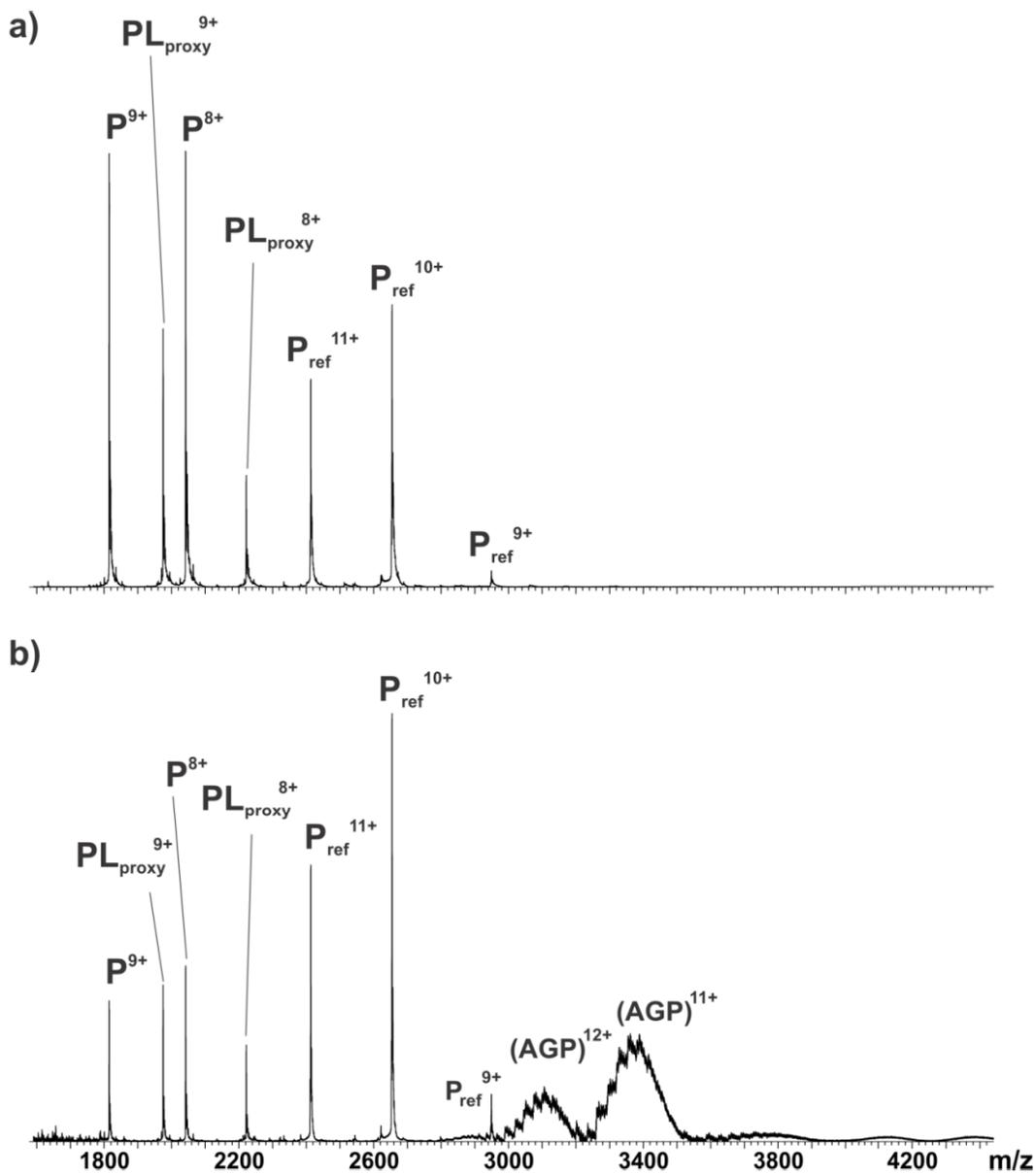


**Figure 2.10.** ESI-MS mass spectra acquired in positive ion mode for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) containing hGal-3C (P, 5.3 µM), L$_{proxy}$ (3.4 µM), scFv (P$_{ref}$, 2.5 µM) in the (a) absence and (b) presence of 1.4 mg mL$^{-1}$ AGP.
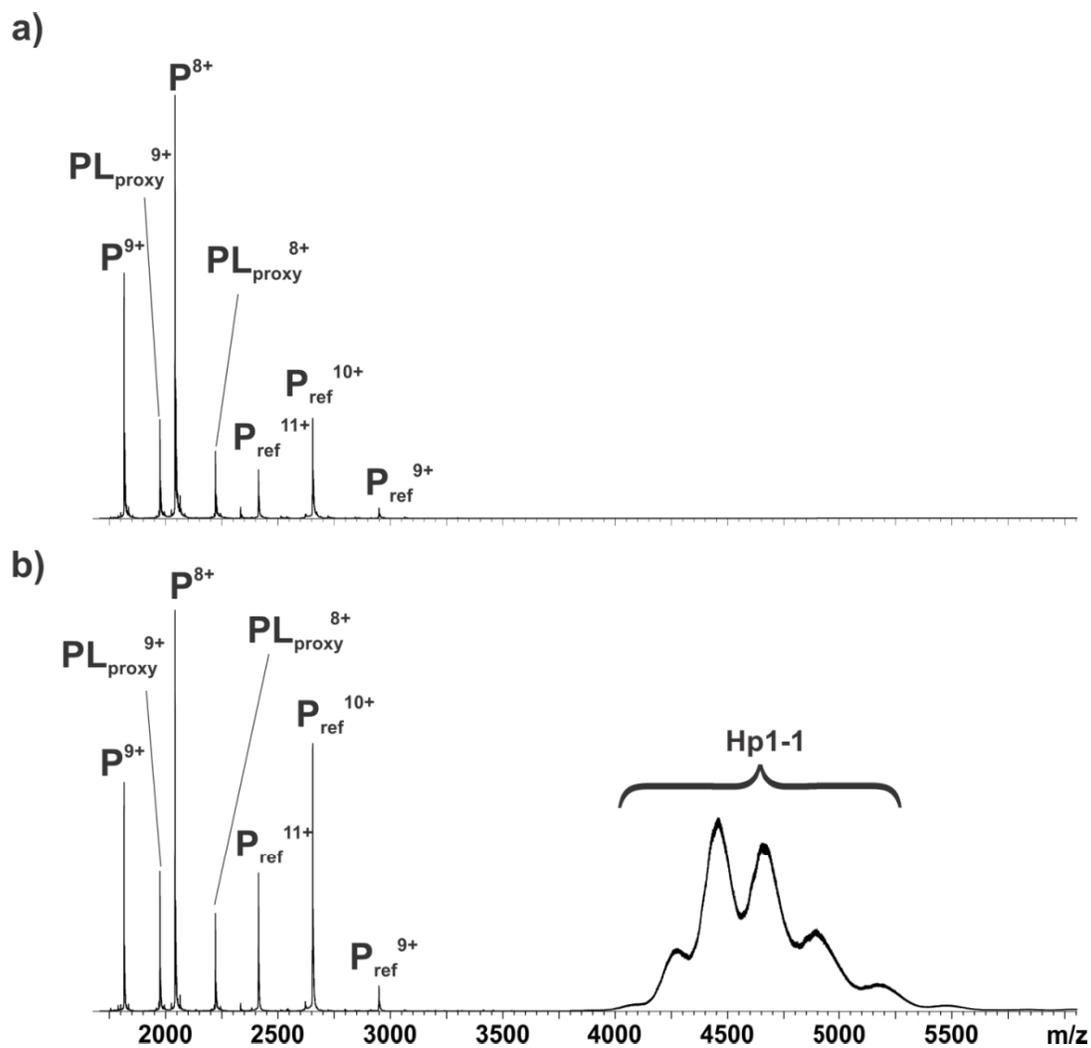
**Figure 2.11.** ESI-MS mass spectra acquired in positive ion mode for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) containing hGal-3C (P, 5.8 µM), $L_{proxy}$ (2.1 µM), scFv ($P_{ref}$, 2.5 µM) in the (a) absence and (b) presence of 2 mg mL$^{-1}$ of Hp1-1.
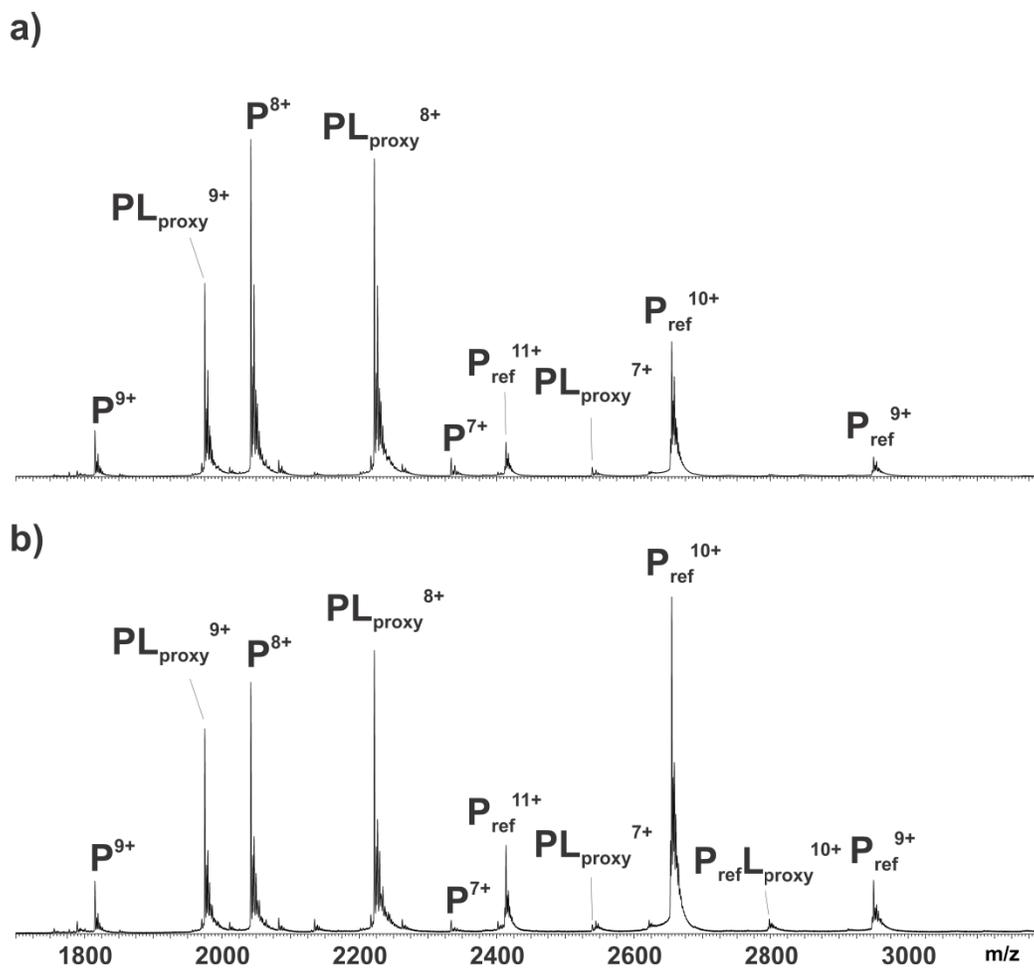
**Figure 2.12.** ESI-MS mass spectra acquired in positive ion mode for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) containing hGal-3C (P, 8.8 µM), $L_{proxy}$ (10.0 µM), scFv ($P_{ref}$, 3.0 µM) in the (a) absence and (b) presence of 6.5 mg mL$^{-1}$ of α2M.
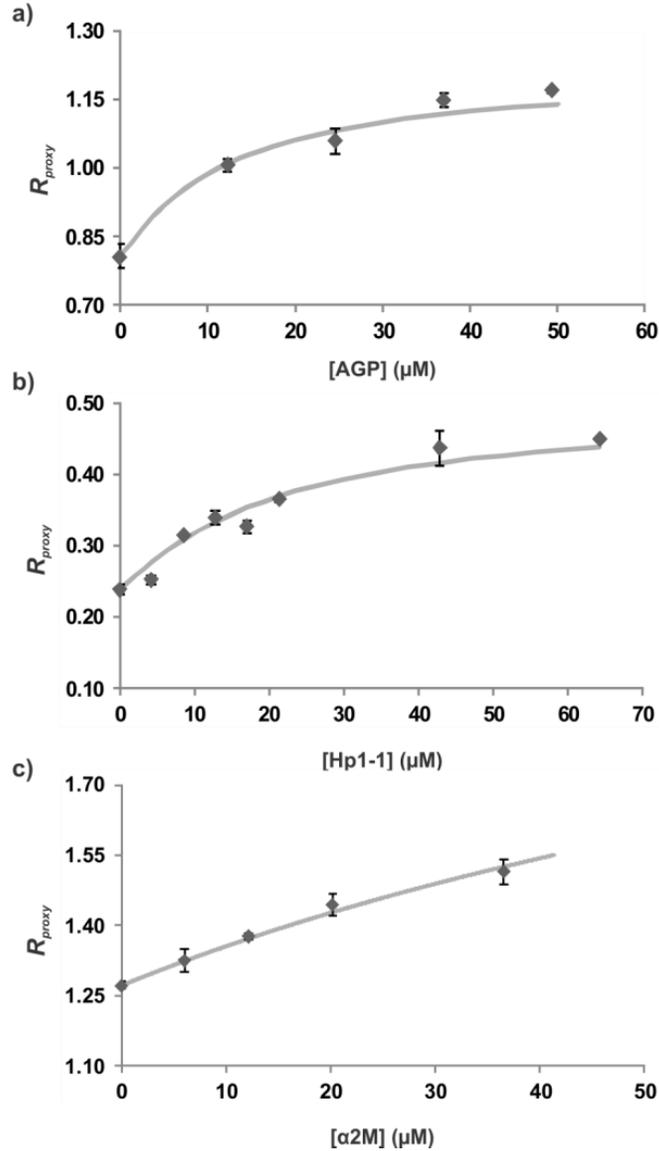
**Figure 2.13**. (a) Plot of $R_{proxy}$ [$\equiv Ab$(hGal-3C + $L_{proxy}$)/$Ab$(hGal-3C)] versus AGP concentration measured for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) of hGal-3C (5.3 µM), $L_{proxy}$ (3.4 µM), scFv ($P_{ref}$, 2.5 µM) and AGP at concentrations ranging from 0 mg mL$^{-1}$ to 1.84 mg mL$^{-1}$ (49.3 µM). (b) Plot of $R_{proxy}$ versus Hp1-1 concentration measured for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) of hGal-3C (5.8 µM), $L_{proxy}$ (2.1 µM), scFv ($P_{ref}$, 2.5 µM) and Hp1-1 at concentrations ranging from 0 mg mL$^{-1}$ to 6.00 mg mL$^{-1}$ (64.2 µM). (c) Plot of $R_{proxy}$ versus α2M concentration measured for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) of hGal-3C (8.8 µM), $L_{proxy}$ (10.0 µM), scFv ($P_{ref}$, 3.0 µM) and α2M

at concentrations ranging from 0 mg mL$^{-1}$ to 6.53 mg mL$^{-1}$ (36.5 μM). Solid lines correspond to best fit of a 1:1 binding model to the experimental data.

**2.3.3 Screening serum GP *N*-glycans against hGal-3C using ESI-MS**

The results obtained using the CaR$^{IMS}$-ESI-MS and *proxy ligand* ESI-MS assays confirm binding of hGal-3C to AGP, Hp1-1 and α2M under the solution conditions investigated. To gain insight into the nature of the glycan structures being recognized, the *N*-glycans were released (as free oligosaccharides) from each of the GPs, and screened, as a library, against hGal-3C by ESI-MS. Shown in Figure 2.14 are representative ESI mass spectra, acquired in negative ion mode, for aqueous solutions of each of the *N*-glycan libraries. By comparing the measured MWs of the ions detected in the library and theoretical MWs calculated using Glyco Mass Calculator (https://www.nist.gov/static/glyco-mass-calc/), the monosaccharide compositions of the released *N*-glycans (NG$_i$, where $i$ = 1 - 21) of each GP were established (Table 2.1). Overall, the *N*-glycans released from the three GPs are consistent with results reported previously.[38,46,51,52] Glycans corresponding to fifteen different MWs (and compositions) and consisting of bi-, tri- and tetra-antennary structures with a range of sialylation and fucosylation, were produced from AGP. According to the ESI mass spectrum, fully sialylated bi-, tri- and tetra-antennary glycans without fucose were the most abundant, followed by sialylated tri- and tetra-antennary glycan with fucose. Tetra-antennary glycans elongated with one or two units of LacNAc were also detected. For Hp1-1, *N*-glycans corresponding to twelve distinct compositions were identified, with fully sialylated di- and tri-antennary structures being the most abundant and tetra-antennary structures being a minor component. Fucosylated bi-, tri- and tetra-antennary structures were also detected. Nine different *N*-glycan compositions were identified for α2M. Mono- and di-sialylated bi-antennary structures with and without fucose dominated; tri-antennary structures with varying degrees of sialylation were also detected.
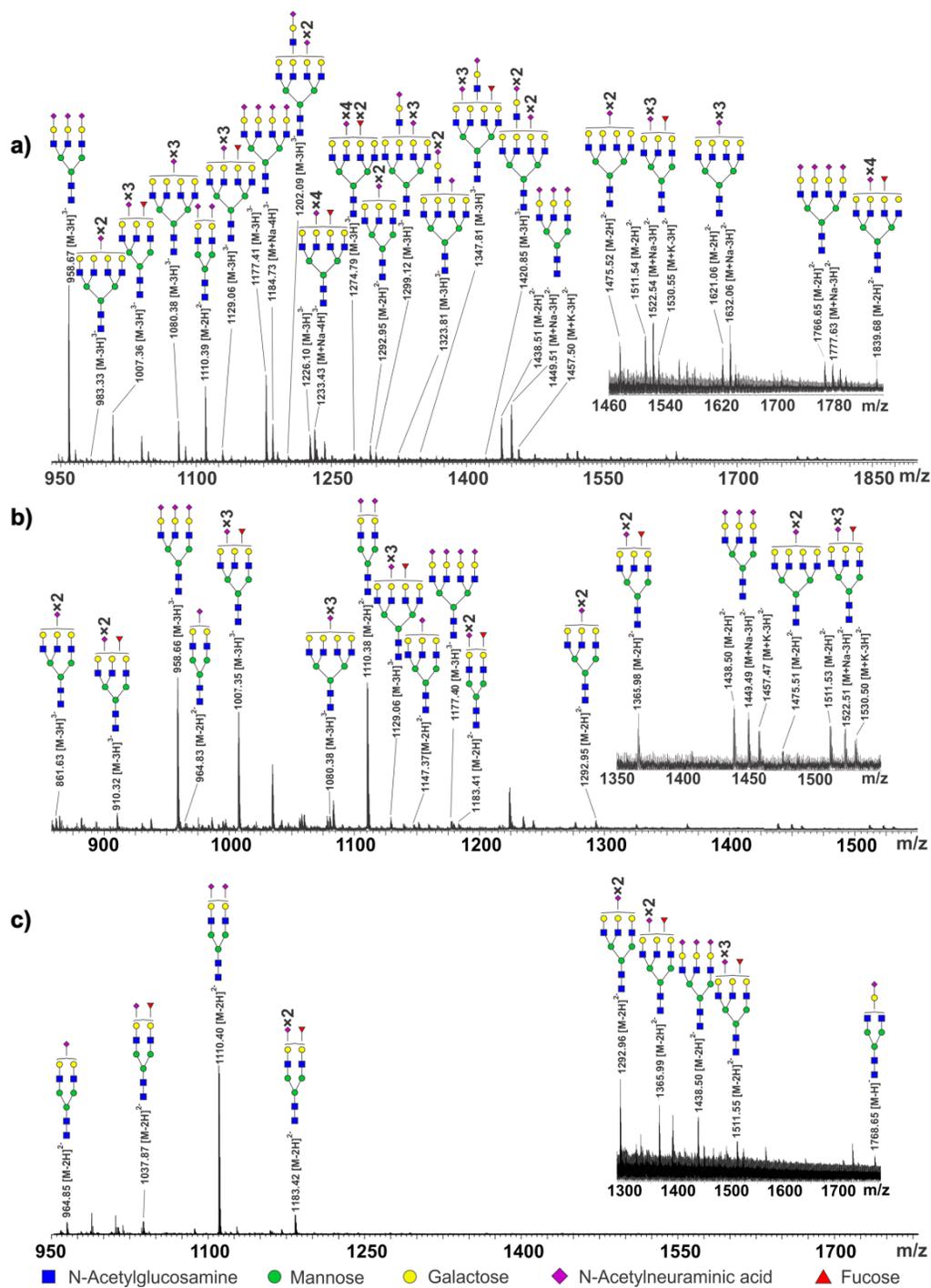
**Figure 2.14.** ESI mass spectra acquired in negative ion mode for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) containing (a) AGP *N*-glycan library, 0.375 µg µL⁻¹ (b) Hp1-1 N-glycan library, 0.145 µg µL⁻¹ and (c) α2M *N*-glycan library, 0.24 µg µL⁻¹.

Shown in Figure 2.15 are representative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, pH 6.8) of hGal-3C, at two different concentrations (12 μM and 40 μM), and each GP $N$-glycan library. A $P_{ref}$ was added to each solution to monitor the occurrence of nonspecific glycan-hGal-3C binding during the ESI process.[34] Results obtained for the AGP library revealed hGal-3C binding to $N$-glycans corresponding to fourteen different MWs (Figures 2.15a and 2.15d). Importantly, only free $P_{ref}$ ions were detected in the mass spectrum, indicating that the occurrence of nonspecific glycan binding to hGal-3C was negligible under the conditions used.[34] Screening of the Hp1-1 library against hGal-3C identified $N$-glycan ligands corresponding to eleven different MWs (Figures 2.15b and 2.15e); eight different $N$-glycan MWs from α2M were found to bind (Figures 15c and 15f). A summary of the $N$-glycan ligands identified from screening is given in Table 2.2. It is also interesting to note that all of the $N$-glycan ligands are sialylated and, in the case of AGP, are composed of both α2,3 and α2,6 sialylation of the terminal Gal residues.[52] The present results, therefore, suggest that α2,6 sialylation does not block binding, as previously suggested.[53] This conclusion is also supported by the recent discovery that sialylated human milk oligosaccharides are recognized by hGal-3C.[10] Interestingly, no binding was measured for $NG_{11}$ from AGP and $NG_6$ from Hp1-1. These observation is likely due to the detrimental effect of α1,3 antenna fucosylation.[54] The absence of $NG_1$ binding can likely be attributed to its relatively low abundance in the α2M library.
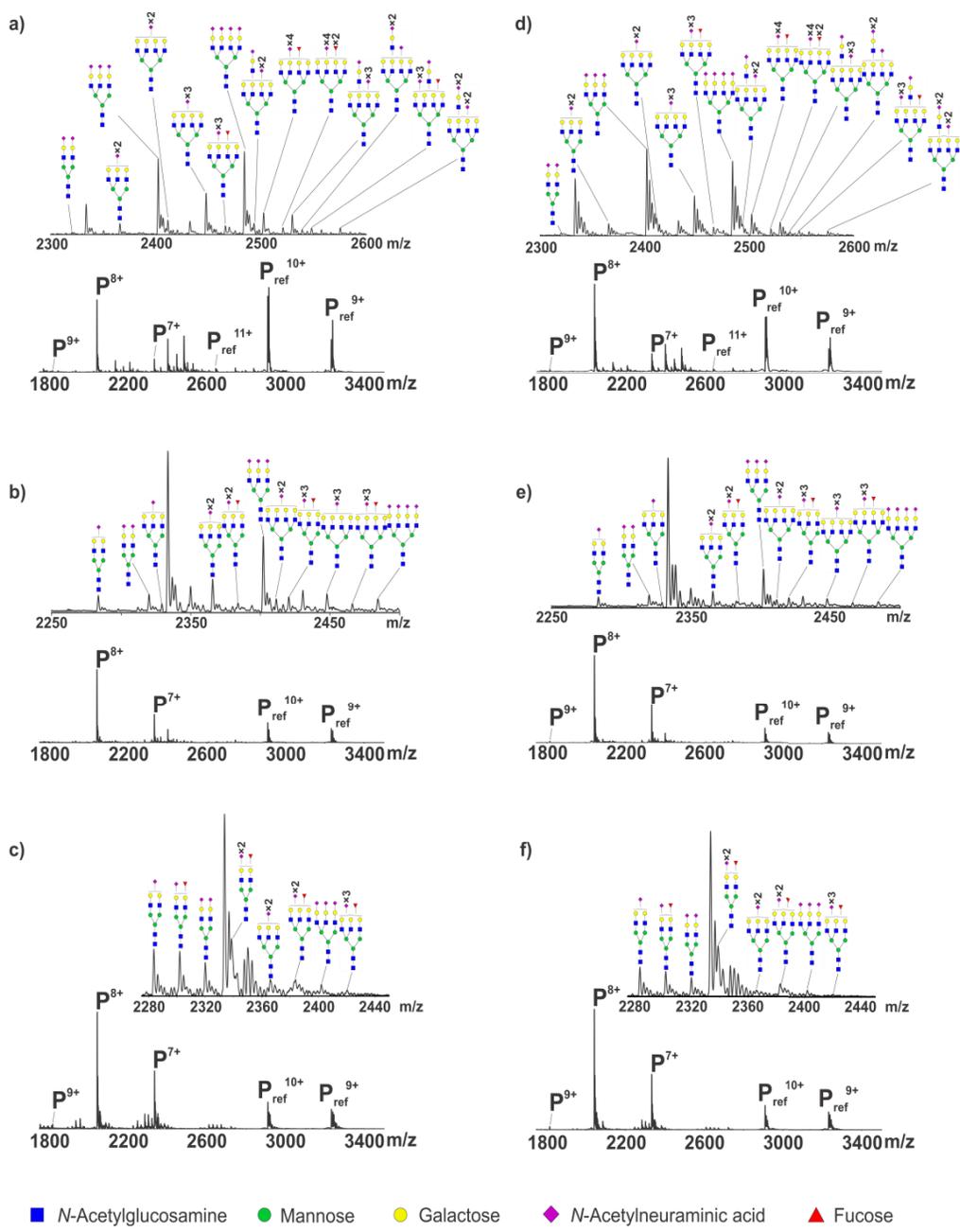
**Figure 2.15**. ESI mass spectra acquired in positive ion mode for 200 mM aqueous ammonium acetate solutions (pH 6.8 and 25 °C) of hGal-3C (≡P), (a), (b), (c) 12 μM, (d), (e), (f) 40 μM, 0.375 μg μL$^{-1}$of AGP $N$-glycan library ((a), (d)), 0.145 μg μL$^{-1}$ of Hp1-1 $N$-glycan library ((b), (e)), and 0.24 μg μL$^{-1}$ of α2M $N$-glycan library ((c), (f)). Complexes of hGal-3C and $N$-glycan ligands at +8 are shown as insets. BCA (P$_{ref}$) (a), (b)

and (c) 6 µM, (d), (e), (f) 20 µM, was added to the solutions to correct for nonspecific oligosaccharide binding.

**Table 2.1.** Relative abundances of *N*-glycans released from the serum glycoproteins AGP, Hp1-1 and α2M, as determined directly by ESI-MS.[a]

| $NG_i$ | Composition Hex_HexNAc_ Fuc_Sia | MW Theoretical | m/z (-3 charge state) | m/z (-2 charge state) | Relative abundance (%) | | |
|---|---|---|---|---|---|---|---|
| | | | | | AGP | Hp1-1 | α2M |
| $NG_1$ | 4_4_0_1 | 1769.63 | 588.88 | 883.82 | ND | ND | 0.2 |
| $NG_2$ | 5_4_0_1 | 1931.69 | 642.90 | 964.84 | ND | 0.2 | 7.4 |
| $NG_3$ | 5_4_1_1 | 2077.75 | 691.58 | 1037.87 | ND | ND | 6.5 |
| $NG_4$ | 5_4_0_2 | 2222.78 | 739.93 | 1110.39 | 13.8 | 50.9 | 67.6 |
| $NG_5$ | 6_5_0_1 | 2296.82 | 764.61 | 1147.41 | ND | 1.1 | ND |
| $NG_6$ | 5_4_1_2 | 2368.84 | 788.61 | 1183.42 | ND | 1.8 | 12.5 |
| $NG_7$ | 6_5_0_2 | 2587.92 | 861.64 | 1292.96 | 2.0 | 5.6 | 2.5 |
| $NG_8$ | 6_5_1_2 | 2733.97 | 910.32 | 1365.99 | ND | 3.4 | 1.5 |
| $NG_9$ | 6_5_0_3 | 2879.01 | 958.67 | 1438.51 | 39.4 | 16.8 | 1.2 |
| $NG_{10}$ | 7_6_0_2 | 2953.05 | 983.35 | 1475.52 | 1.8 | 1.1 | ND |
| $NG_{11}$ | 6_5_1_3 | 3025.07 | 1007.36 | 1511.53 | 10.7 | 14.1 | 0.6 |
| $NG_{12}$ | 7_6_0_3 | 3244.14 | 1080.38 | 1621.07 | 6.3 | 1.5 | ND |
| $NG_{13}$ | 7_6_1_3 | 3390.20 | 1129.07 | 1694.10 | 1.1 | 2.2 | ND |
| $NG_{14}$ | 7_6_0_4 | 3535.24 | 1177.41 | 1766.62 | 14.2 | 1.3 | ND |
| $NG_{15}$ | 8_7_0_3 | 3609.28 | 1202.09 | 1803.64 | 0.7 | ND | ND |
| $NG_{16}$ | 7_6_1_4 | 3681.30 | 1226.10 | 1839.65 | 5.5 | ND | ND |
| $NG_{17}$ | 7_6_2_4 | 3827.35 | 1274.78 | 1912.68 | 0.9 | ND | ND |
| $NG_{18}$ | 8_7_0_4 | 3900.37 | 1299.12 | 1949.19 | 1.8 | ND | ND |
| $NG_{19}$ | 9_8_0_3 | 3974.41 | 1323.80 | 1986.20 | 0.6 | ND | ND |
| $NG_{20}$ | 8_7_1_4 | 4046.43 | 1347.81 | 2022.21 | 0.6 | ND | ND |
| $NG_{21}$ | 9_8_0_4 | 4265.50 | 1420.83 | 2131.75 | 0.5 | ND | ND |

a. The monosaccharide compositions of *N*-glycans correspond to the number of the hexose (Hex), *N*-acetylhexosamine (HexNAc), fucose (Fuc) and sialic acid (Sia) residues. ND ≡ Not detected

**Table 2.2**. List of MWs, compositions and structures of *N*-glycans identified as ligands of hGal-3C by ESI-MS screening of *N*-glycan libraries produced from AGP, Hp1-1 or α2M.[a]

| NG$_i$ | MW Theoretical | Composition [b] Hex_HexNAc_Fuc_Sia | Structure [c] |
|---|---|---|---|
| NG$_2$ | 1931.69 | 5_4_0_1 |  |
| NG$_3$ | 2077.75 | 5_4_1_1 |  |
| NG$_4$ | 2222.78 | 5_4_0_2 |  |
| NG$_5$ | 2296.82 | 6_5_0_1 |  |
| NG$_6$ | 2368.84 | 5_4_1_2 |  |
| NG$_7$ | 2587.92 | 6_5_0_2 |  |
| NG$_8$ | 2733.97 | 6_5_1_2 |  |
| NG$_9$ | 2879.01 | 6_5_0_3 |  |
| NG$_{10}$ | 2953.05 | 7_6_0_2 |  |

| | | | |
|---|---|---|---|
| NG$_{11}$ | **3025.07** | **6_5_1_3** |  |
| NG$_{12}$ | **3244.14** | **7_6_0_3** |  |
| NG$_{13}$ | **3390.20** | **7_6_1_3** |  |
| NG$_{14}$ | **3535.24** | **7_6_0_4** |  |
| NG$_{15}$ | **3609.28** | **8_7_0_3** |  |
| NG$_{16}$ | **3681.30** | **7_6_1_4** |  |
| NG$_{17}$ | **3827.35** | **7_6_2_4** |  |
| NG$_{18}$ | **3900.37** | **8_7_0_4** |  |
| NG$_{19}$ | **3974.41** | **9_8_0_3** |  |
| NG$_{20}$ | **4046.43** | **8_7_1_4** |  |

| | | | |
|---|---|---|---|
| NG$_{21}$ | **4265.50** | **9_8_0_4** |  |

a. All measurements performed in aqueous ammonium acetate solutions (200 mM, pH 6.8).

b. The monosaccharide composition of *N*-glycans is reported as the number of the hexose (**Hex**), *N*-Acetylhexosamine (**HexNAc**), Fucose (**Fuc**) and Sialic acid (**Sia**) residues.

c. Monosaccharides notation: ■, *N*-Acetylgluosamine; ●, Mannose; ●, Galactose; ♦, *N*-Acetylneuraminic acid; ▲, Fucose.

Because the concentrations of the *N*-glycans present in the libraries are not known, it is not possible to extract affinities from the abundance ratios (i.e., *R*) of bound-to-free hGal-3C ions measured in the mass spectrum. However, as described elsewhere, *R* is relatively insensitive to protein concentration for low affinity protein-ligand interactions ($\sim 10^3$ M$^{-1}$).[30] Therefore, the magnitude of the change in *R* (i.e., $\Delta R$) resulting from a change in initial protein concentration can serve to distinguish low affinity ligands from moderate-to-high affinity ($>10^4$ M$^{-1}$) ligands present at unknown concentrations. However, for moderate-to-high affinity ligands present at low concentrations, the magnitude of $\Delta R$ will be limited by the initial *R* value. Consequently, it is more reliable to consider the percentage change in *R* (i.e., $\Delta R\%$) resulting from the change in protein concentration in order to differentiate low affinity ligands from moderate-to-high affinity ones. Based on the theoretical modelling of the $\Delta R\%$ expected for ligand concentrations ranging from 2 μM to 20 μM and affinities from $10 \times 10^3$ M$^{-1}$ to $2 \times 10^5$ M$^{-1}$, it was concluded that ligands with affinities $>5 \times 10^3$ M$^{-1}$ will exhibit a $\Delta R\% > 10\%$ upon increasing the protein concentration from 12 μM to 40 μM (Figure 2.16).
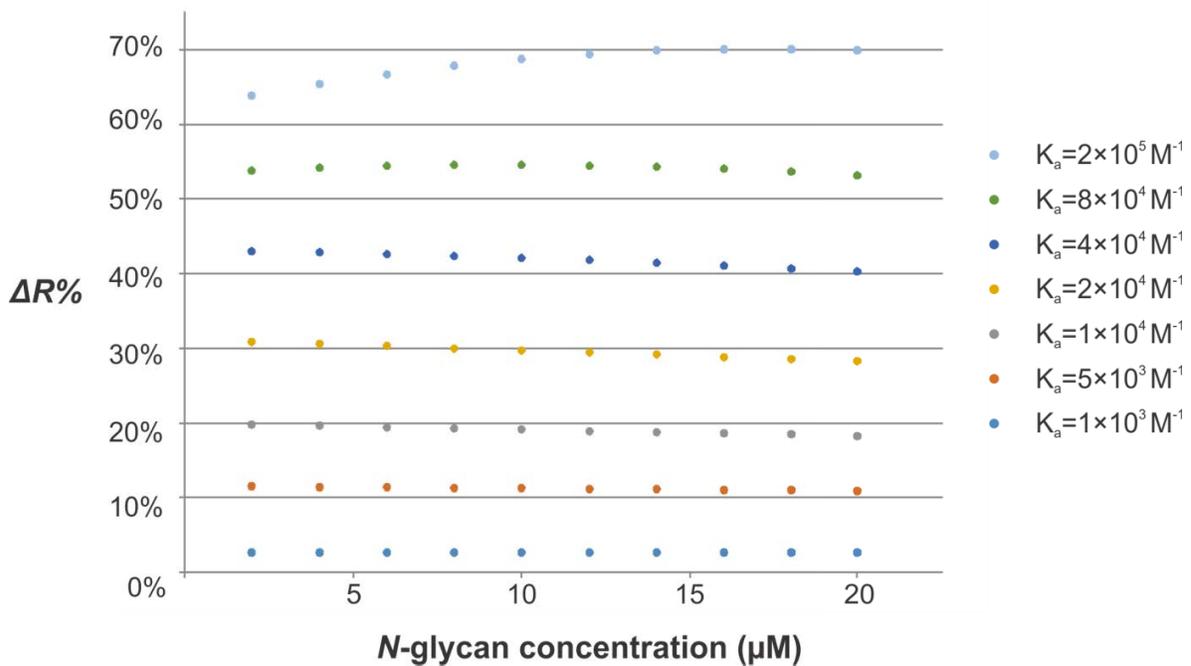
**Figure 2.16.** Plot of $\Delta R\%$ versus $N$-glycan concentrations calculated for $K_a$ of $1 \times 10^3$ M$^{-1}$, $5 \times 10^3$ M$^{-1}$, $1 \times 10^4$ M$^{-1}$, $2 \times 10^4$ M$^{-1}$, $4 \times 10^4$ M$^{-1}$, $8 \times 10^4$ M$^{-1}$ and $2 \times 10^5$ M$^{-1}$ using two different concentrations of hGal-3C (12 µM and 40 µM). $\Delta R\%$ represents percentage change in the ion abundance ratio of $N$-glycan ligand bound-to-free hGal-3C.
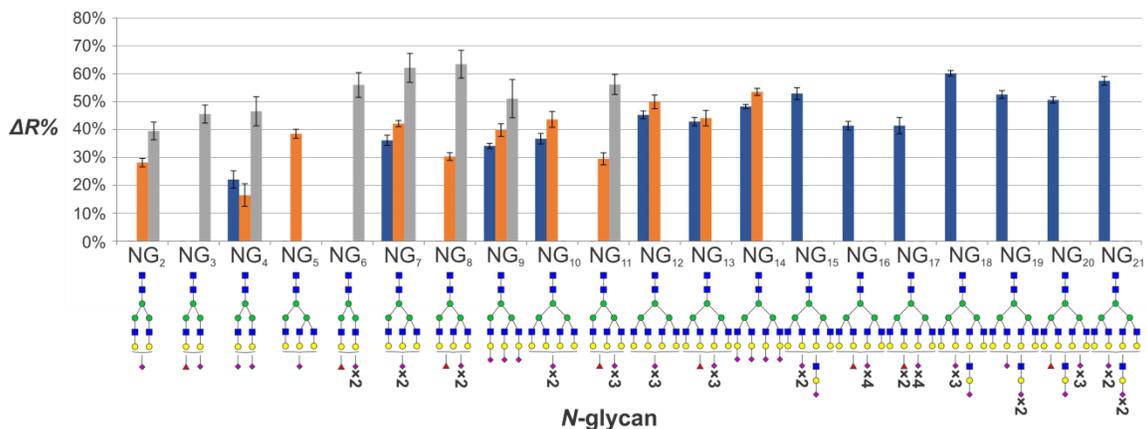


**Figure 2.17.** Plot of $\Delta R\%$, percentage change in the ion abundance ratio of $N$-glycan ligand bound-to-free hGal-3C, measured by ESI-MS for $N$-glycan ligands (grouped as compositional isomers) produced from AGP (blue), Hp1-1 (orange) and α2M (grey),

using two different concentrations of hGal-3C (12 µM and 40 µM) in 200 mM aqueous ammonium acetate solution (pH 6.8 and 25 °C). Errors correspond to one standard deviation of three replicate measurements.

**Table 2.3**. Percentage change ($\Delta R\%$) in the abundance ratio of *N*-glycan ligand bound-to-free hGal-3C measured by ESI-MS for *N*-glycans ($NG_i$), grouped as compositional isomers, produced from AGP, Hp1-1 and α2M at two different concentrations of hGal-3C (12 µM and 40 µM).

| $NG_i$ | $R_{initial}$ (AGP) | $R_{final}$ (AGP) | $\Delta R\%$ (AGP) | $R_{initial}$ (Hp1-1) | $R_{final}$ (Hp1-1) | $\Delta R\%$ (Hp1-1) | $R_{initial}$ (α2M) | $R_{final}$ (α2M) | $\Delta R\%$ (α2M) |
|---|---|---|---|---|---|---|---|---|---|
| $NG_1$ | - | - | - | - | - | - | NB | NB | NB |
| $NG_2$ | - | - | - | 0.037± 0.001 | 0.026± 0.001 | 28.2± 1.5 | 0.127± 0.001 | 0.077± 0.004 | 39.5± 3.2 |
| $NG_3$ | - | - | - | - | - | - | 0.119± 0.001 | 0.065± 0.003 | 45.6± 3.3 |
| $NG_4$ | 0.015± 0.001 | 0.011± 0.001 | 22.1± 3.1 | 0.039± 0.001 | 0.032± 0.001 | 16.5± 4.0 | 0.092± 0.001 | 0.049± 0.04 | 45.5± 5.3 |
| $NG_5$ | - | - | - | 0.019± 0.001 | 0.012± 0.001 | 38.5± 1.6 | - | - | - |
| $NG_6$ | - | - | - | NB | NB | NB | 0.108± 0.002 | 0.048± 0.003 | 56.0± 4.4 |
| $NG_7$ | 0.088± 0.001 | 0.056± 0.001 | 36.1± 1.8 | 0.091± 0.001 | 0.053± 0.001 | 42.1± 1.1 | 0.039± 0.001 | 0.015± 0.001 | 62.1± 5.2 |
| $NG_8$ | - | - | - | 0.018± 0.001 | 0.013± 0.001 | 30.4± 1.4 | 0.043± 0.001 | 0.016± 0.001 | 63.4± 5.0 |
| $NG_9$ | 0.611± 0.002 | 0.403± 0.004 | 34.1± 0.9 | 0.207± 0.003 | 0.124± 0.001 | 39.8± 2.3 | 0.032± 0.001 | 0.015± 0.001 | 51.1± 6.8 |
| $NG_{10}$ | 0.111± 0.001 | 0.071± 0.002 | 36.7± 1.9 | 0.031± 0.001 | 0.017± 0.001 | 43.6± 2.8 | - | - | - |
| $NG_{11}$ | NB | NB | NB | 0.035± 0.000 | 0.024± 0.001 | 29.5± 2.1 | 0.016± 0.001 | 0.007± 0.001 | 56.1± 3.6 |
| $NG_{12}$ | 0.325± 0.002 | 0.178± 0.003 | 45.2± 1.4 | 0.046± 0.001 | 0.023± 0.001 | 49.9± 2.5 | - | - | - |
| $NG_{13}$ | 0.068± 0.001 | 0.039± 0.001 | 42.9± 1.5 | 0.018± 0.001 | 0.010± 0.001 | 44.1± 2.8 | - | - | - |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| NG$_{14}$ | 0.629± 0.001 | 0.326± 0.003 | 48.3± 0.8 | 0.035± 0.001 | 0.016± 0.001 | 53.5± 1.3 | - | - | - |
| NG$_{15}$ | 0.083± 0.001 | 0.039± 0.01 | 52.9± 2.1 | - | - | - | - | - | - |
| NG$_{16}$ | 0.131± 0.001 | 0.077± 0.001 | 41.4± 1.5 | - | - | - | - | - | - |
| NG$_{17}$ | 0.063± 0.001 | 0.037± 0.001 | 41.4± 2.9 | - | - | - | - | - | - |
| NG$_{18}$ | 0.162± 0.001 | 0.065± 0.001 | 60.2± 1.0 | - | - | - | - | - | - |
| NG$_{19}$ | 0.049± 0.001 | 0.023± 0.001 | 52.6± 1.4 | - | - | - | - | - | - |
| NG$_{20}$ | 0.056± 0.001 | 0.027± 0.001 | 50.6± 1.1 | - | - | - | - | - | - |
| NG$_{21}$ | 0.058± 0.001 | 0.025± 0.001 | 57.5± 1.6 | - | - | - | - | - | - |

a. Errors correspond to one standard deviation. NB ≡ No binding detected. '-' ≡ not present in the library.

Plotted in Figure 2.17 and summarized in Table 2.3 are the $\Delta R\%$ values determined for each NG$_i$ from the mass spectra measured at hGal-3C concentrations of 12 μM and 40 μM. In all cases the $\Delta R\%$ values are >15%. Based on the aforementioned analysis, these results suggest that all of the *N*-glycan ligands from the three GPs have affinities >$10^4$ M$^{-1}$ for hGal-3C. This finding can be rationalized by the presence of multiple LacNAc units, which is a known binding motif of hGal-3C.[54]

## 2.4 Conclusions

This work describes a multipronged ESI-MS approach for detecting and quantifying binding between water soluble GBPs and GPs and identifying the underlying glycan-GBP interactions. The CaR$^{IMS}$-ESI-MS method, wherein ion activation (to cause release) is performed post-IMS of the putative complexes, represents a straightforward approach to identify the presence of GBP-GP binding in solution. Competitive binding measurements, carried out using the *proxy ligand*-ESI-MS assay, allow for the apparent affinity the GBP

for the GP to be estimated. ESI-MS screening of the *N*-glycans released enzymatically from the GP against the GBP is used to elucidate the glycan structures that are involved in binding. Performing these measurements at multiple GBP concentrations allows for moderate-to-high affinity glycan (individual or grouped as compositional isomers) ligands to be distinguished from low affinity ligands. To illustrate implementation, the approach was applied to the known interactions between hGal-3C and the three human serum GPs, AGP, Hp1-1 and $\alpha$2M. Specific binding of hGal-3C to each GP was successfully detected by CaR$^{IMS}$-ESI-MS; no binding was detected for the negative control. The $K_{a,app}$ measured by *proxy ligand*-ESI-MS for the three GPs are similar (ranging from 2 x $10^5$ M$^{-1}$ to 4 x $10^5$ M$^{-1}$). These results are consistent with the reported apparent affinities of hGal-3 for serum GPs and their *N*-glycans. Finally, screening of the *N*-glycan libraries against hGal-3C identified ligands corresponding to twenty different saccharide compositions with sialylated bi-, tri- and tetra-antennary structures. Screening results carried out at two different hGal-3C concentrations suggest that all of the *N*-glycan ligands exhibit affinities $\geq 10^4$ M$^{-1}$.

## 2.5 References

(1)     Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49.

(2)     Helenius, A. Intracellular Functions of N-Linked Glycans. *Science (80-. ).* **2002**, *291* (5512), 2364–2369.

(3)     Paulson, J. C.; Blixt, O.; Collins, B. E. Sweet Spots in Functional Glycomics. *Nat. Chem. Biol.* **2006**, *2* (5), 238–248.

(4)     Grigorian, A.; Demetriou, M. Manipulating Cell Surface Glycoproteins by Targeting N-Glycan-Galectin Interactions. In *Methods in Enzymology*; 2010; Vol. 480, pp 245–266.

(5)     Cederfur, C.; Salomonsson, E.; Nilsson, J.; Halim, A.; Öberg, C. T.; Larson, G.; Nilsson, U. J.; Leffler, H. Different Affinity of Galectins for Human Serum Glycoproteins: Galectin-3 Binds Many Protease Inhibitors and Acute Phase Proteins. *Glycobiology* **2008**, *18* (5), 384–394.

(6)     Rillahan, C. D.; Paulson, J. C. Glycan Microarrays for Decoding the Glycome. *Annu. Rev. Biochem.* **2011**, *80* (1), 797–823.

(7)     Song, X.; Heimburg-Molinaro, J.; Smith, D. F.; Cummings, R. D. Glycan Microarrays of Fluorescently-Tagged Natural Glycans. *Glycoconj. J.* **2015**, *32* (7), 465–473.

(8)     Chan, Y. C.; Lin, H. Y.; Tu, Z.; Kuo, Y. H.; Hsu, S. T. D.; Lin, C. H. Dissecting the Structure–Activity Relationship of Galectin–Ligand Interactions. *Int. J. Mol. Sci.* **2018**, *19* (2), 1–20.

(9)     Grant, O. C.; Smith, H. M.; Firsova, D.; Fadda, E.; Woods, R. J. Presentation, Presentation, Presentation! Molecular-Level Insight into Linker Effects on Glycan Array Screening Data. *Glycobiology* **2014**, *24* (1), 17–25.

(10)    Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; St-Pierre, Y.; Klassen, J. S. Human Milk Oligosaccharide Specificities of Human Galectins. Comparison of Electrospray Ionization Mass Spectrometry and Glycan Microarray Screening

Results. *Anal. Chem.* **2017**, *89* (9), 4914–4921.

(11)    He, X. G., Gerona-Navarro, G., Jaffrey, S. R. Ligand Discovery Using Small Molecule Microarrays. *J. Pharmacol. Exp. Ther.* **2004**, *313* (1), 1–7.

(12)    Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein-Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441.

(13)    El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Carbohydrate-Protein Interactions by Electrospray Ionization Mass Spectrometry Analysis. *Biochemistry* **2012**, *51* (21), 4244–4253.

(14)    Haramija, M.; Peter-Katalinić, J. Quantitative Characterization of Galectin-3-C Affinity Mass Spectrometry Measurements: Comprehensive Data Analysis, Obstacles, Shortcuts and Robustness. *Rapid Commun. Mass Spectrom.* **2017**, *31* (20), 1709–1719.

(15)    Jovanović, M.; Peter-Katalinić, J. Preliminary Mass Spectrometry Characterization Studies of Galectin-3 Samples, Prior to Carbohydrate-Binding Studies Using Affinity Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (1), 129–136.

(16)    Fermas, S.; Gonnet, F.; Varenne, A.; Gareil, P.; Daniel, R. Frontal Analysis Capillary Electrophoresis Hyphenated to Electrospray Ionization Mass Spectrometry for the Characterization of the Antithrombin/Heparin Pentasaccharide Complex. *Anal. Chem.* **2007**, *79* (13), 4987–4993.

(17)    Maple, H. J.; Garlish, R. A.; Rigau-Roca, L.; Porter, J.; Whitcombe, I.; Prosser, C. E.; Kennedy, J.; Henry, A. J.; Taylor, R. J.; Crump, M. P.; et al. Automated Protein-Ligand Interaction Screening by Mass Spectrometry. *J. Med. Chem.* **2012**, *55* (2), 837–851.

(18)    Zhang, S.; Van Pelt, C. K.; Wilson, D. B. Quantitative Determination of Noncovalent Binding Interactions Using Automated Nanoelectrospray Mass

Spectrometry. *Anal. Chem.* **2003**, *75* (13), 3010–3018.

(19)  El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Applications of a Catch and Release Electrospray Ionization Mass Spectrometry Assay for Carbohydrate Library Screening. *Anal. Chem.* **2012**, *84* (1), 50–58.

(20)  Han, L.; Morales, L. C.; Richards, M. R.; Kitova, E. N.; Sipione, S.; Klassen, J. S. Investigating the Influence of Membrane Composition on Protein-Glycolipid Binding Using Nanodiscs and Proxy Ligand Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2017**, *89* (17), 9330–9338.

(21)  Han, L.; Kitova, E. N.; Li, J.; Nikjah, S.; Lin, H.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. Protein-Glycolipid Interactions Studied in Vitro Using ESI-MS and Nanodiscs: Insights into the Mechanisms and Energetics of Binding. *Anal. Chem.* **2015**, *87* (9), 4888–4896.

(22)  El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. Quantifying Ligand Binding to Large Protein Complexes Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2012**, *84* (9), 3867–3870.

(23)  Han, L.; Tan, M.; Xia, M.; Kitova, E. N.; Jiang, X.; Klassen, J. S. Gangliosides Are Ligands for Human Noroviruses. *J. Am. Chem. Soc.* **2014**, *136* (36), 12631–12637.

(24)  El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Protein Interactions with Isomeric Carbohydrate Ligands Using a Catch and Release Electrospray Ionization-Mass Spectrometry Assay. *Anal. Chem.* **2013**, *85* (16), 7637–7644.

(25)  El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; Bode, L.; Hage, N.; Falcone, F. H.; Klassen, J. S. Screening Natural Libraries of Human Milk Oligosaccharides against Lectins Using CaR-ESI-MS. *Analyst* **2018**, *143* (2), 536–548.

(26)  El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; St-Pierre, Y.; Klassen, J. S. High-Throughput Label- and Immobilization-Free Screening of Human Milk Oligosaccharides Against Lectins. *Anal. Chem.* **2017**, *89* (17), 8713–8722.

(27)   Li, J.; Fan, X.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Eugenio, L.; Ng, K. K. S.; Xiong, Z. J.; Privé, G. G.; Klassen, J. S. Screening Glycolipids Against Proteins in Vitro Using Picodiscs and Catch-and-Release Electrospray Ionization-Mass Spectrometry. *Anal. Chem.* **2016**, *88* (9), 4742–4750.

(28)   Fatunmbi, O.; Abzalimov, R. R.; Savinov, S. N.; Gershenson, A.; Kaltashov, I. A. Interactions of Haptoglobin with Monomeric Globin Species: Insights from Molecular Modeling and Native Electrospray Ionization Mass Spectrometry. *Biochemistry* **2016**, *55* (12), 1918–1928.

(29)   Engel, N. Y.; Weiss, V. U.; Marchetti-Deschmann, M.; Allmaier, G. NES GEMMA Analysis of Lectins and Their Interactions with Glycoproteins – Separation, Detection, and Sampling of Noncovalent Biospecific Complexes. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 77–86.

(30)   Kitova, E. N.; El-Hawiet, A.; Klassen, J. S. Screening Carbohydrate Libraries for Protein Interactions Using the Direct ESI-MS Assay. Applications to Libraries of Unknown Concentration. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (11), 1908–1916.

(31)   Kovak, M. R.; Saraswati, S.; Goddard, S. D.; Diekman, A. B. Proteomic Identification of Galectin-3 Binding Ligands and Characterization of Galectin-3 Proteolytic Cleavage in Human Prostasomes. *Andrology* **2013**, *1* (5), 682–691.

(32)   Bresalier, R. S.; Byrd, J. C.; Tessler, D.; Lebel, J.; Koomen, J.; Hawke, D.; Half, E.; Liu, K. F.; Mazurek, N. A Circulating Ligand for Galectin-3 Is a Haptoglobin-Related Glycoprotein Elevated in Individuals with Colon Cancer. *Gastroenterology* **2004**, *127* (3), 741–748.

(33)   Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. Structure of a Single-Chain Antibody Variable Domain (Fv) Fragment Complexed with a Carbohydrate Antigen at 1.7-A Resolution. *Proc. Natl. Acad. Sci.* **1994**, *91* (14), 6423–6427.

(34)   Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. Method for Distinguishing Specific

from Nonspecific Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2006**, *78* (9), 3010–3018.

(35)  Morelle, W.; Michalski, J.-C. Analysis of Protein Glycosylation by Mass Spectrometry. *Nat. Protoc.* **2007**, *2* (7), 1585–1602.

(36)  Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87* (8), 4370–4376.

(37)  Imre, T.; Kremmer, T.; Héberger, K.; Molnár-Szöllosi, É.; Ludányi, K.; Pócsfalvi, G.; Malorni, A.; Drahos, L.; Vékey, K. Mass Spectrometric and Linear Discriminant Analysis of N-Glycans of Human Serum Alpha-1-Acid Glycoprotein in Cancer Patients and Healthy Individuals. *J. Proteomics* **2008**, *71* (2), 186–197.

(38)  Clerc, F.; Reiding, K. R.; Jansen, B. C.; Kammeijer, G. S. M.; Bondt, A.; Wuhrer, M. Human Plasma Protein N-Glycosylation. *Glycoconj. J.* **2016**, *33* (3), 309–343.

(39)  Kawasaki, T., Koyama, J., Yamashina, I. Isolation and Characterization of A1-Acid Glycoprotein from Rat Serum. *J. Biochem.* **1966**, *60*, 554–560.

(40)  Luo, Z.; Lei, H.; Sun, Y.; Liu, X.; Su, D. F. Orosomucoid, an Acute Response Protein with Multiple Modulating Activities. *Journal of Physiology and Biochemistry*. 2015, pp 329–340.

(41)  Giménez, E.; Benavente, F.; Barbosa, J.; Sanz-Nebot, V. Towards a Reliable Molecular Mass Determination of Intact Glycoproteins by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21* (16), 2555–2563.

(42)  Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (35), 8763–8768.

(43)  Felitsyn, N.; Kitova, E. N.; Klassen, J. S. Thermal Decomposition of a Gaseous

Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behavior. *Anal. Chem.* **2001**, *73* (19), 4647–4661.

(44)   Yan, J.; Zhou, M.; Gilbert, J. D.; Wolff, J. J.; Somogyi, Á.; Pedder, R. E.; Quintyn, R. S.; Morrison, L. J.; Easterling, M. L.; Paša-Tolić, L.; et al. Surface-Induced Dissociation of Protein Complexes in a Hybrid Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Anal. Chem.* **2017**, *89* (1), 895–901.

(45)   Lipiski, M.; Deuel, J. W.; Baek, J. H.; Engelsberger, W. R.; Buehler, P. W.; Schaer, D. J. Human Hp1-1 and Hp2-2 Phenotype-Specific Haptoglobin Therapeutics Are Both Effective In Vitro and in Guinea Pigs to Attenuate Hemoglobin Toxicity. *Antioxid. Redox Signal.* **2013**, *19* (14), 1619–1633.

(46)   Lee, S. H.; Jeong, S.; Lee, J.; Yeo, I. S.; Oh, M. J.; Kim, U.; Kim, S.; Kim, S. H.; Park, S. Y.; Kim, J. H.; et al. Glycomic Profiling of Targeted Serum Haptoglobin for Gastric Cancer Using Nano LC/MS and LC/MS/MS. *Mol. Biosyst.* **2016**, *12* (12), 3611–3621.

(47)   Mollan, T. L.; Jia, Y.; Banerjee, S.; Wu, G.; Timothy Kreulen, R.; Tsai, A. L.; Olson, J. S.; Crumbliss, A. L.; Alayash, A. I. Redox Properties of Human Hemoglobin in Complex with Fractionated Dimeric and Polymeric Human Haptoglobin. *Free Radic. Biol. Med.* **2014**, *69*, 265–277.

(48)   Jensen, P. E. H.; Sottrup-Jensens, L. Primary Structure of Human Alpha 2-Macroglobulin. Complete Disulfide Bridge Assignment and Localization of Two Interchain Bridges in the Dimeric Proteinase Binding Unit. 1986, pp 15863–15869.

(49)   Rehman, A. A.; Ahsan, H.; Khan, F. H. Alpha-2-Macroglobulin: A Physiological Guardian. *J. Cell. Physiol.* **2013**, *228* (8), 1665–1675.

(50)   Hirabayashi, J. Oligosaccharide Specificity of Galectins: A Search by Frontal Affinity Chromatography. *Biochim. Biophys. Acta - Gen. Subj.* **2002**, *1572* (2–3),

232–254.

(51)   Arnold, J. N.; Wallis, R.; Willis, A. C.; Harvey, D. J.; Royle, L.; Dwek, R. A.; Rudd, P. M.; Sim, R. B. Interaction of Mannan Binding Lectin with A2 Macroglobulin via Exposed Oligomannose Glycans: A Conserved Feature of the Thiol Ester Protein Family? *J. Biol. Chem.* **2006**, *281* (11), 6955–6963.

(52)   Mancera-Arteu, M.; Giménez, E.; Barbosa, J.; Sanz-Nebot, V. Identification and Characterization of Isomeric N-Glycans of Human Alfa-Acid-Glycoprotein by Stable Isotope Labelling and ZIC-HILIC-MS in Combination with Exoglycosidase Digestion. *Anal. Chim. Acta* **2016**, *940*, 92–103.

(53)   Stowell, S. R.; Arthur, C. M.; Mehta, P.; Slanina, K. A.; Blixt, O.; Leffler, H.; Smith, D. F.; Cummings, R. D. Galectin-1, -2, and -3 Exhibit Differential Recognition of Sialylated Glycans and Blood Group Antigens. *J. Biol. Chem.* **2008**, *283* (15), 10109–10123.

(54)   Patnaik, S. K.; Potvin, B.; Carlsson, S.; Sturm, D.; Leffler, H.; Stanley, P. Complex N-Glycans Are the Major Ligands for Galectin-1, -3, and -8 on Chinese Hamster Ovary Cells. *Glycobiology* **2006**, *16* (4), 305–317.

# Chapter 3

# Influence of *N*-glycosylation on Glycoprotein Interactions Evaluated by Electrospray Ionization Mass Spectrometry

## 3.1 Introduction

Glycosylation is one of the most prominent and complex form of post-translation modifications (PTMs), impacting structures and biological functions of secretory and membrane-anchored proteins.[1] Glycans play important roles in modulating the stability and folding of glycoproteins (GPs), as well as regulating glycoprotein intracellular and interspecies recognition.[2,3] Changes in glycosylation occur in many diseases and pathophysiological events (such as autoimmune responses; inflammation and cancer metastasis)[3,4]. Aberrant glycosylation can result from conformational alterations in peptide and early-stage glycan chains, the expression and localization of glycosyltransferases, and is also associated with the activity of nucleotide sugar transporters.[5] Altered glycosylation contributes to the perturbation of the conformation, oligomerization, turnover and stability of GPs, and significantly influences both intermolecular and intramolecular (including both homotypic and heterotypic) interactions in a wide variety of fundamental biological events.[3,6]

*N*-glycosylation (where glycans are linked to Asn residues of a consensus sequon Asn-Xxx-Ser/Thr through a β *N*-glycosidic bond; Xxx can be any amino acid but Pro) and *O*-glycosylation (where glycans are attached to Ser or Thr residues) are the major types of protein glycosylation in humans.[7] *C*-glycosylation, where mannose residues are attached to Trp residues, is less common.[6] The biosynthesis of *N*-glycans, a complicated process involving a series of glycosyltransferases and glycosylhydrolases, occurs initially in the endoplasmic recticulum (ER) at a glycophospholipid, dolichol-PP-GlcNAc$_2$Man$_9$Glc$_3$ (GlcNAc for *N*-acetylglucosamine, Man for mannose, Glc for glucose) and it is then translocated to the Asn residues of a nascent peptide chain

and further processed and modified in the ER and Golgi apparatus in eukaryotic cells.[6]
Three types of *N*-glycans sharing a common core structure
(Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1–4GlcNAcβ1–Asn-Xxx-Ser/Thr) are generated
from this process, high mannose (where only Man is added to the core), complex (where
*N*-acetyllactosamine units (LacNAc), GlcNAc-β1-4Gal, extend the core, which can be
further modified by a bisecting GlcNAc residue linked to the mannosyl core, terminal
sialic acid residues (Sia) and fucose residues (Fuc) at antenna or core GlcNAc) and
hybrid (a combination of both features of high mannose and complex types).[6] Seven
types of *O*-glycans have been found in human, among which the mucin-type *O*-glycan
with an α GlcNAc residue attached to a Ser or Thr is the most abundant form.[8] Unlike the
biosynthesis of *N*-glycans, there is no consensus sequence for the formation of
α-GlcNAc-Ser/Thr linkage.[6] Through the action of numerous Golgi resident
glycosyltransferases after protein posttranslation, mature mucin-type *O*-glycans
consisting of eight core structures (core 7 is not found in human) and the Tn
(GalNAcα1-Ser/Thr) and sialyl Tn (Siaα2-6GalNAcα1-Ser/Thr) are formed and can be
further modified.[6]

Glycan chains generally encompass up to 20% of the mass of GPs.[6] Owing to the
diversity of carbohydrates (microheterogenity - different types of monosaccharides,
noncarbohydrate modifications such as acetylation, sulfation, and phosphorylation,
various glycosidic bonds and branching) and the variable sites of glycosylation
(macroheterogenity), GPs are highly heterogenous.[1] Current analytical methododologies
for characterizing protein glysosylation mainly focus on three aspects: GP, glycopeptide,
and enzymatically or chemically released glycans.

Capillary electrophoresis (CE), in which charged compounds are separated by an electric
field based on different electrophoretic mobility, has been employed as a high-resolution
analytical method for the separation and detection of glycoforms of intact GPs with
different extents of sialylation, sulfation and phosphorylation.[9] It can be achieved by

various modes of CE, including capillary zone electrophoresis (CZE – cationic and anionic analytes are separated in a running buffer by electroosmotic flow, EOF), micellar electrokinetic capillary chromatography (MEKC – both neutral and charged solutes partition between the aqueous phase and the micelles formed by the aggregation of surfactant monomers), and capillary isoelectrofocusing (CIEF – proteins and peptides are separated based on their different isoelectric points, pI).[10,11] Although it has been successfully applied in the separation of protein isoforms (glycoforms and genetic variants), it doesn't provide insights into the attached glycan chains, thus coupling with mass spectrometry (MS) is typically required for the further analysis.

Lectin affinity chromatography, where lectins are immobilized on a separation matrices such as agarose, is an efficient technique for the isolation, purification and characterization of intact GPs due to the specificities of lectins for a variety of oligosaccharide epitopes.[12] GPs from biological fluids can be enriched either by lectins with broad specificities or by multiple specific lectins (e.g., multi-lectin affinity chromatography (M-LAC))[13] and then analyzed by SDS-PAGE. Lectin microarrays have also been employed for the profiling of protein glycosylation. Distinct lectins with different specificities towards glycans are immobilized on a substrate, and the binding of GPs are typically detected by fluorescence using either direct (labelling of the GP) or antibody-overlay (labelling of the GP specific antibody) assay.[14,15] However, due to the weak interactions ($K_a = 10^3$-$10^7$ $M^{-1}$) between lectins and glycans, a number of assays have been developed to enhance the sensitivity.[14] Most importantly, the identification of GPs still relies on (glyco)peptide analysis by (Liquid Chromatography (LC))-MS/MS.

MS represents a versatile and powerful technique for the identification and characterization of protein glycosylation at the intact GP, glycopeptide and glycan levels. The "bottom-up" MS based workflows are the most commonly applied strategy for the mapping of glycosylation sites and profiling of glycans. Generally, *N*-glycans are released by *N*-glycosidase (such as peptide *N*-glycosidase F and A, endoglycosidases)

while *O*-glycans are released chemically (such as reductive β elimination and hydrazinolysis). The purified glycans are then analyzed directly by MS and tandem MS or separation based methods hyphenated with MS ((LC)-MS), including reversed phase (RP) chromatography, hydrophilic interaction chromatography (HILIC) and porous graphitized carbon chromatography (PGC) or other separation techniques such as CE and high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD)) to distinct structural isomers.[16,17] Derivatization of glycans (such as reductive amination and permethylation) facilitates the ionization, detection and identification in MS.[15]

Information on the glycosylation sites can be obtained by LC-MS analysis of peptides or direct analysis of glycopeptides (which is more widely used). Glycopeptides are generated by digestion with specific exo-proteinases (such as trypsin, Lys-C, Arg-C, Glu-C, Asn-N and chymotrypsin) or nonspecific proteases (such as pronase).[11] Enrichment of the glycopeptides prior to the analysis is often performed to enhance the signal intensity due to their low ionization efficiency (decreased hydrophobicity resulted from glycosylation) and low abundance compared to peptides.[18] And in the past 20 years, lectin affinity based, covalent interactions based (hydrazide and boronic acid chemistry) and chromatographic separation and solid-phase extraction (RP chromatography, cation exchange chromatography, size exclusion chromatography (SEC), PGC, and HILIC) based enrichment approaches have been developed and extensively employed.[19]

Tandem MS assists the localization of glycosylation sites as well as peptide sequences and glycan structures. In dissociation techniques such as collisional-induced dissociation (CID) and infrared multiphoton dissociation (IRMPD), cleavage of peptide backbone CO-N, which occurs via vibrational excitation, predominantly forms *b* and *y'* ions.[20] However, fragmentation of glycans is preferentially generated, as cleavages of the peptide backbone require much higher collisional energies than the cleavages of the glycosidic linkages.[21] Therefore, minimal information is provided for the peptide

backbone sequence. The limitation of labile glycan fragmentation can be overcome by high energy collisional dissociation (HCD), a new technique exclusive to Orbitrap mass spectrometers. Unlike CID, where glycan oxonium ions (e.g., HexNAc (*N*-acetylhexosamine) at *m/z* 204.09; Sia at *m/z* 292.10; and $HexNAc_1Hex_1$ (Hex for hexose) at *m/z* 366.14 ) from glycopeptide precursor ions at high *m/z* are not detectable in an ion trap, trap efficiency of those ions is enhanced in HCD.[22] Alternatively, radical induced electron-based techniques, such as electron-capture dissociation (ECD) and electron-transfer dissociation (ETD), cleave disulfide bonds as well as the $N-C_\alpha$ bonds, producing even electron *c* and odd electron $z^\bullet$ ions while labile PTMs remain.[20]

Although "bottom up" MS strategy is mature and widely employed in the present glycoproteomics research, a downside of this approach is that information of the intact GP, such as GP isoforms and the entire PTMs, is lost.[23] Moreover, probing sources of peptides from a mixture of proteins with assorted PTMs is difficult.[24] More importantly, due to the digestion of the protein or the removal of glycans, binding properties of intact GPs are completely missing. These problems can be overcome by applying the "top down" MS approach. Profiling of GP isoforms can be achieved without labelling, and correlations between glycopeptides and intact GPs can be obtained in combination with tandem MS such as ECD and ETD.[20,24] And recently, a "top down" high-resolution MS based method using an Orbitrap mass spectrometer has been developed for the elucidation of microhetergenity and glycoproteoform related ligands interactions of GPs (i.e., human alpha-1-acid GP (AGP)-warfarin, haptoglobin-hemoglobin) showing the prominence of native high-resolution MS in investigating protein glycosylation and binding.[25] However, influence of glycosylation on GP interactions was studied on asialo-GPs, where Sia residues were removed enzymatically from GPs. Affinities of warfarin, an anticoagulant drug, to different glycoforms of asialo-AGP were determined by ESI-MS titration experiments, where initial warfarin concentrations were estimated to be the same as they were at equilibrium. It was established that fucosylation and

*N*-glycan branching have measurable effects on the binding constants, but their affinities are all within a factor of two.[25] Most importantly, the average association constant ($K_a$) value ($2.8 \times 10^6$ M$^{-1}$) of asialo-AGP binding to warfarin is ten times larger than the previous reported values ($2.3 \times 10^5$ M$^{-1}$ and $1.2 \times 10^5$ M$^{-1}$)[26,27] for AGP binding to warfarin, suggesting that Sia residues significantly decrease AGP-warfarin binding.

Here, high-resolution native ESI-MS is employed to investigate the AGP-warfarin binding system with the objective of testing some of the reported conclusions and extending them to a more comprehensive study of the impact of glycosylation on AGP-warfarin binding. First, high resolution MS was applied to elucidate the microheterogeneity of AGP, and following treatments with sialidase and fucosidase. Then, *N*-glycans were completely removed from AGP by PNGase F. Warfarin binding to major glycoforms of sialidase treated AGP was then quantified using estimated concentrations of glycoforms and compared to the $K_a$ of warfarin for intact AGP determined by isothermal titration calorimetry (ITC) and the reported results. Effect of glcosylation on GP interactions was studied more extensively by comparing differences in warfarin binding to *N*-glycans partially- and completely- removed AGP.

## 3.2 Experimental

### 3.2.1 Materials

Human plasma AGP, neuraminidase from *Clostridium perfringens*, α-L-fucosidase from bovine kidney, warfarin, imidazole and 2-methylimidazole were purchased from Sigma-Aldrich Canada (Oakville, Canada). α1-3,4 fucosidase from *Prunus dulcis* was purchased from New England Biolabs Canada (Whitby, Canada). His-tagged PNGase F was a gift from Prof. Matthew Macauley (University of Alberta). Dimethyl sulphoxide (DMSO) was purchased from Caledon Laboratory Chemicals (Georgetown, Canada). Nickel chelating resin was purchased from G-Biosiences (St. Louis, US).

### 3.2.2 Sample preparation

Asialo-AGP was obtained by digestion of 400 µg AGP with 0.2 units of neuraminidase

from *Clostridium perfringens* in 1× GlycoBuffer 1 (5 mM $CaCl_2$, 20mM sodium acetate, pH 5.5) at 37 °C overnight. 100 μg of asialo-AGP was then treated with 16 units of α1-3,4 fucosidase and BSA (final concentration at 0.1 mg mL$^{-1}$) in 1× GlycoBuffer 1 (5 mM $CaCl_2$, 20mM sodium acetate, pH 5.5) at 37 °C overnight. deglycosylated AGP was obtained by digestion of 6 mg AGP with 2 mg of His-tagged PNGase F in 1× GlycoBuffer 2 (50 mM sodium phosphate, pH 7.5) at 37 °C overnight. His-tagged PNGase F was removed using nickel chelating resin. All AGP samples were dialyzed against an aqueous solution of 200 mM ammonium acetate (pH 6.8) using an Amicon 0.5 mL microconcentrator (EMD Millipore, Billerica, MA) with a MW cut-off of 10 kDa (AGP, asialo-AGP, fucosidase treated asialo-AGP) or 3 kDa (PNGase F treated AGP) and stored at -20 °C until needed. The concentrations of the protein stock solutions were estimated by UV absorption (280 nm). Warfarin was dissolved in 100% DMSO as a stock solution. Binding measurements of asialo-AGP and warfarin was carried out at 200 mM ammonium acetate (pH 7.6) in order to compare with the reported results.

### 3.2.3 Mass spectrometry

ESI-MS measurements were performed using a Q Exactive UHMR hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, US) equipped with a nanoflow ESI (nanoESI) source in positive ion mode. The nanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.78 mm i.d.), pulled to ~5 μm outer-diameter using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). To perform ESI, a platinum wire was inserted into the solution and a voltage of 0.9-1.1 kV was applied. The capillary temperature was maintained at 100 ºC for all experiments. Backing pressure was maintained at ~$3.03×10^{-9}$ mbar. Data acquisition and processing were performed using Thermo Fisher Xcalibur software.

### 3.2.4 Isothermal titration calorimetry

The ITC measurements were carried out using a VP-ITC (MicroCal, Inc. USA). For the ITC experiment, the AGP solution (50 μM) in the sample cell was titrated with a solution

of warfarin (500 μM); both the protein and ligand solutions were 200 mM aqueous ammonium acetate (pH 6.8) at 25 °C.

### 3.2.5 AGP structure and warfarin docking

Crystal structure of AGP was obtained from Protein Data Bank (PDB ID: 3KQ0). The N-glycans were generated in Glycam molecular builder and appended to the structure of AGP using Pymol. The warfarin structure was retrieved from PubChem database (compound ID: 54678486). Molecule docking was performed using AutoDock Vina.

### 3.2.6 Direct ESI-MS analysis of AGP-warfarin interactions

The affinities of warfarin (L) binding to different glycoforms ($P_i$, $i$ for different glycoforms) of native AGP and asialo-AGP were measured using direct ESI-MS assay. This assay is based on the direct detection and quantification of the gas phase ions of free and ligand-bound protein. The magnitude of $K_{a,i}$ is determined from the ratio ($R_i$) of total abundance ($Ab$) of ligand-bound and free protein ions, as measured by ESI-MS for solutions with known initial concentrations of ligand and protein ($[L]_o$ and $[Pi]_o$, respectively). For a 1:1 ligand-protein complex (Eq. 3.1), $K_{a,i}$ is calculated using Eq. 3.2:

$$P_i + L \rightleftharpoons P_iL \tag{3.1}$$

$$K_{a,i} = \frac{[P_iL]}{[P_i][L]} = \frac{R_i}{[L]_0 - \frac{R_i}{1+R_i}[P_i]_0} \tag{3.2}$$

$$R_i = \frac{\sum Ab(P_iL)}{\sum Ab(P_i)} = \frac{[P_iL]}{[P_i]} \tag{3.3}$$

Initial average AGP concentration ($[P_T]_0$) was calculated based on the MW of the most abundant glycoform. $[P_i]_0$ was estimated based on relative abundances measured by ESI-MS.

By performing titration experiments where $[P_T]_0$ is fixed and $[L]_0$ is varied, $K_{a,i}$ values are established from nonlinear regression analysis of the experimentally determined concentration dependence of the fraction of ligand-bound protein, $R_i/(R_i+1)$ (Eq. 3.4):

$$\frac{R_i}{R_i+1} = \frac{1+K_{a,i}[L]_0+K_{a,i}[P_i]_0-\sqrt{(1-K_{a,i}[L]_0+K_{a,i}[P_i]_0)^2+4K_{a,i}[L]_0}}{2K_{a,i}[P_i]_0} \quad (3.4)$$

Alternatively, as $K_{a,i}$ can be expressed as Eq. 3.5, and the ratio of affinities (e.g., $K_{a,u}/K_{a,i}$) for two different forms of AGP (e.g., $P_u$ (unglycosylated AGP) and $P_i$) can be determined from the corresponding abundance ratio (e.g., $R_u/R_i$).

$$K_{a,i} = \frac{[P_iL]}{[P_i][L]} = \frac{R_i}{[L]_0 - \frac{R_i}{1+R_i}[P_i]_0} \quad (3.5)$$

$$\frac{[P_uL]/[P_u]}{[P_iL]/[P_i]} = \frac{Ab(P_uL)/Ab(P_u)}{Ab(P_iL)/Ab(P_i)} = \frac{R_u}{R_i} = \frac{K_{a,u}}{K_{a,i}} \quad (3.6)$$

## 3.3 Results and Discussion

### 3.3.1 Characterization of microheterogeneity of AGP

#### 3.3.1.1 Characterization of microheterogeneity of native AGP

Human alpha 1 acid glycoprotein (AGP) is an acute phase serum glycoprotein produced mainly by hepatic cells and it is highly glycosylated with five *N*-glycosylation sites (Asn 38, Asn 54, Asn 75, and Asn85)[28] occupied by complex-type *N*-glycans. AGP belongs to the lipocalin family, which is able to transport small hydrophobic molecules, and it is responsible for binding and transport of numerous ligands in plasma, including basic, neutral and acidic pharmaceutical compounds such as dipyridamole, progesterone and warfarin.[29] It has been found that not only the concentration but also the glycosylation micoheterogenity (i.e., degree of branching and fucosylation) of AGP alters during inflammation as well as other pathophysiological states (e.g., rheumatoid arthritis, carcinoma).[30–33]

Human AGP is produced from three adjacent genes: AGP-A, which encodes ORM1 proteins (F1, F2 and S variants); whereas AGP-B/B' encode ORM 2 protein (A variant).[34,35] Distributed worldwide, F1 and S variants, which differ in one amino acid (F1 has Gln-38 and S has Arg-38) out of 183 residues, are the most common proteoforms

of ORM 1.[35] The native mass spectrum of AGP (Figure 3.1) is complicated due to the fact that $N$-glycans of AGP are sialylated across all the five glycosylation sites. Additionally, partially overlapping charge states (+10 to +12) and the similarity in masses of two Fuc residues (292.28 Da) and one Sia residue (291.25 Da) increase the difficulties in accurate assignments.

Shown in Figure 3.1 is a representative mass spectrum acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of AGP (0.189 mg mL$^{-1}$) and DMSO (1%). A peak envelope from $m/z$ at 4300 to 4800 at charge states from +8 to +10 was observed. The most abundant glycoform was assigned as P4 (experimental MW: 35999.59 ± 0.01 Da), i.e., $Hex_{32}HexNAc_{27}Fuc_aSia_b$ (a/2+b=16), which is within 1 Da difference from the theoretical MW of AGP (MW of AGP F1 variant peptide backbone: 21539.06 Da; mass of $Hex_{32}HexNAc_{27}Fuc_0Sia_{16}$: 15334.77 Da). Glycoform analysis was performed by searching the mass of glycan in an in-house software with two assumptions: the difference between the number of Hex and HexNAc is five as no bisecting GlcNAc has been observed for AGP $N$-glycans;[36,37] the difference between the number of Hex and Sia is more than ten (5 glycosylation sites × 2 core GlcNAc) since Sia are only linked to LacNAc. Thirteen main peaks (P0-P6) with $Hex_1HexNAc_1$ and $Fuc_aSia_b$ (a/2+b=1) differences were detected and both F1 and S variants were detected, with F1 variant showing higher relative abundances than S variant. In total, 105 glycoproteoforms were identified (Table 3.1), which are 62 more than the number of AGP glycoforms reported,[25] with the assumption that at most four Fuc residues are detected on five glycosylation sites. However, the mass differences of one Sia and three Fuc residues (729.68Da) or five Fuc residues (730.71 Da) and $Hex_2HexNAc_2$ (730.67 Da) are 0.99 Da and 0.04 Da respectively, and thus glycoforms with more (or less) branches and less (more) Fuc and Sia residues might contribute to the relative abundances of each annotated glycoform.
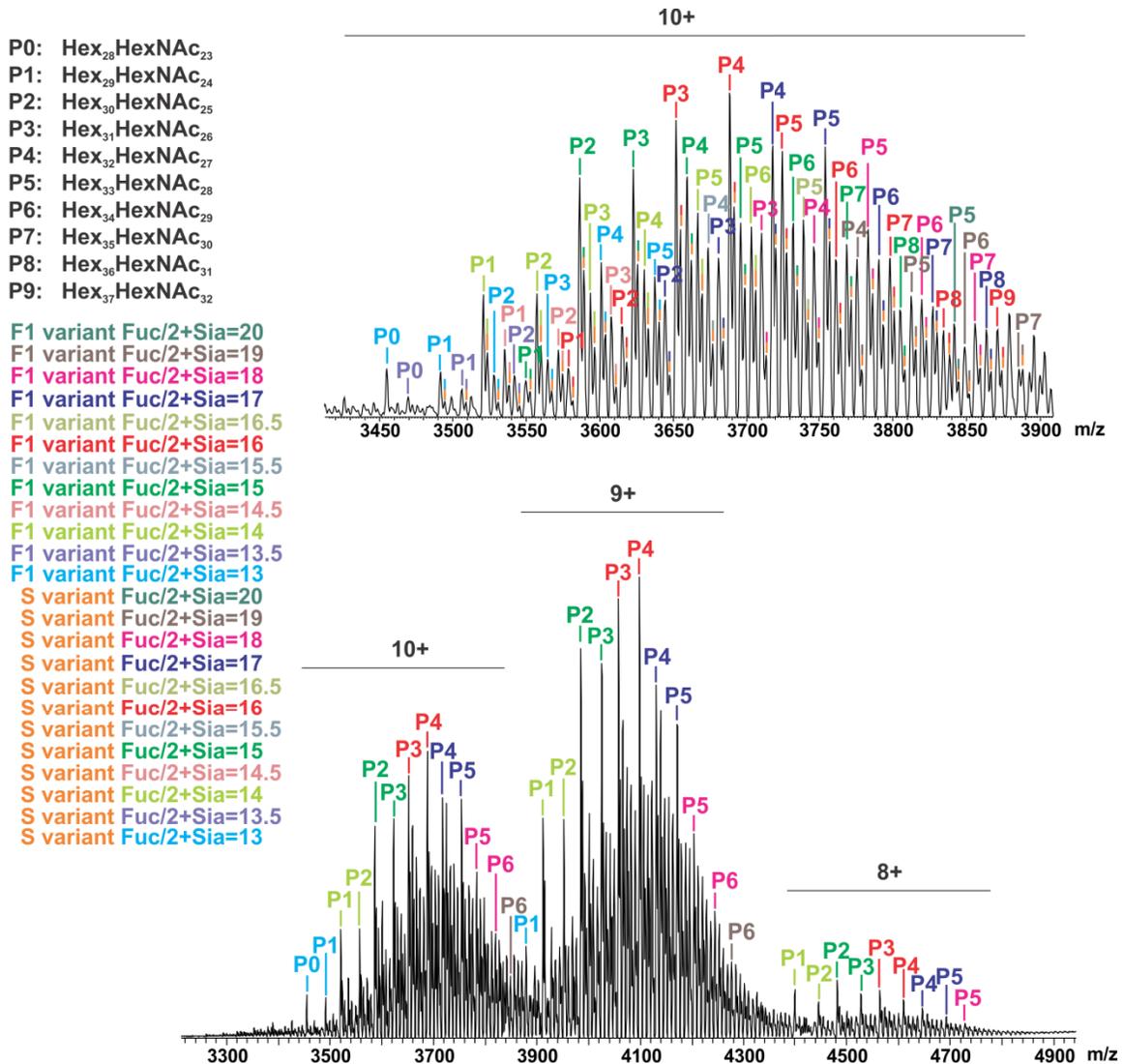
**Figure 3.1** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of AGP (0.189 mg mL$^{-1}$) and DMSO (1%). Glycoform assignments of AGP at +10 are shown as an inset, and S variants that share the same compositions with the F variants are labelled beside in orange.

**Table 3.1** Annotation of AGP glycoproteoforms

| Num. | Genetic Variant | Composition[a] Hex_HexNAc_Fuc/2+Sia | Experimental[b] Mass (Da) | Theoretical[c] Mass (Da) | Mass Difference (Da) | Annotation |
|------|------|------|------|------|------|------|
| 1 | F1 | 28_23_13 | 34541.13 ± 0.01 | 34539.76 | 1.37 | P0 |
| 2 | F1 | 28_23_13.5 | 34684.13 ± 0.01 | 34685.90 | 1.77 | P0 |
| 3 | S | 28_23_13.5 | 34712.63 ± 1.84 | 34713.95 | 1.32 | |
| 4 | F1 | 29_24_13 | 34906.23 ± 0.01 | 34906.12 | 0.11 | P1 |
| 5 | S | 29_24_13 | 34933.63 ± 0.01 | 34934.17 | 0.54 | |
| 6 | F1 | 29_24_13.5 | 35051.53 ± 0.01 | 35050.21 | 1.32 | P1 |
| 7 | S | 29_24_13.5 | 35079.13 ± 0.01 | 35078.26 | 0.87 | |
| 8 | F1 | 29_24_14 | 35197.04 ± 0.10 | 35196.35 | 0.69 | P1 |
| 9 | S | 29_24_14 | 35225.42 ± 0.01 | 35224.40 | 1.02 | |
| 10 | F1 | 30_25_13 | 35269.71 ± 0.01 | 35270.43 | 0.72 | P2 |
| 11 | S | 30_25_13 | 35296.76 ± 0.01 | 35298.48 | 1.71 | |
| 12 | F1 | 29_24_14.5 | 35342.92 ± 0.63 | 35342.49 | 0.43 | P1 |
| 13 | S | 29_24_14.5 | 35372.51 ± 0.01 | 35370.54 | 1.97 | |
| 14 | F1 | 30_25_13.5 | 35410.81 ± 1.00 | 35415.54 | 4.73 | P2 |
| 15 | F1 | 29_24_15 | 35485.48 ± 0.01 | 35487.60 | 2.13 | P1 |
| 16 | S | 29_24_15 | 35515.02 ± 0.97 | 35515.65 | 0.64 | |
| 17 | F1 | 30_25_14 | 35562.43 ± 0.01 | 35561.68 | 0.75 | P2 |
| 18 | S | 30_25_14 | 35589.33 ± 0.01 | 35589.73 | 0.40 | |
| 19 | F1 | 31_26_13 | 35634.25 ± 0.70 | 35634.73 | 0.48 | P3 |
| 20 | S | 31_26_13 | 35660.68 ± 0.70 | 35662.78 | 2.10 | |
| 21 | F1 | 30_25_14.5 | 35707.30 ± 0.01 | 35707.82 | 0.53 | P2 |
| 22 | S | 30_25_14.5 | 35734.73 ± 0.70 | 35735.87 | 1.14 | |
| 23 | F1 | 29_24_16 | 35775.73 ± 0.01 | 35779.89 | 4.16 | P1 |

| 24 | S | 29_24_16 | 35806.06 ± 0.97 | 35807.94 | 1.88 | |
|---|---|---|---|---|---|---|
| 25 | F1 | 30_25_15 | 35852.97 ± 0.01 | 35852.94 | 0.03 | **P2** |
| 26 | S | 30_25_15 | 35883.16 ± 0.01 | 35880.99 | 2.17 | |
| 27 | F1 | 31_26_14 | 35926.20 ± 0.01 | 35927.02 | 0.81 | **P3** |
| 28 | S | 31_26_14 | 35955.44 ± 0.72 | 35955.07 | 0.37 | |
| 29 | F1 | 32_27_13 | 35999.59 ± 0.01 | 36000.07 | 0.48 | **P4** |
| 30 | S | 32_27_13 | 36028.44 ± 0.01 | 36028.12 | 0.32 | |
| 31 | F1 | 31_26_14.5 | 36068.43 ± 0.64 | 36072.13 | 3.70 | **P3** |
| 32 | F1 | 30_25_16 | 36141.82 ± 0.66 | 36143.16 | 1.35 | **P2** |
| 33 | S | 30_25_16 | 36172.78 ± 0.73 | 36171.21 | 1.57 | |
| 34 | F1 | 31_26_15 | 36216.85 ± 0.01 | 36218.27 | 1.42 | **P3** |
| 35 | S | 31_26_15 | 36247.48 ± 0.01 | 36246.32 | 1.16 | |
| 36 | F1 | 32_27_14 | 36291.13 ± 0.01 | 36292.35 | 1.22 | **P4** |
| 37 | S | 32_27_14 | 36318.96 ± 0.01 | 36320.40 | 1.44 | |
| 38 | F1 | 33_28_13 | 36364.33 ± 0.01 | 36365.40 | 1.07 | **P5** |
| 39 | S | 33_28_13 | 36393.57 ± 0.01 | 36393.45 | 0.12 | |
| 40 | F1 | 30_25_17 | 36434.13 ± 0.73 | 36436.47 | 2.35 | **P2** |
| 41 | S | 30_25_17 | 36465.06 ± 0.73 | 36464.52 | 0.54 | |
| 42 | F1 | 31_26_16 | 36509.64 ± 0.01 | 36509.53 | 0.11 | **P3** |
| 43 | S | 31_26_16 | 36537.61 ± 0.01 | 36537.58 | 0.04 | |
| 44 | F1 | 32_27_15 | 36584.87 ± 0.01 | 36583.60 | 1.26 | **P4** |
| 45 | S | 32_27_15 | 36611.53 ± 0.01 | 36611.65 | 0.12 | |
| 46 | F1 | 33_28_14 | 36655.89 ± 0.01 | 36656.65 | 0.76 | **P5** |
| 47 | S | 33_28_14 | 36684.02 ± 0.01 | 36684.70 | 0.69 | |
| 48 | F1 | 32_27_15.5 | 36727.06 ± 0.01 | 36728.72 | 1.65 | **P4** |
| 49 | S | 32_27_15.5 | 36757.30 ± 0.74 | 36756.77 | 0.54 | |
| 50 | F1 | 31_26_17 | 36797.07 ± 0.01 | 36800.78 | 3.71 | **P3** |

| 51 | S | 31_26_17 | 36828.80 ± 0.74 | 36828.83 | 0.03 | |
|---|---|---|---|---|---|---|
| 52 | F1 | 32_27_16 | 36873.17 ± 0.01 | 36873.83 | 0.66 | **P4** |
| 53 | S | 32_27_16 | 36903.11 ± 0.03 | 36901.88 | 1.23 | |
| 54 | F1 | 33_28_15 | 36947.50 ± 0.75 | 36947.91 | 0.41 | **P5** |
| 55 | S | 33_28_15 | 36976.58 ± 0.01 | 36975.96 | 0.62 | |
| 56 | F1 | 34_27_14 | 37020.07 ± 0.01 | 37021.99 | 1.92 | **P6** |
| 57 | S | 34_27_14 | 37050.14 ± 0.01 | 37050.04 | 0.10 | |
| 58 | F1 | 31_26_18 | 37089.31 ± 0.01 | 37089.52 | 3.75 | **P3** |
| 59 | S | 31_26_18 | 37121.04 ± 0.01 | 37121.11 | 0.07 | |
| 60 | F1 | 32_27_17 | 37166.33 ± 0.01 | 37165.09 | 1.24 | **P4** |
| 61 | S | 32_27_17 | 37194.15 ± 0.75 | 37193.14 | 1.01 | |
| 62 | F1 | 33_28_16 | 37240.62 ± 0.01 | 37239.16 | 1.45 | **P5** |
| 63 | S | 33_28_16 | 37268.47 ± 0.68 | 37267.21 | 1.26 | |
| 64 | F1 | 34_29_15 | 37313.51 ± 0.01 | 37313.24 | 0.27 | **P6** |
| 65 | S | 34_29_15 | 37340.47 ± 0.77 | 37341.29 | 0.82 | |
| 66 | F1 | 33_28_16.5 | 37382.62 ± 0.75 | 37385.30 | 2.68 | **P5** |
| 67 | S | 33_28_16.5 | 37413.11 ± 0.75 | 37413.35 | 0.24 | |
| 68 | F1 | 32_27_18 | 37453.97 ± 0.01 | 37457.37 | 3.39 | **P4** |
| 69 | S | 32_27_18 | 37483.98 ± 0.77 | 37485.42 | 1.44 | |
| 70 | F1 | 33_28_17 | 37530.04 ± 0.77 | 37530.42 | 0.37 | **P5** |
| 71 | S | 33_28_17 | 37561.29 ± 0.01 | 37558.47 | 2.82 | |
| 72 | F1 | 34_29_16 | 37601.26 ± 0.01 | 37604.50 | 3.24 | **P6** |
| 73 | S | 34_29_16 | 37632.05 ± 0.01 | 37632.55 | 0.50 | |
| 74 | F1 | 35_30_15 | 37675.20 ± 0.01 | 37678.58 | 3.37 | **P7** |
| 75 | S | 35_30_15 | 37705.57 ± 0.72 | 37706.63 | 1.05 | |
| 76 | F1 | 32_27_19 | 37745.86 ± 0.78 | 37749.65 | 3.79 | **P4** |
| 77 | S | 32_27_19 | 37776.28 ± 0.78 | 37777.70 | 1.42 | |

| 78 | F1 | 33_28_18 | 37820.80 ± 0.01 | 37821.67 | 0.87 | **P5** |
|-----|-----|-----------|-------------------|-----------|-------|--------|
| 79 | S | 33_28_18 | 37851.86 ± 0.01 | 37849.72 | 2.13 | |
| 80 | F1 | 34_29_17 | 37894.89 ± 0.77 | 37895.75 | 0.86 | **P6** |
| 81 | S | 34_29_17 | 37922.44 ± 0.78 | 37923.80 | 1.36 | |
| 82 | F1 | 35_30_16 | 37969.79 ± 0.71 | 37969.83 | 0.04 | **P7** |
| 83 | S | 35_30_16 | 37996.79 ± 0.01 | 37997.88 | 1.09 | |
| 84 | F1 | 36_31_15 | 38039.16 ± 0.01 | 38043.91 | 4.75 | **P8** |
| 85 | S | 36_31_15 | 38068.88 ± 1.32 | 38071.96 | 3.08 | |
| 86 | F1 | 33_28_19 | 38113.32 ± 0.79 | 38114.98 | 1.67 | **P5** |
| 87 | S | 33_28_19 | 38143.23 ± 1.50 | 38143.03 | 0.20 | |
| 88 | F1 | 34_29_18 | 38185.18 ± 0.01 | 38187.01 | 1.82 | **P6** |
| 89 | S | 34_29_18 | 38214.54 ± 0.79 | 38215.06 | 0.52 | |
| 90 | F1 | 35_30_17 | 38261.44 ± 0.79 | 38261.08 | 0.36 | **P7** |
| 91 | S | 35_30_17 | 38288.86 ± 0.79 | 38289.13 | 0.28 | |
| 92 | F1 | 36_31_16 | 38332.61 ± 0.79 | 38335.16 | 2.55 | **P8** |
| 93 | S | 36_31_16 | 38361.34 ± 1.19 | 38363.21 | 1.88 | |
| 94 | F1 | 33_18_20 | 38406.18 ± 1.57 | 38406.24 | 0.05 | **P5** |
| 95 | S | 33_18_20 | 38435.05 ± 0.01 | 38434.29 | 0.76 | |
| 96 | F1 | 34_29_16 | 38476.15 ± 0.71 | 38478.26 | 2.11 | **P6** |
| 97 | S | 34_29_16 | 38507.07 ± 0.01 | 38506.31 | 0.76 | |
| 98 | F1 | 35_30_18 | 38550.80 ± 0.79 | 38552.34 | 1.54 | **P7** |
| 99 | S | 35_30_18 | 38580.86 ± 1.08 | 38580.39 | 0.47 | |
| 100 | F1 | 36_31_17 | 38624.77 ± 0.75 | 38626.42 | 1.64 | **P8** |
| 101 | S | 36_31_17 | 38654.03 ± 0.01 | 38654.47 | 0.44 | |
| 102 | F1 | 37_32_16 | 38699.02 ± 0.01 | 38700.50 | 1.48 | **P9** |
| 103 | S | 37_32_16 | 38729.10 ± 0.01 | 38728.55 | 0.55 | |
| 104 | F1 | 35_30_19 | 38844.20 ± 1.39 | 38844.62 | 0.42 | **P7** |

| 105 | S | 35_30_19 | $38871.03 \pm 0.01$ | 38872.67 | 1.64 | |

a. The compositions of glycoproteoforms correspond to the number of hexose (Hex), *N*-Acetylhexosamine (HexNAc), the sum of half of the number of fucose (Fuc) and the number of sialic acid (Sia) residues.

b. The experimental masses were calculated based on the average MW of glycoproteoforms at three charge states (+8 to +10), and errors correspond to one standard deviation of three replicate measurements.

c. The theoretical masses were calculated from the mass of peptide backbone of AGP (F1 variant: 21539.06 Da and S variant: 21567.11 Da) and the closest average masses of monosaccharides (Hex: 162.1406 Da, HexNAc: 203.1925 Da, Fuc: 146.1412 Da and Sia: 291.2546 Da) to the experimental values.

### 3.3.1.2 Characterization of microheterogeneity of asialo-AGP

Desialylation substantially reduces the heterogeneity and the ambiguity in assignments of AGP glycoforms. Shown in Figure 3.2 is a representative mass spectrum acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of asialo-AGP (0.628 mg mL$^{-1}$) and DMSO (1%). The native mass spectrum of asialo-AGP exhibits a peak envelope from *m/z* at 3050 to 4350 at charge states from +8 to +10 with the most abundant glycoform assigned as P4 (experimental MW: $32214.32 \pm 0.01$ Da) , i.e., $Hex_{32}HexNAc_{27}Fuc_0$, which is consistent with the most abundant glycoform (i.e., $Hex_{32}HexNAc_{27}Fuc_aSia_b$, a/2+b=16) detected in AGP and within 1 Da difference from the theoretical MW of asialo-AGP (MW of AGP F1 variant peptide backbone: 21539.06 Da; mass of $Hex_{32}HexNAc_{27}Fuc_0$: 10674.70 Da). Both F1 and S variants (labelled in orange), with the relative abundances of F1 variant higher than that of S variant, were detected with mass differences of around $28.57 \pm 2.56$ Da, which is in agreement with the findings in the glycoproteoform analysis of AGP. Eleven main peaks (P0-P10) and 98 glycoproteoforms in total were identified ranging from unfucosylated to tetrafucosylated

(Table 3.2), which is in agreement with the number (98) of reported asialo-AGP glycoform.[25] Notably, mass errors of tetrafucosylated asialo-AGP were higher than the other ones, which could stem from the overlapping between glycoforms with $Hex_1HexNAc_2$ (568.53 Da) and $Fuc_4$ (584.56 Da) differences at charge states from +8 to +10.



P0:  $Hex_{28}HexNAc_{23}$
P1:  $Hex_{29}HexNAc_{24}$
P2:  $Hex_{30}HexNAc_{25}$
P3:  $Hex_{31}HexNAc_{26}$
P4:  $Hex_{32}HexNAc_{27}$
P5:  $Hex_{33}HexNAc_{28}$
P6:  $Hex_{34}HexNAc_{29}$
P7:  $Hex_{35}HexNAc_{30}$
P8:  $Hex_{36}HexNAc_{31}$
P9:  $Hex_{37}HexNAc_{32}$
P10: $Hex_{38}HexNAc_{33}$
F1 variant $Fuc_0$
F1 variant $Fuc_1$
F1 variant $Fuc_2$
F1 variant $Fuc_3$
F1 variant $Fuc_4$
S variant $Fuc_0$
S variant $Fuc_1$
S variant $Fuc_2$
S variant $Fuc_3$
S variant $Fuc_4$
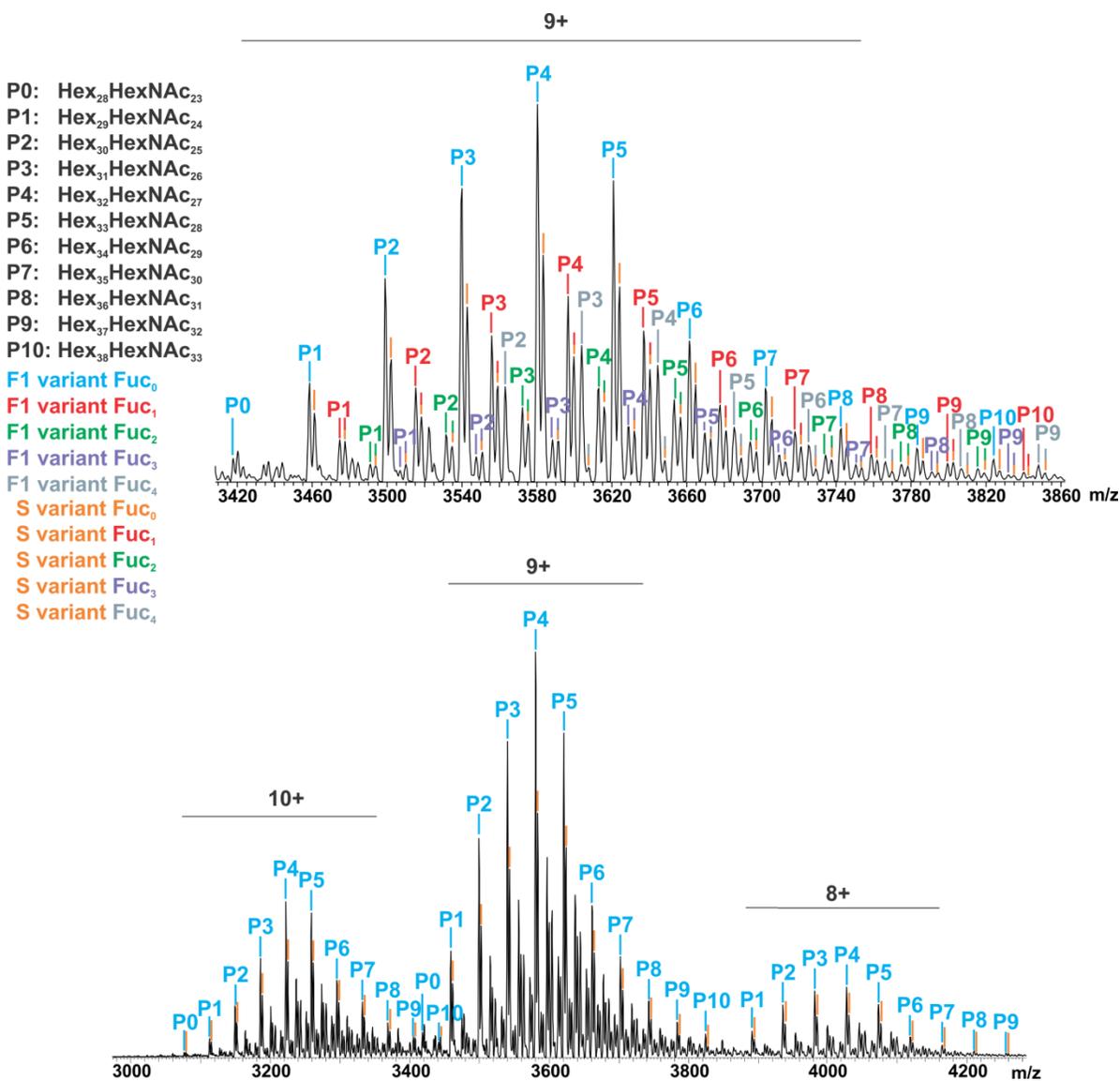
**Figure 3.2** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of asialo-AGP (0.628 mg mL$^{-1}$) and DMSO (1%). Glycoform assignments of asialo-AGP at +9 are shown as an inset, and S variants that share the same compositions with the F variants are labelled beside in orange.

**Table 3.2** Annotation of asialo-AGP glycoproteoforms

| Num. | Genetic Variant | Composition[a] Hex_HexNAc_Fuc | Experimental[b] Mass (Da) | Theoretical[c] Mass (Da) | Mass Difference (Da) | Annotation |
|---|---|---|---|---|---|---|
| 1 | F1 | 28_23_0 | 30753.36 ± 0.01 | 30752.42 | 0.93 | **P0** |
| 2 | S | 28_23_0 | 30774.87 ± 0.01 | 30780.47 | 5.60 | |
| 3 | F1 | 29_24_0 | 31118.05 ± 0.52 | 31117.76 | 0.30 | **P1** |
| 4 | S | 29_24_0 | 31143.20 ± 0.01 | 31145.81 | 2.60 | |
| 5 | F1 | 29_24_1 | 31264.15 ± 0.01 | 31263.90 | 0.25 | **P1** |
| 6 | S | 29_24_1 | 31288.68 ± 0.01 | 31291.95 | 3.27 | |
| 7 | F1 | 29_25_0 | 31323.76 ± 0.01 | 31320.95 | 2.81 | |
| 8 | S | 29_25_0 | 31351.90 ± 0.01 | 31349.00 | 2.90 | |
| 9 | F1 | 29_24_2 | 31411.38 ± 0.52 | 31410.04 | 1.34 | **P1** |
| 10 | S | 29_24_2 | 31435.26 ± 0.04 | 31438.09 | 2.83 | |
| 11 | F1 | 30_25_0 | 31482.41 ± 0.01 | 31483.09 | 0.68 | **P2** |
| 12 | S | 30_25_0 | 31509.53 ± 0.01 | 31511.14 | 1.61 | |
| 13 | F1 | 29_24_3 | 31554.10 ± 0.56 | 31556.18 | 2.08 | **P1** |
| 14 | S | 29_24_3 | 31580.90 ± 0.52 | 31584.23 | 3.33 | |
| 15 | F1 | 30_25_1 | 31629.16 ± 0.01 | 31629.23 | 0.07 | **P2** |
| 16 | S | 30_25_1 | 31656.07 ± 0.53 | 31657.28 | 1.22 | |
| 17 | F1 | 30_26_0 | 31692.53 ± 0.59 | 31686.28 | 6.25 | |
| 18 | S | 30_26_0 | 31717.04 ± 0.02 | 31714.33 | 2.71 | |
| 19 | F1 | 30_25_2 | 31774.93 ± 0.59 | 31775.37 | 0.44 | **P2** |
| 20 | S | 30_25_2 | 31803.27 ± 0.01 | 31803.42 | 0.15 | |
| 21 | F1 | 31_26_0 | 31848.84 ± 0.01 | 31848.42 | 0.42 | **P3** |
| 22 | S | 31_26_0 | 31875.25 ± 0.01 | 31876.47 | 1.22 | |
| 23 | F1 | 30_25_3 | 31918.48 ± 0.01 | 31921.51 | 3.03 | **P2** |

| 24 | S | 30_25_3 | 31947.85 ± 0.55 | 31949.56 | 1.72 | |
|----|----|---------|-----------------|----------|------|------|
| 25 | F1 | 31_26_1 | 31994.51 ± 0.01 | 31994.57 | 0.05 | **P3** |
| 26 | S | 31_26_1 | 32022.30 ± 0.01 | 32022.62 | 0.32 | |
| 27 | F1 | 30_25_4 | 32059.84 ± 0.01 | 32067.66 | 7.81 | **P2** |
| 28 | S | 30_25_4 | 32092.05 ± 2.15 | 32095.71 | 3.65 | |
| 29 | F1 | 31_26_2 | 32139.99 ± 0.01 | 32140.71 | 0.71 | **P3** |
| 30 | S | 31_26_2 | 32168.38 ± 0.55 | 32168.76 | 0.38 | |
| 31 | F1 | 32_27_0 | 32214.32 ± 0.01 | 32213.76 | 0.56 | **P4** |
| 32 | S | 32_27_0 | 32241.24 ± 0.01 | 32241.81 | 0.57 | |
| 33 | F1 | 31_26_3 | 32285.24 ± 0.01 | 32286.85 | 1.61 | **P3** |
| 34 | S | 31_26_3 | 32312.59 ± 0.58 | 32314.90 | 2.30 | |
| 35 | F1 | 32_27_1 | 32359.23 ± 0.55 | 32359.90 | 0.67 | **P4** |
| 36 | S | 32_27_1 | 32387.50 ± 0.55 | 32387.95 | 0.45 | |
| 37 | F1 | 31_26_4 | 32425.31 ± 0.01 | 32432.99 | 7.68 | **P3** |
| 38 | S | 31_26_4 | 32460.14 ± 0.63 | 32461.04 | 0.89 | |
| 39 | F1 | 32_27_2 | 32505.55 ± 0.01 | 32506.04 | 0.49 | **P4** |
| 40 | S | 32_27_2 | 32533.20 ± 0.62 | 32534.09 | 0.89 | |
| 41 | F1 | 33_28_0 | 32578.69 ± 0.01 | 32579.09 | 0.40 | **P5** |
| 42 | S | 33_28_0 | 32607.23 ± 0.01 | 32607.14 | 0.09 | |
| 43 | F1 | 32_27_3 | 32650.42 ± 0.59 | 32652.18 | 1.76 | **P4** |
| 44 | S | 32_27_3 | 32678.20 ± 0.01 | 32680.23 | 2.03 | |
| 45 | F1 | 33_28_1 | 32724.46 ± 0.01 | 32725.23 | 0.77 | **P5** |
| 46 | S | 33_28_1 | 32753.17 ± 0.01 | 32753.28 | 0.11 | |
| 47 | F1 | 32_27_4 | 32792.03 ± 0.01 | 32798.32 | 6.29 | **P4** |
| 48 | S | 32_27_4 | 32825.41 ± 0.63 | 32826.37 | 0.97 | |
| 49 | F1 | 33_28_2 | 32869.87 ± 0.01 | 32871.37 | 1.50 | **P5** |
| 50 | S | 33_28_2 | 32899.24 ± 0.63 | 32899.42 | 0.18 | |

| 51 | F1 | 34_29_0 | 32941.68 ± 0.01 | 32944.42 | 2.74 | **P6** |
|----|----|---------|-----------------|----------|------|--------|
| 52 | S | 34_29_0 | 32972.08 ± 0.01 | 32972.47 | 0.39 | |
| 53 | F1 | 33_28_3 | 33015.84 ± 0.57 | 33017.51 | 1.67 | **P5** |
| 54 | S | 33_28_3 | 33043.80 ± 0.63 | 33045.56 | 1.76 | |
| 55 | F1 | 34_29_1 | 33089.48 ± 0.63 | 33090.56 | 1.09 | **P6** |
| 56 | S | 34_29_1 | 33118.71 ± 0.59 | 33118.61 | 0.10 | |
| 57 | F1 | 33_28_4 | 33158.61 ± 0.01 | 33163.66 | 5.04 | **P5** |
| 58 | S | 33_28_4 | 33191.39 ± 0.62 | 33191.71 | 0.31 | |
| 59 | F1 | 34_29_2 | 33233.92 ± 0.01 | 33236.71 | 2.78 | **P6** |
| 60 | S | 34_29_2 | 33264.71 ± 0.01 | 33264.76 | 0.05 | |
| 61 | F1 | 35_30_0 | 33308.28 ± 0.01 | 33309.76 | 1.48 | **P7** |
| 62 | S | 35_30_0 | 33337.85 ± 0.01 | 33337.81 | 0.04 | |
| 63 | F1 | 34_29_3 | 33378.89 ± 1.09 | 33382.85 | 3.96 | **P6** |
| 64 | S | 34_29_3 | 33409.89 ± 0.03 | 33410.90 | 1.01 | |
| 65 | F1 | 35_30_1 | 33453.77 ± 0.01 | 33455.90 | 2.13 | **P7** |
| 66 | S | 35_30_1 | 33483.93 ± 0.65 | 33483.95 | 0.02 | |
| 67 | F1 | 34_29_4 | 33521.93 ± 0.58 | 33528.99 | 7.06 | **P6** |
| 68 | S | 34_29_4 | 33553.16 ± 0.87 | 33557.04 | 3.88 | |
| 69 | F1 | 35_30_2 | 33601.08 ± 0.65 | 33602.04 | 0.96 | **P7** |
| 70 | S | 35_30_2 | 33629.69 ± 0.65 | 33630.09 | 0.40 | |
| 71 | F1 | 36_31_0 | 33675.76 ± 0.01 | 33675.09 | 0.67 | **P8** |
| 72 | S | 36_31_0 | 33703.22 ± 0.01 | 33703.14 | 0.08 | |
| 73 | F1 | 35_30_3 | 33745.86 ± 0.65 | 33748.18 | 2.32 | **P7** |
| 74 | S | 35_30_3 | 33776.69 ± 0.65 | 33776.23 | 0.46 | |
| 75 | F1 | 36_31_1 | 33821.08 ± 0.01 | 33821.23 | 0.15 | **P8** |
| 76 | S | 36_31_1 | 33849.99 ± 0.01 | 33849.28 | 0.71 | |
| 77 | F1 | 35_30_4 | 33888.13 ± 0.01 | 33894.32 | 6.19 | **P7** |

| 78 | S | 35_30_4 | 33920.65 ± 2.46 | 33922.37 | 1.72 | |
|----|---|---------|-----------------|----------|------|-----|
| 79 | F1 | 36_31_2 | 33964.60 ± 1.02 | 33967.37 | 2.77 | **P8** |
| 80 | S | 36_31_2 | 33993.37 ± 0.63 | 33995.42 | 2.05 | |
| 81 | F1 | 37_32_0 | 34040.51 ± 0.65 | 34040.42 | 0.09 | **P9** |
| 82 | S | 37_32_0 | 34068.86 ± 0.64 | 34068.47 | 0.38 | |
| 83 | F1 | 36_31_3 | 34110.93 ± 1.29 | 34113.51 | 2.59 | **P8** |
| 84 | S | 36_31_3 | 34140.70 ± 0.65 | 34141.56 | 0.86 | |
| 85 | F1 | 37_32_1 | 34188.96 ± 0.01 | 34186.56 | 2.39 | **P9** |
| 86 | S | 37_32_1 | 34216.98 ± 1.09 | 34214.61 | 2.36 | |
| 87 | F1 | 36_31_4 | 34254.57 ± 0.01 | 34259.65 | 5.08 | **P8** |
| 88 | S | 36_31_4 | 34284.10 ± 0.01 | 34287.70 | 3.60 | |
| 89 | F1 | 37_32_2 | 34331.89 ± 0.60 | 34332.71 | 0.81 | **P9** |
| 90 | S | 37_32_2 | 34359.72 ± 0.71 | 34360.76 | 1.03 | |
| 91 | F1 | 38_33_0 | 34405.47 ± 0.68 | 34405.76 | 0.29 | **P10** |
| 92 | S | 38_33_0 | 34432.91 ± 0.67 | 34433.81 | 0.90 | |
| 93 | F1 | 37_32_3 | 34476.53 ± 1.18 | 34478.85 | 2.32 | **P9** |
| 94 | S | 37_32_3 | 34505.72 ± 2.24 | 34506.90 | 1.18 | |
| 95 | F1 | 38_33_1 | 34554.89 ± 2.04 | 34551.90 | 2.99 | **P10** |
| 96 | S | 38_33_1 | 34578.52 ± 0.96 | 34579.95 | 1.43 | |
| 97 | F1 | 37_32_4 | 34621.14 ± 0.01 | 34624.99 | 3.85 | **P9** |
| 98 | S | 37_32_4 | 34653.35 ± 0.99 | 34653.04 | 0.31 | |

a. The compositions of glycoproteoforms correspond to the number of hexose (Hex), *N*-Acetylhexosamine (HexNAc), fucose (Fuc).

b. The experimental masses were calculated based on the average MW of glycoproteoforms at three charge states (+8 to +10), and errors correspond to one standard deviation of three replicate measurements.

c. The theoretical masses were calculated from the mass of peptide backbone of AGP (F1 variant: 21539.06 Da and S variant: 21567.11 Da) and the average masses of monosaccharides (Hex: 162.1406 Da, HexNAc: 203.1925 Da, Fuc: 146.1412 Da).

### 3.3.1.3 Characterization of microheterogeneity of α1-3,4 fucosidase treated asialo-AGP

Glycoform assignments of asialo-AGP can be further simplified by treatment of asialo-AGP with fucosidase. Asialo-AGP was first treated with an α-L-fucosidase from bovine kidney with a broad specificity (which cleaves α1-6 linked Fuc more efficiently than other α-Fuc linkages), but it failed to cleave Fuc either α1-6 linked to the core GlcNAc (core Fuc) or α1-3 linked to the outer GlcNAc (antenna Fuc). This is probably due to the low accessibility of the core Fuc on highly branched *N*-glycans of native asialo-AGP and decreased enzyme efficiency in the hydrolysis of antenna Fuc. Therefore, a fucosidase from *Prunus dulcis* which is specific for the cleavage of α1-3,4 linked Fuc was employed to release the antenna Fuc residues on asialo-AGP. Shown in Figure 3.3 is a representative mass spectrum acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of asialo-AGP treated with α1-3,4 fucosidase (0.36 mg mL$^{-1}$) and DMSO (1%). The mass spectrum shows a similar peak envelope as asialo-AGP, i.e., from *m/z* at 3050 to 4350 at charge states from +8 to +10, with the most abundant glycoform assigned as P4, i.e., Hex$_{32}$HexNAc$_{27}$Fuc$_0$, which is within 1 Da difference from the theoretical MW of the corresponding glycoform (MW of AGP F1 variant peptide backbone: 21539.06 Da; mass of Hex$_{32}$HexNAc$_{27}$Fuc$_0$: 10674.70 Da). As with AGP and asialo-AGP, both F1 and S variants (labelled in orange) were detected with mass differences of around 27.32 ± 3.01 Da. Eleven non-fucosylated main peak series (P0-P10) and fewer glycoproteoforms (59 in total) were identified, ranging from unfucosylated, mono, di- and tetrafucosylated (Table 3.3). Interestingly, only four difucosylated glyproteoforms of AGP were detected and trifucosylated glycoproteoforms

were completely absent, suggesting that Fuc of those glycoproteoforms are mostly α1-3 attached to the GlcNAc on the external branches of *N*-glycans.
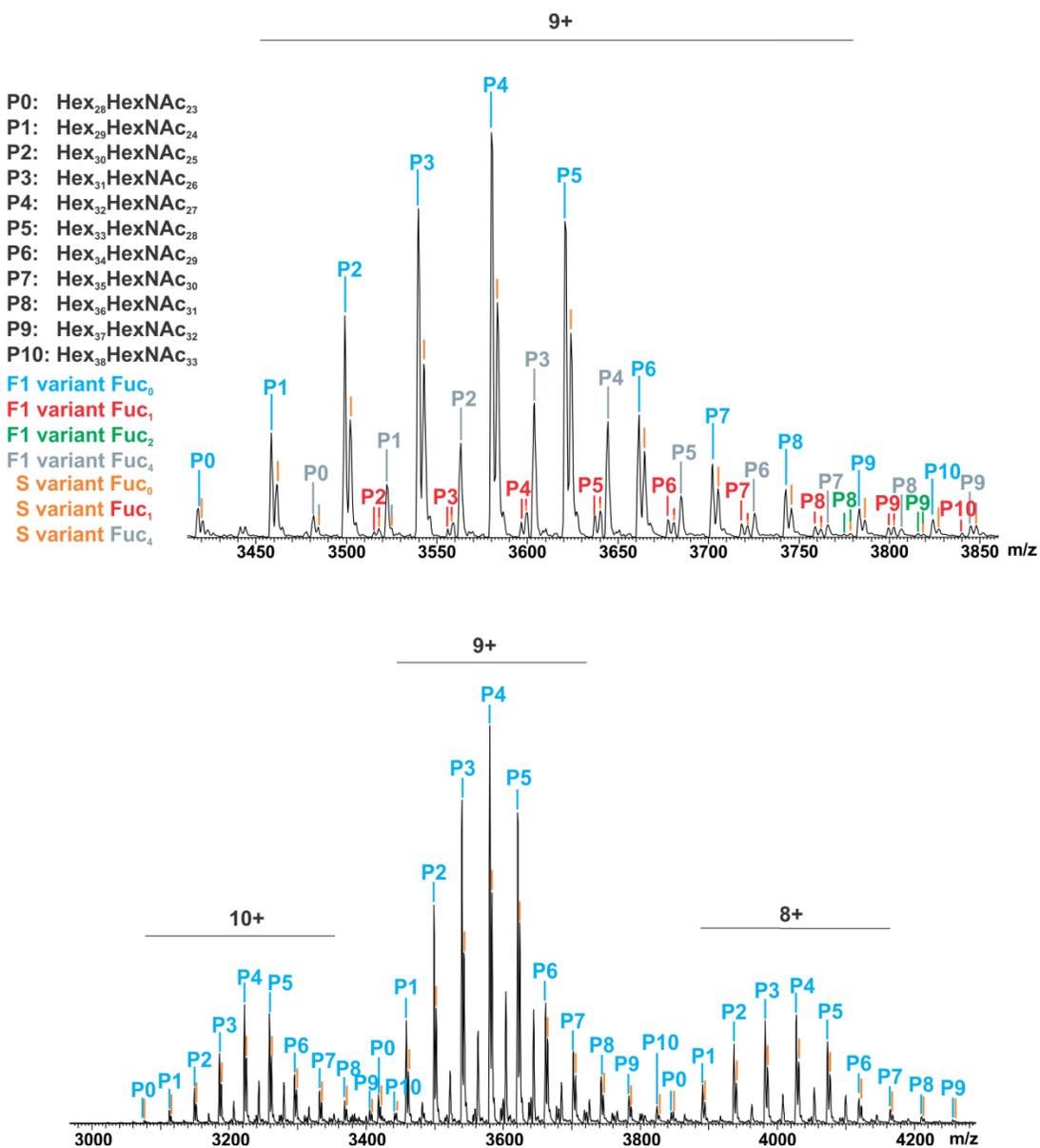


**Figure 3.3** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of asialo-AGP treated with α1-3,4 fucosidase (0.36 mg mL$^{-1}$) and DMSO (1%). Glycoform assignments of asialo-AGP at +9 are shown as an inset, and S variants that share the same compositions with the F variants are labelled beside in orange.

**Table 3.3** Annotation of α1-3,4 fucosidase treated asialo-AGP glycoproteoforms

| Num. | Genetic Variant | Composition[a] Hex_HexNAc_Fuc | Experimental[b] Mass (Da) | Theoretical[c] Mass (Da) | Mass Difference (Da) | Annotation |
|---|---|---|---|---|---|---|
| 1 | F1 | 28_23_0 | 30752.30 ± 0.01 | 30752.42 | 0.12 | **P0** |
| 2 | S | 28_23_0 | 30776.24 ± 0.01 | 30780.47 | 4.24 | |
| 3 | F1 | 29_24_0 | 31118.79 ± 0.01 | 31117.76 | 1.03 | **P1** |
| 4 | S | 29_24_0 | 31143.58 ± 0.54 | 31145.81 | 2.22 | |
| 5 | F1 | 29_24_1 | ND | - | - | |
| 6 | S | 29_24_1 | ND | - | - | |
| 7 | F1 | 29_25_0 | 31327.25 ± 0.01 | 31320.95 | 6.30 | |
| 8 | S | 29_25_0 | 31350.32 ± 0.59 | 31349.00 | 1.32 | |
| 9 | F1 | 29_24_2 | ND | - | - | |
| 10 | S | 29_24_2 | ND | - | - | |
| 11 | F1 | 30_25_0 | 31482.41 ± 0.01 | 31483.09 | 0.68 | **P2** |
| 12 | S | 30_25_0 | 31509.96 ± 0.59 | 31511.14 | 1.18 | |
| 13 | F1 | 29_24_3 | ND | - | - | |
| 14 | S | 29_24_3 | ND | - | - | |
| 15 | F1 | 30_25_1 | 31628.10 ± 0.01 | 31629.23 | 1.14 | **P2** |
| 16 | S | 30_25_1 | 31653.03 ± 0.01 | 31657.28 | 4.26 | |
| 17 | F1 | 30_26_0 | 31693.37 ± 0.01 | 31686.28 | 7.08 | |
| 18 | S | 30_26_0 | 31710.26 ± 0.01 | 31714.33 | 4.08 | |
| 19 | F1 | 30_25_2 | ND | - | - | |
| 20 | S | 30_25_2 | ND | - | - | |
| 21 | F1 | 31_26_0 | 31849.22 ± 0.53 | 31848.42 | 0.80 | **P3** |
| 22 | S | 31_26_0 | 31875.25 ± 0.01 | 31876.47 | 1.22 | |
| 23 | F1 | 30_25_3 | ND | - | - | |
| 24 | S | 30_25_3 | ND | - | - | |

| 25 | F1 | 31_26_1 | 31994.22 ± 0.01 | 31994.57 | 0.35 | **P3** |
|----|----|---------|-----------------|----------|------|--------|
| 26 | S  | 31_26_1 | 32021.49 ± 0.79 | 32022.62 | 1.12 |        |
| 27 | F1 | 30_25_4 | 32059.84 ± 0.01 | 32067.66 | 7.81 | P2     |
| 28 | S  | 30_25_4 | ND              | -        | -    |        |
| 29 | F1 | 31_26_2 | ND              | -        | -    |        |
| 30 | S  | 31_26_2 | ND              | -        | -    |        |
| 31 | F1 | 32_27_0 | 32214.32 ± 0.01 | 32213.76 | 0.56 | P4     |
| 32 | S  | 32_27_0 | 32242.01 ± 0.55 | 32241.81 | 0.21 |        |
| 33 | F1 | 31_26_3 | ND              | -        | -    |        |
| 34 | S  | 31_26_3 | ND              | -        | -    |        |
| 35 | F1 | 32_27_1 | 32359.62 ± 0.55 | 32359.90 | 0.28 | **P4** |
| 36 | S  | 32_27_1 | 32388.27 ± 0.01 | 32387.95 | 0.33 |        |
| 37 | F1 | 31_26_4 | 32425.31 ± 0.01 | 32432.99 | 7.68 | P3     |
| 38 | S  | 31_26_4 | ND              | -        | -    |        |
| 39 | F1 | 32_27_2 | ND              | -        | -    |        |
| 40 | S  | 32_27_2 | ND              | -        | -    |        |
| 41 | F1 | 33_28_0 | 32578.70 ± 0.01 | 32579.09 | 0.39 | P5     |
| 42 | S  | 33_28_0 | 32607.23 ± 0.01 | 32607.14 | 0.09 |        |
| 43 | F1 | 32_27_3 | ND              | -        | -    |        |
| 44 | S  | 32_27_3 | ND              | -        | -    |        |
| 45 | F1 | 33_28_1 | 32725.30 ± 0.59 | 32725.23 | 0.07 | **P5** |
| 46 | S  | 33_28_1 | 32753.17 ± 0.01 | 32753.28 | 0.11 |        |
| 47 | F1 | 32_27_4 | 32792.03 ± 0.01 | 32798.32 | 6.29 | P4     |
| 48 | S  | 32_27_4 | 32832.05 ± 1.89 | 32826.37 | 5.68 |        |
| 49 | F1 | 33_28_2 | ND              | -        | -    |        |
| 50 | S  | 33_28_2 | ND              | -        | -    |        |
| 51 | F1 | 34_29_0 | 32945.50 ± 0.01 | 32944.42 | 1.08 | P6     |

| 52 | S | 34_29_0 | 32972.48 ± 0.57 | 32972.47 | 0.01 | |
|---|---|---|---|---|---|---|
| 53 | F1 | 33_28_3 | ND | - | - | |
| 54 | S | 33_28_3 | ND | - | - | |
| 55 | F1 | 34_29_1 | 33089.92 ± 0.01 | 33090.56 | 0.64 | **P6** |
| 56 | S | 34_29_1 | 33119.13 ± 0.01 | 33118.61 | 0.52 | |
| 57 | F1 | 33_28_4 | 33157.41 ± 0.01 | 33163.66 | 6.24 | **P5** |
| 58 | S | 33_28_4 | 33191.04 ± 0.96 | 33191.71 | 0.67 | |
| 59 | F1 | 34_29_2 | ND | - | - | |
| 60 | S | 34_29_2 | ND | - | - | |
| 61 | F1 | 35_30_0 | 33310.80 ± 0.01 | 33309.76 | 1.04 | **P7** |
| 62 | S | 35_30_0 | 33337.85 ± 0.01 | 33337.81 | 0.04 | |
| 63 | F1 | 34_29_3 | ND | - | - | |
| 64 | S | 34_29_3 | ND | - | - | |
| 65 | F1 | 35_30_1 | 33455.93 ± 0.57 | 33455.90 | 0.03 | **P7** |
| 66 | S | 35_30_1 | 33484.85 ± 0.01 | 33483.95 | 0.90 | |
| 67 | F1 | 34_29_4 | 33521.10 ± 0.01 | 33528.99 | 7.88 | **P6** |
| 68 | S | 34_29_4 | ND | - | - | |
| 69 | F1 | 35_30_2 | ND | - | - | |
| 70 | S | 35_30_2 | ND | - | - | |
| 71 | F1 | 36_31_0 | 33675.76 ± 0.01 | 33675.09 | 0.67 | **P8** |
| 72 | S | 36_31_0 | 33704.46 ± 0.01 | 33703.14 | 1.32 | |
| 73 | F1 | 35_30_3 | ND | - | - | |
| 74 | S | 35_30_3 | ND | - | - | |
| 75 | F1 | 36_31_1 | 33821.08 ± 0.01 | 33821.23 | 0.15 | **P8** |
| 76 | S | 36_31_1 | 33849.99 ± 0.01 | 33849.28 | 0.71 | |
| 77 | F1 | 35_30_4 | 33888.13 ± 0.01 | 33894.32 | 6.19 | **P7** |
| 78 | S | 35_30_4 | 33919.01 ± 0.01 | 33922.37 | 3.36 | |

| 79 | F1 | 36_31_2 | 33964.92 ± 1.87 | 33967.37 | 2.46 | **P8** |
|----|----|---------|-----------------|----------|------|--------|
| 80 | S | 36_31_2 | 33993.52 ± 0.93 | 33995.42 | 1.90 | |
| 81 | F1 | 37_32_0 | 34041.37 ± 0.01 | 34040.42 | 0.94 | **P9** |
| 82 | S | 37_32_0 | 34069.31 ± 0.01 | 34068.47 | 0.83 | |
| 83 | F1 | 36_31_3 | ND | - | - | |
| 84 | S | 36_31_3 | ND | - | - | |
| 85 | F1 | 37_32_1 | 34186.16 ± 1.00 | 34186.56 | 0.40 | **P9** |
| 86 | S | 37_32_1 | 34213.63 ± 0.01 | 34214.61 | 0.99 | |
| 87 | F1 | 36_31_4 | 34254.48 ± 0.01 | 34259.65 | 5.18 | **P8** |
| 88 | S | 36_31_4 | ND | - | - | |
| 89 | F1 | 37_32_2 | 34333.53 ± 1.87 | 34332.71 | 0.82 | **P9** |
| 90 | S | 37_32_2 | 34363.11 ± 0.01 | 34360.76 | 2.35 | |
| 91 | F1 | 38_33_0 | 34405.47 ± 0.68 | 34405.76 | 0.29 | **P10** |
| 92 | S | 38_33_0 | 34433.38 ± 0.01 | 34433.81 | 0.43 | |
| 93 | F1 | 37_32_3 | ND | - | - | |
| 94 | S | 37_32_3 | ND | - | - | |
| 95 | F1 | 38_33_1 | 34552.01 ± 0.01 | 34551.90 | 0.11 | **P10** |
| 96 | S | 38_33_1 | ND | - | - | |
| 97 | F1 | 37_32_4 | 34624.02 ± 2.04 | 34624.99 | 0.97 | **P9** |
| 98 | S | 37_32_4 | 34655.07 ± 0.01 | 34653.04 | 2.03 | |

a. The compositions of glycoproteoforms correspond to the number of hexose (Hex/H), *N*-Acetylhexosamine (HexNAc), fucose (Fuc).

b. The experimental masses were calculated based on the average MW of glycoproteoforms at three charge states (+8 to +10), and errors correspond to one standard deviation of three replicate measurements. ND ≡ Not Detected.

c. The theoretical masses were calculated from the mass of peptide backbone of AGP (F1 variant: 21539.06 Da and S variant: 21567.11 Da) and the average masses of monosaccharides (Hex: 162.1406 Da, HexNAc: 203.1925 Da, Fuc: 146.1412 Da).

**Table 3.4** Fractions of glycoproteoforms of asialo AGP and α1-3,4 fucosidase treated asialo-AGP (asialo-AGP-Fuc)

| Num. | Genetic Variant | Composition Hex_HexNAc_Fuc | Fraction (%) asialo-AGP | Fraction (%) asialo-AGP-Fuc | Annotation |
|---|---|---|---|---|---|
| 1 | F1 | 28_23_0 | 0.30 | 0.80 | P0 |
| 2 | S | 28_23_0 | 0.37 | 0.56 | |
| 3 | F1 | 29_24_0 | 1.48 | 2.87 | P1 |
| 4 | S | 29_24_0 | 1.05 | 1.50 | |
| 5 | F1 | 29_24_1 | 0.67 | 0.00 | P1 |
| 6 | S | 29_24_1 | 0.62 | 0.00 | |
| 7 | F1 | 29_25_0 | 0.38 | 0.59 | |
| 8 | S | 29_25_0 | 0.31 | 0.28 | |
| 9 | F1 | 29_24_2 | 0.30 | 0.00 | P1 |
| 10 | S | 29_24_2 | 0.27 | 0.00 | |
| 11 | F1 | 30_25_0 | 3.12 | 6.22 | P2 |
| 12 | S | 30_25_0 | 1.94 | 3.28 | |
| 13 | F1 | 29_24_3 | 0.21 | 0.00 | P1 |
| 14 | S | 29_24_3 | 0.30 | 0.00 | |
| 15 | F1 | 30_25_1 | 1.48 | 0.11 | P2 |
| 16 | S | 30_25_1 | 1.01 | 0.16 | |
| 17 | F1 | 30_26_0 | 0.86 | 1.51 | |
| 18 | S | 30_26_0 | 0.25 | 0.26 | |
| 19 | F1 | 30_25_2 | 0.75 | 0.00 | P2 |

| 20 | S | 30_25_2 | 0.57 | 0.00 | |
|----|----|---------|------|------|------|
| 21 | F1 | 31_26_0 | 4.66 | 9.28 | **P3** |
| 22 | S | 31_26_0 | 2.82 | 4.97 | |
| 23 | F1 | 30_25_3 | 0.42 | 0.00 | **P2** |
| 24 | S | 30_25_3 | 0.49 | 0.00 | |
| 25 | F1 | 31_26_1 | 2.33 | 0.16 | **P3** |
| 26 | S | 31_26_1 | 1.55 | 0.36 | |
| 27 | F1 | 30_25_4 | 1.51 | 2.69 | **P2** |
| 28 | S | 30_25_4 | 0.10 | 0.00 | |
| 29 | F1 | 31_26_2 | 1.19 | 0.00 | **P3** |
| 30 | S | 31_26_2 | 0.94 | 0.00 | |
| 31 | F1 | 32_27_0 | 6.12 | 11.68 | **P4** |
| 32 | S | 32_27_0 | 3.70 | 6.69 | |
| 33 | F1 | 31_26_3 | 0.70 | 0.00 | **P3** |
| 34 | S | 31_26_3 | 0.67 | 0.00 | |
| 35 | F1 | 32_27_1 | 3.02 | 0.45 | **P4** |
| 36 | S | 32_27_1 | 2.07 | 0.73 | |
| 37 | F1 | 31_26_4 | 2.23 | 3.93 | **P3** |
| 38 | S | 31_26_4 | 0.29 | 0.00 | |
| 39 | F1 | 32_27_2 | 1.59 | 0.00 | **P4** |
| 40 | S | 32_27_2 | 1.25 | 0.00 | |
| 41 | F1 | 33_28_0 | 5.05 | 9.40 | **P5** |
| 42 | S | 33_28_0 | 3.28 | 5.94 | |
| 43 | F1 | 32_27_3 | 0.95 | 0.00 | **P4** |
| 44 | S | 32_27_3 | 0.86 | 0.00 | |
| 45 | F1 | 33_28_1 | 2.54 | 0.64 | **P5** |
| 46 | S | 33_28_1 | 1.89 | 0.75 | |

| 47 | F1 | 32_27_4 | 1.94 | 3.42 | P4 |
|----|----|---------|------|------|-----|
| 48 | S  | 32_27_4 | 0.40 | 0.05 |    |
| 49 | F1 | 33_28_2 | 1.39 | 0.00 | P5 |
| 50 | S  | 33_28_2 | 1.11 | 0.00 |    |
| 51 | F1 | 34_29_0 | 2.42 | 3.63 | P6 |
| 52 | S  | 34_29_0 | 1.69 | 2.55 |    |
| 53 | F1 | 33_28_3 | 0.85 | 0.00 | P5 |
| 54 | S  | 33_28_3 | 0.74 | 0.00 |    |
| 55 | F1 | 34_29_1 | 1.31 | 0.54 | P6 |
| 56 | S  | 34_29_1 | 0.91 | 0.44 |    |
| 57 | F1 | 33_28_4 | 0.93 | 1.23 | P5 |
| 58 | S  | 33_28_4 | 0.42 | 0.11 |    |
| 59 | F1 | 34_29_2 | 0.72 | 0.00 | P6 |
| 60 | S  | 34_29_2 | 0.56 | 0.00 |    |
| 61 | F1 | 35_30_0 | 1.63 | 2.22 | P7 |
| 62 | S  | 35_30_0 | 1.13 | 1.44 |    |
| 63 | F1 | 34_29_3 | 0.57 | 0.00 | P6 |
| 64 | S  | 34_29_3 | 0.40 | 0.00 |    |
| 65 | F1 | 35_30_1 | 0.88 | 0.41 | P7 |
| 66 | S  | 35_30_1 | 0.62 | 0.34 |    |
| 67 | F1 | 34_29_4 | 0.62 | 0.71 | P6 |
| 68 | S  | 34_29_4 | 0.20 | 0.00 |    |
| 69 | F1 | 35_30_2 | 0.46 | 0.00 | P7 |
| 70 | S  | 35_30_2 | 0.40 | 0.00 |    |
| 71 | F1 | 36_31_0 | 0.95 | 1.43 | P8 |
| 72 | S  | 36_31_0 | 0.68 | 0.87 |    |
| 73 | F1 | 35_30_3 | 0.30 | 0.00 | P7 |

| 74 | S | 35_30_3 | 0.31 | 0.00 | |
|---|---|---|---|---|---|
| 75 | F1 | 36_31_1 | 0.61 | 0.40 | **P8** |
| 76 | S | 36_31_1 | 0.40 | 0.27 | |
| 77 | F1 | 35_30_4 | 0.34 | 0.39 | **P7** |
| 78 | S | 35_30_4 | 0.19 | 0.11 | |
| 79 | F1 | 36_31_2 | 0.28 | 0.05 | **P8** |
| 80 | S | 36_31_2 | 0.33 | 0.20 | |
| 81 | F1 | 37_32_0 | 0.59 | 0.85 | **P9** |
| 82 | S | 37_32_0 | 0.44 | 0.54 | |
| 83 | F1 | 36_31_3 | 0.23 | 0.00 | **P8** |
| 84 | S | 36_31_3 | 0.17 | 0.00 | |
| 85 | F1 | 37_32_1 | 0.54 | 0.20 | **P9** |
| 86 | S | 37_32_1 | 0.38 | 0.20 | |
| 87 | F1 | 36_31_4 | 0.24 | 0.15 | **P8** |
| 88 | S | 36_31_4 | 0.14 | 0.00 | |
| 89 | F1 | 37_32_2 | 0.33 | 0.06 | **P9** |
| 90 | S | 37_32_2 | 0.33 | 0.05 | |
| 91 | F1 | 38_33_0 | 0.41 | 0.55 | **P10** |
| 92 | S | 38_33_0 | 0.34 | 0.38 | |
| 93 | F1 | 37_32_3 | 0.17 | 0.00 | **P9** |
| 94 | S | 37_32_3 | 0.15 | 0.00 | |
| 95 | F1 | 38_33_1 | 0.10 | 0.08 | **P10** |
| 96 | S | 38_33_1 | 0.05 | 0.00 | |
| 97 | F1 | 37_32_4 | 0.17 | 0.23 | **P9** |
| 98 | S | 37_32_4 | 0.09 | 0.08 | |
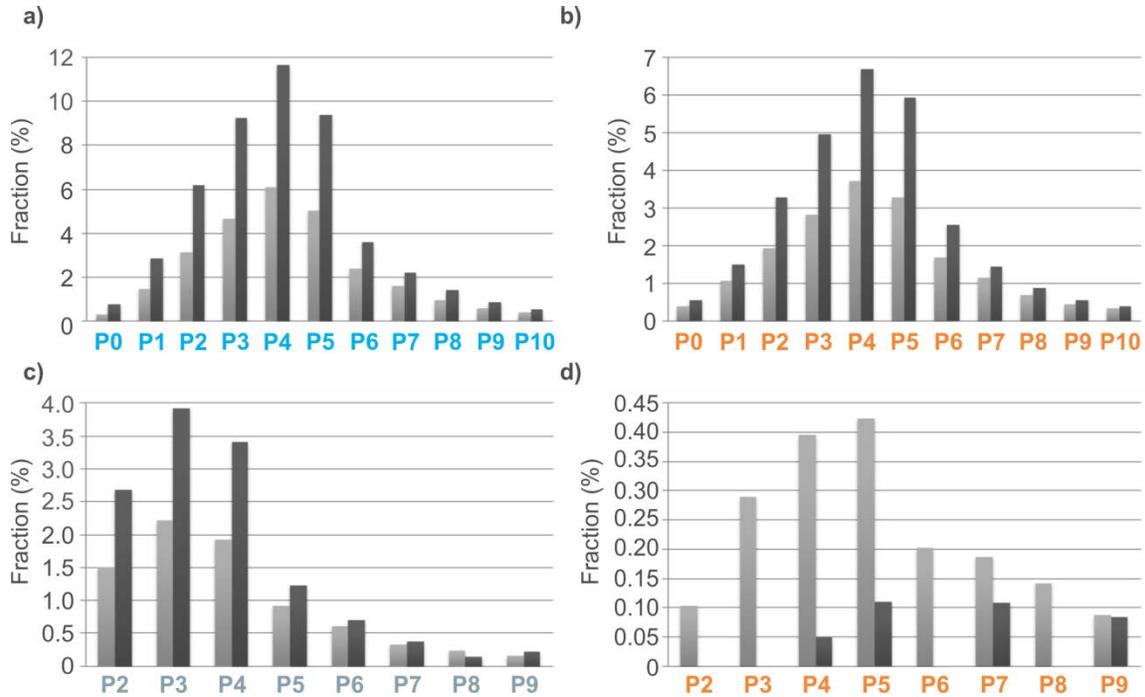
**Figure 3.4** Fractions of non-glycosylated ((a), (b)) and tetraglycosylated ((c), (d)) glycoproteoforms of asialo-AGP (light grey) and α1-3,4 fucosidase treated asialo-AGP (dark grey) F1 ((a), (c)) and S variants ((b), (d)) to the total abundance of glycoforms.

Shown in Table 3.4 and Figure 3.4 are fractions of detected glycoproteoforms of asialo-AGP and α1-3,4 fucosidase treated asialo-AGP to the total abundance of glycoforms based on three charge states (+8 to +10). Fractions of unfucosylated glycoforms of AGP F1 and S variants significantly increased after treatment with α1-3,4 fucosidase as a result of the contribution from difucosylated and trifucosylated glycoproteoforms. The increase in the fractions of tetrafucosylated glycoforms of AGP F1 variants might come from the loss of one α1-3 Fuc residue from hexafucosylated glycoforms, which are not included in the current study due to the overlapping with other glycoforms. Interestingly, fractions of tetrafucosylated glycoforms of AGP S variants significantly decreased after treatment with α1-3,4 fucosidase, indicating these glycoforms are predominantly α1-3 fucosylated.

**3.3.1.4 Characterization of PNGase F treated AGP**

*N*-glycans can be completely removed from glycoproteins by PNGase F, an amidase which cleaves the linkage between asparagine and GlcNAc from high mannose, hybrid, and complex type *N*-glycans. However, under non-denaturing condition where the disulfide bridges of the polypeptide are maintained, fully deglycosylation of AGP was only partially achieved by exposure to PNGase F, which indicates GlcNAc residues at certain glycosylation sites are probably not easily accessible to the enzyme due to the high degree of branching/fucosylation and the steric hindrance introduced by adjacent loops. As shown in Figure 3.5 and Table 3.5, a mixture of completely deglycosylated and partially deglycosylated AGP where glycans are attached to one to three of the five glycosylation sites (in total 57 glycoforms of F1 variants) were identified. The average masses of non-glycosylated AGP are $21543.28 \pm 0.01$ Da (F1 variant) and $21571.01 \pm 0.01$ Da (S variant), which is within 1 Da mass differences of the theoretical masses (21539.06 Da for F1 variant and 21567.11 Da for S variant) after subtraction of 5.16 Da as a result of the PNGase F hydrolysis of five asparagine residues into aspartic acids. Due to the multiple overlapping of glycoforms of F and S variants, annotations of S variants were not included in this study. Relative abundances of each individual glycoform might not represent their real abundances in solution as a result of the similar masses of three Fuc and Sia residues (729.68 Da) and two Hex and HexNAc residues (730.67 Da).
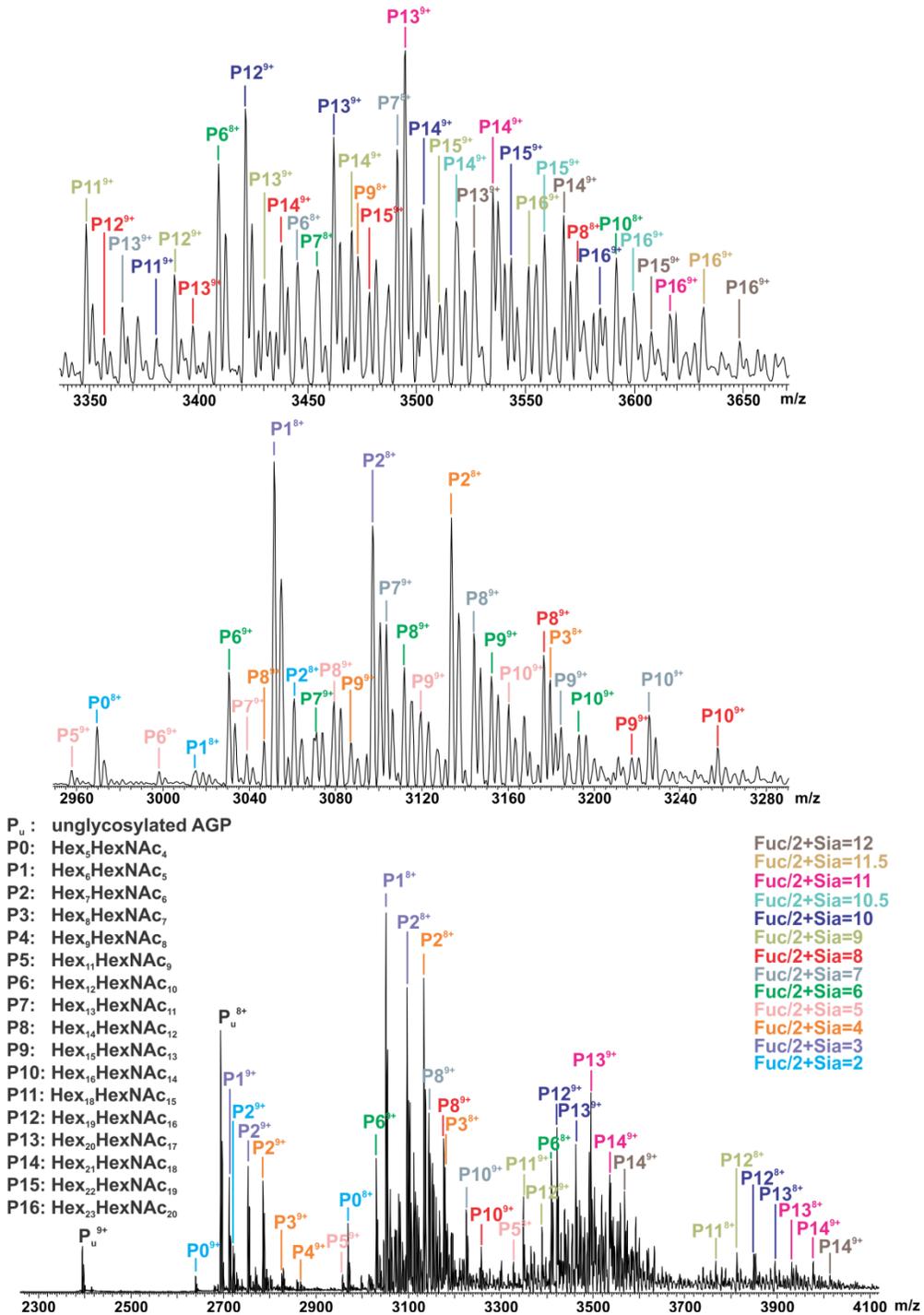
**Figure 3.5** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of PNGase F treated AGP (0.253 mg mL$^{-1}$) and DMSO (1%). Glycoform assignments of PNGase F treated AGP (F1 variant) at $m/z$ from approximately 2960 to 3280 and 3350 to 3650 are shown as insets.

**Table 3.5** Annotation of PNGase F treated AGP F1 variant glycoforms

| Num. | Composition[a] Hex_HexNAc_Fuc/2+Sia | Experimental[b] Mass (Da) | Theoretical[c] Mass (Da) | Mass Difference (Da) | Annotation |
|---|---|---|---|---|---|
| 1 | 5_4_2 | 23747.53 ± 0.01 | 23745.04 | 2.49 | P0 |
| 2 | 6_5_2 | 24112.83 ± 0.01 | 24110.38 | 2.46 | P1 |
| 3 | 6_5_3 | 24404.48 ± 0.01 | 24401.63 | 2.85 | P1 |
| 4 | 7_6_2 | 24478.35 ± 0.57 | 24475.71 | 2.64 | P2 |
| 5 | 7_6_3 | 24769.85 ± 0.01 | 24766.96 | 2.88 | P2 |
| 6 | 7_6_4 | 25061.90 ± 0.01 | 25058.22 | 3.68 | P2 |
| 7 | 8_7_4 | 25427.73 ± 0.01 | 25423.55 | 4.17 | P3 |
| 8 | 9_8_4 | 25792.18 ± 0.01 | 25788.88 | 3.29 | P4 |
| 9 | 11_9_5 | 26610.65 ± 0.66 | 26608.64 | 2.01 | P5 |
| 10 | 12_10_5 | 26972.35 ± 0.01 | 26973.97 | 1.63 | P6 |
| 11 | 12_10_6 | 27266.24 ± 0.01 | 27266.26 | 0.02 | P6 |
| 12 | 13_11_5 | 27340.14 ± 0.01 | 27338.28 | 1.86 | P7 |
| 13 | 14_12_4 | 27411.98 ± 0.01 | 27412.36 | 0.38 | P8 |
| 14 | 13_11_6 | 27630.54 ± 0.01 | 27631.59 | 1.05 | P7 |
| 15 | 14_12_5 | 27704.10 ± 0.01 | 27703.61 | 0.49 | P8 |
| 16 | 15_13_4 | 27774.30 ± 0.01 | 27777.69 | 3.39 | P9 |
| 17 | 13_11_7 | 27923.01 ± 0.01 | 27921.82 | 1.20 | P7 |
| 18 | 14_12_6 | 27997.64 ± 0.01 | 27995.89 | 1.75 | P8 |
| 19 | 15_13_5 | 28066.53 ± 0.01 | 28068.94 | 2.42 | P9 |
| 20 | 14_12_7 | 28288.02 ± 0.01 | 28287.15 | 0.87 | P8 |
| 21 | 15_13_6 | 28361.24 ± 0.01 | 28361.23 | 0.01 | P9 |
| 22 | 16_14_5 | 28431.28 ± 0.01 | 28434.28 | 3.00 | P10 |
| 23 | 14_12_8 | 28580.32 ± 0.01 | 28579.43 | 0.89 | P8 |

| 24 | 15_13_7 | 28652.70 ± 0.01 | 28652.48 | 0.22 | P9 |
|---|---|---|---|---|---|
| 25 | 16_14_6 | 28725.54 ± 0.01 | 28727.59 | 2.05 | P10 |
| 26 | 15_13_8 | 28945.65 ± 0.01 | 28944.76 | 0.88 | P9 |
| 27 | 16_14_7 | 29016.32 ± 0.01 | 29018.84 | 2.53 | P10 |
| 28 | 16_14_8 | 29309.98 ± 0.01 | 29310.10 | 0.12 | P10 |
| 29 | 18_15_8 | 29835.98 ± 0.76 | 29835.52 | 0.46 | P11 |
| 30 | 18_15_9 | 30128.52 ± 0.81 | 30128.83 | 0.30 | P11 |
| 31 | 19_16_8 | 30203.37 ± 1.57 | 30201.88 | 1.49 | P12 |
| 32 | 20_17_7 | 30276.08 ± 0.01 | 30274.93 | 1.15 | P13 |
| 33 | 18_15_10 | 30419.05 ± 0.01 | 30418.02 | 1.02 | P11 |
| 34 | 19_16_9 | 30491.72 ± 0.01 | 30493.13 | 1.42 | P12 |
| 35 | 20_17_8 | 30566.46 ± 0.01 | 30566.18 | 0.28 | P13 |
| 36 | 19_16_10 | 30783.93 ± 0.85 | 30785.41 | 1.49 | P12 |
| 37 | 20_17_9 | 30860.82 ± 0.01 | 30858.46 | 2.36 | P13 |
| 38 | 21_18_8 | 30933.35 ± 0.01 | 30931.51 | 1.84 | P14 |
| 39 | 20_17_10 | 31149.52 ± 0.87 | 31149.72 | 0.19 | P13 |
| 40 | 21_18_9 | 31226.05 ± 0.01 | 31222.77 | 3.28 | P14 |
| 41 | 22_19_8 | 31296.26 ± 0.01 | 31296.85 | 0.59 | P15 |
| 42 | 20_17_11 | 31442.78 ± 0.01 | 31442.00 | 0.77 | P13 |
| 43 | 21_18_10 | 31517.38 ± 0.01 | 31515.05 | 2.33 | P14 |
| 44 | 22_19_9 | 31586.81 ± 0.01 | 31588.10 | 1.30 | P15 |
| 45 | 21_18_10.5 | 31658.20 ± 0.01 | 31660.16 | 1.97 | P14 |
| 46 | 20_17_12 | 31731.42 ± 0.01 | 31732.23 | 0.81 | P13 |
| 47 | 21_18_11 | 31807.26 ± 0.01 | 31807.33 | 0.07 | P14 |
| 48 | 22_19_10 | 31880.62 ± 0.91 | 31880.38 | 0.23 | P15 |
| 49 | 23_20_9 | 31953.07 ± 0.91 | 31953.44 | 0.37 | P16 |
| 50 | 22_19_10.5 | 32019.51 ± 0.01 | 32025.50 | 5.99 | P15 |

| 51 | 21_18_12 | 32099.33 ± 0.01 | 32097.56 | 1.76 | P14 |
|----|----------|------------------|----------|------|-----|
| 52 | 23_20_10 | 32246.47 ± 0.89 | 32246.75 | 0.28 | P16 |
| 53 | 23_20_10.5 | 32389.04 ± 0.91 | 32390.83 | 1.79 | P16 |
| 54 | 22_19_12 | 32458.75 ± 0.92 | 32461.87 | 3.11 | P15 |
| 55 | 23_20_11 | 32536.35 ± 0.01 | 32535.94 | 0.40 | P16 |
| 56 | 23_20_11.5 | 32678.40 ± 0.01 | 32682.09 | 3.69 | P16 |
| 57 | 23_20_12 | 32828.83 ± 0.01 | 32827.20 | 1.63 | P16 |

a. The compositions of glycoforms correspond to the number of hexose (Hex), *N*-Acetylhexosamine (HexNAc), and the sum of half of the number of fucose (Fuc) and the number of sialic acid (Sia) residues.

b. The experimental masses were calculated based on the average MW of glycoforms at three charge states (+8 to +9), and errors correspond to one standard deviation of three replicate measurements.

c. The theoretical masses were calculated from the mass of peptide backbone of AGP (F1 variant: 21539.06 Da) and the average masses of monosaccharides (Hex: 162.1406 Da, HexNAc: 203.1925 Da, Fuc: 146.1412 Da, and Sia: 291.2546 Da).

### 3.3.2 Influence of *N*-glycosylation on AGP-warfarin Interaction

Shown in Figure 3.6b is the structure of AGP with warfarin docking to the drug binding site. A previous study revealed that the ligand binding pocket of AGP is composed of three distinct lobes (lobe I, II and III, shown in Figure 3.6a).[38] Lobe I is the deepest and largest of the three lobes, and its non-polar character enables the binding of hydrophobic molecules. Lobes II and III are smaller sized and negatively charged and located on each side of lobe I. Notably, of the four loops (which connect the 8 β-strands) forming the entrance to the binding pocket, loop I contains one of the glycosylation sites (Asn 38) within an α-helix and shows high structural flexibility.[38] Together with the other glycosylation site (Asn 75) that is in proximity to the entrance to the binding pocket,

these features explain the diverse properties of ligand binding to AGP and suggest the possibility of altered binding activity due to the glycosylation changes.[29,38,39] Here, the binding of warfarin to AGP at different degrees of glycosylation was studied to investigate the influence of *N*-glycans on AGP-drug interactions.
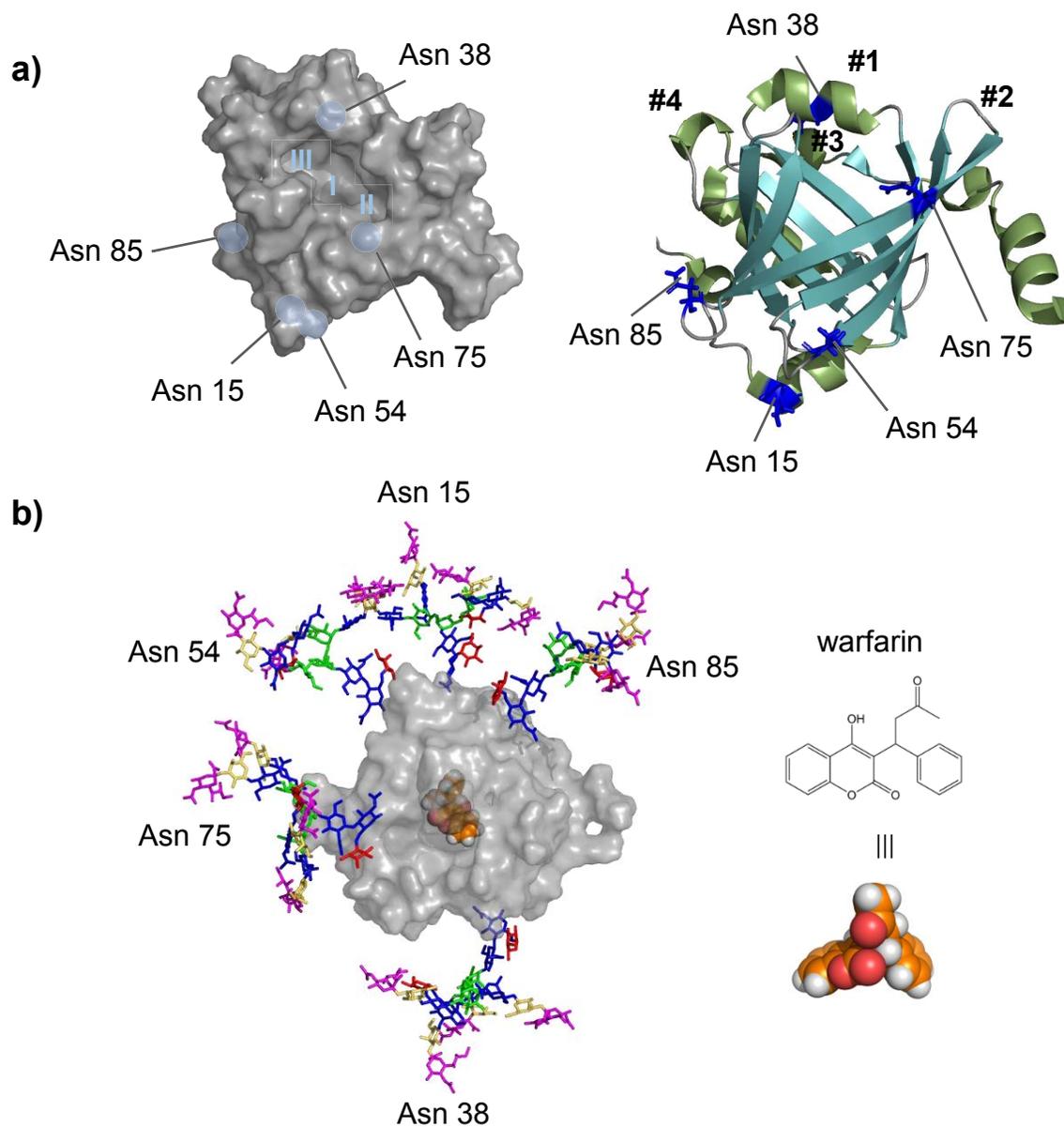


**Figure 3.6** a) AGP structure and the ligand-binding site with its three lobes I, II, and III, and four loops connecting the β-strands (light blue), adapted from Ref 38. b) AGP (blue

sticks: GlcNAc, red sticks: Fuc, green sticks: Man, pink sticks: Sia) with warfarin (orange spheres) docking to the ligand binding pocket.

ITC is considered as the gold standard for measuring thermodynamics of protein-ligand interactions. Shown in Figure 3.7 are the raw and integrated ITC data measured for warfarin binding to native AGP in aqueous ammonium acetate (200 mM, pH 6.8) at 25 °C. The affinity of warfarin to native AGP is $(2.5 \pm 0.1) \times 10^5$ M$^{-1}$ according to the best fit of a 1:1 binding model, which is in good agreement with the previously reported values ($2.3 \times 10^5$ M$^{-1}$ and $1.2 \times 10^5$ M$^{-1}$), [26,27] but is ten times smaller than the previously reported average $K_a$ value ($2.8 \times 10^6$ M$^{-1}$) of asialo-AGP binding to warfarin.[25] $\Delta H$ and $\Delta S$ are $-1.3 \times 10^4$ cal/mol and $-16.8$ cal/mol/K, respectively. However, ITC does not provide any insights into the interactions of ligand to each glycoform.

Quantifying interactions of ligands to individual glycoforms of glycoproteins is challenging due to the unknown concentrations of each glycoform and potential overlapping of the complexes with the unbound species. In the case of intact AGP binding to warfarin, complexes of major glycoforms of AGP F1 variants overlapped with free AGP S variants (Figure 3.8), and thus the abundances of complexes in mass spectra cannot represent their real abundances in solution.

Peak overlapping of AGP-warfarin complexes with free AGP was substantially reduced by desialylation of AGP. To test the affinities of warfarin for glycoforms of asialo-AGP measured by ESI-MS titration experiments, warfarin was titrated into a solution containing asialo-AGP (5.6 μM) and DMSO (1%). Shown in Figure 3.9 is a representative ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, pH 7.6, 25 °C) of asialo-AGP ($[P_T]_0$=0.18 mg mL$^{-1}$, ~5.6 μM), warfarin (1.5 μM) and DMSO (1%). Ion signal corresponding to both free and warfarin-bound AGP was detected.
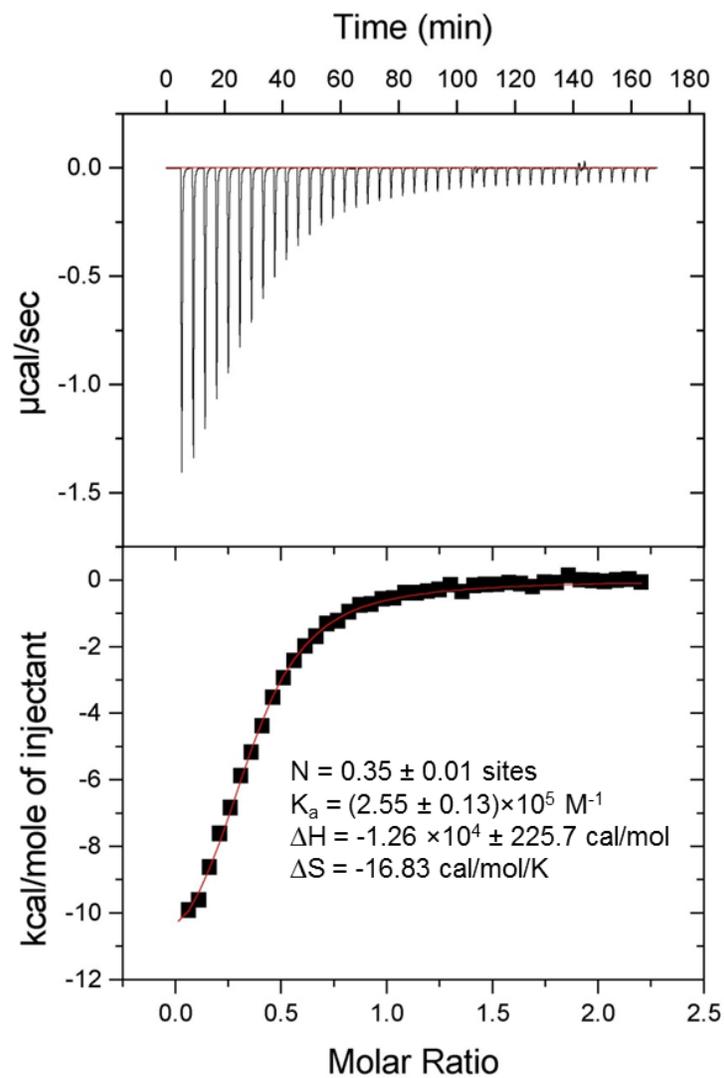
**Figure 3.7** ITC data measured for the binding of AGP (50 μM) to warfarin (500 μM) in aqueous ammonium acetate (200 mM, pH 6.8 ) at 25 °C.
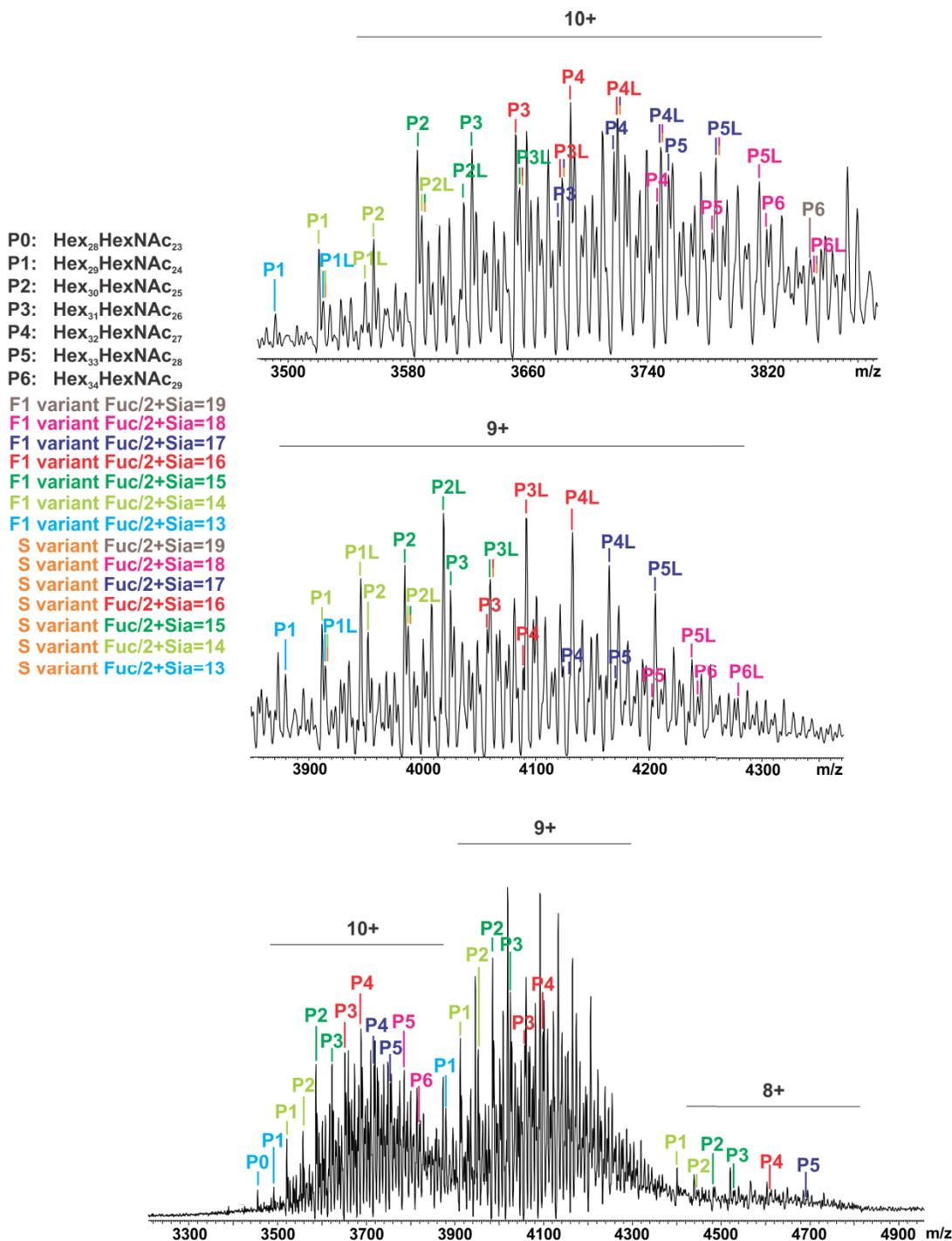
**Figure 3.8** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 7.6 and 25 °C) of AGP (0.48 mg mL-1), warfarin (L, 3.0 μM) and DMSO (1%). Glycoform assignments of AGP and its complexes with warfarin at +9 and +10 are shown as insets. Peak overlapping is shown as two vertical lines.
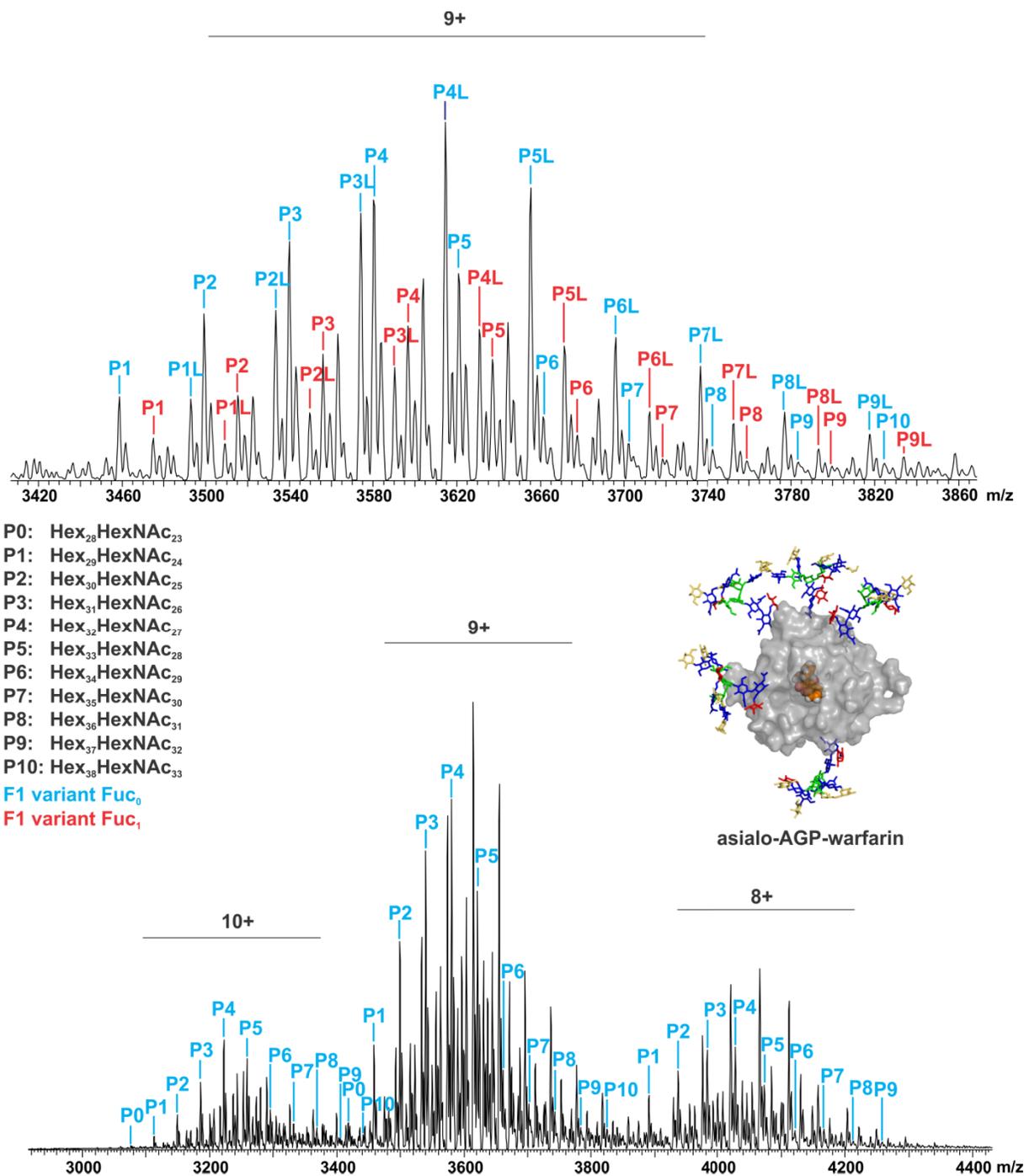
**Figure 3.9** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 7.6 and 25 °C) of asialo-AGP (0.18 mg mL$^{-1}$), warfarin (L, 1.5 μM) and DMSO (1%). Glycoform assignments of asialo-AGP its complexes with warfarin at +9 are shown as an inset.

Due to the low abundance of some glycoforms and the multiple overlapping between free AGP glycoproteoforms and their complexes, warfarin affinities for unfucosylated glycoforms (P1 to P8) were determined by direct ESI-MS assay. Initial concentration of each glycoform was estimated based on their relative abundances ($[P1]_0$=0.28%$[P_T]_0$, $[P2]_0$=0.63%$[P_T]_0$, $[P3]_0$=0.96%$[P_T]_0$, $[P4]_0$=1.26%$[P_T]_0$, $[P5]_0$=1.09%$[P_T]_0$, $[P6]_0$=0.56%$[P_T]_0$, $[P7]_0$=0.38%$[P_T]_0$, $[P8]_0$=0.23%$[P_T]_0$ ) measured by ESI-MS. Plotted in Figure 3.10 are the fractions of warfarin-bound asialo-AGP measured at warfarin concentrations ranging from 1 μM to 6 μM. $K_{a,i}$ values for all major unfucosylated glycoforms of asialo-AGP determined by fitting Eq. 3.4 to the experimental data are in the range of $(4{\sim}8){\times}10^5$ $M^{-1}$ (Figure 3.11), which are in good agreement with the value measured for AGP binding warfarin using ITC( $(2.5 \pm 0.1) \times 10^5$ $M^{-1}$). However, they are around 5 times on average lower than the reported values ($\sim3.1{\times}10^6$ $M^{-1}$) [25] for the same glycoforms of asialo-AGP. This is possibly due to the different distributions of warfarin enantiomers (i.e., S-warfarin and R-warfarin) from different sources, as affinities of S-warfarin to AGP F1 and S variants significantly higher than that of R-warfarin.[40] Interestingly, affinities of warfarin for different glycoforms (unfucosylated P1 to P8) are all within a factor of two, suggesting that *N*-glycan branching does not have a significant effect on asialo-AGP recognition of warfarin. Moreover, the influence of the distribution of *N*-glycans on different glycosylation sites may also contribute to the measured affinities of warfarin for each glycoform of asialo-AGP.
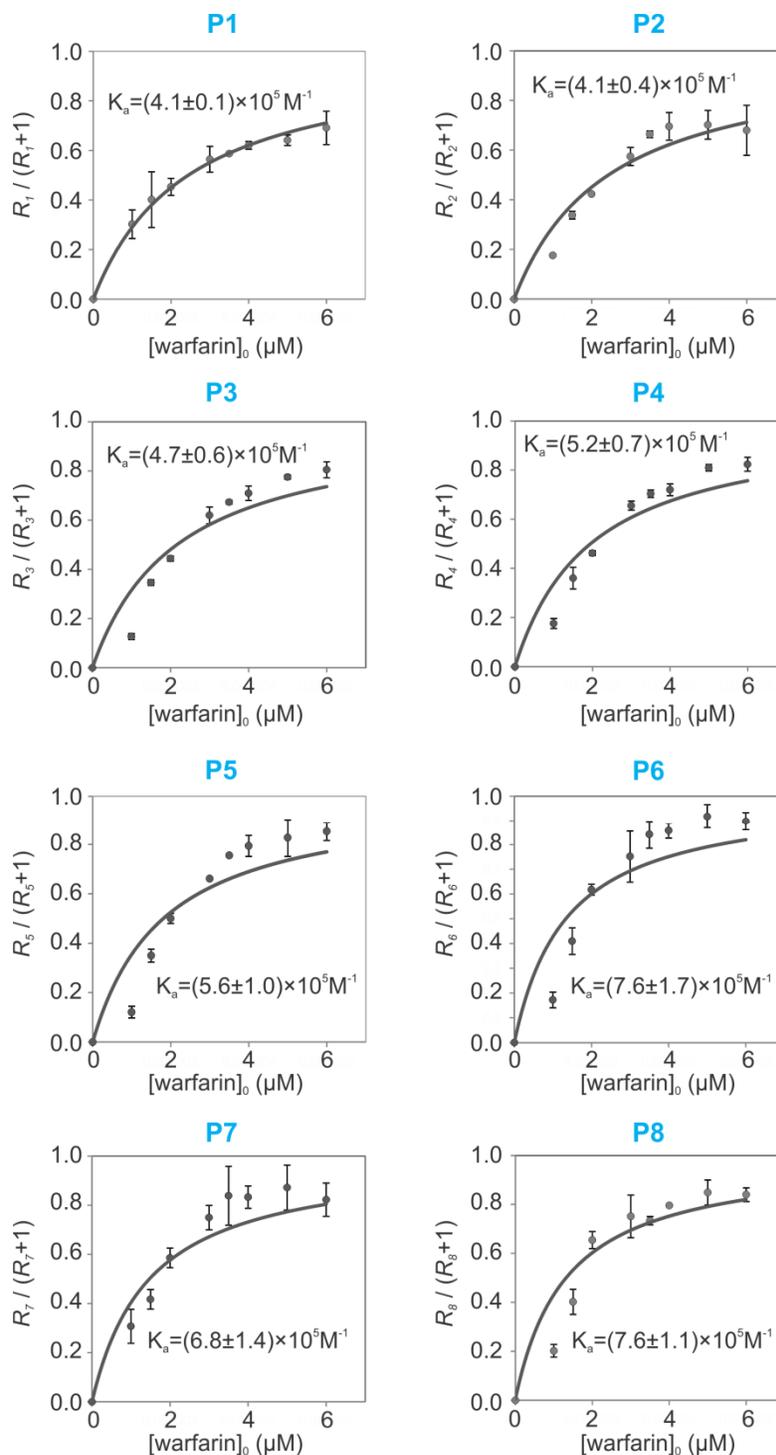
**Figure 3.10** Dependence of the fractions of warfarin-bound asialo-AGP ($R_i/(R_i+1)$) on warfarin concentration measured by ESI-MS. The solid line represents the nonlinear fit of Eq. 3.4 to the experimental binding data. Errors correspond to one standard deviation of three replicate measurements.
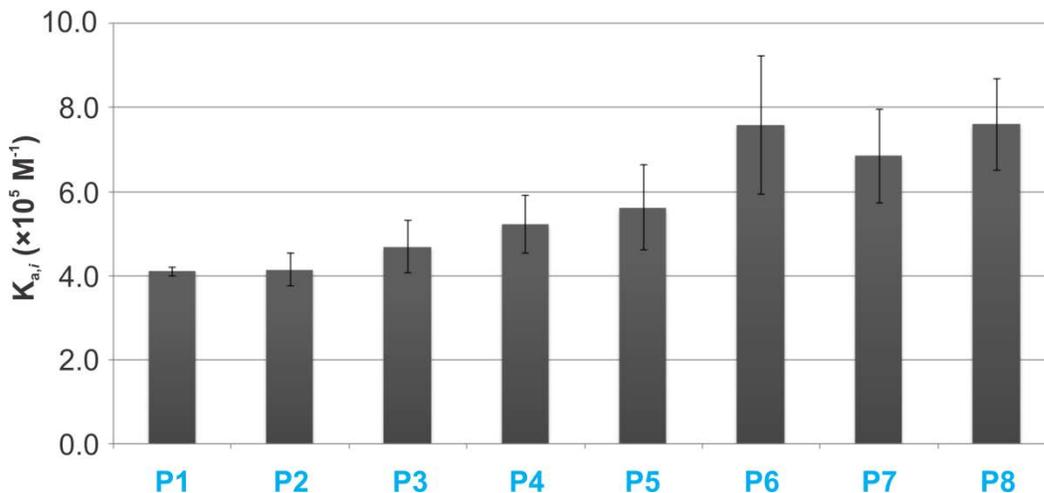
**Figure 3.11** A summary of $K_{a,i}$ values of warfarin to eight major glycoforms of asialo-AGP.

The complexity of the glycoforms can be completely eliminated by enzymatic removal of all the *N*-glycans from AGP. Theoretically, a deglycosylated AGP can be obtained by treatment of an endoglycosidase (i.e PNGase F), and association constants can be measured by the direct ESI-MS assay. As described elsewhere,[41] the association constants ratio of two proteins (e.g., $K_{a,u}/K_{a,i}$) binding to a mutual ligand (warfarin) in the same solution is equal to the corresponding abundance ratios of ligand-bound to free protein ($R_u/R_i$), and thus $K_{a,i}$ of warfarin binding to different glycoforms of AGP can be determined using Eq. 3.6. Shown in Figure 3.12 b is a representative ESI mass spectrum acquired in positive ion mode for acqueous ammonium acetate solutions (200 mM, pH 6.8) of PNGase F treated AGP (0.253 mg mL$^{-1}$) and warfarin (50 μM). Interestingly, completely deglycosylated and partially deglycosylated AGP with glycosylation sites occupied by different numbers of *N*-glycans showed strikingly different binding properties to warfarin. Shown in Figure 3.13 and Table 3.6 are relative abundances of the gaseous warfarin-bound AGP to free AGP ions (i.e., $R_u$ for unglycosylated AGP, $R_i$ for partially glycosylated AGP) determined for completely deglycosylated and partially deglycosylated AGP at warfarin (L) concentrations of 3.5 μM, 10 μM and 50 μM.

It should be noted that $R$ values were not calculated for all of the detected species due to the overlapping of AGP (F1 variant)-warfarin complexes with glycoforms of AGP F1 and S variants. Besides, $R$ values of glycoforms that bind stronger to warfarin change more rapidly and the binding sites are effectively saturated below 10 µM of warfarin, making it extremely difficult to directly compare $R$ values of glycoforms with various affinities to warfarin in a wide range of concentrations. Almost no binding was observed for non-glycosylated AGP with increasing concentrations of warfarin, suggesting that complete removal of glycan dramatically reduced the binding of AGP to warfarin. This further limited the application of ratios of the abundance ratios in determining relative affinities of different forms of protein (i.e., $P_u$ and $P_i$) for warfarin. And due to the multiple overlaps of the isoforms, $R_u/R_i$ is not independent of the ligand concentrations as multiple species with the same masses can contribute to measurable changes in $R_i$. Interestingly, $R$ values are found to increase with degree of AGP glycosylation, indicating the significant roles $N$-glycans play in promoting interactions of AGP and warfarin. One possible explanation is that $N$-glycans at Asn38 and Asn75 that bordering the binding of warfarin expand the binding cavity of AGP.[42] Notably, AGP attached to three $N$-glycans generally showed higher initial $R$ values than those attached to two $N$-glycans. It is likely that the possibility that Asn38 and Asn75 remain glycosylated after PNGase F treatment is higher in AGP with a higher degree of glycosylation, and binding of AGP to warfarin is more likely to increase with the assistance of glycans in close proximity to the binding site.
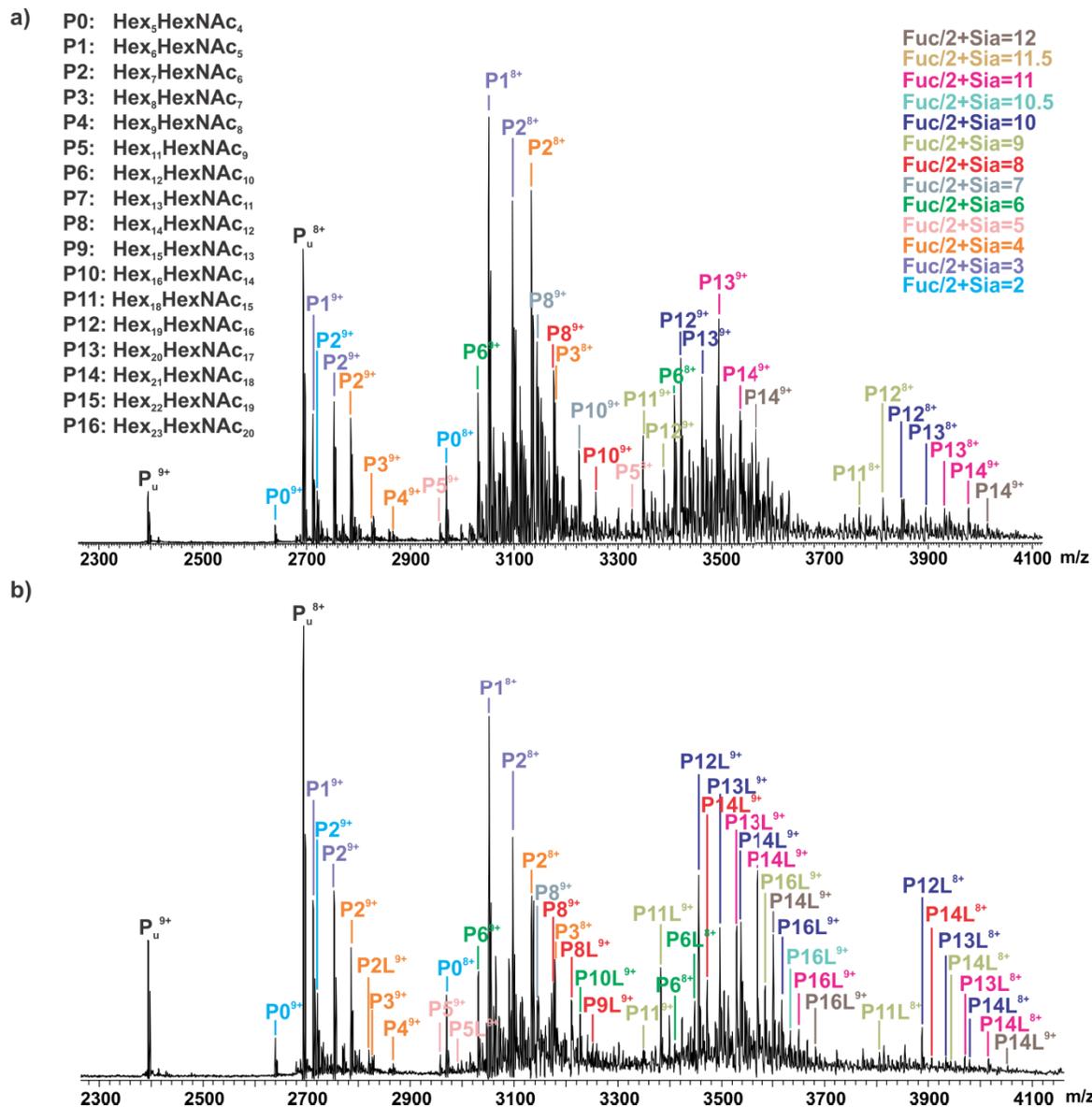
**Figure 3.12** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of PNGase F treated AGP (0.253 mg mL$^{-1}$) (a) alone and (b) with warfarin (50 µM).
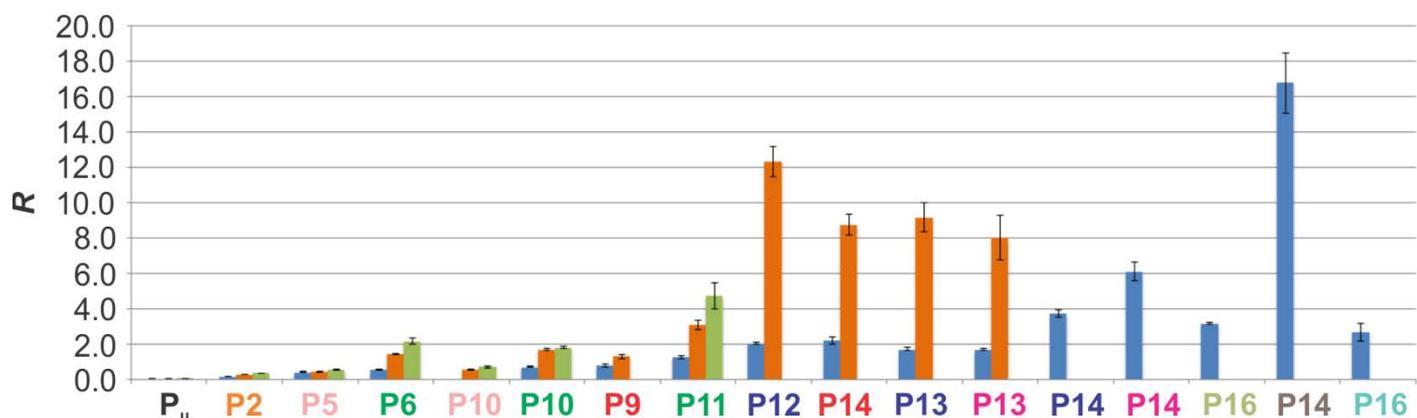
**Figure 3.13** Plot of ion abundance ratios ($R$) of warfarin-bound to free PNGase F treated AGP (0.253 mg mL$^{-1}$), measured by ESI-MS at three different concentrations of warfarin (blue: 3.5 µM; orange: 10 µM; green: 50 µM) in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C). Errors correspond to one standard deviation of three replicate measurements.

**Table 3.6** Abundance ratios ($R$) of warfarin bound-to-free AGP measured by ESI-MS at three different concentrations (3.5 µM, 10 µM and 50 µM) of warfarin.

| Annotation | Composition Hex_HexNAc_Fuc/2+Sia | $R$ (3.5 µM) | $R$ (10 µM) | $R$ (50 µM) |
|---|---|---|---|---|
| $P_u$ | | 0.06±0.01 | 0.07±0.01 | 0.08±0.01 |
| P2 | 7_6_4 | 0.16±0.01 | 0.30±0.01 | 0.38±0.01 |
| P5 | 11_9_5 | 0.43±0.03 | 0.44±0.03 | 0.56±0.04 |
| P6 | 12_10_6 | 0.58±0.04 | 1.44±0.03 | 2.19±0.16 |
| P10 | 16_14_5 | OL | 0.57±0.03 | 0.72±0.05 |
| P10 | 16_14_6 | 0.72±0.03 | 1.69±0.06 | 1.82±0.06 |
| P9 | 15_13_8 | 0.79±0.07 | 1.32±0.12 | SR |
| P11 | 18_15_9 | 1.27±0.07 | 3.10±0.24 | 4.73±0.72 |
| P12 | 19_16_10 | 2.07±0.08 | 12.33±0.88 | SR |
| P14 | 21_18_8 | 2.20±0.19 | 8.75±0.58 | SR |

| P13 | 20_17_10 | 1.72±0.09 | 9.18±0.83 | SR |
|---|---|---|---|---|
| P13 | 20_17_11 | 1.69±0.06 | 8.05±1.27 | SR |
| P14 | 21_18_10 | 3.74±0.18 | SR | SR |
| P14 | 21_18_11 | 6.12±0.53 | SR | SR |
| P16 | 23_20_9 | 3.17±0.53 | SR | SR |
| P14 | 21_18_12 | 16.80±1.71 | SR | SR |
| P16 | 23_20_10.5 | 2.70±0.51 | SR | SR |

a. Errors correspond to one standard deviation. OL≡ Overlapped. SR≡ saturated, where no free protein ions are detected.

## 3.4 Conclusions

This work demonstrates the capability of high-resolution MS for characterizing heterogenous glycoproteins and correlating glycoprotein-drug interactions to glycan compositions. Glycoform profiling of native and enzymatically treated GPs unveiled the heterogeneity arising from glycan occupancy of the glycosylation sites and different proteoforms. It offers a direct view of the relative abundance of glycoproteoforms with distinguishable masses while maintaining the disulfide bridges of the peptides, enabling assessment the impact of *N*-glycosylation on glycoprotein-drug binding. The direct ESI-MS assay was employed to evaluate the affinities of warfarin for major glycoforms of asialo-AGP, Results obtained by direct ESI-MS assay using estimated concentrations of glycoforms are on the same order of magnitude with measurements carried out by ITC for AGP and warfarin, suggesting that sialylation has no effect on warfarin binding to AGP. Furthermore, semi-quantitative comparison of warfarin-binding properties of AGP with various degrees of glycosylation showed that binding of warfarin to AGP was enhanced with increasing degree of AGP glycosylation, revealing the indispensable participation of *N*-glycans in promoting AGP recognition of warfarin.

## 3.5 References

(1)    Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. *Nat. Chem. Biol.* **2010**, *6* (10), 713–723.

(2)    Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49.

(3)    Zhao, Y. Y.; Takahashi, M.; Gu, J. G.; Miyoshi, E.; Matsumoto, A.; Kitazume, S.; Taniguchi, N. Functional Roles of N-Glycans in Cell Signaling and Cell Adhesion in Cancer. *Cancer Science*. **2008**, 1304–1310.

(4)    Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell*. **2006**, 855–867.

(5)    Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nature Reviews Cancer*. **2015**, 540–555.

(6)    Stowell, S. R.; Ju, T.; Cummings, R. D. Protein Glycosylation in Cancer. *Annu. Rev. Pathol. Mech. Dis.* **2015**, *10* (1), 473–510.

(7)    Stick, R. V.; Williams, S. J. Glycoproteins and Proteoglycans. In *Carbohydrates: The Essential Molecules of Life*; **2009**, 369–412.

(8)    Wopereis, S.; Lefeber, D. J.; Morava, É.; Wevers, R. A. Mechanisms in Protein O-Glycan Biosynthesis and Clinical and Molecular Aspects of Protein O-Glycan Biosynthesis Defects: A Review. *Clinical Chemistry*. **2006**, 574–600.

(9)    Hooker, A. D.; James, D. C. Analysis of Glycoprotein Heterogeneity by Capillary Electrophoresis and Mass Spectrometry. *Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol.* **2000**, *14* (3), 241–249.

(10)   Pantazaki, A.; Taverna, M.; Vidal-Madjar, C. Recent Advances in the Capillary Electrophoresis of Recombinant Glycoproteins. *Analytica Chimica Acta*. **1999**, 137–156.

(11)   Thormann, W.; Molteni, S.; Caslavska, J.; Schmutz, A. Clinical and Forensic

Applications of Capillary Electrophoresis. *Electrophoresis* **1994**, *15* (1), 3–12.

(12) Geyer, H.; Geyer, R. Strategies for Analysis of Glycoprotein Glycosylation. *Biochimica et Biophysica Acta - Proteins and Proteomics*. **2006**, 1853–1869.

(13) Yang, Z.; Hancock, W. S. Monitoring Glycosylation Pattern Changes of Glycoproteins Using Multi-Lectin Affinity Chromatography. *J. Chromatogr. A* **2005**, *1070* (1–2), 57–64.

(14) Hu, S.; Wong, D. T. Lectin Microarray. *Proteomics - Clin. Appl.* **2009**, *3* (2), 148–154.

(15) Wang, H.; Li, H.; Zhang, W.; Wei, L.; Yu, H.; Yang, P. Multiplex Profiling of Glycoproteins Using a Novel Bead-Based Lectin Array. *Proteomics* **2014**, *14* (1), 78–86.

(16) Wuhrer, M. Glycomics Using Mass Spectrometry. *Glycoconj. J.* **2013**, *30* (1), 11–22.

(17) Wuhrer, M.; Deelder, A. M.; Hokke, C. H. Protein Glycosylation Analysis by Liquid Chromatography-Mass Spectrometry. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. **2005**, 124–133.

(18) Kolarich, D.; Jensen, P. H.; Altmann, F.; Packer, N. H. Determination of Site-Specific Glycan Heterogeneity on Glycoproteins. *Nat. Protoc.* **2012**, *7* (7), 1285–1298.

(19) Ongay, S.; Boichenko, A.; Govorukhina, N.; Bischoff, R. Glycopeptide Enrichment and Separation for Protein Glycosylation Analysis. *J. Sep. Sci.* **2012**, *35* (18), 2341–2372.

(20) Zamfir, A. D.; Flangea, C.; Serb, A.; Zagrean, A. M.; Rizzi, A. M.; Sisu, E. *Mass Spectrometry of Glycoproteins*; Kohler, J. J., Patrie, S. M., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, **2013**; Vol. 951.

(21) Zhang, N.; Liu, S.; Liu, Z.; Wan, D.; Cui, M.; Liu, H. Mass Spectrometry-Based Analysis of Glycoproteins and Its Clinical Applications in Cancer Biomarker

Discovery. *Clin. Proteomics* **2014**, *11* (1), 14.

(22)   Yu, A.; Zhao, J.; Peng, W.; Banazadeh, A.; Williamson, S. D.; Goli, M.; Huang, Y.; Mechref, Y. Advances in Mass Spectrometry-Based Glycoproteomics. *Electrophoresis*. **2018**, 3104–3122.

(23)   Li, F.; Glinskii, O. V.; Glinsky, V. V. Glycobioinformatics: Current Strategies and Tools for Data Mining in MS-Based Glycoproteomics. *Proteomics* **2013**, *13* (2), 341–354.

(24)   Cui, W.; Rohrs, H. W.; Gross, M. L. Top-down Mass Spectrometry: Recent Developments, Applications and Perspectives. *Analyst* **2011**, *136* (19), 3854.

(25)   Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (35), 8763–8768.

(26)   Urien, S.; Albengres, E.; Zini, R.; Tillement, J. P. Evidence for Binding of Certain Acidic Drugs to A1-Acid Glycoprotein. *Biochem. Pharmacol.* **1982**, *31* (22), 3687–3689.

(27)   Urien, S.; Albengres, E.; Pinquier, J.-L.; Tillement, J.-P. Role of Alpha-1 Acid Glycoprotein, Albumin, and Nonesterified Fatty Acids in Serum Binding of Apazone and Warfarin. *Clin. Pharmacol. Ther.* **1986**, *39* (6), 683–689.

(28)   Ongay, S.; Neusüß, C. Isoform Differentiation of Intact AGP from Human Serum by Capillary Electrophoresis-Mass Spectrometry. *Anal. Bioanal. Chem.* **2010**, *398* (2), 845–855.

(29)   Israili, Z. H.; Dayton, P. G. Human Alpha-1-Glycoprotein and Its Interactions with Drugs. *Drug Metabolism Reviews*. **2001**, 161–235.

(30)   Saroha, A.; Biswas, S.; Chatterjee, B. P.; Das, H. R. Altered Glycosylation and Expression of Plasma Alpha-1-Acid Glycoprotein and Haptoglobin in Rheumatoid Arthritis. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2011**, *879* (20), 1839–1843.

(31)   Imre, T.; Kremmer, T.; Héberger, K.; Molnár-Szöllosi, É.; Ludányi, K.; Pócsfalvi, G.; Malorni, A.; Drahos, L.; Vékey, K. Mass Spectrometric and Linear Discriminant Analysis of N-Glycans of Human Serum Alpha-1-Acid Glycoprotein in Cancer Patients and Healthy Individuals. *J. Proteomics* **2008**, *71* (2), 186–197.

(32)   Tanabe, K.; Kitagawa, K.; Kojima, N.; Iijima, S. Multifucosylated Alpha-1-Acid Glycoprotein as a Novel Marker for Hepatocellular Carcinoma. *J. Proteome Res.* **2016**, *15* (9), 2935–2944.

(33)   Ayyub, A.; Saleem, M.; Fatima, I.; Tariq, A.; Hashmi, N.; Musharraf, S. G. Glycosylated Alpha-1-Acid Glycoprotein 1 as a Potential Lung Cancer Serum Biomarker. *Int. J. Biochem. Cell Biol.* **2016**, *70*, 68–75.

(34)   Colombo, S.; Buclin, T.; Décosterd, L. A.; Telenti, A.; Furrer, H.; Lee, B. L.; Biollaz, J.; Eap, C. B. Orosomucoid (A1-Acid Glycoprotein) Plasma Concentration and Genetic Variants: Effects on Human Immunodeficiency Virus Protease Inhibitor Clearance and Cellular Accumulation. *Clin. Pharmacol. Ther.* **2006**, *80* (4), 307–318.

(35)   Taguchi, K.; Nishi, K.; Giam Chuang, V. T.; Maruyama, T.; Otagiri, M. Molecular Aspects of Human Alpha-1 Acid Glycoprotein — Structure and Function. In *Acute Phase Proteins*; InTech, 2013; Vol. i, p 13.

(36)   Nakano, M. Detailed Structural Features of Glycan Chains Derived from 1-Acid Glycoproteins of Several Different Animals: The Presence of Hypersialylated, O-Acetylated Sialic Acids but Not Disialyl Residues. *Glycobiology* **2004**, *14* (5), 431–441.

(37)   Mancera-Arteu, M.; Gim Enez, E.; Barbosa, J. E.; Oria Sanz-Nebot, V. Identification and Characterization of Isomeric N-Glycans of Human Alfa-Acid-Glycoprotein by Stable Isotope Labelling and ZIC-HILIC-MS in Combination with Exoglycosidase Digestion. *Anal. Chim. Acta* **2016**, *940*, 92–103.

(38)   Schönfeld, D. L.; Ravelli, R. B. G.; Mueller, U.; Skerra, A. The 1.8-Å Crystal

Structure of A1-Acid Glycoprotein (Orosomucoid) Solved by UV RIP Reveals the Broad Drug-Binding Activity of This Human Plasma Lipocalin. *J. Mol. Biol.* **2008**, *384* (2), 393–405.

(39)  Nishi, K.; Ono, T.; Nakamura, T.; Fukunaga, N.; Izumi, M.; Watanabe, H.; Suenaga, A.; Maruyama, T.; Yamagata, Y.; Curry, S.; et al. Structural Insights into Differences in Drug-Binding Selectivity between Two Forms of Human A1-Acid Glycoprotein Genetic Variants, the A and F1*S Forms. *J. Biol. Chem.* **2011**, *286* (16), 14427–14434.

(40)  Nakagawa, T.; Kishino, S.; Itoh, S.; Sugawara, M.; Miyazaki, K. Differential Binding of Disopyramide and Warfarin Enantiomers to Human A1-Acid Glycoprotein Variants. *Br. J. Clin. Pharmacol.* **2003**, *56* (6), 664–669.

(41)  Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; St-Pierre, Y.; Klassen, J. S. Human Milk Oligosaccharide Specificities of Human Galectins. Comparison of Electrospray Ionization Mass Spectrometry and Glycan Microarray Screening Results. *Anal. Chem.* **2017**, *89* (9), 4914–4921.

(42)  Fernandes, C. L.; Ligabue-Braun, R.; Verli, H. Structural Glycobiology of Human A1-Acid Glycoprotein and Its Implications for Pharmacokinetics and Inflammation. *Glycobiology* **2015**, *25* (10), 1125–1133.

# Chapter 4

## Conclusions and Future Work

## 4.1 Conclusions

This work demonstrates the application of ESI-MS in characterizing glycoproteins (GPs) and investigating GP interactions. The first research project was focused on the development of a multipronged approach to detect and quantify water soluble GBP-glycoprotein interactions, and to elucidate glycan ligands of the GBP. The second project characterized the microheterogeneity of GPs and studied the influence of glycosylation on GP-drug interactions.

Chapter 2 describes a multipronged ESI-MS approach for detecting and quantifying binding between water soluble GBPs and GPs and identifying the underlying glycan-GBP interactions. Specific binding of hGal-3C to three human serum GPs (AGP, Hp1-1 and $\alpha$2M), which are known to be recognized by hGal-3C, was successfully detected by CaR$^{IMS}$-ESI-MS. The $K_{a,app}$ determined by competitive binding measurements, *proxy ligand*-ESI-MS assay, for the three GPs (ranging from 2 x $10^5$ M$^{-1}$ to 4 x $10^5$ M$^{-1}$) are consistent with the reported apparent affinities of hGal-3 for serum GPs and their *N*-glycans. ESI-MS screening of the *N*-glycan libraries released enzymatically from the GsP against hGal-3C identified ligands corresponding to twenty different saccharide compositions with sialylated bi-, tri- and tetra-antennary structures. And screening results carried out at two different hGal-3C concentrations suggest that all of the *N*-glycan ligands exhibit affinities $\geq 10^4$ M$^{-1}$.

Chapter 3 is focused on using a direct ESI-MS approach to examine binding properties of warfarin, an anticoagulant drug, to asialo-AGP determined by saturation binding experiments and extend them to more extensively investigate the impact of glycosylation on AGP-warfarin interaction by removal of *N*-glycans. High-resolution mass spectrometry is employed to characterize the microheterogeneity of a human serum

glycoprotein. AGP, and following treatments with sialidase, α1-3,4 fucosidase and PNGase F. Complexity of the spectrum and ambiguity of glycoproteoform assignments were substantially reduced with decreased levels of fucosylation and sialylation. Direct analysis of AGP with different degrees of glycosylation revealed 105 glycoproteoforms for AGP, 98 for asialo-AGP and 59 for α1-3,4 fucosidase treated asialo-AGP corresponding to F1 and S variants. Affinities of warfarin for major unfucosylated glycoforms of asialo-AGP measured by direct ESI-MS are within a factor of two and are around five times less than the reported values, but are in good agreements with values determined by ITC. Therefore, *N*-glycan branching does not have measurable effects on asialo-AGP-warfarin binding. Influence of *N*-glycosylation on GP-ligand binding was further studied on interactions between PNGase F treated AGP and warfarin, where completely deglycosylated AGP showed no binding to warfarin. Semi-quantitative comparison of warfarin-binding properties of AGP with various degrees of glycosylation revealed the indispensable participation of *N*-glycans in promoting the ligand recognition of AGP.

## 4.2 Future work

There are various possible extensions of the current work. Structural characterization using LC-MS of *N*-glycans released from GPs from commercial and natural sources can provide further insights into glycan compositions and isomers. Direct screening of *N*-glycan libraries against other GBPs in combination with CaR-ESI-MS/MS can correlate glycan structures to their binding properties. The assignment of glycoproteoforms can be also improved by exploring charge reducing agents (e.g. triethylammonium acetate (TEAA), imidazole derivatives) and by applying ECD /ETD-MS. Affinities of warfarin for native AGP can be determined by direct ESI-MS assay and compared to results measured by ITC. Binding of warfarin to different glycoforms of α1-3,4 fucosidase treated asialo-AGP can be determined and compared to that of asialo-AGP to further investigate the influence of Fuc on AGP-warfarin

interaction. Overlapping in *m/z* of free AGP glycoforms and their complexes can potentially be overcome by exploring other binding partners with high MW, such as vitamin $B_{12}$[1] with a molecular weight of 1355.4 Da.

High-resolution MS has recently emerged as a promising tool for direct glycoform profiling of intact glycoproteins.[2,3] Isomeric glycoforms can be obtained by treatment of highly specific enzymes that catalyze the hydrolysis of specific linkages and ligand binding properties to different isomers can be determined. An example study is the development of a direct ESI-MS based method to quantify α2,6 and α2,3 sialylated prostate specific antigen (PSA) for the diagnosis of prostate cancer.

Prostate cancer is one of the most prevalent cancers diagnosed in men worldwide.[4] PSA, or human kallikrein 3 (hK3), an androgen-regulated protease and a member of the tissue kallikrein family (a subgroup of serine proteases) mainly produced by prostate epithelial cells[5], represents a powerful biomarker for the diagnosis and prognosis of prostate cancer.[6,7] Measurements of serum PSA concentrations have been considered as the gold standard for the detection of prostate cancer, as levels of serum PSA are elevated in prostate cancer patients due to the disruption of the prostate gland.[8] However, PSA-based screening assays remains controversial due to sensitivity (20%) and specificity (93%) in differentiating benign prostatic hyperplasia (BPH) from cancer especially in the PSA 'grey zone' (4-10 ng/mL),[9,10] leading to over-diagnosis, unnecessary biopsies and over-treatment of indolent cancers. Numerous strategies have been proposed to improve the specificity, such as the PSA index (free PSA (fPSA) / total PSA (tPSA)) which is lower (20% for the grey zone) in aggressive forms of prostate cancer,[11] and FDA-approved prostate health index (PHI) where increasing values indicate a higher risk of prostate cancer.[12] Although other PSA-related variables such as PSA density, PSA velocity and age-adjusted PSA, no substantial improvement has been made and false positives cannot be completely avoided.

Alterations of the cellular glycosylation pattern are known to be the principal

characteristic of cancer progression.[13] PSA is a 28.4 kDa glycoprotein (237 amino acids) comprising five interchain disulphide bonds with a single glycosylation site at Asn 45 occupied by N-linked oligosaccharide constituting 8% of the total weight.[14] Monitoring changes of PSA glycosylation represents another non-invasive assay for the diagnosis of prostate cancer. An increase in α2,3 linked sialic acids from serum PSA has been reported in prostate cancer patients compared to BPH/non-prostate cancer patients using lectin-based methods (e.g., α2,3-sialic acid binding lectin *Maackia amurensis* agglutinin (MAA), α2,6-sialic acid binding lectin *Sambucus nigra* agglutinin (SNA)),[15,16] magnetic microbead-based immunoassay,[17] and high-performance liquid chromatography (HPLC)-based N-glycan analysis.[18] Differences in core fucosylation in serum PSA were also observed in aggressive and non-aggressive prostate cancer patients. While core fucosylation of multiantennary glycans showed a decrease and more core-fucosylated bi- and tri-antennary glycans were detected in prostate cancer samples,[18] the overall core fucosylation levels in serum PSA quantified by *Pholiota squarrosa* lectin (PhoSL), which binds exclusively to core fucose, were decreased in high-risk prostate cancer patients.[16] Additionally, studies using α1,2-fucose specific lectin *Ulex europaeus* agglutinin I (UEA-I) and *Trichosanthes japonica* agglutinin II (TJA-II) showed an increase in α1,2-fucosylation of serum PSA from prostate cancer patients.[19,20] However, drawbacks of lectin and antibody based assays for quantification of PSA glycopattern changes includes the requirement of labelling, external calibration and the availability of lectin and anti-carbohydrate antibodies with high binding specificity and affinity.[21] While glycan analysis provides an alternative way for detection altered glycosylation in PSA with high sensitivity, lack of internal standards or external calibration is necessary for absolute glycan quantification.[21]

Time-resolved ESI-MS analysis combined with exoglycosidase treatment has the potential as a method to quantify α2,3 sialylated PSA, which could serve as a biomarker. Preliminary experiments were performed on PSA extracted from semen (Lee

Biosolutions, Inc), which revealed two major species (28430.50 Da and 28283.93 Da), corresponding to di-antennary fucosylated ($Hex_5HexNAc_4Fuc_1Sia_2$) and non-fucosylated ($Hex_5HexNAc_4Fuc_0Sia_2$) PSA (Figure 4.1a). Shown in Figure 4.2b is a representative mass spectrum acquired in positive ion mode for aqueous 200 mM ammonium acetate solutions (pH 6.8, 25 °C) containing PSA (0.0534 mg mL$^{-1}$) and human neuraminidase 2 (NEU 2), which preferentially catalyzes the removal of α2,3-linked sialic acid.[22] Abundance ratio of unreacted substrates to initial substrates in the time-resolved ESI-MS gradually decreased and reached a plateau, indicating that all PSA with α2,3-linked sialic acid has been consumed and that enzymatic activity of NEU2 on α2,6-linked sialic acid is negligible. The relative amounts of α2,3- and α2,6-linked sialic acid can be determined by linear extrapolation of the reaction progress plot (Figure 4.2), which were 30.4% and 69.6% in the PSA standard, respectively. This method can be further applied to the quantification of serum PSA isomers extracted from patients for prostate cancer diagnosis.
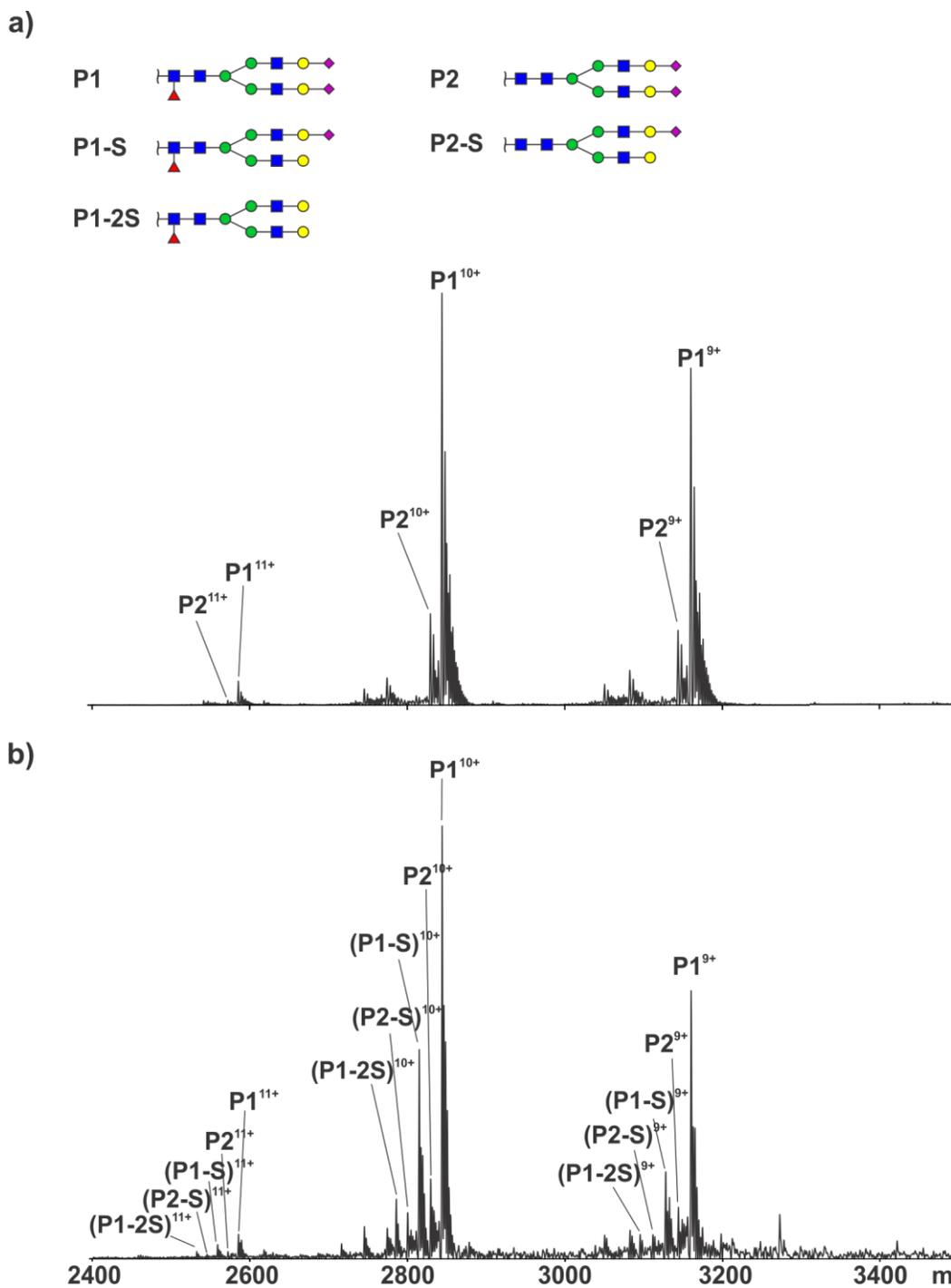
**Figure 4.1** ESI mass spectra acquired in positive ion mode for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of PSA (0.0534 mg mL$^{-1}$) (a) alone and (b) with NEU 2 acquired at 298-299 min.
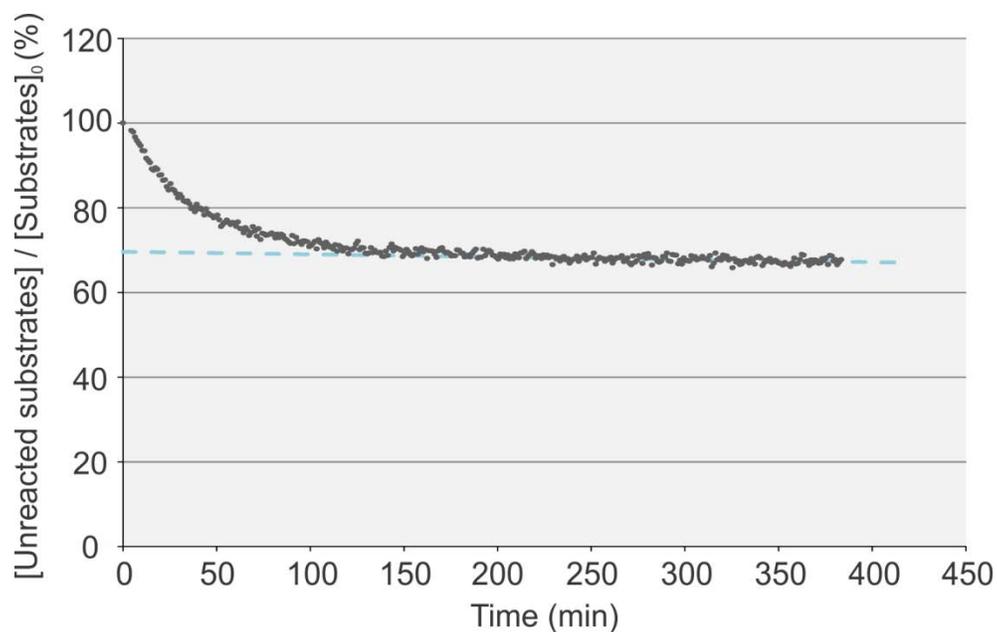
**Figure 4.2** A plot of the abundance ratio of unreacted substrates to total substrates (i.e., P1, P2, P1-S, P1-2S, P2-S) versus reaction time measured by ESI-MS for 200 mM ammonium acetate solutions (pH 6.8 and 25 °C) of PSA (0.0534 mg mL$^{-1}$) with NEU 2.

## 4.3 References

(1)     Chiu, K. M.; Mortensen, R. F.; Osmand, A. P.; Gewurzt, H.; Gewurz, H. Interactions of Alpha1 Acid Glycoprotein with the Immune System. I Purification and Effects upon Lymphocyte Responsiveness. *Immunology* **1977**, *32* (6), 997–1005.

(2)     Yang, Y.; Liu, F.; Franc, V.; Halim, L. A.; Schellekens, H.; Heck, A. J. R. Hybrid Mass Spectrometry Approaches in Glycoprotein Analysis and Their Usage in Scoring Biosimilarity. *Nat. Commun.* **2016**, *7*.

(3)     Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (35), 8763–8768.

(4)     Siegel, R. L.; Miller, K. D.; Jemal, A. Cancer Statistics, 2016. *CA. Cancer J. Clin.* **2016**, *66* (1), 7–30.

(5)     Watt, K. W.; Lee, P. J.; M'Timkulu, T.; Chan, W. P.; Loor, R. Human Prostate-Specific Antigen: Structural and Functional Similarity with Serine Proteases. *Proc. Natl. Acad. Sci.* **1986**, *83* (10), 3166–3170.

(6)     Stamey, T. A.; Yang, N.; Hay, A. R.; McNeal, J. E.; Freiha, F. S.; Redwine, E. Prostate-Specific Antigen as a Serum Marker for Adenocarcinoma of the Prostate. *N. Engl. J. Med.* **1987**, *317* (15), 909–916.

(7)     Shariat, S. F.; Semjonow, A.; Lilja, H.; Savage, C.; Vickers, A. J.; Bjartell, A. Tumor Markers in Prostate Cancer I: Blood-Based Markers. *Acta Oncologica*. **2011**, 61–75.

(8)     Gilgunn, S.; Conroy, P. J.; Saldova, R.; Rudd, P. M.; O'Kennedy, R. J. Aberrant PSA Glycosylation - A Sweet Predictor of Prostate Cancer. *Nature Reviews Urology*. **2013**, 99–107.

(9)     Andriole, G. L.; Crawford, E. D.; Grubb, R. L.; Buys, S. S.; Chia, D.; Church, T. R.; Fouad, M. N.; Gelmann, E. P.; Kvale, P. A.; Reding, D. J.; et al. Mortality

Results from a Randomized Prostate-Cancer Screening Trial. *N. Engl. J. Med.* **2009**, *360* (13), 1310–1319.

(10)   Li, Y.; Tian, Y.; Rezai, T.; Prakash, A.; Lopez, M. F.; Chan, D. W.; Zhang, H. Simultaneous Analysis of Glycosylated and Sialylated Prostate-Specific Antigen Revealing Differential Distribution of Glycosylated Prostate-Specific Antigen Isoforms in Prostate Cancer Tissues. *Anal. Chem.* **2011**, *83* (1), 240–245.

(11)   Pelzer, A. E.; Volgger, H.; Bektic, J.; Berger, A. P.; Rehder, P.; Bartsch, G.; Horninger, W. The Effect of Percentage Free Prostate-Specific Antigen (PSA) Level on the Prostate Cancer Detection Rate in a Screening Population with Low PSA Levels. *BJU Int.* **2005**, *96* (7), 995–998.

(12)   Stephan, C.; Ralla, B.; Jung, K. Prostate-Specific Antigen and Other Serum and Urine Markers in Prostate Cancer. *Biochimica et Biophysica Acta - Reviews on Cancer*. August **2014**, 99–112.

(13)   Meany, D. L.; Chan, D. W. Aberrant Glycosylation Associated with Enzymes as Cancer Biomarkers. *Clinical Proteomics*. **2011**, 7.

(14)   Bélanger, A.; Van Halbeek, H.; Graves, H. C. B.; Grandbois, K.; Stamey, T. A.; Huang, L.; Poppe, I.; Labrie, F. Molecular Mass and Carbohydrate Structure of Prostate Specific Antigen: Studies for Establishment of an International PSA Standard. *Prostate* **1995**, *27* (4), 187–197.

(15)   Ohyama, C.; Hosono, M.; Nitta, K.; Oh-eda, M.; Yoshikawa, K.; Habuchi, T.; Arai, Y.; Fukuda, M. Carbohydrate Structure and Differential Binding of Prostate Specific Antigen to Maackia Amurensis Lectin between Prostate Cancer and Benign Prostate Hypertrophy. *Glycobiology* **2004**, *14* (8), 671–679.

(16)   Llop, E.; Ferrer-Batallé, M.; Barrabés, S.; Guerrero, P. E.; Ramírez, M.; Saldova, R.; Rudd, P. M.; Aleixandre, R. N.; Comet, J.; de Llorens, R.; et al. Improvement of Prostate Cancer Diagnosis by Detecting PSA Glycosylation-Specific Changes. *Theranostics* **2016**, *6* (8), 1190–1204.

(17) Yoneyama, T.; Ohyama, C.; Hatakeyama, S.; Narita, S.; Habuchi, T.; Koie, T.; Mori, K.; Hidari, K. I. P. J.; Yamaguchi, M.; Suzuki, T.; et al. Measurement of Aberrant Glycosylation of Prostate Specific Antigen Can Improve Specificity in Early Detection of Prostate Cancer. *Biochem. Biophys. Res. Commun.* **2014**, *448* (4), 390–396.

(18) Saldova, R.; Fan, Y.; Fitzpatrick, J. M.; Watson, R. W. G.; Rudd, P. M. Core Fucosylation and A2-3 Sialylation in Serum N-Glycome Is Significantly Increased in Prostate Cancer Comparing to Benign Prostate Hyperplasia. *Glycobiology* **2011**, *21* (2), 195–205.

(19) Dwek, M. V.; Jenks, A.; Leathem, A. J. C. A Sensitive Assay to Measure Biomarker Glycosylation Demonstrates Increased Fucosylation of Prostate Specific Antigen (PSA) in Patients with Prostate Cancer Compared with Benign Prostatic Hyperplasia. *Clin. Chim. Acta* **2010**, *411* (23–24), 1935–1939.

(20) Fukushima, K.; Satoh, T.; Baba, S.; Yamashita, K. A1,2-Fucosylated and β-N-Acetylgalactosaminylated Prostate-Specific Antigen as an Efficient Marker of Prostatic Cancer. *Glycobiology* **2009**, *20* (4), 452–460.

(21) Etxebarria, J.; Reichardt, N. C. Methods for the Absolute Quantification of N-Glycan Biomarkers. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (8), 1676–1687.

(22) Hunter, C. D.; Khanna, N.; Richards, M. R.; Rezaei Darestani, R.; Zou, C.; Klassen, J. S.; Cairo, C. W. Human Neuraminidase Isoenzymes Show Variable Activities for 9- O -Acetyl-Sialoside Substrates. *ACS Chem. Biol.* **2018**, *13* (4), 922–932.

# List of References

## Chapter 1

(1)    Ghazarian, H.; Idoni, B.; Oppenheimer, S. B. A Glycobiology Review: Carbohydrates, Lectins and Implications in Cancer Therapeutics. *Acta Histochemica*. **2011**, 236–247.

(2)    Herget, S.; Toukach, P. V; Ranzinger, R.; Hull, W. E.; Knirel, Y. A.; Von Der Lieth, C. W. Statistical Analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): Characteristics and Diversity of Bacterial Carbohydrates in Comparison with Mammalian Glycans. *BMC Struct. Biol.* **2008**, *8*.

(3)    Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. *Nat. Chem. Biol.* **2010**, *6* (10), 713–723.

(4)    Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49.

(5)    Zhao, Y. Y.; Takahashi, M.; Gu, J. G.; Miyoshi, E.; Matsumoto, A.; Kitazume, S.; Taniguchi, N. Functional Roles of N-Glycans in Cell Signaling and Cell Adhesion in Cancer. *Cancer Science*. **2008**, 1304–1310.

(6)    Apweiler, R.; Hermjakob, H.; Sharon, N. On the Frequency of Protein Glycosylation, as Deduced from Analysis of the SWISS-PROT Database. *Biochim. Biophys. Acta* **1999**, *1473* (1), 4–8.

(7)    Cummings, R. D.; Pierce, J. M. The Challenge and Promise of Glycomics. *Chem. Biol.* **2014**, *21* (1), 1–15.

(8)    Stowell, S. R.; Ju, T.; Cummings, R. D. Protein Glycosylation in Cancer. *Annu. Rev. Pathol.* **2015**, *10*, 473–510.

(9)    Zeng, X.; Andrade, C. A. S.; Oliveira, M. D. L.; Sun, X. L. Carbohydrate-Protein Interactions and Their Biosensing Applications. *Analytical and Bioanalytical Chemistry*. **2012**, 3161–3176.

(10)   Paulson, J. C.; Blixt, O.; Collins, B. E. Sweet Spots in Functional Glycomics.

*Nature Chemical Biology*. **2006**, 238–248.

(11)   Zhong, X.; Somers, W. Recent Advances in Glycosylation Modifications in the Context of Therapeutic Glycoproteins. In *Integrative Proteomics*; InTech, **2012**; Vol. i, 13.

(12)   Costa, A. R.; Rodrigues, M. E.; Henriques, M.; Oliveira, R.; Azeredo, J. Glycosylation: Impact, Control and Improvement during Therapeutic Protein Production. *Critical Reviews in Biotechnology*. Informa Healthcare USA, Inc **2014**, 281–299.

(13)   Liang, Y. Applications of Isothermal Titration Calorimetry in Protein Science. *Acta Biochimica et Biophysica Sinica*. **2008**, 565–576.

(14)   Dam, T. K.; Brewer, C. F. Thermodynamic Studies of Lectin−Carbohydrate Interactions by Isothermal Titration Calorimetry. *Chem. Rev.* **2002**, *102* (2), 387–430.

(15)   Torres, F. E.; Recht, M. I.; Coyle, J. E.; Bruce, R. H.; Williams, G. Higher Throughput Calorimetry: Opportunities, Approaches and Challenges. *Current Opinion in Structural Biology*. **2010**, 598–605.

(16)   Rajarathnam, K.; Rösgen, J. Isothermal Titration Calorimetry of Membrane Proteins - Progress and Challenges. *Biochimica et Biophysica Acta - Biomembranes*. **2014**, 69–77.

(17)   Huber, W.; Mueller, F. Biomolecular Interaction Analysis in Drug Discovery Using Surface Plasmon Resonance Technology. *Curr. Pharm. Des.* **2007**, *12* (31), 3999–4021.

(18)   De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. Kinetic Characterization by Surface Plasmon Resonance-Based Biosensors: Principle and Emerging Trends. *Cell. Mol. Bioeng.* **2008**, *1* (4), 204–215.

(19)   Tudos, A. J.; Schasfoort, R. B. M. *Handbook of Surface Plasmon Resonance*; Schasfoort, R. B. M., Ed.; Royal Society of Chemistry: Cambridge, **2017**.

(20) Tateno, H.; Nakamura-Tsuruta, S.; Hirabayashi, J. Frontal Affinity Chromatography: Sugar-Protein Interactions. *Nat. Protoc.* **2007**, *2* (10), 2529–2537.

(21) Kasai, K. *Lectins*; Hirabayashi, J., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, **2014**; Vol. 1200.

(22) Iwaki, J.; Hirabayashi, J. Carbohydrate-Binding Specificity of Human Galectins: An Overview by Frontal Affinity Chromatography. *Trends Glycosci. Glycotechnol.* **2018**, *30* (172), 137–153.

(23) Chen, Z.; Weber, S. G. Determination of Binding Constants by Affinity Capillary Electrophoresis, Electrospray Ionization Mass Spectrometry and Phase-Distribution Methods. *TrAC - Trends Anal. Chem.* **2008**, *27* (9), 738–748.

(24) Smith, D. F.; Song, X.; Cummings, R. D. Use of Glycan Microarrays to Explore Specificity of Glycan-Binding Proteins. In *Methods in Enzymology*; **2010**; Vol. 480, 417–444.

(25) Patwa, T. H.; Zhao, J.; Andersen, M. A.; Simeone, D. M.; Lubman, D. M. Screening of Glycosylation Patterns in Serum Using Natural Glycoprotein Microarrays and Multi-Lectin Fluorescence Detection. *Anal. Chem.* **2006**, *78* (18), 6411–6421.

(26) Patwa, T.; Li, C.; Simeone, D. M.; Lubman, D. M. Glycoprotein Analysis Using Protein Microarrays and Mass Spectrometry. *Mass Spectrom. Rev.* **2010**, *29* (5), 830–844.

(27) Stowell, S. R.; Cummings, R. D. *Galectins*; Stowell, S. R., Cummings, R. D., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, **2015**; Vol. 1207.

(28) Kilcoyne, M.; Gerlach, J. Q.; Kane, M.; Joshi, L. Surface Chemistry and Linker Effects on Lectin-Carbohydrate Recognition for Glycan Microarrays. *Anal. Methods* **2012**, *4* (9), 2721–2728.

(29)  Han, L.; Shams-Ud-Doha, K.; Kitova, E. N.; Klassen, J. S. Screening Oligosaccharide Libraries against Lectins Using the Proxy Protein Electrospray Ionization Mass Spectrometry Assay. *Anal. Chem.* **2016**, *88* (16), 8224–8231.

(30)  Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; St-Pierre, Y.; Klassen, J. S. Human Milk Oligosaccharide Specificities of Human Galectins. Comparison of Electrospray Ionization Mass Spectrometry and Glycan Microarray Screening Results. *Anal. Chem.* **2017**, *89* (9), 4914–4921.

(31)  El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Carbohydrate-Protein Interactions by Electrospray Ionization Mass Spectrometry Analysis. *Biochemistry* **2012**, *51* (21), 4244–4253.

(32)  Liu, L.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Fatty Acid Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (2), 310–318.

(33)  Cubrilovic, D.; Biela, A.; Sielaff, F.; Steinmetzer, T.; Klebe, G.; Zenobi, R. Quantifying Protein-Ligand Binding Constants Using Electrospray Ionization Mass Spectrometry: A Systematic Binding Affinity Study of a Series of Hydrophobically Modified Trypsin Inhibitors. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (10), 1768–1777.

(34)  Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441.

(35)  Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. Affinities of Recombinant Norovirus P Dimers for Human Blood Group Antigens. *Glycobiology* **2013**, *23* (3), 276–285.

(36)  El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; St-Pierre, Y.; Klassen, J. S. High-Throughput Label- and Immobilization-Free Screening of Human Milk Oligosaccharides Against Lectins. *Anal. Chem.* **2017**, *89* (17), 8713–8722.

(37)   El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; Bode, L.; Hage, N.; Falcone, F. H.; Klassen, J. S. Screening Natural Libraries of Human Milk Oligosaccharides against Lectins Using CaR-ESI-MS. *Analyst* **2018**, *143* (2), 536–548.

(38)   Wang, Y.; Park, H.; Lin, H.; Kitova, E. N.; Klassen, J. S. Multipronged ESI-MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins. *Anal. Chem.* **2019**, *91* (3), 2140–2147.

(39)   Yao, Y.; Shams-Ud-Doha, K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Carbohydrate Interactions Using Liquid Sample Desorption Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2014**, *26* (1), 98–106.

(40)   Rademacher, C.; Shoemaker, G. K.; Kim, H. S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T.; et al. Ligand Specificity of CS-35, a Monoclonal Antibody That Recognizes Mycobacterial Lipoarabinomannan: A Model System for Oligofuranoside-Protein Recognition. *J. Am. Chem. Soc.* **2007**, *129* (34), 10489–10502.

(41)   Liuni, P.; Wilson, D. J. Understanding and Optimizing Electrospray Ionization Techniques for Proteomic Analysis. *Expert Rev. Proteomics* **2011**, *8* (2), 197–209.

(42)   Gross, J. H. *Mass Spectrometry*; Springer International Publishing: Cham, **2017**; Vol. 56.

(43)   Rayleigh, Lord. XX. On the Equilibrium of Liquid Conducting Masses Charged with Electricity. *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **1882**, *14* (87), 184–186.

(44)   Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9.

(45)   Karas, M.; Bahr, U.; Dülcks, T. Nano-Electrospray Ionization Mass Spectrometry: Addressing Analytical Problems beyond Routine. *Fresenius. J. Anal. Chem.* **2000**,

*366* (6–7), 669–676.

(46)    Iribarne, J. V. On the Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.* **1976**, *64* (6), 2287.

(47)    Kaltashov, I. A.; Mohimen, A. Estimates of Protein Surface Areas in Solution by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2005**, *77* (16), 5370–5379.

(48)    McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. Release of Native-like Gaseous Proteins from Electrospray Droplets via the Charged Residue Mechanism: Insights from Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2015**, *137* (39), 12667–12676.

(49)    Metwally, H.; Duez, Q.; Konermann, L. Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry. *Anal. Chem.* **2018**, *90* (16), 10069–10077.

(50)    Miller, P. E.; Denton, M. B. The Quadrupole Mass Filter: Basic Operating Concepts. *J. Chem. Educ.* **1986**, *63* (7), 617.

(51)    Kero, F. A.; Pedder, R. E.; Yost, R. A. Quadrupole Mass Analyzers: Theoretical and Practical Considerations. In *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*; **2009**.

(52)    Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a Travelling Wave-Based Radio-Frequency-Only Stacked Ring Ion Guide. *Rapid Commun. Mass Spectrom.* **2004**, *18* (20), 2401–2414.

(53)    Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. An Investigation of the Mobility Separation of Some Peptide and Protein Ions Using a New Hybrid Quadrupole/Travelling Wave IMS/Oa-ToF Instrument. *Int. J. Mass Spectrom.* **2007**, *261* (1), 1–12.

(54)    Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in Travelling Wave Ion

Mobility Resolution. *Rapid Commun. Mass Spectrom.* **2011**, *25* (11), 1559–1566.

(55)  Shukla, A. K.; Futrell, J. H. Tandem Mass Spectrometry: Dissociation of Ions by Collisional Activation. *J. Mass Spectrom.* **2000**, *35* (9), 1069–1090.

(56)  Guilhaus, M. Special Feature: Tutorial. Principles and Instrumentation in Time-of-Flight Mass Spectrometry. Physical and Instrumental Concepts. *J. Mass Spectrom.* **1995**, *30* (11), 1519–1532.

(57)  Marshall, A. G. Fourier Transform Mass Spectrometry. *Spectrosc. Biomed. Sci.* **2018**, 87–105.

(58)  Perry, R. H.; Cooks, R. G.; Noll, R. J. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. *Mass Spectrom. Rev.* **2008**, *27* (6), 661–699.

(59)  Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **2000**, *72* (6), 1156–1162.

(60)  Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.

(61)  Wang, W.; Kitova, E. N.; Klassen, J. S. Influence of Solution and Gas Phase Processes on Protein-Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2003**, *75* (19), 4945–4955.

(62)  El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Protein Interactions with Isomeric Carbohydrate Ligands Using a Catch and Release Electrospray Ionization-Mass Spectrometry Assay. *Anal. Chem.* **2013**, *85* (16), 7637–7644.

(63)  El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Applications of a Catch and Release Electrospray Ionization Mass Spectrometry Assay for Carbohydrate Library Screening. *Anal. Chem.* **2012**, *84* (1), 50–58.

(64) Wigger, M.; Eyler, J. R.; Benner, S. A.; Li, W.; Marshall, A. G. Fourier Transform-Ion Cyclotron Resonance Mass Spectrometric Resolution, Identification, and Screening of Non-Covalent Complexes of Hck Src Homology 2 Domain Receptor and Ligands from a 324-Member Peptide Combinatorial Library. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (10), 1162–1169.

(65) Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C.; Jia, F.; Cairo, C. W.; Klassen, J. S. Protein-Glycosphingolipid Interactions Revealed Using Catch-and-Release Mass Spectrometry. *Anal. Chem.* **2012**, *84* (18), 7618–7621.

(66) Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Klassen, J. S. Identifying Carbohydrate Ligands of a Norovirus P Particle Using a Catch and Release Electrospray Ionization Mass Spectrometry Assay. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (1), 111–119.

(67) Rezaei Darestani, R.; Winter, P.; Kitova, E. N.; Tuszynski, J. A.; Klassen, J. S. Screening Anti-Cancer Drugs against Tubulin Using Catch-and-Release Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (5), 876–885.

(68) Li, J.; Fan, X.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Eugenio, L.; Ng, K. K. S.; Xiong, Z. J.; Privé, G. G.; Klassen, J. S. Screening Glycolipids Against Proteins in Vitro Using Picodiscs and Catch-and-Release Electrospray Ionization-Mass Spectrometry. *Anal. Chem.* **2016**, *88* (9), 4742–4750.

(69) Han, L.; Kitova, E. N.; Li, J.; Nikjah, S.; Lin, H.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. Protein-Glycolipid Interactions Studied in Vitro Using ESI-MS and Nanodiscs: Insights into the Mechanisms and Energetics of Binding. *Anal. Chem.* **2015**, *87* (9), 4888–4896.

(70) Wang, Y.; Park, H.; Lin, H.; Kitova, E. N.; Klassen, J. S. Multipronged ESI-MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins. *Anal. Chem.* **2019**, *91* (3), 2140–2147.

(71)    Lane, L. A.; Ruotolo, B. T.; Robinson, C. V; Favrin, G.; Benesch, J. L. P. A Monte Carlo Approach for Assessing the Specificity of Protein Oligomers Observed in Nano-Electrospray Mass Spectra. *Int. J. Mass Spectrom.* **2009**, *283* (1–3), 169–177.

(72)    Van Dongen, W. D.; Heck, A. J. R. Binding of Selected Carbohydrates to Apo-Concanavalin A Studied by Electrospray Ionization Mass Spectrometry. *Analyst* **2000**, *125* (4), 583–589.

(73)    El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. Quantifying Labile Protein-Ligand Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (11), 1893–1899.

(74)    Sun, J.; Kitova, E. N.; Klassen, J. S. Method for Stabilizing Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2007**, *79* (2), 416–425.

(75)    Liu, L.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Fatty Acid Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (2), 310–318.

(76)    Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Gas Phase Stabilization of Noncovalent Protein Complexes Formed by Electrospray Ionization. *Anal. Chem.* **2009**, *81* (18), 7801–7806.

(77)    Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. Top-down ESI-ECD-FT-ICR Mass Spectrometry Localizes Noncovalent Protein-Ligand Binding Sites. *J. Am. Chem. Soc.* **2006**, *128* (45), 14432–14433.

(78)    Zhang, H.; Lu, H.; Chingin, K.; Chen, H. Stabilization of Proteins and Noncovalent Protein Complexes during Electrospray Ionization by Amino Acid Additives. *Anal. Chem.* **2015**, *87* (14), 7433–7438.

(79)    Landreh, M.; Alvelius, G.; Johansson, J.; Jörnvall, H. Protective Effects of Dimethyl Sulfoxide on Labile Protein Interactions during Electrospray Ionization.

*Anal. Chem.* **2014**, *86* (9), 4135–4139.

(80)    Hossain, B. M.; Konermann, L. Pulsed Hydrogen/Deuterium Exchange MS/MS for Studying the Relationship between Noncovalent Protein Complexes in Solution and in the Gas Phase after Electrospray Ionization. *Anal. Chem.* **2006**, *78* (5), 1613–1619.

(81)    Wang, W.; Kitova, E. N.; Klassen, J. S. Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Factors Influencing Their Formation and Stability. *Anal. Chem.* **2005**, *77* (10), 3060–3071.

(82)    Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. Method for Distinguishing Specific from Nonspecific Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2006**, *78* (9), 3010–3018.

(83)    Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. Nonspecific Interactions Between Proteins and Charged Biomolecules in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (3), 472–481.

(84)    Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. Method for Identifying Nonspecific Protein-Protein Interactions in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2007**, *79* (21), 9301–9311.

(85)    Daubenfeld, T.; Bouin, A. P.; van der Rest, G. A Deconvolution Method for the Separation of Specific Versus Nonspecific Interactions in Noncovalent Protein-Ligand Complexes Analyzed by ESI-FT-ICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (9), 1239–1248.

(86)    Shimon, L.; Sharon, M.; Horovitz, A. A Method for Removing Effects of Nonspecific Binding on the Distribution of Binding Stoichiometries: Application to Mass Spectroscopy Data. *Biophys. J.* **2010**, *99* (5), 1645–1649.

(87)    Sun, N.; Sun, J.; Kitova, E. N.; Klassen, J. S. Identifying Nonspecific Ligand Binding in Electrospray Ionization Mass Spectrometry Using the Reporter Molecule Method. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (7), 1242–1250.

(88) Kitova, E. N.; Soya, N.; Klassen, J. S. Identifying Specific Small-Molecule Interactions Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2011**, *83* (13), 5160–5167.

(89) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. Direct Quantification of Protein-Metal Ion Affinities by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2010**, *82* (6), 2170–2174.

(90) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. Influence of Response Factors on Determining Equilibrium Association Constants of Non-Covalent Complexes by Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2003**, *38* (5), 491–501.

(91) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. Affinities of Shiga Toxins 1 and 2 for Univalent and Oligovalent Pk-Trisaccharide Analogs Measured by Electrospray Ionization Mass Spectrometry. *Glycobiology* **2007**, *17* (10), 1127–1137.

(92) Chitta, R. K.; Rempel, D. L.; Gross, M. L. Determination of Affinity Constants and Response Factors of the Noncovalent Dimer of Gramicidin by Electrospray Ionization Mass Spectrometry and Mathematical Modeling. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (7), 1031–1038.

(93) Lin, H.; Kitova, E. N.; Klassen, J. S. Quantifying Protein-Ligand Interactions by Direct Electrospray Ionization-Ms Analysis: Evidence of Nonuniform Response Factors Induced by High Molecular Weight Molecules and Complexes. *Anal. Chem.* **2013**, *85* (19), 8919–8922.

## Chapter 2

(1) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49.

(2) Helenius, A. Intracellular Functions of N-Linked Glycans. *Science (80-. ).* **2002**, *291* (5512), 2364–2369.

(3) Paulson, J. C.; Blixt, O.; Collins, B. E. Sweet Spots in Functional Glycomics. *Nat.*

*Chem. Biol.* **2006**, *2* (5), 238–248.

(4)     Grigorian, A.; Demetriou, M. Manipulating Cell Surface Glycoproteins by Targeting N-Glycan-Galectin Interactions. In *Methods in Enzymology*; 2010; Vol. 480, pp 245–266.

(5)     Cederfur, C.; Salomonsson, E.; Nilsson, J.; Halim, A.; Öberg, C. T.; Larson, G.; Nilsson, U. J.; Leffler, H. Different Affinity of Galectins for Human Serum Glycoproteins: Galectin-3 Binds Many Protease Inhibitors and Acute Phase Proteins. *Glycobiology* **2008**, *18* (5), 384–394.

(6)     Rillahan, C. D.; Paulson, J. C. Glycan Microarrays for Decoding the Glycome. *Annu. Rev. Biochem.* **2011**, *80* (1), 797–823.

(7)     Song, X.; Heimburg-Molinaro, J.; Smith, D. F.; Cummings, R. D. Glycan Microarrays of Fluorescently-Tagged Natural Glycans. *Glycoconj. J.* **2015**, *32* (7), 465–473.

(8)     Chan, Y. C.; Lin, H. Y.; Tu, Z.; Kuo, Y. H.; Hsu, S. T. D.; Lin, C. H. Dissecting the Structure–Activity Relationship of Galectin–Ligand Interactions. *Int. J. Mol. Sci.* **2018**, *19* (2), 1–20.

(9)     Grant, O. C.; Smith, H. M.; Firsova, D.; Fadda, E.; Woods, R. J. Presentation, Presentation, Presentation! Molecular-Level Insight into Linker Effects on Glycan Array Screening Data. *Glycobiology* **2014**, *24* (1), 17–25.

(10)    Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; St-Pierre, Y.; Klassen, J. S. Human Milk Oligosaccharide Specificities of Human Galectins. Comparison of Electrospray Ionization Mass Spectrometry and Glycan Microarray Screening Results. *Anal. Chem.* **2017**, *89* (9), 4914–4921.

(11)    He, X. G., Gerona-Navarro, G., Jaffrey, S. R. Ligand Discovery Using Small Molecule Microarrays. *J. Pharmacol. Exp. Ther.* **2004**, *313* (1), 1–7.

(12)    Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein-Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet?

*J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441.

(13) El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Carbohydrate-Protein Interactions by Electrospray Ionization Mass Spectrometry Analysis. *Biochemistry* **2012**, *51* (21), 4244–4253.

(14) Haramija, M.; Peter-Katalinić, J. Quantitative Characterization of Galectin-3-C Affinity Mass Spectrometry Measurements: Comprehensive Data Analysis, Obstacles, Shortcuts and Robustness. *Rapid Commun. Mass Spectrom.* **2017**, *31* (20), 1709–1719.

(15) Jovanović, M.; Peter-Katalinić, J. Preliminary Mass Spectrometry Characterization Studies of Galectin-3 Samples, Prior to Carbohydrate-Binding Studies Using Affinity Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (1), 129–136.

(16) Fermas, S.; Gonnet, F.; Varenne, A.; Gareil, P.; Daniel, R. Frontal Analysis Capillary Electrophoresis Hyphenated to Electrospray Ionization Mass Spectrometry for the Characterization of the Antithrombin/Heparin Pentasaccharide Complex. *Anal. Chem.* **2007**, *79* (13), 4987–4993.

(17) Maple, H. J.; Garlish, R. A.; Rigau-Roca, L.; Porter, J.; Whitcombe, I.; Prosser, C. E.; Kennedy, J.; Henry, A. J.; Taylor, R. J.; Crump, M. P.; et al. Automated Protein-Ligand Interaction Screening by Mass Spectrometry. *J. Med. Chem.* **2012**, *55* (2), 837–851.

(18) Zhang, S.; Van Pelt, C. K.; Wilson, D. B. Quantitative Determination of Noncovalent Binding Interactions Using Automated Nanoelectrospray Mass Spectrometry. *Anal. Chem.* **2003**, *75* (13), 3010–3018.

(19) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. Applications of a Catch and Release Electrospray Ionization Mass Spectrometry Assay for Carbohydrate Library Screening. *Anal. Chem.* **2012**, *84* (1), 50–58.

(20) Han, L.; Morales, L. C.; Richards, M. R.; Kitova, E. N.; Sipione, S.; Klassen, J. S.

Investigating the Influence of Membrane Composition on Protein-Glycolipid Binding Using Nanodiscs and Proxy Ligand Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2017**, *89* (17), 9330–9338.

(21)    Han, L.; Kitova, E. N.; Li, J.; Nikjah, S.; Lin, H.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. Protein-Glycolipid Interactions Studied in Vitro Using ESI-MS and Nanodiscs: Insights into the Mechanisms and Energetics of Binding. *Anal. Chem.* **2015**, *87* (9), 4888–4896.

(22)    El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. Quantifying Ligand Binding to Large Protein Complexes Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2012**, *84* (9), 3867–3870.

(23)    Han, L.; Tan, M.; Xia, M.; Kitova, E. N.; Jiang, X.; Klassen, J. S. Gangliosides Are Ligands for Human Noroviruses. *J. Am. Chem. Soc.* **2014**, *136* (36), 12631–12637.

(24)    El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. Quantifying Protein Interactions with Isomeric Carbohydrate Ligands Using a Catch and Release Electrospray Ionization-Mass Spectrometry Assay. *Anal. Chem.* **2013**, *85* (16), 7637–7644.

(25)    El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; Bode, L.; Hage, N.; Falcone, F. H.; Klassen, J. S. Screening Natural Libraries of Human Milk Oligosaccharides against Lectins Using CaR-ESI-MS. *Analyst* **2018**, *143* (2), 536–548.

(26)    El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E. N.; St-Pierre, Y.; Klassen, J. S. High-Throughput Label- and Immobilization-Free Screening of Human Milk Oligosaccharides Against Lectins. *Anal. Chem.* **2017**, *89* (17), 8713–8722.

(27)    Li, J.; Fan, X.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Eugenio, L.; Ng, K. K. S.; Xiong, Z. J.; Privé, G. G.; Klassen, J. S. Screening Glycolipids Against Proteins in Vitro Using Picodiscs and Catch-and-Release Electrospray Ionization-Mass Spectrometry. *Anal. Chem.* **2016**, *88* (9), 4742–4750.

(28)    Fatunmbi, O.; Abzalimov, R. R.; Savinov, S. N.; Gershenson, A.; Kaltashov, I. A.

Interactions of Haptoglobin with Monomeric Globin Species: Insights from Molecular Modeling and Native Electrospray Ionization Mass Spectrometry. *Biochemistry* **2016**, *55* (12), 1918–1928.

(29)    Engel, N. Y.; Weiss, V. U.; Marchetti-Deschmann, M.; Allmaier, G. NES GEMMA Analysis of Lectins and Their Interactions with Glycoproteins – Separation, Detection, and Sampling of Noncovalent Biospecific Complexes. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 77–86.

(30)    Kitova, E. N.; El-Hawiet, A.; Klassen, J. S. Screening Carbohydrate Libraries for Protein Interactions Using the Direct ESI-MS Assay. Applications to Libraries of Unknown Concentration. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (11), 1908–1916.

(31)    Kovak, M. R.; Saraswati, S.; Goddard, S. D.; Diekman, A. B. Proteomic Identification of Galectin-3 Binding Ligands and Characterization of Galectin-3 Proteolytic Cleavage in Human Prostasomes. *Andrology* **2013**, *1* (5), 682–691.

(32)    Bresalier, R. S.; Byrd, J. C.; Tessler, D.; Lebel, J.; Koomen, J.; Hawke, D.; Half, E.; Liu, K. F.; Mazurek, N. A Circulating Ligand for Galectin-3 Is a Haptoglobin-Related Glycoprotein Elevated in Individuals with Colon Cancer. *Gastroenterology* **2004**, *127* (3), 741–748.

(33)    Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. Structure of a Single-Chain Antibody Variable Domain (Fv) Fragment Complexed with a Carbohydrate Antigen at 1.7-A Resolution. *Proc. Natl. Acad. Sci.* **1994**, *91* (14), 6423–6427.

(34)    Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. Method for Distinguishing Specific from Nonspecific Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2006**, *78* (9), 3010–3018.

(35)    Morelle, W.; Michalski, J.-C. Analysis of Protein Glycosylation by Mass Spectrometry. *Nat. Protoc.* **2007**, *2* (7), 1585–1602.

(36)    Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L.

P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87* (8), 4370–4376.

(37)   Imre, T.; Kremmer, T.; Héberger, K.; Molnár-Szöllosi, É.; Ludányi, K.; Pócsfalvi, G.; Malorni, A.; Drahos, L.; Vékey, K. Mass Spectrometric and Linear Discriminant Analysis of N-Glycans of Human Serum Alpha-1-Acid Glycoprotein in Cancer Patients and Healthy Individuals. *J. Proteomics* **2008**, *71* (2), 186–197.

(38)   Clerc, F.; Reiding, K. R.; Jansen, B. C.; Kammeijer, G. S. M.; Bondt, A.; Wuhrer, M. Human Plasma Protein N-Glycosylation. *Glycoconj. J.* **2016**, *33* (3), 309–343.

(39)   Kawasaki, T., Koyama, J., Yamashina, I. Isolation and Characterization of A1-Acid Glycoprotein from Rat Serum. *J. Biochem.* **1966**, *60*, 554–560.

(40)   Luo, Z.; Lei, H.; Sun, Y.; Liu, X.; Su, D. F. Orosomucoid, an Acute Response Protein with Multiple Modulating Activities. *Journal of Physiology and Biochemistry*. 2015, pp 329–340.

(41)   Giménez, E.; Benavente, F.; Barbosa, J.; Sanz-Nebot, V. Towards a Reliable Molecular Mass Determination of Intact Glycoproteins by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21* (16), 2555–2563.

(42)   Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (35), 8763–8768.

(43)   Felitsyn, N.; Kitova, E. N.; Klassen, J. S. Thermal Decomposition of a Gaseous Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behavior. *Anal. Chem.* **2001**, *73* (19), 4647–4661.

(44)   Yan, J.; Zhou, M.; Gilbert, J. D.; Wolff, J. J.; Somogyi, Á.; Pedder, R. E.; Quintyn, R. S.; Morrison, L. J.; Easterling, M. L.; Paša-Tolić, L.; et al. Surface-Induced

Dissociation of Protein Complexes in a Hybrid Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Anal. Chem.* **2017**, *89* (1), 895–901.

(45) Lipiski, M.; Deuel, J. W.; Baek, J. H.; Engelsberger, W. R.; Buehler, P. W.; Schaer, D. J. Human Hp1-1 and Hp2-2 Phenotype-Specific Haptoglobin Therapeutics Are Both Effective In Vitro and in Guinea Pigs to Attenuate Hemoglobin Toxicity. *Antioxid. Redox Signal.* **2013**, *19* (14), 1619–1633.

(46) Lee, S. H.; Jeong, S.; Lee, J.; Yeo, I. S.; Oh, M. J.; Kim, U.; Kim, S.; Kim, S. H.; Park, S. Y.; Kim, J. H.; et al. Glycomic Profiling of Targeted Serum Haptoglobin for Gastric Cancer Using Nano LC/MS and LC/MS/MS. *Mol. Biosyst.* **2016**, *12* (12), 3611–3621.

(47) Mollan, T. L.; Jia, Y.; Banerjee, S.; Wu, G.; Timothy Kreulen, R.; Tsai, A. L.; Olson, J. S.; Crumbliss, A. L.; Alayash, A. I. Redox Properties of Human Hemoglobin in Complex with Fractionated Dimeric and Polymeric Human Haptoglobin. *Free Radic. Biol. Med.* **2014**, *69*, 265–277.

(48) Jensen, P. E. H.; Sottrup-Jensens, L. Primary Structure of Human Alpha 2-Macroglobulin. Complete Disulfide Bridge Assignment and Localization of Two Interchain Bridges in the Dimeric Proteinase Binding Unit. 1986, pp 15863–15869.

(49) Rehman, A. A.; Ahsan, H.; Khan, F. H. Alpha-2-Macroglobulin: A Physiological Guardian. *J. Cell. Physiol.* **2013**, *228* (8), 1665–1675.

(50) Hirabayashi, J. Oligosaccharide Specificity of Galectins: A Search by Frontal Affinity Chromatography. *Biochim. Biophys. Acta - Gen. Subj.* **2002**, *1572* (2–3), 232–254.

(51) Arnold, J. N.; Wallis, R.; Willis, A. C.; Harvey, D. J.; Royle, L.; Dwek, R. A.; Rudd, P. M.; Sim, R. B. Interaction of Mannan Binding Lectin with A2 Macroglobulin via Exposed Oligomannose Glycans: A Conserved Feature of the Thiol Ester Protein Family? *J. Biol. Chem.* **2006**, *281* (11), 6955–6963.

(52)  Mancera-Arteu, M.; Giménez, E.; Barbosa, J.; Sanz-Nebot, V. Identification and Characterization of Isomeric N-Glycans of Human Alfa-Acid-Glycoprotein by Stable Isotope Labelling and ZIC-HILIC-MS in Combination with Exoglycosidase Digestion. *Anal. Chim. Acta* **2016**, *940*, 92–103.

(53)  Stowell, S. R.; Arthur, C. M.; Mehta, P.; Slanina, K. A.; Blixt, O.; Leffler, H.; Smith, D. F.; Cummings, R. D. Galectin-1, -2, and -3 Exhibit Differential Recognition of Sialylated Glycans and Blood Group Antigens. *J. Biol. Chem.* **2008**, *283* (15), 10109–10123.

(54)  Patnaik, S. K.; Potvin, B.; Carlsson, S.; Sturm, D.; Leffler, H.; Stanley, P. Complex N-Glycans Are the Major Ligands for Galectin-1, -3, and -8 on Chinese Hamster Ovary Cells. *Glycobiology* **2006**, *16* (4), 305–317.

## Chapter 3

(1)  Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. *Nat. Chem. Biol.* **2010**, *6* (10), 713–723.

(2)  Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49.

(3)  Zhao, Y. Y.; Takahashi, M.; Gu, J. G.; Miyoshi, E.; Matsumoto, A.; Kitazume, S.; Taniguchi, N. Functional Roles of N-Glycans in Cell Signaling and Cell Adhesion in Cancer. *Cancer Science*. **2008**, 1304–1310.

(4)  Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell*. **2006**, 855–867.

(5)  Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nature Reviews Cancer*. **2015**, 540–555.

(6)  Stowell, S. R.; Ju, T.; Cummings, R. D. Protein Glycosylation in Cancer. *Annu. Rev. Pathol. Mech. Dis.* **2015**, *10* (1), 473–510.

(7)  Stick, R. V.; Williams, S. J. Glycoproteins and Proteoglycans. In *Carbohydrates: The Essential Molecules of Life*; **2009**, 369–412.

(8)     Wopereis, S.; Lefeber, D. J.; Morava, É.; Wevers, R. A. Mechanisms in Protein O-Glycan Biosynthesis and Clinical and Molecular Aspects of Protein O-Glycan Biosynthesis Defects: A Review. *Clinical Chemistry*. **2006**, 574–600.

(9)     Hooker, A. D.; James, D. C. Analysis of Glycoprotein Heterogeneity by Capillary Electrophoresis and Mass Spectrometry. *Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol.* **2000**, *14* (3), 241–249.

(10)    Pantazaki, A.; Taverna, M.; Vidal-Madjar, C. Recent Advances in the Capillary Electrophoresis of Recombinant Glycoproteins. *Analytica Chimica Acta*. **1999**, 137–156.

(11)    Thormann, W.; Molteni, S.; Caslavska, J.; Schmutz, A. Clinical and Forensic Applications of Capillary Electrophoresis. *Electrophoresis* **1994**, *15* (1), 3–12.

(12)    Geyer, H.; Geyer, R. Strategies for Analysis of Glycoprotein Glycosylation. *Biochimica et Biophysica Acta - Proteins and Proteomics*. **2006**, 1853–1869.

(13)    Yang, Z.; Hancock, W. S. Monitoring Glycosylation Pattern Changes of Glycoproteins Using Multi-Lectin Affinity Chromatography. *J. Chromatogr. A* **2005**, *1070* (1–2), 57–64.

(14)    Hu, S.; Wong, D. T. Lectin Microarray. *Proteomics - Clin. Appl.* **2009**, *3* (2), 148–154.

(15)    Wang, H.; Li, H.; Zhang, W.; Wei, L.; Yu, H.; Yang, P. Multiplex Profiling of Glycoproteins Using a Novel Bead-Based Lectin Array. *Proteomics* **2014**, *14* (1), 78–86.

(16)    Wuhrer, M. Glycomics Using Mass Spectrometry. *Glycoconj. J.* **2013**, *30* (1), 11–22.

(17)    Wuhrer, M.; Deelder, A. M.; Hokke, C. H. Protein Glycosylation Analysis by Liquid Chromatography-Mass Spectrometry. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. **2005**, 124–133.

(18)    Kolarich, D.; Jensen, P. H.; Altmann, F.; Packer, N. H. Determination of

Site-Specific Glycan Heterogeneity on Glycoproteins. *Nat. Protoc.* **2012**, *7* (7), 1285–1298.

(19) Ongay, S.; Boichenko, A.; Govorukhina, N.; Bischoff, R. Glycopeptide Enrichment and Separation for Protein Glycosylation Analysis. *J. Sep. Sci.* **2012**, *35* (18), 2341–2372.

(20) Zamfir, A. D.; Flangea, C.; Serb, A.; Zagrean, A. M.; Rizzi, A. M.; Sisu, E. *Mass Spectrometry of Glycoproteins*; Kohler, J. J., Patrie, S. M., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, **2013**; Vol. 951.

(21) Zhang, N.; Liu, S.; Liu, Z.; Wan, D.; Cui, M.; Liu, H. Mass Spectrometry-Based Analysis of Glycoproteins and Its Clinical Applications in Cancer Biomarker Discovery. *Clin. Proteomics* **2014**, *11* (1), 14.

(22) Yu, A.; Zhao, J.; Peng, W.; Banazadeh, A.; Williamson, S. D.; Goli, M.; Huang, Y.; Mechref, Y. Advances in Mass Spectrometry-Based Glycoproteomics. *Electrophoresis*. **2018**, 3104–3122.

(23) Li, F.; Glinskii, O. V.; Glinsky, V. V. Glycobioinformatics: Current Strategies and Tools for Data Mining in MS-Based Glycoproteomics. *Proteomics* **2013**, *13* (2), 341–354.

(24) Cui, W.; Rohrs, H. W.; Gross, M. L. Top-down Mass Spectrometry: Recent Developments, Applications and Perspectives. *Analyst* **2011**, *136* (19), 3854.

(25) Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (35), 8763–8768.

(26) Urien, S.; Albengres, E.; Zini, R.; Tillement, J. P. Evidence for Binding of Certain Acidic Drugs to A1-Acid Glycoprotein. *Biochem. Pharmacol.* **1982**, *31* (22), 3687–3689.

(27) Urien, S.; Albengres, E.; Pinquier, J.-L.; Tillement, J.-P. Role of Alpha-1 Acid Glycoprotein, Albumin, and Nonesterified Fatty Acids in Serum Binding of

Apazone and Warfarin. *Clin. Pharmacol. Ther.* **1986**, *39* (6), 683–689.

(28) Ongay, S.; Neusüß, C. Isoform Differentiation of Intact AGP from Human Serum by Capillary Electrophoresis-Mass Spectrometry. *Anal. Bioanal. Chem.* **2010**, *398* (2), 845–855.

(29) Israili, Z. H.; Dayton, P. G. Human Alpha-1-Glycoprotein and Its Interactions with Drugs. *Drug Metabolism Reviews*. **2001**, 161–235.

(30) Saroha, A.; Biswas, S.; Chatterjee, B. P.; Das, H. R. Altered Glycosylation and Expression of Plasma Alpha-1-Acid Glycoprotein and Haptoglobin in Rheumatoid Arthritis. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2011**, *879* (20), 1839–1843.

(31) Imre, T.; Kremmer, T.; Héberger, K.; Molnár-Szöllosi, É.; Ludányi, K.; Pócsfalvi, G.; Malorni, A.; Drahos, L.; Vékey, K. Mass Spectrometric and Linear Discriminant Analysis of N-Glycans of Human Serum Alpha-1-Acid Glycoprotein in Cancer Patients and Healthy Individuals. *J. Proteomics* **2008**, *71* (2), 186–197.

(32) Tanabe, K.; Kitagawa, K.; Kojima, N.; Iijima, S. Multifucosylated Alpha-1-Acid Glycoprotein as a Novel Marker for Hepatocellular Carcinoma. *J. Proteome Res.* **2016**, *15* (9), 2935–2944.

(33) Ayyub, A.; Saleem, M.; Fatima, I.; Tariq, A.; Hashmi, N.; Musharraf, S. G. Glycosylated Alpha-1-Acid Glycoprotein 1 as a Potential Lung Cancer Serum Biomarker. *Int. J. Biochem. Cell Biol.* **2016**, *70*, 68–75.

(34) Colombo, S.; Buclin, T.; Décosterd, L. A.; Telenti, A.; Furrer, H.; Lee, B. L.; Biollaz, J.; Eap, C. B. Orosomucoid (A1-Acid Glycoprotein) Plasma Concentration and Genetic Variants: Effects on Human Immunodeficiency Virus Protease Inhibitor Clearance and Cellular Accumulation. *Clin. Pharmacol. Ther.* **2006**, *80* (4), 307–318.

(35) Taguchi, K.; Nishi, K.; Giam Chuang, V. T.; Maruyama, T.; Otagiri, M. Molecular Aspects of Human Alpha-1 Acid Glycoprotein — Structure and Function. In *Acute*

*Phase Proteins*; InTech, 2013; Vol. i, p 13.

(36)   Nakano, M. Detailed Structural Features of Glycan Chains Derived from 1-Acid Glycoproteins of Several Different Animals: The Presence of Hypersialylated, O-Acetylated Sialic Acids but Not Disialyl Residues. *Glycobiology* **2004**, *14* (5), 431–441.

(37)   Mancera-Arteu, M.; Gim Enez, E.; Barbosa, J. E.; Oria Sanz-Nebot, V. Identification and Characterization of Isomeric N-Glycans of Human Alfa-Acid-Glycoprotein by Stable Isotope Labelling and ZIC-HILIC-MS in Combination with Exoglycosidase Digestion. *Anal. Chim. Acta* **2016**, *940*, 92–103.

(38)   Schönfeld, D. L.; Ravelli, R. B. G.; Mueller, U.; Skerra, A. The 1.8-Å Crystal Structure of A1-Acid Glycoprotein (Orosomucoid) Solved by UV RIP Reveals the Broad Drug-Binding Activity of This Human Plasma Lipocalin. *J. Mol. Biol.* **2008**, *384* (2), 393–405.

(39)   Nishi, K.; Ono, T.; Nakamura, T.; Fukunaga, N.; Izumi, M.; Watanabe, H.; Suenaga, A.; Maruyama, T.; Yamagata, Y.; Curry, S.; et al. Structural Insights into Differences in Drug-Binding Selectivity between Two Forms of Human A1-Acid Glycoprotein Genetic Variants, the A and F1*S Forms. *J. Biol. Chem.* **2011**, *286* (16), 14427–14434.

(40)   Nakagawa, T.; Kishino, S.; Itoh, S.; Sugawara, M.; Miyazaki, K. Differential Binding of Disopyramide and Warfarin Enantiomers to Human A1-Acid Glycoprotein Variants. *Br. J. Clin. Pharmacol.* **2003**, *56* (6), 664–669.

(41)   Shams-Ud-Doha, K.; Kitova, E. N.; Kitov, P. I.; St-Pierre, Y.; Klassen, J. S. Human Milk Oligosaccharide Specificities of Human Galectins. Comparison of Electrospray Ionization Mass Spectrometry and Glycan Microarray Screening Results. *Anal. Chem.* **2017**, *89* (9), 4914–4921.

(42)   Fernandes, C. L.; Ligabue-Braun, R.; Verli, H. Structural Glycobiology of Human A1-Acid Glycoprotein and Its Implications for Pharmacokinetics and

Inflammation. *Glycobiology* **2015**, *25* (10), 1125–1133.

## Chapter 4

(1)   Chiu, K. M.; Mortensen, R. F.; Osmand, A. P.; Gewurzt, H.; Gewurz, H. Interactions of Alpha1 Acid Glycoprotein with the Immune System. I Purification and Effects upon Lymphocyte Responsiveness. *Immunology* **1977**, *32* (6), 997–1005.

(2)   Yang, Y.; Liu, F.; Franc, V.; Halim, L. A.; Schellekens, H.; Heck, A. J. R. Hybrid Mass Spectrometry Approaches in Glycoprotein Analysis and Their Usage in Scoring Biosimilarity. *Nat. Commun.* **2016**, *7*.

(3)   Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci.* **2018**, *115* (35), 8763–8768.

(4)   Siegel, R. L.; Miller, K. D.; Jemal, A. Cancer Statistics, 2016. *CA. Cancer J. Clin.* **2016**, *66* (1), 7–30.

(5)   Watt, K. W.; Lee, P. J.; M'Timkulu, T.; Chan, W. P.; Loor, R. Human Prostate-Specific Antigen: Structural and Functional Similarity with Serine Proteases. *Proc. Natl. Acad. Sci.* **1986**, *83* (10), 3166–3170.

(6)   Stamey, T. A.; Yang, N.; Hay, A. R.; McNeal, J. E.; Freiha, F. S.; Redwine, E. Prostate-Specific Antigen as a Serum Marker for Adenocarcinoma of the Prostate. *N. Engl. J. Med.* **1987**, *317* (15), 909–916.

(7)   Shariat, S. F.; Semjonow, A.; Lilja, H.; Savage, C.; Vickers, A. J.; Bjartell, A. Tumor Markers in Prostate Cancer I: Blood-Based Markers. *Acta Oncologica*. **2011**, 61–75.

(8)   Gilgunn, S.; Conroy, P. J.; Saldova, R.; Rudd, P. M.; O'Kennedy, R. J. Aberrant PSA Glycosylation - A Sweet Predictor of Prostate Cancer. *Nature Reviews Urology*. **2013**, 99–107.

(9)   Andriole, G. L.; Crawford, E. D.; Grubb, R. L.; Buys, S. S.; Chia, D.; Church, T.

R.; Fouad, M. N.; Gelmann, E. P.; Kvale, P. A.; Reding, D. J.; et al. Mortality Results from a Randomized Prostate-Cancer Screening Trial. *N. Engl. J. Med.* **2009**, *360* (13), 1310–1319.

(10) Li, Y.; Tian, Y.; Rezai, T.; Prakash, A.; Lopez, M. F.; Chan, D. W.; Zhang, H. Simultaneous Analysis of Glycosylated and Sialylated Prostate-Specific Antigen Revealing Differential Distribution of Glycosylated Prostate-Specific Antigen Isoforms in Prostate Cancer Tissues. *Anal. Chem.* **2011**, *83* (1), 240–245.

(11) Pelzer, A. E.; Volgger, H.; Bektic, J.; Berger, A. P.; Rehder, P.; Bartsch, G.; Horninger, W. The Effect of Percentage Free Prostate-Specific Antigen (PSA) Level on the Prostate Cancer Detection Rate in a Screening Population with Low PSA Levels. *BJU Int.* **2005**, *96* (7), 995–998.

(12) Stephan, C.; Ralla, B.; Jung, K. Prostate-Specific Antigen and Other Serum and Urine Markers in Prostate Cancer. *Biochimica et Biophysica Acta - Reviews on Cancer*. August **2014**, 99–112.

(13) Meany, D. L.; Chan, D. W. Aberrant Glycosylation Associated with Enzymes as Cancer Biomarkers. *Clinical Proteomics*. **2011**, 7.

(14) Bélanger, A.; Van Halbeek, H.; Graves, H. C. B.; Grandbois, K.; Stamey, T. A.; Huang, L.; Poppe, I.; Labrie, F. Molecular Mass and Carbohydrate Structure of Prostate Specific Antigen: Studies for Establishment of an International PSA Standard. *Prostate* **1995**, *27* (4), 187–197.

(15) Ohyama, C.; Hosono, M.; Nitta, K.; Oh-eda, M.; Yoshikawa, K.; Habuchi, T.; Arai, Y.; Fukuda, M. Carbohydrate Structure and Differential Binding of Prostate Specific Antigen to Maackia Amurensis Lectin between Prostate Cancer and Benign Prostate Hypertrophy. *Glycobiology* **2004**, *14* (8), 671–679.

(16) Llop, E.; Ferrer-Batallé, M.; Barrabés, S.; Guerrero, P. E.; Ramírez, M.; Saldova, R.; Rudd, P. M.; Aleixandre, R. N.; Comet, J.; de Llorens, R.; et al. Improvement of Prostate Cancer Diagnosis by Detecting PSA Glycosylation-Specific Changes.

*Theranostics* **2016**, *6* (8), 1190–1204.

(17)   Yoneyama, T.; Ohyama, C.; Hatakeyama, S.; Narita, S.; Habuchi, T.; Koie, T.; Mori, K.; Hidari, K. I. P. J.; Yamaguchi, M.; Suzuki, T.; et al. Measurement of Aberrant Glycosylation of Prostate Specific Antigen Can Improve Specificity in Early Detection of Prostate Cancer. *Biochem. Biophys. Res. Commun.* **2014**, *448* (4), 390–396.

(18)   Saldova, R.; Fan, Y.; Fitzpatrick, J. M.; Watson, R. W. G.; Rudd, P. M. Core Fucosylation and A2-3 Sialylation in Serum N-Glycome Is Significantly Increased in Prostate Cancer Comparing to Benign Prostate Hyperplasia. *Glycobiology* **2011**, *21* (2), 195–205.

(19)   Dwek, M. V.; Jenks, A.; Leathem, A. J. C. A Sensitive Assay to Measure Biomarker Glycosylation Demonstrates Increased Fucosylation of Prostate Specific Antigen (PSA) in Patients with Prostate Cancer Compared with Benign Prostatic Hyperplasia. *Clin. Chim. Acta* **2010**, *411* (23–24), 1935–1939.

(20)   Fukushima, K.; Satoh, T.; Baba, S.; Yamashita, K. A1,2-Fucosylated and β-N-Acetylgalactosaminylated Prostate-Specific Antigen as an Efficient Marker of Prostatic Cancer. *Glycobiology* **2009**, *20* (4), 452–460.

(21)   Etxebarria, J.; Reichardt, N. C. Methods for the Absolute Quantification of N-Glycan Biomarkers. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (8), 1676–1687.

(22)   Hunter, C. D.; Khanna, N.; Richards, M. R.; Rezaei Darestani, R.; Zou, C.; Klassen, J. S.; Cairo, C. W. Human Neuraminidase Isoenzymes Show Variable Activities for 9- O -Acetyl-Sialoside Substrates. *ACS Chem. Biol.* **2018**, *13* (4), 922–932.