

**Quantitative Characterization of Noncovalent Protein-Carbohydrate  
Interactions using Electrospray Ionization Mass Spectrometry**

by

Ling Han

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

Doctor of Philosophy

Department of Chemistry  
University of Alberta

© Ling Han, 2015

## Abstract

The interactions between water-soluble proteins and carbohydrates found on the surfaces of cells play important roles in many physiological and pathological cellular processes. Carbohydrates function as receptors for signaling, cellular recognition and adhesion, and pathogen infections. This thesis focuses on the development and application of electrospray ionization mass spectrometry (ESI-MS) methods to detect and quantify protein-carbohydrate interactions in vitro.

In Chapter 2, the intrinsic affinities (per binding site) of the protruding domain dimer (P dimer, 69 kDa) of the human norovirus (NoV) strain VA387 to a panel of 47 soluble analogs of histo-blood group antigens (HBGAs) were quantified using the direct ESI-MS assay. Our results revealed that the P dimer exhibits a broad specificity for the HBGAs and bind, albeit weakly (intrinsic association constants ( $K_{a,int}$ ) of  $10^2 - 10^3 \text{ M}^{-1}$ ), to all of the oligosaccharides tested. Overall, the A and B antigens exhibit stronger binding than the H and Lewis antigens. In addition, the affinities are also affected by the precursor chain type of HBGAs but not by the chain length.

In Chapter 3, the applicability of the catch-and-release (CaR)-ESI-MS assay for screening carbohydrate libraries against large protein complexes was demonstrated for the first time. Libraries containing as many as 146 compounds were screened against NoV VA387 subviral P particle (24-mer, 865 kDa). Notably, the results of the screening experiments revealed NoV interactions with oligosaccharides with structures found in human milk and the cell wall of mycobacteria. The affinities of these newly discovered ligands are comparable to those of the HBGA receptors.

In Chapters 4 and 6, the direct ESI-MS assay was combined with a competitive binding strategy in order to measure the affinities of protein-carbohydrate interactions that can't be directly quantified by ESI-MS. In Chapter 4, the affinities of the NoV P particle and virus-like

particle (VLP, 180-mer, 10.5 MDa) to HBGA ligands were quantified using the *proxy protein* ESI-MS method, which utilizes competitive protein binding. The results revealed that HBGA ligands exhibit similar affinities for the P particle and P dimer whereas the HBGA affinities for the VLP are consistently higher than those measured for the P dimer, but within a factor of three. In Chapter 6, the *proxy ligand* ESI-MS assay, which is on the basis of competitive ligand binding, combined with nanodisc technology to solubilize glycolipids was used to determine the interactions of cholera toxin B subunit homopentamer (CTB<sub>5</sub>) with GM1, and a family 51 carbohydrate-binding module (CBM) with B type 2 tetrasaccharide neoglycolipid. A notable finding of this study is that the affinities of the glycolipid ligands in the nanodisc are lower, by a factor of  $\leq 5$ , than those of the corresponding oligosaccharides in solution.

In Chapter 5, the screening using CaR-ESI-MS assay revealed the first evidence that human NoVs bind to gangliosides. Moreover, affinities measurements were reported for the NoV VA387 P dimer, P particle and VLP, and VA115 P dimer for a series of ganglioside oligosaccharides. Notably, the ganglioside affinities are similar in magnitude to those of HBGA receptors for NoVs. Additional confirmation of NoV-ganglioside interactions was provided by the binding measurements using the enzyme-linked immunosorbent assays.

Finally, also in Chapter 6, a systematic ESI-MS investigation aimed at elucidating the processes that influence binding of water-soluble proteins to glycolipids incorporated into nanodiscs was described. The interactions of CTB<sub>5</sub> to GM1 nanodiscs studied by ESI-MS indicated that proteins bind reversibly to nanodisc-associated glycolipids, and that proteins possessing multiple ligand binding sites are able to interact with glycolipids originating from different nanodiscs. Moreover, the nature of the protein-glycolipid complexes detected by ESI-MS is likely to be influenced by the diffusion of glycolipids between nanodiscs.

## Preface

My research related to the human noroviruses (Chapters 2 to 5 in this thesis) was conducted in collaboration with Professors Ming Tan and Xi Jiang at the Cincinnati Children's Hospital Medical Center. Chapter 2 has been published as: 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*, 276-285. I was responsible for all the experiments and data analysis, as well as preparation of the manuscript. P. Kitov assisted with the molecular simulations and interpretation of the binding results. Expression and purification of the single chain monoclonal antibody was done by E. Kitova; and she also assisted with my experiments. M. Tan, L. Wang, M. Xia and X. Jiang were responsible for the expression and purification of norovirus P dimers.

Chapter 3 has been published as: Han, L.; Kitova, E.; Tan, M.; Jiang, X.; Klassen, J., "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*, 111-119. I performed all the experiments and prepared the manuscript. E. Kitova provided input on the manuscript. M. Tan and X. Jiang provided the norovirus P particle used in this study.

Chapter 4 has been published as: Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S., "Affinities of human histo-blood group antigens for norovirus capsid protein complexes". *Glycobiology* **2015**, *25*, 170-180. I performed all the ESI-MS measurements, analyzed the data, and prepared the manuscript. The binding models used in this study were developed by me and E. Kitova. I also expressed and purified the glycosyltransferase

(GTA) used in this study. Additionally, expression and purification of the norovirus capsid proteins; and gel filtration analysis of the norovirus virus-like particle were done by M. Tan and X. Jiang. B. Pluinage and A. Boraston were responsible for the expression and purification of the family 51 carbohydrate binding module. The single chain monoclonal antibody was provided by E. Kitova after purification.

Chapter 5 has been published as: 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*, 12631-12637. I was responsible for doing all the ESI-MS measurements and analyzing data. Expression and purification of the norovirus capsid proteins, as well as the enzyme-linked immunosorbent assay were conducted by M. Tan, M. Xia and X. Jiang. In addition, M. Tan, X. Jiang, E. Kitova and I were involved in the manuscript preparation.

Chapter 6 has been published as: Han, L.; Kitova, E. N.; Li, J.; Nikjah, S.; Lin, H.; Pluinage, 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*, 4888-4896. I performed all the experiments, analyzed the data, and prepared the manuscript. The binding models for *proxy ligand* ESI-MS assay were developed by me, with several useful ideas provided by E. Kitova and H. Lin. Nanodiscs were prepared by E. Kitova and S. Nikjah; and picodiscs were made by J. Li. Additionally, B. Pluinage and A. Boraston provided the family 51 carbohydrate binding module.

The supervisory author, J. Klassen, was involved throughout the projects in concept formation and manuscript composition.

## **Acknowledgement**

Five years ago, I arrived at Edmonton and joined Klassen group with my dream of being a successful chemist in the future. Now I am on the stage to finish all the work here. Looking backwards, my PhD studies are accompanied with a lot of help from many people. To all of them I wish to express my sincere gratitude.

First, I would like to express the great gratitude to my supervisor, Professor John Klassen, for his guidance, encouragement, and patience throughout my PhD studies. During these years, he inspires me thinking of the significance and new aspects of the projects; and offers me freedom to carry out my research. His ideas and advice always bring me to the goal, and from this point of view, I am a lucky PhD student.

I also would like to thank other members of my supervisory and examining committee, Prof. Alex Brown, Prof. Jed Harrison, Prof. Nils Petersen and Prof. Robert Campbell for their invaluable comments and advice on my research. Special thanks must go to Prof. Joseph Loo from University of California, Los Angeles, for his participation of my defense and the time spent reviewing my thesis.

I am very grateful to my collaborators, Prof. Ming Tan and Prof. Xi Jiang at Cincinnati Children's Hospital Medical Center for their collaborative work and contributions on the norovirus projects, as well as helpful discussions. I would like to thank Prof. Todd Lowary and his group who generously provided carbohydrate ligands for my library screening projects. I thank Prof. Alisdair Boraston at University of Victoria and Prof. Christopher Cairo for providing us the precious

carbohydrate-binding proteins. I am thankful to Dr. Pavel Kitov for the help with molecular simulations and other meaningful discussions. I also thank Blake Zheng and Gareth Lambkin for their professional trainings and assistance for the biological analyses.

It is my pleasure to have worked with many excellent colleagues in Klassen group. I would like to thank Dr. Elena Kitova, who helps my bench work and is very patient and kind to me. I would like to bring my acknowledgments to Dr. Rambod Daneshfar, Dr. Glen Shoemaker, Dr. Naoto Soya, Dr. Amr El-Hawiet, Dr. Lan Liu and Dr. Lu Deng as they taught me the first lesson in the lab. I also appreciate Jun Li and Sanaz Nikjah, who have been involved in many aspects of the protein-glycolipid interaction projects and shared ideas. Moreover, I would like to acknowledge the support and friendship from other labmates in Klassen group, present or past: Dr. Aneika Leney, Dr. Mickey Richards, Yixuan Zhang, Jingjing Zhang, Km Shams-Ud-Doha, Reza Rezaei Darestani, Xuxin Fan, Yuyu Yao, Yajie Chen, Nian Sun, Hong Lin, Nobar Jalili and Brittany Granoski.

I must also thank the Alberta Glycomics Centre for generous funding and the Alberta Innovates Technology Futures for a graduate scholarship.

Finally, I am deeply thankful to my parents and my wife, Di Wu, who have always been a source of encouragement and motivation for my life and career.

... And farewell! The great time that I spent at University of Alberta!

## Table of Contents

Chapter 1 Quantitative Characterization of Noncovalent Protein-Carbohydrate Interactions using Electrospray Ionization Mass Spectrometry.....	1
1.1 Introduction.....	1
1.1.1 Biological importance of carbohydrates .....	1
1.1.2 Protein-carbohydrate interactions .....	2
1.1.3 Model membrane systems for studying protein-glycolipid interactions.....	6
1.2 Electrospray Ionization Mass Spectrometry .....	8
1.2.1 ESI mechanisms.....	8
1.2.2 Nanoflow ESI-MS .....	13
1.2.3 MS instrumentation.....	13
1.2.3.1 Hybrid quadrupole-ion mobility separation-time of flight mass spectrometer .....	13
1.2.3.2 Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer ...	21
1.3 ESI-MS Based Methods for Quantifying Noncovalent Protein-Carbohydrate Interactions .....	26
1.3.1 <i>Direct</i> ESI-MS assay.....	26
1.3.2 <i>Indirect</i> ESI-MS assays.....	29
1.4 Potential Pitfalls of ESI-MS Binding Assays.....	30
1.4.1 In-source dissociation .....	30
1.4.2 Nonspecific ligand binding.....	32
1.4.3 Non-uniform response factors.....	35
1.4.4 Other challenges.....	36
1.5 The Present Work.....	37
1.6 References.....	40
Chapter 2 Affinities of Recombinant Norovirus P Dimers for Human Histo-Blood Group Antigens .....	53
2.1 Introduction.....	53
2.2 Experimental Section.....	55
2.2.1 Proteins .....	55

2.2.2 Carbohydrates .....	56
2.2.3 Mass spectrometry .....	57
2.2.4 Determination of $K_a$ values.....	58
2.3 Results and Discussion .....	61
2.3.1 HBGA affinities for NoV P dimers.....	61
2.3.2 Interpretation of the binding data.....	73
2.4 Conclusion .....	76
2.5 References.....	77
Chapter 3 Identifying Carbohydrate Ligands of a Norovirus P Particle using a Catch and Release Electrospray Ionization Mass Spectrometry Assay .....	81
3.1 Introduction.....	81
3.2 Experimental .....	84
3.2.1 Proteins .....	84
3.2.2 Carbohydrates .....	85
3.2.3 Mass spectrometry .....	91
3.3 Results and Discussion .....	93
3.3.1 Validating the CaR-ESI-MS assay.....	93
3.3.2 Screening carbohydrate libraries against the NoV P particle.....	100
3.3.2.1 <i>Library1</i> .....	100
3.3.2.2 <i>Library2</i> .....	103
3.4 Conclusions.....	106
3.5 References.....	107
Chapter 4 Affinities of Human Histo-Blood Group Antigens for Norovirus Capsid Protein Complexes.....	112
4.1 Introduction.....	112
4.2 Experimental Section.....	115
4.2.1 Proteins .....	115
4.2.2 Carbohydrates .....	117
4.2.3 Mass spectrometry .....	117
4.2.4 Gel filtration chromatography.....	118

4.2.5 Determination of $K_a$ values.....	119
4.2.5.1 <i>Direct ESI-MS assay</i> .....	119
4.2.5.2 <i>Proxy protein ESI-MS assay</i> .....	119
4.2.5.3 <i>Application of the proxy protein ESI-MS method using <math>P_{proxy}</math> that possesses multiple ligand binding sites</i> .....	120
4.3 Results and Discussion .....	123
4.3.1 ESI-MS analysis of human NoV VA387 P particle and VLP .....	123
4.3.2 Affinities of HBGA oligosaccharides for the human NoV VA387 P particle and VLP .....	126
4.4 Conclusion .....	149
4.5 References.....	149
Chapter 5 Gangliosides are Ligands for Human Noroviruses .....	154
5.1 Introduction.....	154
5.2 Experimental Section.....	157
5.2.1 Proteins .....	157
5.2.2 Ligands.....	159
5.2.3 Mass spectrometry .....	164
5.2.3.1 <i>Catch-and-Release ESI-MS assay</i> .....	165
5.2.3.2 <i>Direct ESI-MS assay</i> .....	166
5.2.3.3 <i>Proxy protein ESI-MS method</i> .....	167
5.2.4 Enzyme-linked immunosorbent assay (ELISA) .....	168
5.3 Results and Discussion .....	168
5.3.1 Ganglioside binding to NoV VA387 P particle.....	168
5.3.2 Ganglioside affinities for NoV VA387 capsid proteins .....	174
5.3.3 Ganglioside affinities for NoV VA115 P dimer .....	183
5.3.4 Binding of sialic acid-containing glycoconjugates to NoV VA387 and VA115 ..	183
5.4 Conclusions.....	185
5.5 References.....	185
Chapter 6 Protein-Glycolipid Interactions Studied in vitro using ESI-MS and Nanodiscs. Insights into the Mechanisms and Energetics of Binding.....	190
6.1 Introduction.....	190

6.2 Experimental Section .....	192
6.2.1 Proteins .....	192
6.2.2 Phospholipids, glycolipids and oligosaccharides.....	194
6.2.3 Preparation of nanodiscs .....	197
6.2.4 Preparation of picodiscs.....	197
6.2.5 Mass spectrometry .....	198
6.2.5.1 <i>Direct ESI-MS assay</i> .....	199
6.2.5.2 <i>Proxy ligand ESI-MS assay</i> .....	199
6.2.5.3 <i>Application of proxy ligand ESI-MS method to quantify CTB<sub>5</sub>-GM1 ND interactions</i> .....	202
6.2.6 Ultracentrifugation and SDS-PAGE .....	207
6.3 Results and Discussion .....	209
6.3.1 CTB <sub>5</sub> binding to GM1 nanodiscs.....	209
6.3.1.1 CTB <sub>5</sub> -GM1 nanodisc interactions revealed by ESI-MS.....	210
6.3.1.2 Reversibility of CTB <sub>5</sub> -GM1 nanodisc interactions.....	217
6.3.1.3 Release of CTB <sub>5</sub> -GM1 complexes from nanodiscs in the gas phase.....	219
6.3.1.4 Mechanism of CTB <sub>5</sub> -GM1 nanodisc binding.....	227
6.3.2 CBM binding to B <sub>2</sub> NGL in nanodiscs .....	234
6.3.3 Protein affinities for glycolipids in nanodiscs – the <i>proxy ligand</i> ESI-MS assay	238
6.4 Conclusions.....	245
6.5 References.....	246
Chapter 7 Conclusions and Future Work .....	249
7.1 Conclusions.....	249
7.2 Future work.....	254
7.2.1 Rapid and quantitative screening of oligosaccharide libraries against carbohydrate-binding proteins using the <i>proxy protein</i> ESI-MS assay.....	255
7.2.2 Quantifying protein-glycolipid interactions in different lipid environments.....	257
7.2.3 Dissecting multivalent protein-glycolipid interactions.....	258
7.3 References.....	260
List of References .....	262

## List of Tables

<b>Table 2.1.</b>	Intrinsic association constants, $K_{a,int}$ ( $M^{-1}$ ) for binding of the HBGA oligosaccharides ( <b>L1-L47</b> ) and HMOs ( <b>L48-L49</b> ) with norovirus VA387 P dimer and truncated (tr)-P dimer, measured at 25 °C and pH 7 by the direct ESI-MS assay. ....	65-67
<b>Table 2.2.</b>	Intrinsic association constants, $K_{a,int}$ ( $M^{-1}$ ) for binding of the HBGA oligosaccharides with norovirus VA387 P dimer measured by the direct ESI-MS assay. All measurements were performed at 25 °C and pH 7 in 10 mM ammonium acetate buffer with /without 10 mM imidazole. ....	72
<b>Table 3.1.</b>	Structures and molecular weights (MWs) of components of carbohydrate library. ....	86-91
<b>Table 4.1.</b>	Association constants ( $K_a$ ) for binding of the HBGA oligosaccharides ( <b>L1 – L14</b> ) with CBM, Gal-3C and GTA measured at 25 °C and pH 7 by the ESI-MS assay. ....	132
<b>Table 4.2.</b>	Intrinsic association constants ( $K_{a,int}$ ) for HBGA oligosaccharides ( <b>L1 – L14</b> ) binding to human NoV VA387 P dimer, P particle and VLP, measured at 25 °C and pH 7 using the ESI-MS <i>proxy protein</i> assay. ....	141

<b>Table 5.1.</b>	Intrinsic (per binding site) association constants ( $K_{a,int}$ ) for P dimer and the oligosaccharides of thirteen gangliosides measured in aqueous ammonium acetate (200 mM) at pH 7 and 25 °C using the direct ESI-MS assay. ....	176
<b>Table 5.2.</b>	Intrinsic association constants ( $K_{a,int}$ ) for the NoV VA387 P dimer and the histo-blood group type 3 oligosaccharides (H3, A3 and B3) in an aqueous ammonium acetate solution (200 mM) at pH 7 and 25 °C measured using the direct ESI-MS assay performed in negative ion mode. ....	177
<b>Table 5.3.</b>	Summary of results from screening of ganglioside oligosaccharide against the P particle of human NoV VA387 in an aqueous ammonium acetate solution (200 mM) at pH 7 and 25 °C using the CaR-ESI-MS assay. ....	178

## List of Figures

<b>Figure 1.1.</b>	Schematic representation of ESI process operated in positive ion mode, adapted from reference 80. ....	9
<b>Figure 1.2.</b>	Different ESI models proposed for the formation of gas phase ions. (a) IEM: Small ion ejection from a charged nanodroplet. (b) CRM: Formation of a folded protein into the gas phase. (c) CEM: Ejection of disordered macromolecule. Figure is adapted from reference 91. ....	12
<b>Figure 1.3.</b>	A schematic diagram of the Synapt G2S Q-IMS-TOF mass spectrometer, adapted from the Waters user's manual. ....	14
<b>Figure 1.4.</b>	(a) Stability diagram (adapted from reference 96) of quadrupole as a function of DC and RF voltages for ions with different $m/z$ values $[(m/z)_1 < (m/z)_2 < (m/z)_3]$ . (b) Schematic diagram of the quadrupole used in the Waters Synapt mass spectrometers. ....	16
<b>Figure 1.5.</b>	Diagrams of (a) stacked ring ion guide (SRIG) and (b) its operational principle. Figure is adapted from reference 100. ....	20
<b>Figure 1.6.</b>	Schematic diagram of the Bruker ApexQe 9.4-Tesla FT-ICR mass spectrometer used in this study. Figure is reproduced from the Bruker user's manual. ....	22
<b>Figure 1.7.</b>	Cyclotron motions of a positive ion and a negative ion in magnetic field (B). .....	24

<b>Figure 1.8.</b>	Illustration of ion excitation, image current detection and generation of mass spectrum by FT-ICR MS. Figure is adapted from the Bruker user's manual. ...24
<b>Figure 1.9.</b>	Illustration of the nonspecific protein-ligand interactions during the ESI process. Figure is adapted from reference 46. ....33
<b>Figure 2.1.</b>	Direct ESI-MS analysis of norovirus VA387 P dimer and truncated-P dimer at pH 7 and 25 °C. Representative mass spectra acquired for 10 mM ammonium acetate solution with (a) P dimer (12 μM) and (b) truncated P dimer (12 μM). Three truncated-P dimer species (tr-P <sub>2</sub> (I), tr-P <sub>2</sub> (II) and tr-P <sub>2</sub> (III)) were resolved. ....62
<b>Figure 2.2.</b>	ESI mass spectra in positive ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing (a) norovirus VA387 P dimer (12 μM) with 50 μM <b>L31</b> (B type 3 tetrasaccharide, MW 801 Da); and (b) truncated-P dimer (12 μM) with 60 μM <b>L31</b> . A P <sub>ref</sub> (10 μM) was added to each solution to correct the mass spectra for the occurrence nonspecific carbohydrate-protein binding during ESI process. Insets: Normalized distribution of <b>L31</b> bound to the P dimer or three truncated-P dimer species (tr-P <sub>2</sub> (I), tr-P <sub>2</sub> (II) and tr-P <sub>2</sub> (III)) after correcting the ESI mass spectra for nonspecific ligand binding. ....64
<b>Figure 2.3.</b>	ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing norovirus VA387 P dimer (12 μM) with (a) 20 μM <b>L21</b> (A type 6 tetrasaccharide, MW 691 Da), (b) 70 μM <b>L21</b> , and (c) 100 μM <b>L21</b> . A P <sub>ref</sub> (10 μM) was added to each solution to correct the mass spectra for the occurrence nonspecific carbohydrate-protein binding

during ESI process. Insets: Normalized distribution of **L21** (at concentrations of (a) 20  $\mu\text{M}$ , (b) 70  $\mu\text{M}$  and (c) 100  $\mu\text{M}$ ) bound to the P dimer after correcting the ESI mass spectra for nonspecific ligand binding. ....69

**Figure 2.4.** Fraction of ligand-bound P dimer (i.e.,  $f = (R_1+2R_2)/(1+R_1+R_2)/2$ , after correction for the nonspecific ligand binding) versus ligand concentration measured for **L8**, **L21** and **L34**, which represent H, A and B type 6 antigens, respectively. The titration experiments were carried out on aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing P dimer (12  $\mu\text{M}$ ),  $P_{\text{ref}}$  (10  $\mu\text{M}$ ) and ligand concentrations of between 10 and 120  $\mu\text{M}$ . The solid curves correspond to the best fit of eq 2.8 to the experimental data for each ligand. The errors bars represent one standard deviation. ....70

**Figure 2.5.** (a) ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate (10 mM) solution at pH 7 and 25 °C containing norovirus VA387 P dimer (12  $\mu\text{M}$ ), **L21** (40  $\mu\text{M}$ ),  $P_{\text{ref}}$  (10  $\mu\text{M}$ ) and imidazole (10 mM). Inset: Normalized distribution of **L21** bound to the P dimer after correction for nonspecific binding determined from the mass spectra. ....71

**Figure 2.6.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing norovirus VA387 P dimer (12  $\mu\text{M}$ ) and 80  $\mu\text{M}$  (a) **L48** (LNT) or (b) **L49** (LNnT). A  $P_{\text{ref}}$  (10  $\mu\text{M}$ ) was added to each solution to correct the mass spectra for the occurrence nonspecific carbohydrate-protein binding during the ESI process. Insets: Normalized distributions of **L48** and **L49** bound to the P dimer before and after correction

for nonspecific binding determined from the corresponding mass spectra. ....73

**Figure 2.7.** Interaction model for the norovirus VA387 P dimer binding to HBGA oligosaccharides adapted from the proposed model for norovirus-HBGA binding (reference 4). (a) Binding model proposed for A/B/H oligosaccharides; illustrated by an A/B/H type 1 antigen. (b) Binding model proposed for Lewis oligosaccharides; illustrated by a Le<sup>b</sup> antigen. This model assumes that VA387 P dimer can accommodate both the H epitope and A/B epitope independently in two nearby binding pockets. In addition, the VA387 P dimer can also accommodate Lewis epitope in the H epitope binding pocket. ....75

**Figure 3.1.** ESI mass spectra acquired in negative ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) of 5 μM P particle using a Trap voltage of (a) 5 V and (b) 80 V (a mild CID condition that facilitates removing solvent adducts but without cause fragmentation of the P particle). ....94

**Figure 3.2.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, at pH 7 and 25 °C) of P particle (5 μM) and L1 (B type 3 tetrasaccharide, 10 μM), (b) CID mass spectrum measured for the same solution as in part (a) using a broad (m/z 200 units) quadrupole isolation window centered at -65 charge state (m/z 13 420) and a Trap voltage of 150 V. ....96

**Figure 3.3.** (a) ESI mass spectrum obtained in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5 μM) and L1, L2 and L3 (10 μM each), at pH 7 and 25 °C. (b) and (c) CID mass spectra

measured with a Trap voltage of 200 V for the P particle and its carbohydrate complexes using a broad ( $m/z$  200 units) quadrupole isolation window centered at the charge state of (b) -65 ( $m/z$  13 420) and (c) -61 ( $m/z$  14 200). .....97

**Figure 3.4.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of P particle (5  $\mu$ M) and **L20** (10  $\mu$ M). (b) CID mass spectrum acquired for solution in part (a) using a broad ( $m/z$  200 units) quadrupole isolation window centered at  $m/z$  13 420 (-65 charge state). (c) CID mass spectrum acquired for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of P particle (5  $\mu$ M) and **L20** (10  $\mu$ M) using a broad ( $m/z$  200 units) quadrupole isolation window centered at  $m/z$  13 420. For (b) and (c), a Trap voltage of 150 V was used. ....99

**Figure 3.5.** (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5  $\mu$ M) and *Library1* (10  $\mu$ M each), at pH 7 and 25 °C. (b) CID mass spectrum of the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 150 V. (c) Arrival time distributions measured for the deprotonated **L1** and **L5** ions ( $m/z$  800.4) following release from the P particle (*post-release*) and the deprotonated **L1** and **L5** ions obtained directly from solution (*reference*). ..... 101

**Figure 3.6.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of P particle (5  $\mu$ M) and **L14**, **L15**, **L16**, **L17** and **L18** (20  $\mu$ M each). (b) CID mass spectrum

acquired for solution in part (a) using a broad ( $m/z$  200 units) quadrupole isolation window centered at  $m/z$  13 420 (-65 charge state) and a Trap voltage of 150 V. .... 103

**Figure 3.7.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5  $\mu$ M) and *Library2* (10  $\mu$ M each), at pH 7 and 25 °C. (b) Corresponding CID mass spectrum of the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 180 V. (c) Arrival time distributions measured for the deprotonated **L11** and **L51** ions ( $m/z$  1014.5) following release from the P particle (*post-release*) and the deprotonated **L11** and **L51** ions obtained directly from solution (*reference*). .... 105

**Figure 4.1.** (a) ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate (200 mM, pH 7 and 25 °C) solution with (a) 3  $\mu$ M P particle (corresponding to 72  $\mu$ M monomer) and (b) 0.2  $\mu$ M VLP (corresponding to 36  $\mu$ M VP1) of human NoV VA387. (b) and (c) ESI mass spectra acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 0.2  $\mu$ M (corresponding to 36  $\mu$ M VP1) NoV VA387 VLP. The measurements in (b) and (c) were carried out on two different days. The measurements were acquired using a Waters Synapt G2S mass spectrometer. .... 125

**Figure 4.2.** Chromatograph of gel filtration of the human NoV VA387 VLP (5 mg mL<sup>-1</sup>) in 1x PBS (pH 7.4, 25°C) buffer. The observation of only a single peak

(corresponding to the void volume at ~45 mL) indicates the capsid protein of NoV VA387 predominantly assembles into large complexes, with molecular weights  $\geq 800$  kDa. The Superdex 200 gel filtration column was calibrated as described previously (reference 7)..... 126

**Figure 4.3.** ESI mass spectra obtained in positive ion mode for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing P dimer (P<sub>2</sub>, 12 μM), (b) GTA (10 μM), (c) CBM (12 μM) and (d) Gal-3C (GL, 5 μM). The measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer. .... 127

**Figure 4.4.** Plot of fraction of ligand-bound CBM versus ligand concentration measured for (a) L1 (B trisaccharide), (b) L2 (A trisaccharide), (c) L3 (B type 2 tetrasaccharide), (d) L4 (A type 2 tetrasaccharide), (e) L5 (B type 5 tetrasaccharide), (f) L6 (A type 5 tetrasaccharide), (g) L7 (B type 6 tetrasaccharide), (h) L8 (A type 6 tetrasaccharide), and (i) L12 (B type 3 tetrasaccharide). The ESI-MS binding measurements were carried out on 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing CBM (12 μM), Ubq (P<sub>ref</sub>, 8 μM) and each ligand at a minimum of eight different concentrations ranging from 2.5 to 100 μM. The solid curves correspond to the best fit of eq 1.10 to the experimental data and the error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer. .... 129

**Figure 4.5.** Plot of fraction of ligand-bound Gal-3C versus ligand concentration measured for (a) L9 (B type 1 tetrasaccharide), (b) L10 (A type 1 tetrasaccharide), (c)

**L11** (H type 6 trisaccharide), and (d) **L14** (type 2 trisaccharide). The ESI-MS binding measurements were carried out on 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing Gal-3C (5 μM), Ubq ( $P_{ref}$ , 3 μM) and each ligand at a minimum of nine different concentrations ranging from 2.5 to 100 μM. The solid curves correspond to the best fit of eq 1.10 to the experimental data and the error bars correspond to one standard derivation. These measurements were performed on Waters Synapt G2S mass spectrometer.

..... 131

**Figure 4.6.** Representative ESI mass spectra obtained in positive ion mode for a 50 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12 μM CBM, 8 μM Ubq ( $P_{ref}$ ), 35 μM **L1** (B trisaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. (d) Plot of  $R_{proxy}$  versus concentration of monomer in the P particle. The solution condition for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to the best fit of eq 4.1b for the experimental data. The error bars correspond to one standard deviation. The measurements were carried out using a Bruker ApexQe FT-ICR mass spectrometer. .... 134

**Figure 4.7.** Representative ESI mass spectra obtained in positive ion mode for a 50 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 10 μM GTA, 10 μM scFv ( $P_{ref}$ ), 60 μM **L1** (B trisaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. (d) Plot of  $\theta$  versus concentration of monomer in the P particle. The solution condition for each measurement was same as (a), but with the addition of 0 to 12 μM P particle.

The solid curve corresponds to the best fit of eq 4.11 to the experimental data. The error bars correspond to one standard deviation. The measurements were carried out using a Bruker ApexQe FT-ICR mass spectrometer. .... 135

**Figure 4.8.** Plots of  $R_{\text{proxy}}$  versus P particle concentration measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C ) containing CBM (12  $\mu\text{M}$ ), Ubq ( $P_{\text{ref}}$ , 8  $\mu\text{M}$ ), (a) **L2** (A trisaccharide, 30  $\mu\text{M}$ ), (b) **L3** (B type 2 tetrasaccharide, 40  $\mu\text{M}$ ), (c) **L4** (A type 2 tetrasaccharide, 40  $\mu\text{M}$ ), (d) **L5** (B type 5 tetrasaccharide, 35  $\mu\text{M}$ ), (e) **L6** (A type 5 tetrasaccharide, 30  $\mu\text{M}$ ), (f) **L7** (B type 6 tetrasaccharide, 50  $\mu\text{M}$ ) and (g) **L8** (A type 6 tetrasaccharide, 40  $\mu\text{M}$ ), and P particle (0 – 16  $\mu\text{M}$ , which corresponds to 0 – 384  $\mu\text{M}$  of monomer). The solid curves correspond to the best fit of eq 4.1b to the experimental data for each ligand. The error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer. .... 136

**Figure 4.9.** Plot of the  $\theta$  versus P particle concentration measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing GTA (10  $\mu\text{M}$ ), scFv ( $P_{\text{ref}}$ , 10  $\mu\text{M}$ ), **L13** (H disaccharide, 40  $\mu\text{M}$ ) and human NoV VA387 P particle (0 – 12  $\mu\text{M}$ , which corresponds to 0 – 288  $\mu\text{M}$  of monomer). The solid curve corresponds to the best fit of eq 4.11 to the experimental data. The error bars correspond to one standard derivation. These measurements were performed on Bruker ApexQe FT-ICR mass spectrometer..... 137

**Figure 4.10.** Representative ESI mass spectra measured for 200 mM aqueous ammonium

acetate solutions (pH 7 and 25 °C) containing 5 μM Gal-3C (GL), 3 μM Ubiquitin (Ubq, P<sub>ref</sub>), 20 μM **L10** (A type 1 tetrasaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. The measurements were performed on a Waters Synapt G2S mass spectrometer.

..... 138

**Figure 4.11.** Plots of  $R_{\text{proxy}}$  versus P particle concentration measured for 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C ) containing Gal-3C (5 μM), Ubq (P<sub>ref</sub>, 3 μM), (a) **L9** (B type 1 tetrasaccharide, 20 μM), (b) **L10** (A type 1 tetrasaccharide, 20 μM), and (c) **L11** (H type 6 trisaccharide, 40 μM), and P particle (0 – 12 μM, which corresponds to 0 – 288 μM of monomer). The solid curves correspond to the best fit of eq 4.1b to the experimental data for each ligand. The error bars correspond to one standard derivation. Measurements were performed on a Waters Synapt G2S mass spectrometer. .... 139

**Figure 4.12.** ESI mass spectra measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 12 μM P dimer (P<sub>2</sub>), 10 μM scFv (P<sub>ref</sub>), 40 μM **L12** (B type 3 tetrasaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. (d) Plot of the occupancy function  $\theta$  versus P particle concentration. The solution conditions for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to the best fit of eq 4.11 to the experimental data. The error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer. .... 140

**Figure 4.13.** ESI mass spectra obtained in positive ion mode for aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12  $\mu\text{M}$  CBM, 8  $\mu\text{M}$  Ubq ( $P_{\text{ref}}$ ), 35  $\mu\text{M}$  L1 (B trisaccharide) with 6  $\mu\text{M}$  P particle (24-mer) of human NoV VA387. The concentration of ammonium acetate in (a) was 200 mM and in (b) 800 mM. These measurements were performed on a Waters Synapt G2S mass spectrometer. .... 142

**Figure 4.14.** A representative ESI mass spectrum measured in positive ion mode for 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12  $\mu\text{M}$  VA387 P dimer ( $P_2$ ), 8  $\mu\text{M}$  scFv ( $P_{\text{ref}}$ ) and 80  $\mu\text{M}$  L14 (type 2 trisaccharide). Inset, normalized distribution of L14 bound P dimer before and after correction for nonspecific ligand binding. The measurement was performed on a Waters Synapt G2S mass spectrometer. .... 143

**Figure 4.15.** ESI mass spectra measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 5  $\mu\text{M}$  Gal-3C (GL), 10  $\mu\text{M}$  Ubiquitin (Ubq,  $P_{\text{ref}}$ ), 40  $\mu\text{M}$  L14 (type 2 trisaccharide) with (a) 0  $\mu\text{M}$ , (b) 4  $\mu\text{M}$  and (c) 8  $\mu\text{M}$  P particle (24-mer) of human NoV VA387. (d) Plot of  $R_{\text{proxy}}$  versus P particle concentration. The solution conditions for each measurement was same as (a), but with the addition of 0 to 8  $\mu\text{M}$  P particle. The error bars correspond to one standard derivation. The measurements were performed on a Waters Synapt G2S mass spectrometer. .... 145

**Figure 4.16.** Representative ESI mass spectra obtained in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12  $\mu\text{M}$  CBM,

4  $\mu\text{M}$  Ubq ( $P_{\text{ref}}$ ), 25  $\mu\text{M}$  **L1** (B trisaccharide) with (a) 0 nM, (b) 380 nM and (c) 760 nM VLP (180-mer) of human NoV VA387. The measurements were carried out using a Waters Synapt G2S mass spectrometer. .... 147

**Figure 4.17.** Plots of  $R_{\text{proxy}}$  versus concentration of monomer in the VLP measured for aqueous ammonium acetate solutions (200 mM, pH 7 and 25 °C) containing CBM (12  $\mu\text{M}$ ), Ubq ( $P_{\text{ref}}$ , 4  $\mu\text{M}$ ), VLP (180-mer) of human NoV VA387 (0 – 760 nM) with (a) **L1** (B trisaccharide, 25  $\mu\text{M}$ ), (b) **L2** (A trisaccharide, 20  $\mu\text{M}$ ), (c) **L7** (B type 6 tetrasaccharide, 40  $\mu\text{M}$ ) and (d) **L8** (A type 6 tetrasaccharide, 25  $\mu\text{M}$ ). The solid curves correspond to the best fit of eq 4.1b to the experimental data for each ligand. The error bars correspond to one standard deviation. The measurements were carried out using a Waters Synapt G2S mass spectrometer. .... 148

**Figure 5.1.** Structures of the twenty-component carbohydrate library consisting of the oligosaccharides of gangliosides, globosides and HBGAs and the structures of polyacrylamide (PAA)-conjugated sialic acid-containing oligosaccharides used for the ELISA binding measurements. .... 159-164

**Figure 5.2.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of NoV VA387 P particle (3  $\mu\text{M}$ ) and a twenty-component (10  $\mu\text{M}$  each) carbohydrate library consisting of the oligosaccharides of GM1a, GM1b, GM2, GM3, GD1a, GD1b, GD2, GD3, GT1a, GT1c, GT2, GT3, Fuc-GM1, asialo GM1, asialo GM2, Gb3 and Gb4, as well as the H3, B3 and A3 oligosaccharides. (b) and (c) CID mass

spectrum measured for the -61 and -63 charge states, respectively, of the free and ligand-bound P particle. A Trap voltage of 200 V was used. .... 171

**Figure 5.3.** (a) Arrival time distributions measured for the (a) deprotonated ions of GM1a/GM1b ( $m/z$  997.3) following release from the P particle (post-release) and the deprotonated GM1a and GM1b ions obtained directly from solution (reference) and (b) double deprotonated GD1a/GD1b ions ( $m/z$  644.1) following release from the P particle (post-release) and the GD1a and GD1b ions obtained directly from solution (reference). .... 172

**Figure 5.4.** CID mass spectra acquired in negative ion mode for an aqueous ammonium acetate solutions (200 mM, pH 7 and 25 °C) of P particle (3  $\mu$ M) and 10  $\mu$ M of (a) GD1a and (b) GD1b using a broad (200  $m/z$ ) quadrupole isolation window centered at  $m/z$  14,350. A Trap voltage of 200 V was used. .... 173

**Figure 5.5.** ESI mass spectra acquired in negative ion mode for aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of (a) NoV VA387 P dimer ( $P_2$ , 12  $\mu$ M), GM3 trisaccharide (80  $\mu$ M) and  $P_{ref}$  (4  $\mu$ M) and (b) NoV VA115 P dimer ( $P_2$ , 12  $\mu$ M, MW 67,712 Da), GM3 trisaccharide (80  $\mu$ M) and  $P_{ref}$  (4  $\mu$ M). Insets, normalized distribution of GM3 bound to  $P_2$  of (a) VA387 and (b) VA115) after correction for nonspecific ligand binding. .... 175

**Figure 5.6.** Representative ESI mass spectra measured in positive ion mode for aqueous ammonium acetate solutions (160 mM, pH 7 and 25 °C) of  $P_{proxy}$  (Gal-3C, 3.0  $\mu$ M),  $P_{ref}$  (Ubq, 1.0  $\mu$ M), GM3 trisaccharide (40  $\mu$ M) without (a) or with (b) NoV VA387 VLP (570 nM, 180-mer). Insets show the normalized distributions

of free and GM3-bound  $P_{\text{proxy}}$ , after correction for nonspecific ligand binding. (c) Plot of the abundance ratio of GM3-bound  $P_{\text{proxy}}$  to free  $P_{\text{proxy}}$  ( $R_{\text{proxy}}$ ) versus VLP concentration. The solution conditions for each measurement were the same as in (a), but with the addition of VLP. The curve represents the best fit of eq 5.5 to the experimental data. .... 181

**Figure 5.7.** (a) Representative ESI mass spectrum measured in positive ion mode for aqueous ammonium acetate solution (160 mM, pH 7 and 25 °C) of  $P_{\text{proxy}}$  (Galectin-3, 3.0  $\mu\text{M}$ ),  $P_{\text{ref}}$  (ubiquitin, 1.0  $\mu\text{M}$ ), GM3 trisaccharide (40  $\mu\text{M}$ ) and NoV P particle (6.0  $\mu\text{M}$ , 24-mer). Inset corresponds to normalized distribution of GM3 bound to  $P_{\text{proxy}}$  after correction for nonspecific ligand binding. (b) Plot of abundance ratio of GM3 trisaccharide-bound  $P_{\text{proxy}}$  to free  $P_{\text{proxy}}$  ( $R_{\text{proxy}}$ ) versus P particle concentration. The solution conditions for each measurement were the same as in (a), but with the addition of P particle. The curve represents the best fit of eq 5.5 to the experimental data. ....182

**Figure 5.8.** Binding of NoV VLP, P particles and GST-P fusion protein of VA387, as well as GST-P fusion protein of VA115 to PAA-conjugated GM3 trisaccharide, 6'-sialylacNAc and Neu5Ac in 1x PBS (pH 7.4). GST, which does not show binding to any of the three glycoconjugates, served as a negative control. ... 184

**Figure 6.1.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of (a) CBM (12  $\mu\text{M}$ ) alone or (b) CBM (12  $\mu\text{M}$ ) with B trisaccharide (B-tri, 40  $\mu\text{M}$ ). Inset shows the normalized distributions of free and B-tri-bound CBM measured for the three isoforms

	(CBM-I, -II and -III). .....	193
<b>Figure 6.2.</b>	Amino acid sequence of the recombinant fragment of the family 51 carbohydrate binding module (CBM). .....	194
<b>Figure 6.3.</b>	Structures of the phospholipids, glycolipids and oligosaccharides. ....	195-196
<b>Figure 6.4.</b>	ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3 μM CTB <sub>5</sub> with (a) 0.6 μM and (b) 1.4 μM 10% GM1 ND (corresponding to 12 and 28 μM GM1, respectively). Insets show normalized distributions of free and GM1-bound CTB <sub>5</sub> ; theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1 <sub>os</sub> to CTB <sub>5</sub> . .....	211
<b>Figure 6.5.</b>	ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3 μM CTB <sub>5</sub> with (a) 3 μM and (b) 24.4 μM 0.5% GM1 ND (corresponding to 3 and 24.4 μM GM1, respectively). Insets show normalized distributions of free and GM1-bound CTB <sub>5</sub> ; theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1 <sub>os</sub> to CTB <sub>5</sub> . .....	212
<b>Figure 6.6.</b>	ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB <sub>5</sub> (3 μM) with (a) 6.8 μM and (b) 10.2 μM 1% GM1 ND (corresponding to 13.6 and 20.4 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB <sub>5</sub> . The theoretical distributions were calculated using	

association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>. ..... 213

**Figure 6.7.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) with (a) 2.1 μM and (b) 4.3 μM 2.5% GM1 ND (corresponding to 10.5 and 21.5 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB<sub>5</sub>. The theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>. ..... 214

**Figure 6.8.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) with (a) 1.2 μM and (b) 2.0 μM 5% GM1 ND (corresponding to 12 and 20 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB<sub>5</sub>. The theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>. ..... 215

**Figure 6.9.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 12 μM 0.5% GM1 ND (corresponding to 12 μM GM1) with (a) 3 μM and (c) 6 μM CTB<sub>5</sub>. (b) ESI mass spectrum acquired upon addition of another 3 μM CTB<sub>5</sub> to the solution in (a). (d) Normalized distributions of free and GM1-bound CTB<sub>5</sub> measured from mass spectra in (a) ■, (b) ■ and (c) ▨. .... 218

**Figure 6.10.** Plot of fraction of occupied ligand binding sites in CTB<sub>5</sub> (*f*) versus GM1 concentration measured by ESI-MS in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3 μM CTB<sub>5</sub> and varying concentrations of (a) 0.5%, (b) 1%, (c) 2.5%, (d) 5% and (e) 10% GM1 ND. Dashed lines represent the molar ratio of GM1 to the CTB<sub>5</sub> binding sites. (f) Plot of *f* versus GM1<sub>os</sub> concentration measured by ESI-MS in positive ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 3.8 μM CTB<sub>5</sub> with GM1<sub>os</sub> (1.0 μM – 60 μM). The solid curve corresponds to the theoretical plot calculated using affinities for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub> reported in reference 34. The error bars correspond to one standard deviation. .... 220

**Figure 6.11.** (a) and (b) SDS-PAGE analysis of aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> alone, CTB<sub>5</sub> with empty ND and CTB<sub>5</sub> with 0.5% GM1 NDs subjected to ultracentrifugation using a filter with a MWCO of 100 kDa.; supernatant (MW ≥100 kDa) and filtrate (MW ≤100 kDa). Molecular weight markers (lane 1 and 8); supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) alone (lanes 2 and 3, respectively); supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) with ND (11 μM) containing no GM1 (lanes 4 and 5, respectively); supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) with 0.5% GM1 ND (24 μM) (lanes 6 and 7, respectively); and supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) with 3 μM 0.5% GM1 ND (lanes 9 and 10, respectively). .... 222

**Figure 6.12.** ESI mass spectra acquired in positive ion mode for an aqueous ammonium

acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (5 μM) and 0.5% GM1 ND (24 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa.; (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa). ..... 223

**Figure 6.13.** ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (5 μM) and 0.5% GM1 ND (3 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa: (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa). (c) CID mass spectrum acquired in negative ion mode for ions of  $m/z > 2,500$  produced in (b) using a collision energy of 120 V in the Trap. .. 224

**Figure 6.14.** (a) ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) and 0.5% GM1 ND (14 μM). (b) CID mass spectrum for ions produced in (a) and centred at  $m/z$  11,000 (which correspond to GM1 ND). (c) CID mass spectrum for ions centred at  $m/z$  11,000 produced from aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 0.5% GM1 ND (14 μM). A collision energy of 200 V in Trap was used for the CID experiments. .... 226

**Figure 6.15.** Possible mechanisms for the stepwise binding of CTB<sub>5</sub> to GM1 ND. (a) *Nanodisc recruitment mechanism* - CTB<sub>5</sub> binds irreversibly to GM1 ligands from multiple NDs. (b) *Glycolipid extraction mechanism* - CTB<sub>5</sub> interacts with GM1 in one ND, followed by dissociation of the resulting (CTB<sub>5</sub> + *q*GM1) complex from the original ND and rapid re-binding to GM1 in another ND. (c)

*Glycolipid diffusion mechanism* - GM1 rapidly redistribute between NDs and can be recruited by CTB<sub>5</sub>. Note: to facilitate visualizing multivalent binding, a linear arrangement of subunits is used to represent the CTB<sub>5</sub> homopentamer.  
..... 228

**Figure 6.16.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 0.5% A<sub>2</sub>NGL ND (12 μM) and picodisc (PD, 54 μM) containing POPC and GM1. (b) – (e) CID mass spectra acquired for ions centred at m/z 11,500 produced from: (b) the same solution as in (a); aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of (c) 0.5% A<sub>2</sub>NGL ND (12 μM); (d) PD (54 μM) containing POPC and GM1; and (e) 0.5% A<sub>2</sub>NGL ND (12 μM) and 0.5% GM1 ND (12 μM). For all measurements a collision energy of 200 V in the Trap was used. .... 231-232

**Figure 6.17.** CID mass spectra acquired in negative ion mode for ions centred at m/z 5,500 produced by ESI from an aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of (a) 1% GM1 ND (16 μM) with picodisc (PD, 60 μM) containing POPC; (b) PD (60 μM) containing POPC; and (c) 1% GM1 ND (16 μM). For all measurements, a collisional energy of 120 V in the Trap was used. .... 233

**Figure 6.18.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 μM) with (a) 8 μM and (b) 30.8 μM 10% B<sub>2</sub>NGL ND (corresponding to 160 μM and 616 μM B<sub>2</sub>NGL, respectively). (c) Plots of fraction of ligand-bound CBM (*f*) versus B<sub>2</sub>NGL concentration. The experimental conditions were the same as in (a) and (b), but

with addition of 3.2 – 30.8  $\mu\text{M}$  ND containing 2.5% (●) or 10% (■)  $\text{B}_{2\text{NGL}}$ . The dashed curve represents the theoretical plot calculated from the association constant reported in reference 38 for CBM binding to  $\text{B}_{2\text{os}}$ . The error bars correspond to one standard deviation. .... 236

**Figure 6.19.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12  $\mu\text{M}$ ) and (a) 2.5%  $\text{B}_{2\text{NGL}}$  ND (14  $\mu\text{M}$ ) (corresponding to 70  $\mu\text{M}$   $\text{B}_{2\text{NGL}}$ ) or (b) 2.5%  $\text{B}_{2\text{NGL}}$  ND (30.8  $\mu\text{M}$ ) (corresponding to 154  $\mu\text{M}$   $\text{B}_{2\text{NGL}}$ ). .... 237

**Figure 6.20.** ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12  $\mu\text{M}$ ) and 10%  $\text{B}_{2\text{NGL}}$  ND (25  $\mu\text{M}$ ) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa: (a) supernatant solution ( $\text{MW} \geq 100$  kDa) and (b) filtrate solution ( $\text{MW} \leq 100$  kDa). .... 238

**Figure 6.21.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 12  $\mu\text{M}$  CBM, 40  $\mu\text{M}$  B-tri ( $\text{L}_{\text{proxy}}$ ) with (a) 24  $\mu\text{M}$  15%  $\text{B}_{2\text{NGL}}$  ND and (b) 21  $\mu\text{M}$  10%  $\text{B}_{2\text{NGL}}$  ND (corresponding to 720  $\mu\text{M}$  and 420  $\mu\text{M}$   $\text{B}_{2\text{NGL}}$ , respectively); 5  $\mu\text{M}$   $\text{P}_{\text{ref}}$  (Ubq) was added to each solution to correct for the nonspecific ligand binding during ESI process. (c) Plots of  $R_{\text{proxy}}$  ( $\equiv \text{Ab}(\text{CBM} + \text{B-tri})/\text{Ab}(\text{CBM})$ ) versus  $\text{B}_{2\text{NGL}}$  concentration. The experimental conditions were the same as in (a), but with addition of 0 – 24  $\mu\text{M}$  15%  $\text{B}_{2\text{NGL}}$  ND (●) or 0 – 28  $\mu\text{M}$  10%  $\text{B}_{2\text{NGL}}$  ND (■). The error bars correspond to one standard deviation. .... 240

**Figure 6.22.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 4.4 μM CTB<sub>5</sub> and 20 μM GM1<sub>os</sub> with (a) 0 μM and (b) 2.5 μM 0.5% GM1 ND. Inset shows the normalized distributions of free and GM1<sub>os</sub>-bound CTB<sub>5</sub>. (c) (■) Normalized distributions of free and ligand-bound CTB<sub>5</sub> measured from the mass spectrum in (b); (▨) theoretical distributions were calculated using association constants determined from the *proxy ligand* method and values reported in reference 34 for the stepwise binding of GM1<sub>os</sub> and GM1 to CTB<sub>5</sub>. (d) Plot of  $R_{\text{proxy},5} (\equiv Ab(\text{CTB}_5 + 5\text{GM1}_{\text{os}})/Ab(\text{CTB}_5 + 4\text{GM1}_{\text{os}}))$  versus GM1 concentration. The experimental conditions were the same as in (a) and (b), but with addition of 0 – 2.5 μM 0.5% GM1 ND. The error bars correspond to one standard deviation. .... 243

**Figure 6.23.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (4.0 μM) and GM1<sub>os</sub> (20 μM) with (a) 0 μM and (b) 1.75 μM 1% GM1 ND (corresponding to 3.5 μM GM1). Insets show the normalized distributions of free and GM1<sub>os</sub>-bound CTB<sub>5</sub>. (c) (■) Normalized distributions of free and ligand-bound CTB<sub>5</sub> measured from mass spectrum in (b); (▨) theoretical distributions calculated from the association constants reported for the stepwise binding of GM1<sub>os</sub> (from reference 34) and 1% GM1 ND (determined in the present study) to CTB<sub>5</sub>. (d) Plot of  $R_{\text{proxy},5} (\equiv Ab(\text{CTB}_5 + 5 \text{GM1}_{\text{os}})/Ab(\text{CTB}_5 + 4 \text{GM1}_{\text{os}}))$  versus GM1 concentration. The experimental conditions of are same as in (a) and (b), but with addition of 0 – 1.75 μM 1% GM1 ND. The error bars correspond to one standard deviation. .... 244

**Figure 7.1.** Schematic illustration of the *proxy protein* ESI-MS assay applied for quantifying bindings of a carbohydrate library containing 3 ligands (L1, L2 and L3) to the target protein (P). Simulation mass spectra of the proxy protein ( $P_{\text{proxy}}$ ) with the carbohydrates in the absence (a), and presence (b) of P. .... 256

**Figure 7.2.** Schematic representation of the multivalent CTB<sub>5</sub>-GM1-nanodisc interaction based on the cooperative binding model<sup>10</sup> proposed by Homan and coworkers. The multivalent binding is described by intermolecular ( $K_{a,\text{inter},1}$ ) and intramolecular ( $K_{a,\text{intra},1}$ ,  $K_{a,\text{intra},2}$ ,  $K_{a,\text{intra},3}$ ) association constants. The subscripts 1, 2 and 3 represent the case of GM1 binding to a subunit with zero, one or two ligand-bound nearest neighbour subunits, respectively. Eight distinct species are present in solution. .... 259

## List of Schemes

- Scheme 6.1.** Graphical representation of possible protein-ligand interactions involving a monovalent ligand (L), present in a ND, and a monovalent proxy ligand ( $L_{\text{proxy}}$ ) that binding competitively with a protein (P) with five binding sites. Based on the experimental data described in Figures 6.21b and 6.22b, *vide infra*, only one L is bound to P under the experimental conditions used in this study. The binding model is based on the cooperative model for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub> proposed by Homans,<sup>33</sup> wherein ligand affinity is enhanced when a neighbouring binding site is occupied. The  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  labels are used to distinguish the positional isomers. The intrinsic association constants ( $K_{a,\text{proxy},1}$ ,  $K_{a,\text{proxy},2}$ , and  $K_{a,\text{proxy},3}$  and  $K_{a,1}$ ,  $K_{a,2}$  and  $K_{a,3}$ ), along with the statistic coefficients, are given for each interaction. .... 208-209

## List of Abbreviations

A <sub>2</sub> NGL	A type 2 neoglycolipid
<i>Ab</i>	Abundance of gas-phase ions
ATD	Arrival time distribution
B <sub>2</sub> NGL	B type 2 neoglycolipid
B <sub>2os</sub>	B type 2 tetrasaccharide
BCA assay	Bicinchoninic acid assay
BIRD	Blackbody infrared radiative dissociation
CaR	Catch-and-release
CBM	Carbohydrate-binding module (Family 51)
CEM	Chain ejection model
CID	Collision induced dissociation
CRM	Charge residue model
CTB <sub>5</sub>	Cholera toxin B subunit homopentamer
CV	Calicivirus
Da	Dalton
DC	Direct current
DESI	Desorption electrospray ionization
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DTIMS	Drift time ion mobility spectrometry
ECD	Electron capture dissociation
ELISA	Enzyme-linked immunosorbent assay

ESI	Electrospray ionization
ESI-MS	Electrospray ionization mass spectrometry
ETD	Electron transfer dissociation
FAIMS	Field asymmetric ion mobility spectrometry
FCV	Feline calicivirus
FPLC	Fast performance liquid chromatography
FT	Fourier transform
Fuc	Fucose
Gal	Galactose
Gal-3C	C-terminus of recombinant fragment of human galectin-3
GalNAc	N-acetyl galactosamine
GL	Glycolipid
Glc	Glucose
GlcNAc	N-acetyl glucosamine
GM1 <sub>os</sub>	GM1 pentasaccharide
GST	Glutathione S-transferase
GTA	$\alpha$ -(1 $\rightarrow$ 3)-N-acetyl galactosaminyltransferase
HBGA	Histo-blood group antigen
H-bond	Hydrogen bond
HMO	Human milk oligosaccharide
HRP	Horseradish peroxidase
IC50	Half maximal inhibitory concentration
ICR	Ion cyclotron resonance

IEM	Ion ejection model
IM	Ion mobility
IMS	Ion mobility separation
IRMPD	Infrared radiative multiphoton dissociation
IS	Internal standard
ITC	Isothermal titration calorimetry
$K_a$	Association constant
$K_{a,app}$	Apparent association constant
$K_{a,int}$	Intrinsic association constant
$K_{a,inter}$	Intermolecular association constant
$K_{a,intra}$	Intramolecular association constant
L	Ligand
LC	Liquid Chromatography
$L_{proxy}$	Proxy ligand
m/z	Mass-to-charge ratio
MNV	Murine norovirus
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSP	Membrane scaffold protein
MW	Molecular weight
MWCO	Molecular weight cutoff
nanoESI	Nanoflow electrospray ionization
ND	Nanodisc

Neu5Ac	N-acetyl neuraminic acid
NMR	Nuclear magnetic resonance
NoV	Norovirus
ORFs	Open reading frames
P dimer	Dimer of the norovirus P domain protein
P domain	Protruding domain of the norovirus capsid protein
P particle	Subviral particle assembled by the norovirus P domain protein
P	Protein
P <sub>2</sub>	Norovirus P dimer
PAA	Polyacrylamide
PBS	Phosphate-buffered saline
PL	Protein-ligand complex
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
P <sub>proxy</sub>	Proxy protein
P <sub>ref</sub>	Reference protein
PSaV	Porcine sapovirus
<i>R</i>	Abundance ratio
<i>RF</i>	Response factor
RF	Radiofrequency
scFv	Single chain variable fragment
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SID	Surface induced dissociation
SPR	Surface plasmon resonance spectroscopy

SRIG	Stacked ring ion guide
STD-NMR	Saturation transfer difference-nuclear magnetic resonance
TLC	Thin layer chromatography
TOF	Time-of-flight
tr-P <sub>2</sub>	C-terminus truncated norovirus P dimer
Tris	Tris(hydroxymethyl) aminomethane
TrisHCl	Tris(hydroxymethyl) aminomethane hydrochloride
T-wave	Travelling wave
TWIMS	Travelling wave ion mobility spectrometry
Ubq	Ubiquitin
UV	Ultraviolet
UVPD	Ultraviolet photodissociation
VLP	Virus-like particle

## Chapter 1

### Quantitative Characterization of Noncovalent Protein-Carbohydrate

#### Interactions using Electrospray Ionization Mass Spectrometry

## 1.1 Introduction

### 1.1.1 Biological importance of carbohydrates

Carbohydrates are the most abundant biomolecules in nature. Despite their roles as energy storage compounds such as starch and glycogen, and cellular structural components in plant cell walls (cellulose) and animal exoskeletons (chitin), carbohydrates are also commonly found on the surface of cells of many living organisms, in the form of membrane associated glycolipids and glycoproteins, and in secreted body fluids like milk and saliva, in the form of free oligosaccharides as well as glycolipids and glycoproteins.<sup>1,2</sup> Importantly, the cell-surface carbohydrates serve as receptors and are involved in a number of biological roles such as cellular recognition and adhesion, pathogen infection, signal transduction, trafficking, and immunological response.<sup>3,4</sup> The glycolipids present on the surface of cells are amphipathic, with the hydrophobic lipid moiety embedded into the cellular membranes and the hydrophilic saccharide head group protruding into the aqueous environment, while cell-surface glycoproteins are often integral membrane proteins and can be difficult to obtain in their native conformations. The diverse roles of the carbohydrate receptors are highly dependent on their structures, which are regulated by carbohydrate-modifying enzymes (e.g. glycosyltransferases and glycosidases) in vivo. Indeed, cells in different tissues or organs synthesize different carbohydrate structures, and

such carbohydrate phenotypes are also affected by the differentiation state, physiological environment and tumor development of living cells. Based on the structures of saccharide moieties, carbohydrate receptors are divided into several classes, such as gangliosides, globosides and histo-blood group antigens (HBGAs).

### **1.1.2 Protein-carbohydrate interactions**

The carbohydrate receptors are specifically recognized by other biomolecules (e.g. proteins). For example, lectins, a class of carbohydrate-binding proteins prevalent in plants and animals, are known to bind to cell-surface carbohydrates. In addition, it is known that many bacteria and viruses possess carbohydrate-binding proteins specific for the carbohydrate receptor of the target cell and the binding is a prerequisite for infection to occur. In general, the association of noncovalent protein-carbohydrate complex is driven by the formation of both polar (hydrogen bond) and nonpolar (van der Waals) contacts.<sup>5</sup> Consequently, the interactions are strongly affected by solvent effects. In their unbound form, the protein residues, as well as carbohydrate ligands involved in the binding, are generally well-solvated in an aqueous environment. Hence, the complex formation requires a complete or partial dehydration of the binding partners, which is energetically costly. Therefore, the affinities of proteins for individual carbohydrates are typically weak (with association constants ( $K_a$ ) in the range of  $10^3 - 10^6 \text{ M}^{-1}$ ).<sup>6</sup> However, higher apparent affinities can be achieved through multivalent bindings. For example, virus particles, bacterial toxins and lectins possess multiple ligand binding sites and they are able to simultaneously interact with multiple copies of cell-surface carbohydrates.<sup>7</sup>

Due to the fundamental importance in drug discovery and disease diagnosis, significant attention has been devoted to elucidating the molecular basis of protein-carbohydrate recognition, and binding specificity and affinity. Consequently, developing analytical methods capable of identifying and quantifying biologically or therapeutically relevant protein-carbohydrate interactions in vitro is indispensable. One strategy is based on studying protein interactions to the water-soluble analogs (i.e., oligosaccharides and neoglycoproteins). Alternatively, study of protein-carbohydrate interactions should involve the use of real cellular receptors such as glycolipids.

Among a variety of analytical techniques available for probing protein-carbohydrate interactions, X-ray crystallography provides three dimensional structure information of the protein-carbohydrate complex, and insights into the location of binding sites and the nature of the interaction.<sup>8-10</sup> Unfortunately, the strength of the binding cannot be measured by this technique. Also, X-ray crystallography is time-consuming and costly, and only a small portion of proteins and protein-carbohydrate complexes are readily crystallized, which limits its application. Additionally, it is also possible that the interactions present in solution are different from those in the crystalline state.<sup>11</sup>

Protein-carbohydrate interactions can be analyzed by many surface-based techniques, such as enzyme-linked immunosorbent assay (ELISA), glycan microarray screening, and surface plasmon resonance (SPR) spectroscopy. ELISA is an extensively used tool for measuring protein-carbohydrate interactions with moderately high sensitivity and good reproducibility<sup>12</sup>. While there are many ways of implementing ELISA, the typical setup involves the

immobilization of carbohydrate ligands to the surface of a microplate, which are then incubated with solutions containing the protein, often in the presence of a soluble inhibitor/competitive binder.

As a new technique, glycan microarray assay has been frequently adopted for high-throughput screening of carbohydrate libraries against protein targets.<sup>13,14</sup> To generate the microarrays, a library of oligosaccharides is printed onto a solid support through a covalent linker. Proteins are incubated with the array and specific interactions are detected using fluorescence or immunoassays. The development of glycan microarrays has greatly improved current understanding of protein-protein recognition. Generally, qualitative or at best, semi-quantitative binding data can be provided.

SPR represents a highly sensitive technique capable of measuring both the real-time association and dissociation rate constants (kinetic), and the affinities (thermodynamic) of protein-carbohydrate interactions.<sup>15-19</sup> This technique also requires one of the binding partners (usually the ligand of smaller size) to be immobilized on a sensor chip while the other binding component in solution is flowed over the sensor surface. The protein-carbohydrate interaction is analyzed by measuring small changes in refractive index at the sensor surface.

To some degree, the immobilized glycans, as used by these surface-based techniques, mimic the multivalent display of cell-surface carbohydrates.<sup>20</sup> However, the potential influences of the orientation of the carbohydrate, the nature of coupling (immobilization), ligand density, and the loss of mobility of the immobilized glycans on binding are always in debate,<sup>21,22</sup> which sometimes results in false negatives. Moreover, the binding measured on the surface may not

match the affinities obtained from the solution-based assays.<sup>23,24</sup>

In addition, there are several solution based techniques that are extensively used to study protein-carbohydrate interactions. Among them, isothermal titration calorimetry (ITC) is generally considered the “gold standard” technique for quantifying binding thermochemistry in solution, with the ability of determining the enthalpy and entropy changes, and thereby the Gibbs free energy change and  $K_a$ . However, due to the low sensitivity and low throughput of conventional ITC instruments, large amounts (~mg) of protein and ligand are usually required for a single analysis and each measurement takes 4 – 6 hours.<sup>25-27</sup>

Nuclear magnetic resonance (NMR) spectroscopy is a solution based method for characterizing the structure of small proteins in solution<sup>28</sup> and for accessing interactions between proteins and various ligands.<sup>29,30</sup> Particularly, the saturation transfer difference (STD)-NMR method,<sup>31-34</sup> which focuses on the nuclear Overhauser effect and the observation of the ligand resonance signals, has earned a place for studying protein-carbohydrate interactions, and for library screening. Notably, both the binding affinity and the binding epitopes (i.e., hydrogens of the ligand that are closest to the protein upon binding) can be determined by STD-NMR assay. However, this method is time consuming and typically requires ~mg of samples.

Within the last two decades, electrospray ionization mass spectrometry (ESI-MS) has emerged as a promising technique for identifying and quantifying protein-carbohydrate interactions.<sup>35-44</sup> As a soft ionization method, the noncovalent protein-ligand complexes formed in solution are retained after transfer to the gas-phase by the ESI process. Compared with other techniques, ESI-MS assay offers a number of advantages, such as speed (1–2 min/measurement),

sensitivity (~10 pmol for nanoflow ESI-MS), simplicity (immobilization and label free), and specificity (ability of measuring binding stoichiometry and multiple binding equilibriums directly). The first example of a protein-carbohydrate interaction detected using ESI-MS (i.e., lysozyme binding to its carbohydrate substrates, oligosaccharides of N-acetylglucosamine) was demonstrated by Henion and Ganem in the early 1990s,<sup>36</sup> which was followed by a large number of studies. Overall, the  $K_a$  values measured by ESI-MS assays are in a good agreement with the values obtained by other techniques (as reviewed in references 45 – 48). A detailed description of the implementation of ESI-MS assay along with some of the limitations of this assay will be given in Sections 1.3 and 1.4.

### **1.1.3 Model membrane systems for studying protein-glycolipid interactions**

Due to the insoluble nature of glycolipids, together with the lack of a native membrane environment when studied *in vitro*, protein-glycolipid interactions are considerably underinvestigated and many biologically relevant interactions have yet to be discovered.

Actually, glycolipids can be readily immobilized on hydrophobic surfaces, and their interactions with the protein target can be probed using ELISA, SPR spectroscopy, and thin layer chromatography (TLC).<sup>49-52</sup> Particularly, microarrays prepared using glycosphingolipids and synthetic neoglycolipids enable glycolipid-based glycan microarray screening, which has been successfully applied to discover glycolipid receptors.<sup>53-57</sup> However, the removal of glycolipids from a lipid environment is expected to influence the nature of protein interactions.

Protein-glycolipid interactions are known to be affected by the surrounding lipid

environment (e.g. nature and composition of lipids, and membrane models).<sup>49,58</sup> Indeed, it is well established that glycolipids tend to cluster in the membranes, resulting in the formation of microdomains that are rich in glycolipids.<sup>59</sup> Moreover, protein-glycolipid bindings have been shown to depend on the manner in which the glycolipid is displayed in the membrane, i.e., the extent of surface carbohydrate exposure and the accessibility of the aglycone.<sup>60</sup> Additionally, it is interesting to note that protein-glycolipid interactions may be modulated by other molecules such as cholesterol and/or phosphatidylcholine, where the additional component can enhance protein-glycolipid binding either by directly contacting the protein or by inducing the glycolipid to adopt a more favorable conformation for the interaction.<sup>61-64</sup> For example, cholesterol has been demonstrated to lead to conformational changes in glycolipids through the formation of hydrogen bonds with the ceramide moiety, which makes cells more susceptible to infection by pathological proteins.<sup>65</sup>

Extensive research efforts are being made to develop assays that allow protein-glycolipid binding to be studied under conditions similar to that in living cells, i.e., glycolipids are maintained in a membrane environment.<sup>66</sup> To this end, a variety of model membrane systems for solubilization of glycolipids have been developed, including supported lipid bilayer (which is a planar lipid bilayer structure sitting on a solid support); liposome (which is a lipid bilayer rolled up into a hollow spherical shell and enclosing a small region of water); micelle (of which the hydrophilic head groups are exposed to solvent but hydrophobic tails are toward to the centre); bicelle (of which the bilayer centre is shielded by another micelle-like assembly formed by detergent); and nanodisc<sup>67-69</sup> (which is a discoidal lipid bilayer nanostructure surrounded by two

copies of membrane scaffold proteins (MSPs)). While these model membranes provide a native-like environment for the study of protein–glycolipid interactions, caution should be taken since the establishment and precise control of the size and composition of the membranes remain a challenge.<sup>66</sup>

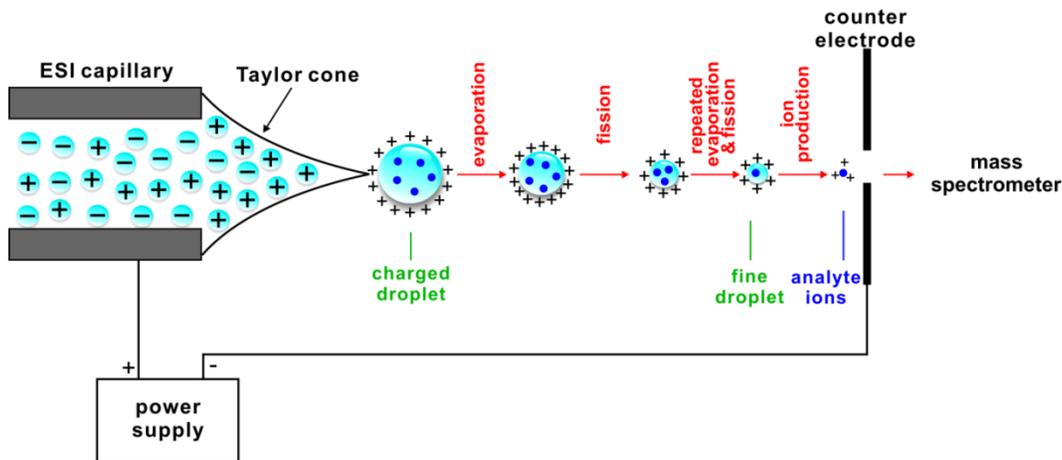
Protein binding to glycolipids incorporated into the model membranes can be probed using diverse spectroscopic- (e.g. fluorescence, NMR and SPR spectroscopy), and microscopic- (e.g. atomic force and total internal reflection fluorescence microscopy) based techniques.<sup>70-75</sup> More recently, the integration of ESI-MS and nanodisc has been reported for studying protein-glycolipid binding in aqueous solution.<sup>76,77</sup>

This thesis focuses on the development and application of ESI-MS methods to detect and characterize protein-carbohydrate interactions. As a starting point, an overview of ESI-MS techniques, including the basic mechanisms of ESI and the instrumentation is given.

## **1.2 Electrospray Ionization Mass Spectrometry**

### **1.2.1 ESI mechanisms**

ESI is a mild ionization technique that not only allows biological molecules to be transferred to the gas-phase without fragmentation, but also retains the noncovalent interactions, such as those involved in protein-carbohydrate complexes. ESI is operated at atmospheric pressure and, in most cases, at room temperature. Three major steps are involved in the process of ESI,<sup>78-81</sup> as shown in Figure 1.1.



**Figure 1.1.** Schematic representation of ESI process operated in positive ion mode, adapted from reference 80.

### (a) Formation of charged droplets

To generate ESI, a high voltage is applied to the ESI capillary containing the analyte solution. The high electric field induces accumulation of charge with the same polarity as the capillary on the solution surface at the capillary tip, leading to the formation of a cone shaped liquid jet referred to as a Taylor cone.<sup>82</sup> At the narrow end of the Taylor cone, where Coulombic repulsion overcomes surface tension, the cone becomes unstable resulting in the emission of  $\mu\text{m}$  or sub- $\mu\text{m}$  sized charged droplets. The radius ( $r_i$ ) and charge ( $Q_i$ ) of initial droplets vary depending on the flow rate and solution conditions, as given in eqs 1.1a and 1.1b:<sup>78,80</sup>

$$r_i \approx \left( \frac{V_f \cdot \epsilon}{K} \right)^{1/3} \quad (1.1a)$$

$$Q_i \approx 0.7[8\pi\sqrt{\epsilon_0\gamma r_i^3}] \quad (1.1b)$$

where  $V_f$  is the volume flow-rate;  $K$ ,  $\varepsilon$  and  $\gamma$  are the conductivity, permittivity and surface tension of the solution, respectively; and  $\varepsilon_0$  is the permittivity of vacuum. In addition, an estimation of the voltage required for the onset of electrospray is given by eq 1.2:<sup>78,80</sup>

$$V_0 = 2 \times 10^5 \sqrt{\gamma r_c} \ln\left(\frac{4d}{r_c}\right) \quad (1.2)$$

where  $r_c$  is the outer radius of the capillary and  $d$  is the distance from capillary tip to the counter-electrode.

### **(b) Repeated shrinkage and fissions of the charged droplets**

The charged droplets ejected from the Taylor cone then shrink over time as a result of solvent evaporation. As the droplet becomes smaller, its radius ( $r$ ) approaches the Rayleigh limit<sup>80,81</sup> (eq 1.3), where all the charges ( $Q_{Ry}$ ) accumulated on the droplet surface lead to increased Coulombic repulsion, which eventually counterbalances the surface tension, and fission of the droplet occurs.

$$Q_{Ry} = 8\pi\sqrt{\varepsilon_0\gamma r^3} \quad (1.3)$$

The Coulombic fission releases a jet of offspring droplets, which have the radii about 1/10 of the parent radius, and carry ~2% of the parent mass as well as ~15% of the parent charge.<sup>78</sup> The droplets undergo a series of evaporation/fission events, and ultimately, the final generation of ~nm sized, highly charged droplets (nanodroplets) are formed. The time scale for evolution to the nanodroplets in the hundreds of microseconds during the conventional ESI process,<sup>78</sup> but much shorter time (~20  $\mu$ s)<sup>81,83,84</sup> is required for nanoflow ESI (*vide infra*).

### **(c) Production of gas-phase ions from highly charged nanodroplets.**

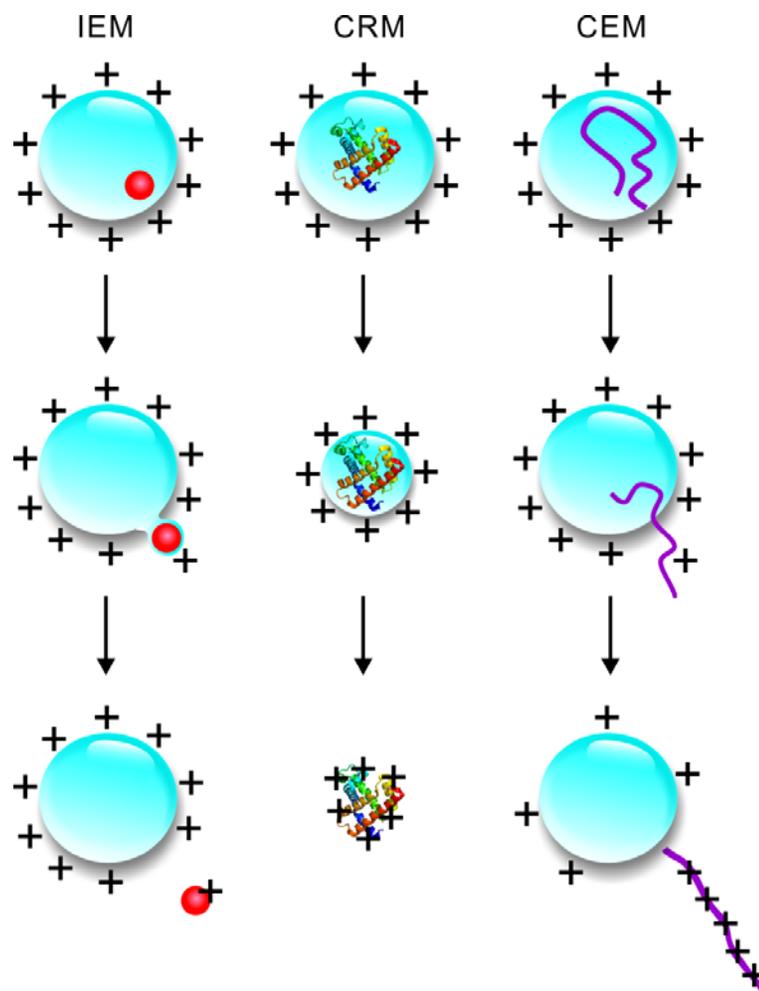
Three theories have been proposed to explain the processes of gas-phase ion formation in ESI, as illustrated in Figure 1.2.

**i) Ion evaporation model (IEM):**<sup>85,86</sup> This model was proposed by Iribarne and Thompson. It assumes that ions emit directly from the surface of highly charged droplets when their sizes become small enough. Charges are acquired locally by analytes as they evaporate from the droplet. IEM is believed to operate for small inorganic and organic ions.

**ii) Charged residue model (CRM):**<sup>87-89</sup> This model, introduced by Dole and coworkers, states that successive fissions of droplets finally yield highly charged nanodroplets containing a single macromolecule. As the final droplets evaporate to dryness, charges on the droplets' surfaces are transferred to the macromolecules and the gas-phase ions are produced. The consequence of CRM is the production of multiple-charged macroions with a narrow charge state distribution, which is thought to be determined by a combination of the accessibility of ionizable residues near the droplet surface and the Rayleigh limit.<sup>80,81,83,84</sup> CRM is experimentally well-supported for natively folded globular proteins.<sup>80,81,83,84,87,88</sup>

**iii) Chain ejection model (CEM):**<sup>89-91</sup> Konermann and coworkers indicated that CEM applies to unfolded proteins where the macromolecular chains are disordered, partially hydrophobic, and capable of carrying excess charge. This mechanism supposes that in a highly charged nanodroplet, the unfolded chains immediately migrate to the droplet surface to minimize solvent interactions with the hydrophobic regions. One chain terminus then gets expelled into the gas-phase. This is followed by stepwise sequential ejection of the remaining chain and separation

from the nanodroplet. The ions produced by CEM are also multiple-charged, but they usually carry more charges and exhibit a wider charge state distribution<sup>91</sup> compared to the ions of folded protein generated via CRM.



**Figure 1.2.** Different ESI models proposed for the formation of gas-phase ions. (a) IEM: Small ion ejection from a highly charged nanodroplet. (b) CRM: Release of a folded protein into the gas-phase. (c) CEM: Ejection of disordered macromolecule. Figure is adapted from reference 91.

## **1.2.2 Nanoflow ESI-MS**

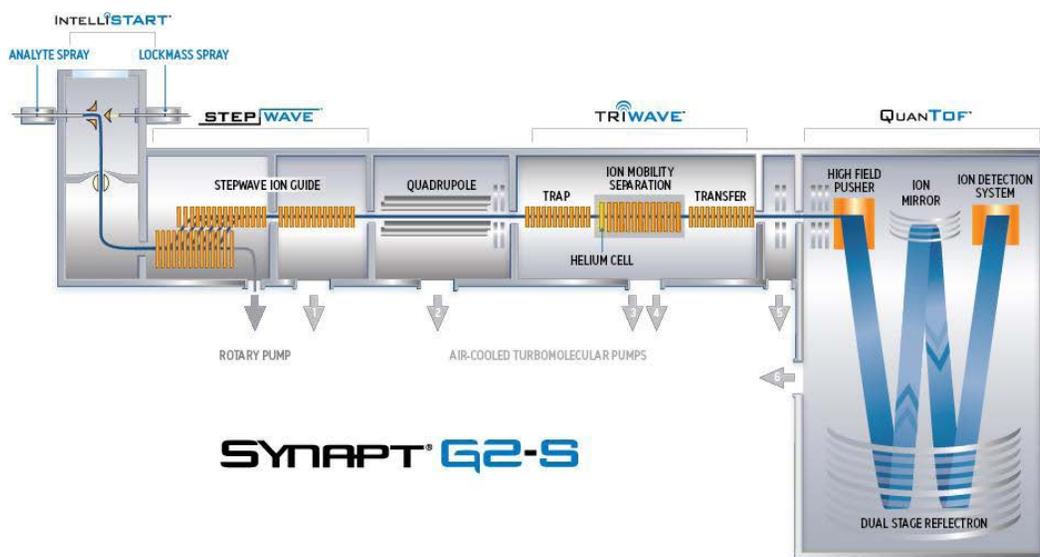
In conventional ESI-MS, analyte solution is introduced into an ESI emitter, typically a stainless steel capillary, through direct infusion using a syringe pump. Also, mixtures can be analyzed by coupling ESI-MS with liquid chromatography (LC-MS). The force-driven flow of conventional ESI has a rate in the  $\mu\text{L}/\text{min}$  range and the droplets produced initially are in  $\mu\text{m}$  sizes. In contrast, nanoflow ESI (nanoESI)<sup>92-94</sup> is performed in fine tips pulled from glass capillaries. The flow rate in nanoESI (without external pumping) is down to  $\leq 10$  nL/min so that a few  $\mu\text{L}$  of solution containing picomoles of analyte is sufficient for an individual nanoESI-MS measurement. In addition, it is known that the lower flow rates should reduce the size of the initially produced droplets (eq 1.1a). Hence, fewer droplets shrinkage/fusion cycles are required before analyte ions are released into the gas-phase, i.e. shorter lifetime, and the original solution composition throughout the formation of gaseous ions is preserved. Moreover, fewer analyte molecules are present in each nanoESI droplet, which limits the nonspecific aggregation that may occur during the ESI process.<sup>43,92-94</sup> Therefore, nanoESI-MS is an ideal choice for characterizing noncovalent protein-carbohydrate interactions, as used in the present thesis.

## **1.2.3 MS instrumentation**

### **1.2.3.1 Hybrid quadrupole-ion mobility separation-time of flight mass spectrometer**

A Synapt G2S quadrupole-ion mobility separation-time of flight (Q-IMS-TOF) mass spectrometer (Waters UK Ltd., Manchester, UK), equipped with a nanoESI source (Figure 1.3) was used for the work described in Chapter 3 – 6. The highly charged nanodroplets generated by

nanoESI are introduced into the mass spectrometer and pass through a “Z-spray source”. The “Z-shaped” trajectory minimizes transfer of neutral molecules and enhances the signal-to-noise ratio. The resulting gaseous ions are then transmitted through the StepWave ion guide, quadrupole mass filter, Trap collision cell, Travelling Wave Ion Mobility cell and Transfer collision cell, and finally reach the orthogonal acceleration reflectron TOF mass analyzer for detection. To perform tandem MS, ions of interest can be isolated by the quadrupole mass filter and subjected to collision-induced dissociation (CID) in either Trap or Transfer collision cell. Moreover, ion mobility separation (IMS) provides another dimension of separation based on the size and shape of the ions. The Synapt G2S mass spectrometer is designed for high sensitivity and wide mass range analyses. A brief overview of the quadrupole, TOF, CID, travelling-wave ion guides, and IMS parts of the instrument will be discussed in the following.

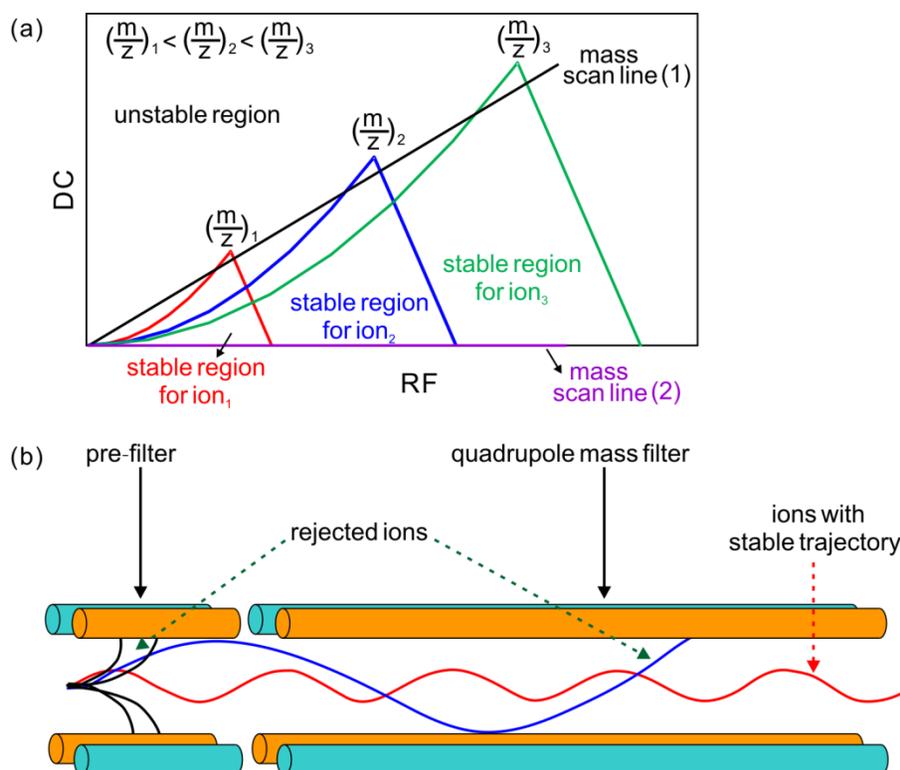


**Figure 1.3.** A schematic diagram of the Synapt G2S Q-IMS-TOF mass spectrometer, adapted from the Waters user’s manual.

### 1.2.3.1.1 Quadrupole mass filter

A quadrupole is a set of four cylindrical metal rods that are accurately arranged in a radial array with each opposing rod pair connected together electrically.<sup>95,96</sup> One opposing rod pair has an applied voltage of  $(U+V_m\cos(\omega t))$  while the other has a voltage of  $-(U+V_m\cos(\omega t))$ , where  $U$  is a direct current (DC) voltage and  $V = V_m\cos(\omega t)$  is a radiofrequency (RF) voltage, with an amplitude of  $V_m$  and frequency of  $\omega$ . By selecting an appropriate combination of the DC and RF, ions of certain mass-to-charge ratio ( $m/z$ ) values can traverse the quadrupole with stable trajectories, whereas other ions outside the  $m/z$  range are expelled by collision with the metal rods (unstable trajectories). Figure 1.4a shows the stability diagram of the quadrupole as a function of DC and RF voltages. Practically, changing DC linearly as a function of RF, a straight mass scan line is obtained, where ions can be isolated by the  $m/z$ . The larger the slope of mass scan line, the higher the resolution of isolation will be. Specifically, a broad  $m/z$  window of ions will have stable trajectories in case when  $DC = 0$  and RF is within the limits of the ions' stability regions. Hence, in MS mode, the quadrupole operates in RF only mode and acts as an excellent ion focusing device to guide ions. In MS/MS mode, both DC and RF are applied, where only ions of desired  $m/z$  values are allowed to pass through.

In the Synapt G2S mass spectrometer, a quadrupole pre-filter is placed in front of the quadrupole mass filter (Figure 1.4b), which can minimize the effects of fringing fields at the entrance to the quadrupole and thus improve the absolute sensitivity.<sup>97</sup>



**Figure 1.4.** (a) Stability diagram (adapted from reference 96) of quadrupole as a function of DC and RF voltages for ions with different  $m/z$  values  $[(m/z)_1 < (m/z)_2 < (m/z)_3]$ . (b) Schematic diagram of the quadrupole used in the Waters Synapt mass spectrometers.

### 1.2.3.1.2 TOF mass analyzer

The TOF mass analyzer measures the flight time that ions take to move through a field-free region (flight tube) between the ion accelerator and the detector.<sup>98</sup> In principle, the  $m/z$  of an ion is related to the flight time ( $t$ ) according to eq 1.4:

$$m/z = t^2 \frac{2eV_s}{L^2} \quad (1.4)$$

where the elementary charge ( $e$ ), acceleration potential ( $V_s$ ) and length of the flight tube ( $L$ ) are constants. From this equation, ions with smaller  $m/z$  will move faster and reach the detector

earlier.

To improve the resolution, the Synapt G2S mass spectrometer adopts an orthogonal acceleration – reflectron TOF mass analyzer.<sup>99</sup> Coupling a TOF mass analyzer with ESI requires the use of the orthogonal acceleration technique. Continuous ions from the ionization source are parceled into packets and filled in the orthogonal accelerator. A pusher is then pulsed to introduce ions into the orthogonally situated flight tube. During the time that the ions continue their flight in the flight tube, the orthogonal accelerator is refilled with a new packet of ions. In addition, the reflectron TOF analyzer uses successive sets of electric grids with increasing potentials, which deflects the ions and reverses their flight direction. Ions with the same  $m/z$  but different initial kinetic energies will penetrate into the field at different depths. Faster ions will penetrate deeper into the field and spend longer time in the reflectron than slower ions. Consequently, the fast and slow ions are focused at the detector and the energy distribution of the ions is compensated. Of particular note, the Waters Synapt G2S instrument can be operated under either single stage reflectron mode (“V mode”) or dual stage reflectron mode (“W mode”, Figure 1.3). The latter can focus the ions twice, leading to higher resolution but to less sensitivity. Taken together, the orthogonal acceleration - reflectron TOF mass analyzer improves the mass resolution, with minimal losses in sensitivity.

#### **1.2.3.1.3 MS/MS**

MS/MS (also known as tandem MS) involves using multiple sequential stages of mass spectrometry segments where ions are selected, energetically activated, dissociated and analyzed.

MS/MS is useful for probing the structure and stability of biomolecules and their complexes through the fragmentation patterns. Ion activation and dissociation can be achieved by a number of techniques, including collision-induced dissociation (CID),<sup>102,103</sup> surface induced dissociation (SID),<sup>104</sup> electron capture dissociation (ECD),<sup>105,106</sup> electron transfer dissociation (ETD)<sup>107</sup> infrared radiative multiphoton dissociation (IRMPD),<sup>108,109</sup> ultraviolet photodissociation (UVPD),<sup>110,111</sup> and blackbody infrared radiative dissociation (BIRD).<sup>108,112,113</sup> For the purpose of this study, only CID, which is the most common ion activation method used in MS/MS, is introduced here.

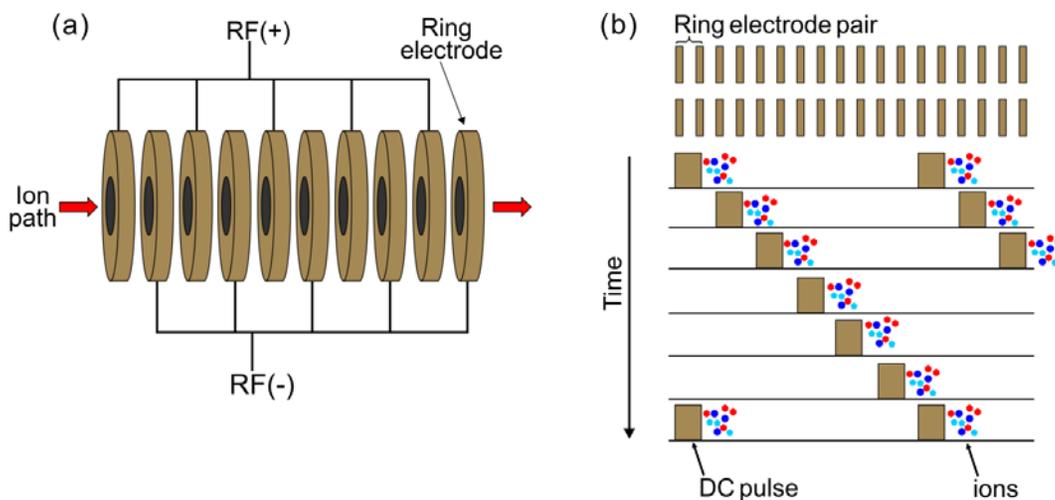
CID can be performed in the Trap and Transfer cells of the Synapt G2S mass spectrometer. In this case, a constant DC voltage (collision energy) was applied to each ring electrode in addition to the transient DC voltage used to propel ions to the next stage of the instrument (*vide infra*). The Trap and/or Transfer cells are filled with neutral background gases (e.g.  $\sim 10^{-2}$  mbar Argon) and the ions of interest (precursor ions) are subjected to energetic collisions with the gas molecules within these regions. The internal energy of precursor ions is accumulated by successive collisions, which eventually induces the decomposition. Finally, the resulting product ions are detected by the mass analyzer.

CID is a “slow-heating” process, where energy randomization is faster than the decomposition. Consequently, the energy will be distributed among all the internal modes of the ion, leading to preferential decomposition at the weakest sites.<sup>114</sup> This feature allows CID to investigate the noncovalent protein-ligand and protein-protein interactions in the gas-phase.<sup>115-125</sup> However, because ions do not reach thermal equilibrium during CID, the kinetic/thermodynamic

parameters for the dissociation process cannot be measured. In addition, since noncovalent interactions are usually broken prior to covalent bond cleavage, the binding site cannot be localized by CID. Examples of CID applied for studying noncovalent protein-carbohydrate bindings can be found in Chapters 3, 5 and 6 in this thesis.

#### **1.2.3.1.4 Travelling-wave ion guides**

Travelling-wave technology is employed for the StepWave ion guide, as well as Trap, Ion Mobility and Transfer cells of the Synapt G2S mass spectrometer. These travelling-wave devices are used to guide (trap, focus, release, fragment and separate) ions. Each device comprises a stacked ring ion guide where a series of ring-shaped electrodes are arranged orthogonally to the ion transmission axis (Figure 1.5a).<sup>100,101</sup> Opposite phases of a RF voltage are applied to adjacent rings and ions are radially confined within the device while motion along the axial direction is allowed. When a DC voltage is applied to a pair of adjacent rings, a potential barrier is produced so that ions within this region cannot cross. The DC voltage is subsequently applied to the next sets of electrodes downstream at regular time intervals providing a continuous sequence of “travelling waves”. The ions are driven away from the potential barriers generated by the travelling waves and consequently are propelled through the device with the waves (Figure 1.5b). Ions can transit with a fast speed in the travelling-wave devices, which allows high data acquisition rates, with sensitivity maintained. Furthermore, ion mobility separation can be performed using the travelling-wave device, which is introduced in the next section.



**Figure 1.5.** Diagrams of (a) stacked ring ion guide (SRIG) and (b) its operational principle.

Figure is adapted from reference 100.

### 1.2.3.1.5 Ion mobility separation

Ion mobility separation (IMS) is a gas-phase electrophoretic technique used to distinguish ions on the basis of their mass, charge and collision cross section (i.e., the effective area for the collision between a target ion and the neutral gas, which is dependent on the size and shape of the target ion).<sup>126-130</sup> In IMS, ions travel through a drift cell with a combination of an electric field that moves the ions towards the drift region and a buffer gas flow that opposes the ions' motion. As an ion passes through the buffer gas, it experiences a number of collisions, which impede its progress towards the detector. Larger ions with greater collision cross sections undergo more collisions than smaller ions and thereby require longer time to migrate through the drift cell. The mobility of an ion ( $K$ ) is determined by eq 1.5:<sup>127</sup>

$$K = \frac{d}{t_d \cdot E} \quad (1.5)$$

where  $t_d$  is the time for an ion traversing the drift cell (i.e., arrival time) of length  $d$ ;  $E$  is the electric field gradient.

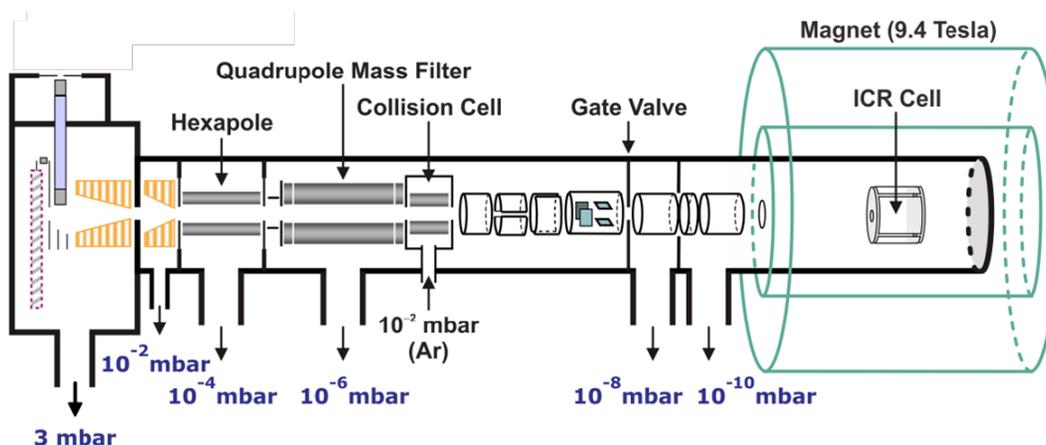
There are several types of ion mobility instrumentation that have been successfully coupled with mass spectrometry, such as drift tube IMS (DTIMS)<sup>131</sup>, field asymmetric waveform IMS (FAIMS)<sup>132</sup>, and travelling-Wave IMS (TWIMS).<sup>100,101,133-136</sup> The detail of TWIMS, which was used in this thesis, is outlined briefly below.

The Synapt G2S mass spectrometer implements a TWIMS technique. In this design, IMS cell is filled with nitrogen gas ( $\sim$  mbar), and a high electric field is applied to one segment of the IMS cell and swept sequentially through the cell one segment at a time in the direction of ion migration. As the waves of the field pass through, ions are moved through the mobility cell in pulses and separated based on their mobilities. Ions of high mobility are able to “surf” the waves and are transported through the IMS cell more quickly while lower-mobility ions slip behind the waves more often and travel more slowly. A particular advantage of the TWIMS device is the substantially enhanced sensitivity through its ion accumulation and radial ion confinement functions.<sup>100,101,133-136</sup> In addition, unlike the DTIMS where the mobility is proportional to the inverse of the arrival time, in case where ions are separated by the TWIMS, the mobility is approximately proportional to the inverse of the square root of the arrival time.<sup>136</sup>

### **1.2.3.2 Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer**

The other mass spectrometer used in this thesis (Chapters 2 and 4) is a Bruker ApexQe (Bruker-Daltonics, Billerica, MA) 9.4-Tesla FT-ICR mass spectrometer (Figure 1.6). Highly

charged nanodroplets containing analytes generated by the nanoESI ion source enter the mass spectrometer through a heated metal capillary, which facilitates the desolvation of the droplets. The resulting gaseous ions are pushed orthogonally by a deflector at the capillary exit and focused by passing through two stages of ion funnel-skimmer. The ions are then accumulated electrodynamically in the hexapole for a certain time period, transmitted through the quadrupole mass filter, and accumulated in the collision cell for another time. In the present work, the quadrupole was solely operated in RF-only mode and acts as a wide band-pass filter for ion transmission. The signal-to-noise ratio can be enhanced through the accumulation processes. After accumulations, the ions are transferred into the ICR cell via a series of ion optics for detection. The ultra-high vacuum required for operating the FT-ICR mass spectrometer (typically  $\sim 10^{-10}$  mbar) is maintained by the differential pumping system.



**Figure 1.6.** Schematic diagram of the Bruker ApexQe 9.4-Tesla FT-ICR mass spectrometer used in this study. Figure is reproduced from the Bruker user's manual.

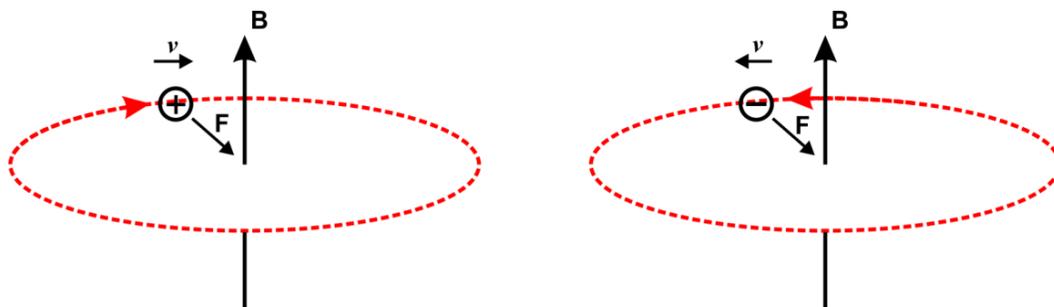
### 1.2.3.2.1 FT-ICR mass analyzer

The FT-ICR mass analyzer<sup>137-141</sup> is based on the principle of ion cyclotron motion in the magnetic field. Ions are stored, mass analyzed and detected in the ICR cell (Figure 1.6), which is composed of three pairs of electrode plates in the roles of ion trapping, excitation and detection, respectively. The ICR cell is located inside a uniform high magnetic field (generated by the superconducting magnet) lying parallel to the axis of ion beam.

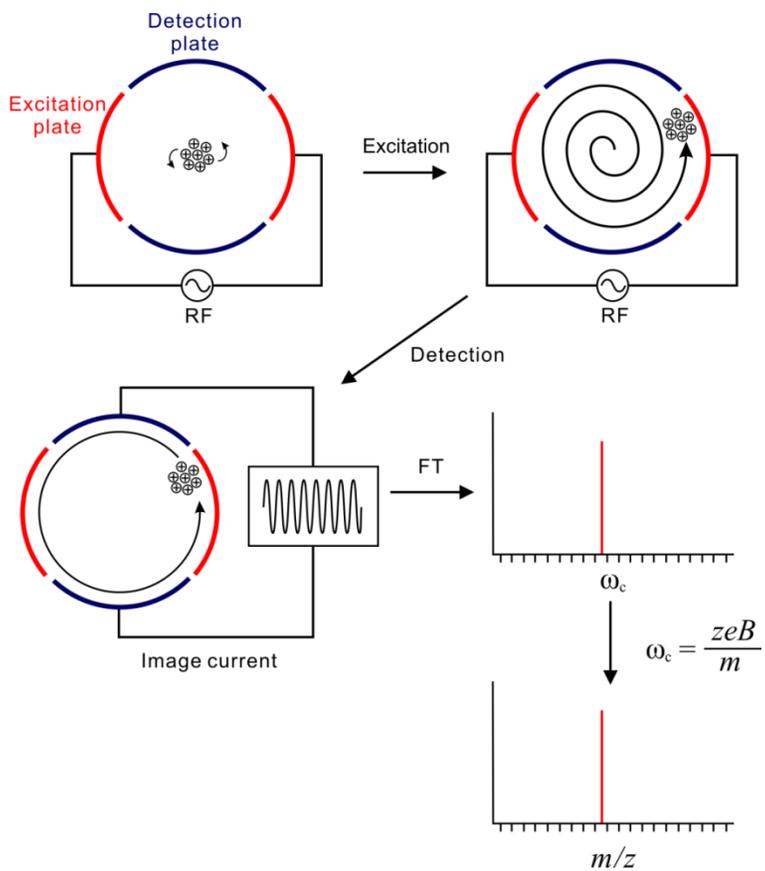
Ions in the ICR cell are trapped by a combination of electric and magnetic fields. The axial motion of an ion is confined by creating a low potential between the trapping plates, which are perpendicular to the magnetic field. More importantly, the ion's radial motion is constrained to a cyclotron orbit perpendicular to the magnetic field (Figure 1.7) with the cyclotron frequency ( $\omega_c$ ) given by eq 1.6:

$$\omega_c = \frac{z \cdot e \cdot B}{m} = \frac{v}{2\pi \cdot r_c} \quad (1.6)$$

where  $z$  is the charge on the ion,  $m$  is the mass of the ion,  $e$  is the elementary charge,  $B$  is the magnetic field strength,  $v$  is the linear velocity of the cyclotron motion, and  $r_c$  is the orbital radius. Consequently, ions with the same  $m/z$  rotate at the same frequency, but more energetic ions (larger  $v$ ) have larger orbital radii.



**Figure 1.7.** Cyclotron motions of a positive ion and a negative ion in magnetic field (B).



**Figure 1.8.** Illustration of ion excitation, image current detection and generation of mass spectrum by FT-ICR MS. Figure is adapted from the Bruker user's manual.

Detection in FT-ICR MS is achieved by applying an oscillating electric field on the excitation plates. Ions of a given  $m/z$  can be coherently excited to a larger radius if the frequency of the field matches  $\omega_c$  of the ions, while ions having different  $\omega_c$  will not be accelerated. As the excited spatially coherent ions pass close to the detection plates, they induce a transient alternating current to the plates called an image current. The frequency of the image current matches the  $\omega_c$  of the ions inducing the current and the amplitude is proportional to the number of these ions in the ICR cell. The transient signal is then amplified and detected while the ions are still trapped in the ICR cell without ever colliding with the electrodes. In practice, ions of individual  $m/z$  are sequentially excited by a RF pulse sweeping the frequency. FT converts the detected image current from the time domain signal into the frequency domain and a mass spectrum can be generated since  $\omega_c$  is related to  $m/z$  (eq 1.6). The whole process of ion excitation, image current detection and generation of mass spectrum by FT-ICR MS is illustrated in Figure 1.8.

A notable feature of FT-ICR MS is the elegant combination of both high resolution and mass accuracy. The resolving power of an FT-ICR mass analyzer is proportional to the magnetic field strength and the acquisition time, which is determined by the dataset size and the frequency of sampling. The resolution is also affected by the signal decay resulting from collisions between excited spatially coherent ions with neutral gas molecules. Therefore, high vacuum ( $10^{-10}$  mbar) is desirable in the ICR cell. Moreover, as  $\omega_c$ , which is insensitive to the kinetic energy of an ion, can be measured with very high precision, the mass accuracy of FT-ICR MS is usually a few ppm.

### 1.3 ESI-MS Based Methods for Quantifying Noncovalent Protein-Carbohydrate Interactions

ESI-MS binding measurements can be operated in the *direct* or *indirect* mode. *Direct* ESI-MS assay measures the abundance ratio of carbohydrate ligand-bound to free protein ions, and allows one to calculate the  $K_a$  value. *Direct* ESI-MS assay is capable of measuring weak affinities (as low as  $10^2 \text{ M}^{-1}$ ). Meanwhile, *indirect* ESI-MS methods (e.g. *proxy protein* ESI-MS<sup>142</sup> and *proxy ligand* ESI-MS assays) have been developed, which combine competitive binding and ESI-MS analysis. *Indirect* ESI-MS methods provide a possibility to extract binding data that cannot be measured directly by ESI-MS (*vide infra*). Moreover, ESI-MS assay is also well-suited for library screening,<sup>40,125,143</sup> where specific interactions between a target protein with a library of carbohydrates can be detected and accurately quantified from a single measurement.

#### 1.3.1 *Direct* ESI-MS assay

The *direct* ESI-MS assay relies on the direct detection of the free and ligand-bound protein ions. For a reversible interaction (eq 1.7) between a monovalent protein (P) and a monovalent ligand (L), the abundance (*Ab*) ratio of ligand-bound to free protein ions (*R*) can be measured from the mass spectrum, and this ratio is expected to reflect the corresponding concentration ratio at equilibrium in solution (eq 1.8).



$$R = \frac{\sum Ab(PL)}{\sum Ab(P)} = \frac{[PL]}{[P]} \quad (1.8)$$

Hence, if the initial concentrations of protein ( $[P]_0$ ) and ligand ( $[L]_0$ ) are known, the association constant ( $K_a$ ) can be calculated from eq 1.9:<sup>46</sup>

$$K_a = \frac{[PL]}{[P][L]} = \frac{R}{[L]_0 - \frac{[P]_0 R}{1+R}} \quad (1.9)$$

where  $[P]$ ,  $[L]$ , and  $[PL]$  in eqs 1.8 and 1.9 are equilibrium concentrations of the protein, ligand and protein-ligand complex in solution, respectively. Normally, the affinity measurements are performed at a number of different concentrations or from a titration experiment, wherein  $[P]_0$  is fixed but  $[L]_0$  is varied. In the latter case,  $K_a$  can be extracted using nonlinear regression analysis of the experimentally determined concentration dependence of the fraction of ligand-bound protein, i.e.,  $R/(R+1)$ , eq 1.10:<sup>46</sup>

$$\frac{[PL]}{[P]_0} = \frac{R}{R+1} = \frac{1 + K_a [P]_0 + K_a [L]_0 - \sqrt{(1 + K_a [P]_0 - K_a [L]_0)^2 + 4K_a [L]_0}}{2K_a [P]_0} \quad (1.10)$$

Experimentally, the  $K_a$  values that can be accurately determined with the *direct* ESI-MS assay range from  $\sim 10^2$  to  $\sim 10^7$   $M^{-1}$ , which suits the study of most protein-carbohydrate interactions.

In cases where a multivalent protein (P) that can sequentially bind up to  $h$  ligand molecules, the following  $h$  interactions are considered, eq 1.11:



.....



The apparent (macroscopic) association constant ( $K_{a,q}$ ) of the  $q^{\text{th}}$  ligand binding to the protein is shown given by eq 1.12:<sup>41</sup>

$$K_{a,q} = \frac{[\text{PL}_q]}{[\text{PL}_{q-1}][\text{L}]} = \frac{R_q / R_{q-1}}{[\text{P}]_0 \frac{\sum_{q=1}^h q R_q}{[\text{L}]_0 - \frac{q-1}{h}} 1 + \sum_{q=1}^h R_q} \quad (1.12)$$

where  $R_q$  is the abundance ratio of  $q$  ligands-bound to free protein ions, which is expected to be equal to the corresponding equilibrium concentration ratio in the solution (eq 1.13):<sup>41</sup>

$$R_q = \frac{\sum Ab(\text{PL}_q)}{\sum Ab(\text{P})} = \frac{[\text{PL}_q]}{[\text{P}]} \quad (1.13)$$

Moreover, the fraction of occupied ligand binding sites ( $f$ ) in P at a given concentration can be expressed, in terms of  $Ab$  or concentration, as given by eq 1.14:

$$f = \frac{\sum_{q=1}^h q \cdot Ab(\text{PL}_q)}{h \left( Ab(\text{P}) + \sum_{q=1}^h Ab(\text{PL}_q) \right)} = \frac{\sum_{q=1}^h q \cdot [\text{PL}_q]}{h \left( [\text{P}] + \sum_{q=1}^h [\text{PL}_q] \right)} \quad (1.14)$$

In the simplest case, where all the  $h$  ligand binding sites are identical and independent, each binding site has the same intrinsic (microscopic) association constant,  $K_{a,\text{int}}$ . The relationship between the  $q^{\text{th}}$  apparent association constant and intrinsic association constant is given by eq 1.15,<sup>41</sup> where the coefficient  $(h-q+1)/q$  is called the statistical factor.

$$K_{a,q} = \frac{h-q+1}{q} K_{a,\text{int}} \quad (1.15)$$

### 1.3.2 *Indirect* ESI-MS assays

In some cases, protein-carbohydrate interactions cannot be studied using the *direct* ESI-MS assay. For example, if one of the binding partners is heterogeneous or too large, *direct* ESI-MS analysis is not feasible due to the instrumental limitations (mass range and mass resolution). In addition, this method is incompatible with the analysis of interactions in which membrane proteins or insoluble cellular receptors are involved. Alternatively, combining the *direct* ESI-MS with competitive protein or ligand binding can make the quantification of these interactions possible. Such *indirect* ESI-MS assays, which rely on introducing a competitor (either a protein or a ligand) and monitoring the extent of complexes formed by the competitor to deduce the strength of binding for the target (ligand or protein), have been described in this thesis. For example, the *proxy protein* ESI-MS method<sup>142</sup>, which was demonstrated to quantify carbohydrate interactions with large proteins, was adapted in Chapters 4 and 5. Moreover, the *proxy ligand* ESI-MS method was developed to measure affinities of proteins to glycolipids incorporated into the nanodiscs (Chapter 6).

There are other ESI-MS based strategies to quantify specific protein-ligand interactions that rely on measuring free ligand concentration at equilibrium in solution. For example, based on ESI-MS analysis of the relative abundance of the ligand (L) to the internal standard (IS), which resembles L, but does not bind to the target protein (P), the dependence of free ligand concentration ( $[L]$ ) in solution on the initial protein concentration ( $[P]_0$ ) can be obtained to deduce the  $K_a$  of P-L interaction.<sup>144</sup> A variation of this method uses a reference ligand ( $L_{ref}$ ), with known affinity for P, to determine the affinities from the relative abundance of the L to  $L_{ref}$  by

ESI-MS.<sup>145</sup>

Like all other techniques, ESI-MS assay also has its own limitations and a more detailed discussion is provided in the following section.

## **1.4 Potential Pitfalls of ESI-MS Binding Assays**

In the ESI-MS binding measurements, free and ligand-bound proteins are transferred from solution to the gas-phase by ESI and detected as ions by the mass spectrometer. Therefore, any physical or chemical process during ESI and in the gas-phase that alters the distributions of these species will lead to incorrect  $K_a$  prediction and, potentially, obscure the true binding stoichiometry. There are three common sources of error associated with ESI-MS assays: i) in-source dissociation; ii) nonspecific ligand-protein binding; and iii) non-uniform response factors. Each of these problems and the available strategies to minimizing their effects are discussed below.

### **1.4.1 In-source dissociation**

Gaseous protein-ligand complex ions may undergo collision-induced dissociation at various stages during the ion sampling process, which reduces the relative abundance of ligand-bound protein to free protein ions and thereby causes a decreased magnitude of  $K_a$ .<sup>146</sup> In the extreme, if no PL complex ions survive for detection, in-source dissociation results in a false negative. The configuration of the ESI source, the choice of instrumental parameters as well as the gas-phase stability of the complex being investigated influence the extent of in-source dissociation. Usually,

“gentle” source conditions such as low desolvation temperatures of the metal sampling capillary and nebulizing gas (if used), low potentials across ion optics, and short accumulation time within the ion storage devices are essential for obtaining reliable  $K_a$  values for protein-ligand complexes susceptible to in-source dissociation. However, there are usually trade-offs between the use of “gentle” source conditions and signal intensity. Thus, a balance must be found between minimizing the effect of in-source dissociation and achieving adequate ion signal. Moreover, the stability of protein-ligand complexes in the gas-phase can be predicted, in part, by the nature of the interactions in solution. For example, complexes formed between protein and small-sized ligand generally exhibit low gas-phase stabilities, because they are stabilized by a limited number of intermolecular interactions.<sup>146-149</sup> In contrast, large-sized ligand-bound protein ions are less susceptible to in-source dissociation. However, the gas-phase stabilities of protein-ligand complex ions are not necessarily parallel with the solution affinities.<sup>146</sup>

In cases where in-source dissociation still occurs under gentle sampling conditions, employment of stabilizing additives may be beneficial. For example, adding high concentration (>1 mM) imidazole to solution has been demonstrated to prevent gas-phase dissociation of protein-ligand complex ions.<sup>146,147</sup> The stabilizing effect of imidazole is attributed to the enhanced evaporative cooling resulting from the dissociation of nonspecifically bound imidazole from the gaseous ions.<sup>146</sup> Additionally, due to its relatively high gas-phase basicity (217 kcal/mol),<sup>150</sup> the use of imidazole may also lead to charge state reduction of the protein-ligand complex ions in positive ion mode. The complex ions of lower charge state are kinetically more stable and thus more resistant to in-source dissociation.<sup>151</sup> Meanwhile, it has been reported that

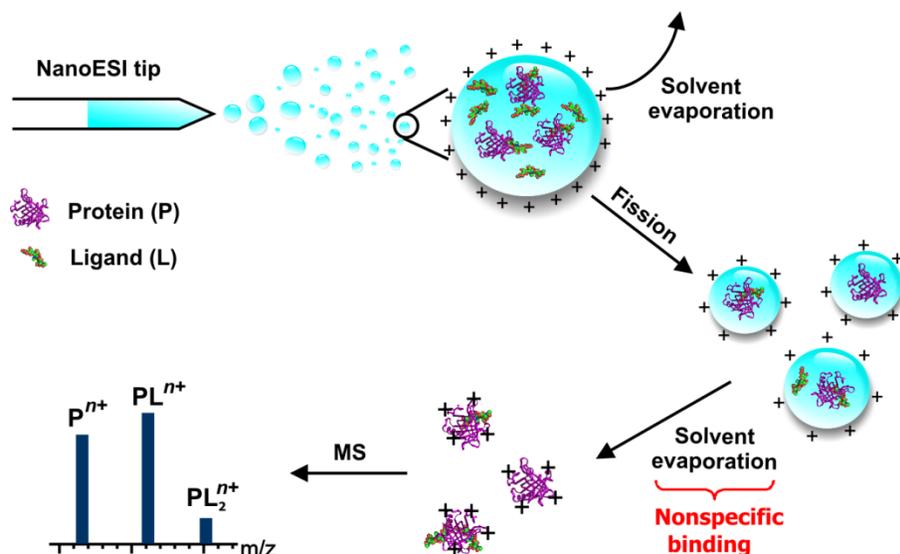
introduction of sulfur hexafluoride<sup>147</sup> or imidazole vapor<sup>147</sup> to the ESI source region can limit the extent of in-source dissociation for some protein-ligand complexes.

#### 1.4.2 Nonspecific ligand binding

As a consequence of the charge residue model (CRM), nonspecific ligand binding<sup>152,153</sup> during the ESI process is another pitfall of ESI-MS assay. As shown in Figure 1.9,<sup>46</sup> based on the CRM, the last stage of solvent evaporation from the highly charged nanodroplets ultimately yields multiple charged gaseous ions. Taking a 1:1 protein-ligand interaction as an example, if one or more free ligand (L) molecules are present with the protein (P) or protein-ligand complex (PL) in a single nanodroplet, L may bind nonspecifically to P or PL through ionic or neutral interactions as the droplet evaporates to dryness, forming a series of  $PL_q$  ( $q \geq 1$ ) ions that are not originated from the solution. Interestingly, these nonspecific complexes exhibit a Poisson-like distribution, which suggests nonspecific ligand binding is a random process.

The occurrence of nonspecific binding results in false positives and misinterpretation of the binding stoichiometry in solution. However, the extent of nonspecific binding decreases with free ligand concentration in solution as well as the size of the ESI droplet.<sup>152</sup> Consequently, nonspecific complex formation is negligible in case of strong protein-ligand binding ( $K_a > 10^6$   $M^{-1}$ ), since nearly all the ligand molecules are bound to the protein in solution. In contrast, when dealing with weak protein-ligand interactions ( $K_a < 10^4$   $M^{-1}$ ), high initial concentrations of ligand ( $\geq 50$   $\mu M$ ) are necessary to produce detectable level of complex signals, and most ligand molecules are unbound in solution at equilibrium. In such cases, nonspecific binding is often

unavoidable.



**Figure 1.9.** Illustration of the nonspecific protein-ligand interactions during the ESI process.

Figure is adapted from reference 46.

In order to identify the occurrence of nonspecific interactions of ESI-MS, a number of strategies have been developed<sup>154-159</sup>, such as the reporter molecule method,<sup>157</sup> nonspecific probe method,<sup>158</sup> and reference protein method.<sup>159</sup> Only a brief overview of the reference protein method is given here, because it is the most straightforward approach that allows ESI mass spectra to be quantitatively corrected for the occurrence of nonspecific ligand binding.

The reference protein method<sup>159</sup> involves the addition of a non-interacting reference protein ( $P_{\text{ref}}$ ) to the ESI solution. The fractions of  $P_{\text{ref}}$  involved in nonspecific binding with the ligand L are used to quantitatively correct the nonspecific complexes of the target protein P to L. For a given  $PL_q$  species, its apparent abundance ( $Ab_{\text{app}}(PL_q)$ ) measured by ESI-MS is composed

of the portion of true abundance of  $PL_q$  ( $Ab(PL_q)$ ) that does not undergo nonspecific binding, as well as the abundances of  $PL_{q-1}$ ,  $PL_{q-2}$ , ...,  $P$  that nonspecifically bind to 1, 2, ...,  $q$  molecules of  $L$ , respectively. Based on the reference protein method,  $Ab_{app}(PL_q)$  can be expressed from  $Ab(PL_q)$  and the distribution of nonspecific  $P_{ref}L_q$  species using eq 1.16:<sup>159</sup>

$$Ab_{app}(PL_q) = f_{0,P_{ref}} Ab(PL_q) - f_{1,P_{ref}} Ab(PL_{q-1}) - f_{2,P_{ref}} Ab(PL_{q-2}) - \dots - f_{q,P_{ref}} Ab(P) \quad (1.16)$$

where  $f_{q,P_{ref}}$  is the fractional abundance of  $P_{ref}L_q$ , eq 1.17:<sup>159</sup>

$$f_{q,P_{ref}} = \frac{Ab(P_{ref}L_q)}{Ab(P_{ref}) + Ab(P_{ref}L) + Ab(P_{ref}L_2) + \dots + Ab(P_{ref}L_q) + Ab(P_{ref}L_{q+1}) + \dots} \quad (1.17)$$

Hence,  $Ab(PL_q)$  can be calculated from eq 1.18a:

$$Ab(PL_q) = \frac{1}{f_{0,P_{ref}}} Ab_{app}(PL_q) - \frac{f_{1,P_{ref}}}{f_{0,P_{ref}}} Ab(PL_{q-1}) - \frac{f_{2,P_{ref}}}{f_{0,P_{ref}}} Ab(PL_{q-2}) - \dots - \frac{f_{q,P_{ref}}}{f_{0,P_{ref}}} Ab(P) \quad (1.18a)$$

In a more comprehensible way, eq 1.18a can be converted into eq 1.18b:

$$R_q = R_{q,app} - R_{1,P_{ref}} R_{q-1} - R_{2,P_{ref}} R_{q-2} - \dots - R_{q-1,P_{ref}} R_1 - R_{q,P_{ref}} \quad (1.18b)$$

where  $R_q$  is the “true” abundance ratio of  $q$ -ligand-bound to free protein ions as given in eq 1.13; and  $R_{q,app}$  and  $R_{q,P_{ref}}$ , which are defined in eqs 1.19a and 1.19b, respectively, are the values that can be measured from mass spectra.

$$R_{q,app} = \frac{Ab_{app}(PL_q)}{Ab_{app}(P)} \quad (1.19a)$$

$$R_{q,P_{ref}} = \frac{Ab(P_{ref}L_q)}{Ab(P_{ref})} \quad (1.19b)$$

The underlying assumption of the reference protein method is that nonspecific ligand binding is a random process and equally affects all protein species present in the ESI droplets. This assumption has been rigorously tested and shown to be valid for a variety of ligands,

including neutral and charged carbohydrates, amino acids, peptides and metal ions.<sup>159-162</sup> Notably, the corrected binding constants are in good agreement with what have been reported using other methods.

### 1.4.3 Non-uniform response factors

For the ESI-MS binding assay, the abundances of protein (P) and protein-ligand complex (PL) ions measured by ESI-MS are related to their respective solution concentrations by a response factor ( $RF$ ), as defined in eq 1.20:<sup>46</sup>

$$\frac{[PL]}{[P]} = \frac{RF_P \cdot Ab(PL)}{RF_{PL} \cdot Ab(P)} = RF_{P/PL} \frac{Ab(PL)}{Ab(P)} \quad (1.20)$$

where  $RF_P$  and  $RF_{PL}$  are the response factors for P and PL, respectively, and  $RF_{P/PL}$  is the relative response factor.  $RF$  accounts for the ionization and detection efficiencies of each protein species, which are dependent on the size, structure and surface properties of P and PL, as well as the solution conditions and the instrumental parameters used for the measurements.

Uniform  $RF$ s for P and PL (i.e.,  $RF_{P/PL} \approx 1$ ) is an essential assumption for ESI-MS binding assays, whereby eqs 1.8 and 1.13 hold and thus  $K_a$  can be determined. This assumption is generally good enough if L is small compared to P, such that the size and surface properties of the P and PL are similar.<sup>46</sup> A rule of thumb suggests that the molecular weight ratio of PL to P should be  $\leq 110\%$  to make the assumption valid.<sup>43</sup>

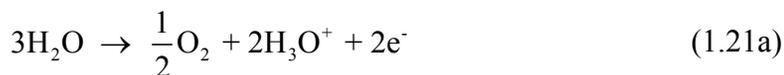
Several strategies have been developed to deal with the non-uniform  $RF$  issue on the affinity determination using ESI-MS assay. One approach involves fitting  $RF_{P/PL}$ , as an

adjustable parameter, to the experimental titration data based on an appropriate binding model.<sup>163-167</sup> However, the potential effect of different analyte concentrations on  $RF_{P/PL}$  is not considered. Moreover, this approach requires high quality experimental data to obtain reliable  $K_a$  values by multiple-parameter regression.<sup>165</sup> In addition, another method proposed involves the use of an internal standard (IS).<sup>167,168</sup> An appropriate IS can be a macromolecule that has similar molecule weight and surface activity to P, but does not bind to L. In this case, by monitoring the abundance of IS, the fluctuations in  $RF_{P/PL}$  due to the concentration changes, instability of ESI and other factors can be reflected.<sup>167</sup>

The uniform  $RF$  assumption is applied in the present work, but it is unlikely to be valid in the specific case in Chapter 6, where glycolipid ligands incorporated in to the nanodiscs were used. To address this challenge, the *proxy ligand* ESI-MS method was developed to reliably measure the binding affinities.

#### 1.4.4 Other challenges

In most cases, protein-ligand affinities are sensitive to pH. However, in the ESI tip, electrochemical reactions are taking place between the solution and the chemically inert electrode, which can alter the solution pH.<sup>39,169</sup> In aqueous solution, oxidation of  $H_2O$  in positive ion mode and reduction of  $H_2O$  in negative ion mode are the major electrochemical reactions occurring at the electrode, which cause the production of  $H_3O^+$  (eq 1.21a) and  $OH^-$  (eq 1.21b), respectively.



At a low solution flow rate used for the nanoESI, the resulting pH changes can be large.<sup>39</sup> Consequently, using ESI solutions with a high buffer capacity and short spraying times (< 10 min) should be a necessary consideration to minimize the effect of pH change on the  $K_a$  values.

Moreover, for protein-carbohydrate binding studies, an appropriate buffer is required to maintain the solution pH and ionic strength, and to keep protein stable. The “physiological” buffers, which contain phosphate, sodium and potassium ions, are generally incompatible with ESI-MS. Instead, ESI-MS assays usually employ a “volatile” buffer, such as aqueous ammonium acetate (1-200 mM). However, the replacement of physiological buffers by ammonium acetate buffer may be a source of discrepancies between ESI-MS and conventional binding assays. Carrying out ESI in the presence of high flow rate gas is a strategy to mitigate the effect of nonvolatile salts added to the solution.<sup>170</sup> Additionally, it is demonstrated that desorption electrospray ionization (DESI)-MS is suitable for quantitative binding measurements performed using phosphate buffered saline (PBS) solution.<sup>171</sup>

## 1.5 The Present Work

This thesis focuses on the development and application of ESI-MS based techniques to expand the versatility of ESI-MS for quantifying noncovalent water soluble protein-carbohydrate interactions. Chapters 2 – 5 investigate the binding of the human norovirus (NoV) capsid proteins to carbohydrate ligands. Chapter 6 extends ESI-MS assays for study of protein

interactions to water insoluble glycolipids.

In order to gain a deeper understanding of the interplay between human NoVs and their hosts, the work described in Chapter 2 focused on quantifying the intrinsic affinities (per binding site) of recombinant human NoVs protruding domain dimers (P dimers, full-length and C-terminus truncated forms, VA387 strain) to a panel of 47 soluble analogs (oligosaccharides) of histo-blood group antigens (HBGAs) using the *direct* ESI-MS assay. The oligosaccharides tested contain the A, B, H and Lewis epitopes, with variable sizes (disaccharide to hexasaccharide) and different precursor chain types (1 – 6). The study showed that the P dimers exhibit a broad specificity for the HBGAs and bind, although weakly (with intrinsic association constant,  $K_{a,int}$   $10^2 - 10^3 \text{ M}^{-1}$ ), to all of the oligosaccharides tested. Overall, A and B antigens bind stronger to the P dimers than H and Lewis antigens. Meanwhile, the affinities are affected by the precursor chain type of HBGAs, but not by the chain length.

Chapter 3 described an ESI-MS method, the catch-and-release (CaR)-ESI-MS assay for screening carbohydrate libraries against human NoVs to rapidly identify NoV ligands and potential inhibitors. In this study, carbohydrate libraries of as many as 146 compounds were screened against the NoV P particle, which is formed by the protruding (P) domain of the NoV capsid protein of VA387 strain. The results revealed that all the HBGA ligands with  $K_{a,int} \geq 300 \text{ M}^{-1}$  are identified as ligands. Furthermore, screening revealed interactions of the P particle with a number of oligosaccharides with structures found in human milk and the cell wall of mycobacteria. The affinities of these newly discovered ligands for the NoV are comparable to those of the HBGA receptors.

In Chapter 4, the intrinsic affinities of HBGA oligosaccharides for the NoV VA387 virus-like particles (VLPs) and the associated subviral P particles were characterized using an *indirect* ESI-MS assay, the *proxy protein* method. The affinities of thirteen HBGA oligosaccharides, containing A, B and H epitopes, with variable sizes (disaccharide to tetrasaccharide) and different precursor chain types (type 1, 2, 3, 5 and 6), were measured for the P particle, while the affinities of the A and B trisaccharides and A and B type 6 tetrasaccharides for the VLP were determined. Comparison of the binding data obtained for the P particle and VLP in this study with those measured for the corresponding P dimer revealed that the HBGA oligosaccharides tested exhibit similar intrinsic affinities for the P dimer and P particle. However, the  $K_{a,int}$  values for the VLP are consistently higher than those for the P dimer, but within a factor of three. While the cause of the subtle differences in HBGA oligosaccharide affinities for the P dimer and P particle and those for the VLP remains unknown, the present data support the use of P dimers or P particles as surrogates to the VLP for human NoV-receptor binding studies.

In Chapter 5, we reported the first experimental evidence that sialic acid-containing glycosphingolipids (gangliosides) are also ligands for human NoVs. ESI-MS based carbohydrate binding measurements performed on assemblies (P dimer, P particle and VLP) of recombinant viral capsid proteins of two human NoV strains, VA387 and VA115, identified their binding to the oligosaccharides of mono-, di- and tri-sialylated gangliosides. The intrinsic (per binding site) affinities measured for these ligands are similar in magnitude ( $10^2 - 10^3 \text{ M}^{-1}$ ) to those of HBGAs. Binding of NoV VLPs, P particles and glutathione S-transferase (GST)-P domain fusion proteins to sialic acid-containing glycoconjugates, observed in enzyme-linked immunosorbent assays,

provided additional confirmation of the NoV-ganglioside interactions.

Chapter 6 described the results of a systematic ESI-MS study aimed at elucidating the processes that influence binding of water soluble proteins to glycolipids (GLs) incorporated into nanodiscs (NDs), and to exploit these insights to quantify the binding energetics. The interactions between the cholera toxin B subunit homopentamer (CTB<sub>5</sub>) and its native ganglioside receptor, GM1, and between a recombinant fragment of family 51 carbohydrate-binding module (CBM), originating from *S. pneumoniae*, with a synthetic B type 2 tetrasaccharide neoglycolipid (B2<sub>NGL</sub>) served as model protein-GL complexes for this study. The results of the ESI-MS measurements revealed that proteins bind reversibly to ND-bound GLs and that proteins possessing multiple ligand binding sites are able to interact with GLs originating from different NDs. Experimental evidence suggested that the diffusion of GLs between NDs is rapid and influences the nature of the protein-GL complexes that are detected. Using a newly developed ESI-MS assay, the *proxy ligand* method, the association constants for the CBM-B2<sub>NGL</sub> and CTB<sub>5</sub>-GM1 interactions were quantified and found to be slightly smaller than those for the corresponding oligosaccharides in solution.

## 1.6 References

- (1) Ghazarian, H.; Idoni, B.; Oppenheimer, S. B. *Acta Histochem.* **2011**, *113*, 236.
- (2) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637.
- (3) Sharon, N.; Lis, H. *Sci. Am.* **1993**, *268*, 82.
- (4) Quijcho, F. A. *Annu. Rev. Biochem.* **1986**, *55*, 287.

- (5) Holgersson, J.; Gustafsson, A.; Breimer, M. E. *Immunol. Cell Biol.* **2005**, *83*, 694.
- (6) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *65*, 441.
- (7) Sacchettini, J. C.; Baum, L. G.; Brewer, C. F. *Biochemistry* **2001**, *40*, 3009.
- (8) Palmer, R. A.; Niwa, H., *Biochem. Soc. Trans.* **2003**, *31*, 973.
- (9) Greco, A.; Ho, J. G. S.; Lin, S.-J.; Palcic, M. M.; Rupnik, M.; Ng, K. K.-S. *Nat. Struct. Mol. Biol.* **2006**, *13*, 460.
- (10) Ho, J. G. S.; Greco, A.; Rupnik, M.; Ng, K. K.-S. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18373.
- (11) Staunton, D.; Owen, J.; Campbell, I. D. *Acc. Chem. Res.* **2002**, *36*, 207.
- (12) Larsen, K.; Thygesen, M. B.; Guillaumie, F.; William, G. T.; Willats, W. G. T.; Jensen, K. J. *Carb. Res.* **2006**, *341*, 1209.
- (13) Wang, D. N.; Liu, S. Y.; Trummer, B. J.; Deng, C.; Wang, A. L. *Nat. Biotechnol.* **2002**, *20*, 275.
- (14) Liu, Y.; Palma, A. S.; Feizi, T. *Biol. Chem.* **2009**, *390*, 647.
- (15) Schuck, P. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 541.
- (16) Homola, J. *Anal. Bioanal. Chem.* **2003**, *377*, 528.
- (17) Daghestani, H. N.; Day, B. W. *Sensors* **2010**, *10*, 9630.
- (18) De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. *Cell. Mol. Bioeng.* **2008**, *1*, 204.
- (19) Karlsson, R. *J. Mol. Recognit.* **2004**, *17*, 151.
- (20) Disney, M. D.; Seeberger, P. H. *Drug Discov. Today: Targets* **2004**, *3*, 151.

- (21) Grant, O. C.; Smith, H. M.; Firsova, D.; Fadda, E.; Woods, R. J. *Glycobiology* **2014**, *24*, 17.
- (22) Fais, M.; Karamanska, R.; Russell, D. A.; Field, R. A. *J. Cereal Sci.* **2009**, *50*, 306–311.
- (23) Day, Y. S. N.; Baird, C. L.; Rich, R. L.; Myszka, D. G. *Protein Sci.* **2002**, *11*, 1017.
- (24) Moll, D.; Schweinsberg, S.; Hammann, C.; Friedrich, W.; Herberg, F. W. *Biol. Chem.* **2007**, *388*, 163.
- (25) Saboury, A. A. *J. Iran. Chem. Soc.* **2006**, *3*, 1.
- (26) Wilcox, D. E. *Inorg. Chim. Acta* **2008**, *361*, 857.
- (27) Velázquez Campoy, A.; Freire, E. *Biophys. Chem.* **2005**, *115*, 115.
- (28) Wishart, D. *Curr. Pharm. Biotechnol.* **2005**, *6*, 105.
- (29) Zuiderweg, E. R. P. *Biochemistry* **2001**, *41*, 1.
- (30) Goldflam, M.; Tarrago, T.; Gairi, M.; Giralt, E., *Methods Mol. Biol.* **2012**, *831*, 233.
- (31) Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1784.
- (32) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108.
- (33) Haselhorst, T.; Lamerz, A. C.; Itzstein, M., *Methods Mol. Biol.* **2009**, *534*, 375.
- (34) Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. *J. Chem. Educ.* **2011**, *88*, 990.
- (35) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294.
- (36) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818.
- (37) Van Dongen, W. D.; Heck, A. J. R. *Analyst* **2000**, *125*, 583.
- (38) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605.
- (39) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945.

- (40) Cederkvist, F. H.; Zamfir, A. D.; Bahrke, S.; Eijnsink, V. G.; Sorlie, M.; Peter-Katalinic, J.; Peter, M. G., *Angew. Chem. Int. Ed.* **2006**, *45*, 2429.
- (41) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (42) Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587.
- (43) Jecklin, M. C.; Touboul, D.; Bovet, C.; Wortmann, A.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 332.
- (44) Jecklin, M. C.; Touboul, D.; Jain, R.; Toole, E. N.; Tallarico, J.; Drueckes, P.; Ramage, P.; Zenobi, R. *Anal. Chem.* **2009**, *81*, 408.
- (45) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1.
- (46) Kitova, E. N.; El-Hawiet, A.; Schnier, P.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.
- (47) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.
- (48) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discov.* **2006**, *5*, 585.
- (49) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.
- (50) Feizi, T.; Chai, W. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 582.
- (51) Lopez, P. H.; Schnaar, R. L. *Methods Enzymol.* **2006**, *417*, 205.
- (52) Kleinschmidt, J. H. Ed. *Lipid-Protein Interactions: Methods and Protocols*, **2014**, *974*, Springer: New York.
- (53) Catimel, B.; Scott, A. M.; Lee, F. T.; Hanai, N.; Ritter, G.; Welt, S.; Old, L. J.; Burgess,

- A. W.; Nice, E. C. *Glycobiology* **1998**, *8*, 927.
- (54) Campanero-Rhodes, M. A.; Smith, A.; Chai, W.; Sonnino, S.; Mauri, L.; Childs, R. A.; Zhang, Y.; Ewers, H.; Helenius, A.; Imberty, A.; Feizi, T. *J. Virol.* **2007**, *81*, 12846.
- (55) Feizi, T. *Ann. N. Y. Acad. Sci.* **2013**, *1292*, 33.
- (56) Song, X.; Heimbürg-Molinari, J.; Cummings, R. D.; Smith, D. F. *Curr. Opin. Chem. Biol.* **2014**, *18*, 70.
- (57) Palma, A. S.; Feizi, T.; Childs, R. A.; Chai, W.; Liu, Y. *Curr. Opin. Chem. Biol.* **2014**, *18*, 87.
- (58) Lingwood, C. A.; Manis, A.; Mahfoud, R.; Khan, F.; Binnington, B.; Mylvaganam, M. *Chem. Phys. Lipids* **2010**, *163*, 27.
- (59) Hooper, N. M. *Curr. Biol.* **1998**, *8*, R114.
- (60) Lingwood, C. A. *Glycoconjugate J.* **1996**, *13*, 495.
- (61) Lingwood, D.; Binnington, B.; Rog, T.; Vattulainen, I.; Grzybek, M.; Coskun, U.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.
- (62) Mahfoud, R.; Manis, A.; Binnington, B.; Ackerley, C.; Lingwood, C. A. *J. Biol. Chem.* **2010**, *285*, 36049.
- (63) Yahi, N.; Aulas, A.; Fantini, J. *PLoS ONE* **2010**, *5*, e9079.
- (64) Gallegos, K. M.; Conrady, D. G.; Karve, S. S.; Gunasekera, T. S.; Herr, A. B.; Weiss, A. *PLoS ONE* **2012**, *7*, e30368.
- (65) Hammache, D.; Piéroni, G.; Yahi, N.; Delézay, O.; Koch, N.; Lafont, H.; Tamalet, C.; Fantini, J. *J. Biol. Chem.* **1998**, *273*, 7967.

- (66) Czogalla, A.; Grzybek, M.; Jones, W.; Coskun, U. *Biochim. Biophys. Acta.* **2014**, *1841*, 1049.
- (67) Nath, A.; Atkins, W. M.; Sligar, S. G. *Biochemistry* **2007**, *46*, 2059.
- (68) Bayburt, T. H.; Sligar, S. G. *FEBS Letters* **2010**, *584*, 1721.
- (69) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477.
- (70) Cho, H.; Wu, M.; Bilgin, B.; Walton, S. P.; Chan, C. *Proteomics* **2012**, *12*, 3273.
- (71) Sanghera, N.; Correia, B. E.; Correia, J. R.; Ludwig, C.; Agarwal, S.; Nakamura, H. K.; Kuwata, K.; Samain, E.; Gill, A. C.; Bonev, B. B.; Pinheiro, T. J. *Chem. Biol.* **2011**, *18*, 1422.
- (72) Shi, J.; Yang, T.; Kataoka, S.; Zhang, Y.; Diaz, A. J.; Cremer, P. S. *J. Am. Chem. Soc.* **2007**, *129*, 5954.
- (73) Rao, C. S.; Lin, X.; Pike, H. M.; Molotkovsky, J. G.; Brown, R. E. *Biochemistry* **2004**, *43*, 13805.
- (74) Chen, W. C.; Kawasaki, N.; Nycholat, C. M.; Han, S.; Pilotte, J.; Crocker, P. R.; Paulson, J. C. *PLoS One* **2012**, *7*, e39039.
- (75) Borch, J.; Torta, F.; Sligar, S. G.; Roepstorff, P. *Anal. Chem.* **2008**, *80*, 6245.
- (76) Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618.
- (77) Leney, A.; Fan, X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271.
- (78) Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 972A.

- (79) Fenn, J. B. *Angew. Chem. Int. Ed.* **2003**, *42*, 3871.
- (80) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804.
- (81) Kebarle, P.; Verkerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898.
- (82) Taylor, G. I. *J. Fluid Mech.* **1965**, *2*, 1.
- (83) Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. *J. Phys. Chem. A* **2002**, *106*, 9957.
- (84) Peschke, M.; Verkerk, U. H.; Kebarle, P. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1424.
- (85) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287;
- (86) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1979**, *71*, 4451.
- (87) Dole, M.; Mack, L. L.; Hines, R. L. *J. Chem. Phys.* **1968**, *49*, 2240.
- (88) de la Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93.
- (89) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2.
- (90) Ahadi, E.; Konermann, L. *J. Phy. Chem. B* **2012**, *116*, 104.
- (91) Konermann, L.; Rodriguez, A. D.; Liu, J. *Anal. Chem.* **2012**, *84*, 6798.
- (92) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1.
- (93) Karas, M.; Bahr, U.; Dulcks, T. *Fresenius J. Anal. Chem.* **2000**, *366*, 669.
- (94) Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300.
- (95) De Hoffmann, E.; Stroobant, V. *Mass Spectrometry Principles and Applications* 3th Ed. **2007**, John Wiley & Sons: New York.
- (96) March, R. E.; Hughes, R. J. *Quadrupole Storage Mass Spectrometry*, **1989**, John Wiley & Sons: New York.
- (97) Pedder, R. E. *Ardara Technologies Technical Note* **2009**, TN\_3005A.

- (98) Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrom. Rev.* **2000**, *19*, 65.
- (99) Mamyryn, B. A. *Int. J. Mass Spectrom.* **2001**, *206*, 251.
- (100) Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinou, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. *Int. J. Mass Spectrom.* **2007**, *261*, 1.
- (101) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401.
- (102) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599.
- (103) Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49.
- (104) Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339.
- (105) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265.
- (106) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857.
- (107) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528.
- (108) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859.
- (109) Polfer, N. C. *Chem. Soc. Rev.* **2011**, *40*, 2211.
- (110) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9*, 4205.
- (111) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 12646.
- (112) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640.

- (113) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194.
- (114) Wytttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2007**, *58*, 511.
- (115) Versluis, C.; Heck, A. J. R. *Int. J. Mass Spectrom.* **2001**, *210*, 637.
- (116) McCammon, M. G.; Hernandez, H.; Sobott, F.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 5950.
- (117) Ilag, L. L.; Ubarretxena-Belandia, I.; Tate, C. G.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 14362.
- (118) Borch, J.; Jorgensen, T. J. D.; Roepstorff, P. *Curr. Opin. Chem. Biol.* **2005**, *9*, 509.
- (119) Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, *13*, 597.
- (120) Sharon, M.; Taverner, T.; Ambroggio, X. I.; Deshaies, R. J.; Robinson, C. V. *PLoS. Biol.* **2006**, *4*, 1314.
- (121) Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V. *Chem. Rev.* **2007**, *107*, 3544.
- (122) Lorenzen, K.; Vannini, A.; Cramer, P.; Heck, A. J. R. *Structure* **2007**, *15*, 1237.
- (123) Damoc, E.; Fraser, C. S.; Zhou, M.; Videler, H.; Mayeur, G. L.; Hershey, J. W. B.; Doudna, J. A.; Robinson, C. V.; Leary, J. A. *Mol. Cell. Proteomics* **2007**, *6*, 1135.
- (124) Tešić, M.; Wicki, J.; Poon, D. K. Y.; Withers, S. G.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 64.
- (125) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50.

- (126) Karasek, F. W. *Anal. Chem.* **1974**, *46*, 710A.
- (127) Creaser, C. S. ; Griffiths, J. R.; Bramwell, C. J.; Noreen, S.; Hill, C. A. ; Paul Thomas, C. L. *Analyst* **2004**, *129*, 984.
- (128) Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H. *J. Mass Spectrom.* **2008**, *43*, 1.
- (129) McCullough, B. J.; Kalapothakis, J.; Eastwood, H.; Kemper, P.; MacMillan, D.; Taylor, K.; Dorin, J.; Barran, P. E. *Anal. Chem.* **2008**, *80*, 6336.
- (130) Sergio, A.; Manel, A.; Marcelo, B. *Anal. Chim. Acta* **2011**, *703*, 114.
- (131) Cohen, M. J.; Karasek, F. W. *J. Chromatogr. Sci.* **1970**, *8*, 330.
- (132) Buryakov, I. A.; Krylov, E. V.; Nazarov, E. G.; Rasulev, U. K. *Int. J. Mass Spectrom. Ion Processes* **1993**, *128*, 143.
- (133) Shvartsburg, A. A.; Smith, R. D. *Anal. Chem.* **2008**, *80*, 9689.
- (134) Wildgoose, J.; McKenna, T.; Hughes, C.; Giles, K.; Pringle, S.; Campuzano, I.; Langridge, J.; Bateman, R. H. *Mol. Cell. Proteomics* **2006**, *5*, S14.
- (135) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protocols* **2008**, *3*, 1139.
- (136) Giles, K.; Williams, J. P.; Campuzano, I. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 1559.
- (137) Amster, I. J. *J. Am. Soc. Mass Spectrom.* **1996**, *31*, 1325.
- (138) Buchanan, M. V.; Hettich, R. L. *Anal. Chem.* **1993**, *65*, A245.
- (139) Marshall, A. G.; Grosshans, P. B. *Anal. Chem.* **1991**, *63*, A215.
- (140) Marshall, A. G.; Hendrickson, C. L. *Int. J. Mass Spectrom.* **2002**, *215*, 59.

- (141) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1.
- (142) El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S., *Anal. Chem.* **2012**, *84*, 3867.
- (143) Cheng, X. H.; Chen, R. D.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J. M.; Sigal, G. B.; Mammen, M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 8859.
- (144) Wortmann, A.; Rossi, F.; Lelais, G.; Zenobi, R. *J. Mass Spectrom.* **2005**, *40*, 777.
- (145) Wortmann, A.; Jecklin, M. C.; Touboul, D.; Badertscher, M.; Zenobi, R. *J. Mass Spectrom.* **2008**, *43*, 600.
- (146) Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 416.
- (147) Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801.
- (148) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646.
- (149) El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1893.
- (150) Hunter, E. P.; Lias, S. G. *Phys. Chem. Ref. Data* **1998**, *27*, 413.
- (151) Kitova, E. N.; Seo, M.; Roy, P. N.; Klassen, J. S. *J. Am. Chem. Soc.* **2008**, *130*, 1214.
- (152) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2005**, *77*, 3060.
- (153) Hossain, B. M.; Konermann, L. *Anal. Chem.* **2006**, *78*, 1613.
- (154) Daubenfeld, T.; Bouin, A. P.; van der Rest, G. *J. Am. Soc. Mass Spectrom.* **2006**, *17*,

1239.

- (155) Lane, L. A.; Ruotolo, B. T.; Robinson, C. V.; Favrin, G.; Benesch, J. L. P. *J. Mass Spectrom.* **2009**, *283*, 169.
- (156) Shimon, L.; Sharon, M.; Horovitz, A. *Biophys. J.* **2010**, *99*, 1645.
- (157) Sun, N.; Sun, J.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1242.
- (158) Kitova, E. N.; Soya, N.; Klassen, J. S. *Anal. Chem.* **2011**, *83*, 5160.
- (159) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (160) Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 8301.
- (161) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472.
- (162) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2010**, *82*, 2170.
- (163) Chitta, R. K.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1031.
- (164) Wilcox, J. M.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **2008**, *80*, 2365.
- (165) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491.
- (166) Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708.
- (167) Mathur, S.; Badertscher, S.; Scott, M.; Zenobi, R. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6187.
- (168) Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708.
- (169) Van Berkel, G. J.; Asano, K. G.; Schnier, P. D. *J. Am. Soc. Mass Spectrom.* **2001**, *12*,

853.

- (170) Yang, P.; Cooks, R. G.; Ouyang, Z.; Hawkridge, A. M.; Muddiman, D. C. *Anal. Chem.* **2005**, *77*, 6174.
- (171) Yao, Y.; Shams-Ud-Doha, K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2015**, *26*, 98.

## Chapter 2

### Affinities of Recombinant Norovirus P Dimers for Human

### Histo-Blood Group Antigens\*

#### 2.1 Introduction

Noroviruses (NoVs), a group of single-stranded, positive-sense RNA viruses in the *Calciviridae* family, are the major viral pathogens responsible for epidemic acute gastroenteritis in both developed and developing countries. Each year, the viruses infect roughly 20 million people,<sup>1,2</sup> resulting in approximately 200 000 deaths.<sup>3</sup> Currently, there is no effective vaccine or antiviral against NoV infections. Human NoVs can be divided into 2 major genogroups (GI and GII), which contain at least 25 different genotypes (GI.1-8 and GII.1-17).<sup>4</sup> The GII.4 is the predominant genotype worldwide causing ~80% of NoV gastroenteritis outbreaks.<sup>2,5-8</sup>

The absence of an *in vitro* cell culture system or a suitable animal model has hindered the characterization of NoVs. Consequently, efforts have focused on the recombinant virus-like particles (VLPs). In vitro expression of NoV VP1, the major capsid protein, through recombinant baculoviruses results in the spontaneous assembly of VLPs that are structurally and antigenically indistinguishable from the authentic viruses.<sup>9</sup> X-ray crystallography analysis of Norwalk virus VLPs revealed that each of VP1 contains two major domains, the N-terminal shell (S) domain and the C-terminal protrusion (P) domain, linked by a flexible hinge.<sup>10</sup> The S domain forms the

---

\* A version of this chapter has been published: Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S., *Glycobiology* **2013**, *23*, 276-285.

interior shell of the capsid, while the P domain is responsible for exterior P dimer formation. The P domain exhibits high sequence variability and is important for host-receptor interactions and the host immune response.<sup>11</sup> On its own, the P domain forms homodimers called the P dimers.<sup>12</sup> The P dimers can further assemble into larger complexes, a 12-mer small P particle<sup>13</sup> and a 24-mer P particle.<sup>14-16</sup> In addition, a soluble P protein in the stool of NoV-infected patients, referred to as P polypeptide, has been reported<sup>17,18</sup> and contains most of the P domain, but lacks the highly conserved arginine cluster at the C-terminus and forms a homodimer.<sup>19,20</sup>

NoVs recognize the human histo-blood group antigens (HBGAs),<sup>4,21,22</sup> which play an important role in host susceptibility of NoV. The HBGAs are complex carbohydrates that consist of oligosaccharides covalently linked to proteins or lipids. They are generally present on red blood cells, mucosal epithelia or as free antigens in body fluids, such as blood, saliva, milk and the intestinal contents.<sup>23</sup> Although the HBGA phenotype is determined by the terminal part of the oligosaccharide chain linked to protein or lipid, the antigen determinants can be associated with different carbohydrate structures, i.e., precursor chain types. There are six possible types of precursor chains.<sup>23</sup> Of these, types 1 – 4 are widely distributed in red blood cells, mucosal epithelia, as well as different organs,<sup>24</sup> whereas the type 6 chain mainly exists in milk and urine.<sup>23</sup> The type 5 structure has not been detected in human tissue or secretions. At present, the biological significance of the different HBGA chain types is not fully understood. The carbohydrate moieties of the HBGAs represent the minimum epitope for NoV recognition.<sup>21,25-27</sup> NoVs recognize human HBGAs in a strain-specific manner and distinct NoV-HBGA binding patterns have been described.<sup>21</sup> The HBGA-binding sites of NoVs are located at the P dimer

interface and the recombinant P dimers are shown to be structurally the same as those of VLP.<sup>11,28,29</sup> Thus, the recombinant P dimer and its complex form, the P particle, have been used as models for NoV-HBGA interactions extensively.<sup>12-15,19,25,30-33</sup>

At present there are few quantitative data available for the interactions between the NoV VLPs and HBGAs. Peters and coworkers recently investigated such interactions using saturation transfer difference nuclear magnetic resonance (STD-NMR) spectroscopy.<sup>27</sup> L-fucose was identified as the minimal structure recognized by a GII.4 VLP (Ast6139) and the association constants ( $K_a$ ) of VLP and the HBGA fragments containing  $\alpha$ -L-Fuc are, at best,  $\sim 10^4 \text{ M}^{-1}$ . In another study, affinities of  $2.6 \times 10^3 \text{ M}^{-1}$  and  $2.2 \times 10^3 \text{ M}^{-1}$  were measured for a GII.10 P dimer binding to H type 2 trisaccharide and L-fucose, respectively.<sup>34</sup>

Here, we describe the first quantitative study of the interactions between GII.4 P dimers, in both their full-length and truncated forms, with HBGA oligosaccharides using the *direct* electrospray ionization mass spectrometry (ESI-MS) assay.<sup>35</sup> The affinities of both P dimers for a library of 41 HBGA oligosaccharides, comprising A, B, H and Lewis antigens, were measured at 25 °C and pH 7.

## 2.2 Experimental Section

### 2.2.1 Proteins

Two forms of P dimers of NoV strain VA387 (GII.4) were studied. The first one, referred to as P dimer ( $P_2$ , MW 69,311 Da), was formed from full-length P domain with an amino acid sequence spanning residues 222 to 539 of VA387 VP1 (AAK84679.2). The second one, referred to as

truncated P dimer (tr-P<sub>2</sub>, MW 69,006 Da), was formed by a truncated P domain lacking the C-terminal arginine-cluster with a sequence spanning residues 222 to 535 of VA387 VP1.<sup>19</sup> Both P dimers were expressed in bacteria through Glutathione S-transferase (GST)-Gene Fusion System (GE-Healthcare Life Sciences, Piscataway, NJ) and purified as described previously.<sup>12,20</sup> A single chain fragment (scFv, MW 26,539 Da) of the monoclonal antibody Se155-4, which served as a reference protein (P<sub>ref</sub>) to correct ESI mass spectra for the occurrence of nonspecific ligand binding, was produced using recombinant technology as described elsewhere.<sup>36</sup> Each protein was concentrated and dialyzed against aqueous 50 mM ammonium acetate (pH 7) using Amicon Ultra 0.5 mL centrifugal filters (Millipore Corp., Billerica, MA) with a MW cutoff of 10 kDa and stored at -20 °C until needed. The concentrations of protein stock solutions were measured by UV absorption.

### 2.2.2 Carbohydrates

A complete list of the HBGA oligosaccharides, which range in size from di- to hexasaccharide, is given in Table 2.1. Compounds **L2–L3**, **L5–L7**, **L15–L20** and **L28–L33**<sup>37-39</sup> were donated by Prof. Todd Lowary (University of Alberta); compound **L27** was donated by Alberta Innovates Technology Futures (Alberta, Canada); compounds **L4**, **L40**, **L42**, **L44** and **L46** were purchased from Dextra (Reading, UK); compounds **L1**, **L8–L14**, **L21–L26**, **L34–L39**, **L41**, **L43**, **L45** and **L47** were purchased from Elicityl SA (Crolles, France). Two compounds, Lacto-N-tetraose (LNT, **L48**) and Lacto-N-neotetraose (LNnT, **L49**), which served as negative controls, were purchased from IsoSep AB (Tullinge, Sweden). Stock solutions of each oligosaccharide were prepared by

dissolving a known amount of the solid sample in ultrafiltered water (Milli-Q, Millipore) to yield a concentration of 1 mM. The solutions were stored at  $-20\text{ }^{\circ}\text{C}$  until needed.

### 2.2.3 Mass spectrometry

All of the binding measurements were carried out in positive ion mode using a 9.4T ApexQe Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker-Daltonics, Billerica, MA) equipped with a modified nanoflow ESI (nanoESI) source. NanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.) pulled to  $\sim 5\text{ }\mu\text{m}$  using a P-97 micropipette puller (Sutter Instruments, Novato, CA). A platinum wire was inserted into the nanoESI tip, and a capillary voltage of  $\sim 1.0\text{ kV}$  was applied to carry out ESI. Each ESI solution was prepared from stock solutions of P dimer (or truncated P dimer), one of HBGA oligosaccharides and Se155-4 scFv, which served as  $P_{\text{ref}}$ . Aqueous ammonium acetate (10 mM) was added to each solution. In all cases, the P dimer (or truncated P dimer) was incubated with the HBGA oligosaccharide for  $\sim 20\text{ min}$  at  $25\text{ }^{\circ}\text{C}$  before ESI-MS analysis. Ions/droplets produced by ESI were introduced into the mass spectrometer through a stainless steel capillary (i.d. 0.43 mm). The capillary voltage was 280 V. The ions were steered by a deflector (250 V) into the first funnel (150 V) and skimmer (20 V) and transmitted through the second funnel (7.5 V) and skimmer (5.0 V), and then accumulated in the first hexapole in-source accumulation cell (h1) for 0.6 sec. The ions were then transferred through the quadrupole (using a low  $m/z$  cut-off of 1500) followed by further accumulation in a second hexapole collision cell (h2) for 0.5 sec. Ions were then transferred to the ICR cell for detection. The pressure in ICR cell region was

$\sim 10^{-10}$  mbar. Data acquisition was performed using the ApexControl software (version 4.0, Bruker-Daltonics, Billerica, MA). The time-domain signal, consisting of the sum of 30 transients containing 32K data points per transient, was subjected to one zero-fill prior to Fourier-transformation.

## 2.2.4 Determination of $K_a$ values

A general description of the *direct* ESI-MS assay can be found in Chapter 1 and in references 35, 40 and 41. In this study, both the P dimer and truncated P dimer of NoV VA387 possess two independent and equivalent HBGA binding sites based on the X-ray crystal structures.<sup>11,20</sup> Specifically, for a protein (P) that has  $h$  independent and equivalent ligand binding sites, the intrinsic (per binding site) association constant  $K_{a,int}$  is taken into account, which describes the addition of the ligand (L) to an available binding site, eq 2.1:



and the concentrations of occupied, free and total binding sites are given by eqs 2.2a – 2.2c:

$$[\text{occupied binding sites}] = \sum_{q=1}^h q[\text{PL}_q] \quad (2.2a)$$

$$[\text{free binding sites}] = h[\text{P}]_0 - \sum_{q=1}^h q[\text{PL}_q] \quad (2.2b)$$

$$[\text{total binding sites}] = h[\text{P}]_0 \quad (2.2c)$$

So  $K_{a,int}$  can be expressed by eq 2.3:

$$K_{a,int} = \frac{[\text{occupied binding sites}]}{[\text{unoccupied binding sites}][L]} = \frac{\sum_{q=1}^h q[PL_q]}{(h[P]_0 - \sum_{q=1}^h q[PL_q])[L]} \quad (2.3)$$

Using  $R_q$  determined from ESI-MS assay (as defined in eq 1.13), and the mass balance of eqs 2.4a and 2.4b,  $K_{a,int}$  can be calculated from eq 2.5a:

$$[P]_0 = [P] + \sum_{q=1}^h [PL_q] \quad (2.4a)$$

$$[L]_0 = [L] + \sum_{q=1}^h q[PL_q] \quad (2.4b)$$

$$K_{a,int} = \frac{\sum_{q=1}^h qR_q}{\left( h(1 + \sum_{q=1}^h R_q) - \sum_{q=1}^h qR_q \right) \left( [L]_0 - \frac{[P]_0 \sum_{q=1}^h qR_q}{1 + \sum_{q=1}^h R_q} \right)} \quad (2.5a)$$

$$= \frac{1}{\left( \frac{h(1 + \sum_{q=1}^h R_q)}{\sum_{q=1}^h qR_q} - 1 \right) \left( [L]_0 - \frac{[P]_0 \sum_{q=1}^h qR_q}{1 + \sum_{q=1}^h R_q} \right)}$$

where  $[P]_0$  and  $[L]_0$  are initial concentrations of P and L, respectively. Using the fraction of binding ( $f$ ) defined in eq 1.14, eq 2.5a can be rewritten as eq 2.5b:

$$K_{a,int} = \frac{1}{\left( \frac{1}{f} - 1 \right) ([L]_0 - h[P]_0 f)} \quad (2.5b)$$

Specifically, if  $h=1$ , the expression of  $K_{a,int}$  reduces to eq 1.10; and if  $h=2$  (for the current situation that P possesses two independent and equivalent binding sites),

$$f = \frac{R_1 + 2R_2}{2(1 + R_1 + R_2)} \quad (2.6)$$

$$K_{a,int} = \frac{R_1 + 2R_2}{(2 + R_1) \left( [L]_0 - \frac{[P]_0 (R_1 + 2R_2)}{1 + R_1 + R_2} \right)} \quad (2.7)$$

Moreover,  $K_{a,int}$  in eq 2.7 is related to the apparent association constant ( $K_{a,app}$ ) by a statistical factor of 2, i.e.,  $K_{a,app} = 2 \times K_{a,int}$ , which comes from the fact that either binding site of free P can be bound by L to form PL.

The  $K_{a,int}$  values reported in Table 2.1 correspond to the average value established from replicate ( $\geq 3$ ) measurements performed on at least three different HBGA concentrations. In all cases, each ESI mass spectrum was corrected for the occurrence of nonspecific HBGA-protein binding during the ESI process using the reference protein method.<sup>42</sup> As described elsewhere, this technique involves the addition of a reference protein ( $P_{ref}$ ), which does not bind specifically to the protein and ligand of interest, to the solution. The “true” abundance (in the absence of nonspecific binding) of the ligand-bound and unbound P-dimer is calculated from the measured abundances of ligand-bound and unbound P-dimer. The underlying assumption with the method, that nonspecific ligand binding is a random process and affects equally all proteins in solution regardless of their size or structure, has been rigorously tested and shown to be generally valid.<sup>42-44</sup>

For a limited number of oligosaccharides,  $K_a$  was determined using a titration approach,<sup>45</sup> where the initial concentration of the P dimer was fixed and the concentration of the HBGA ligand was varied. The value of  $K_{a,int}$  was established from nonlinear regression analysis of the

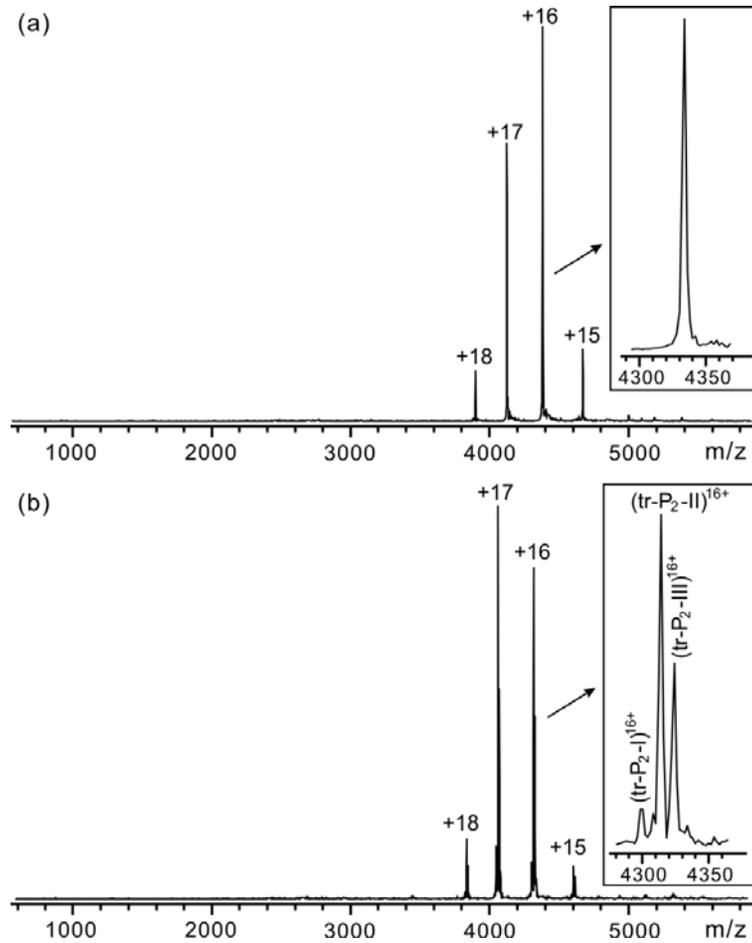
experimentally determined concentration-dependence of the fraction of ligand-bound protein, ( $f$ ), using eq 2.8:

$$f = \frac{1 + K_{a,int} h[P]_0 + K_{a,int} [L]_0 - \sqrt{(1 + K_{a,int} h[P]_0 - K_{a,int} [L]_0)^2 + 4K_{a,int} [L]_0}}{2K_{a,int} h[P]_0} \quad (2.8)$$

## 2.3 Results and Discussion

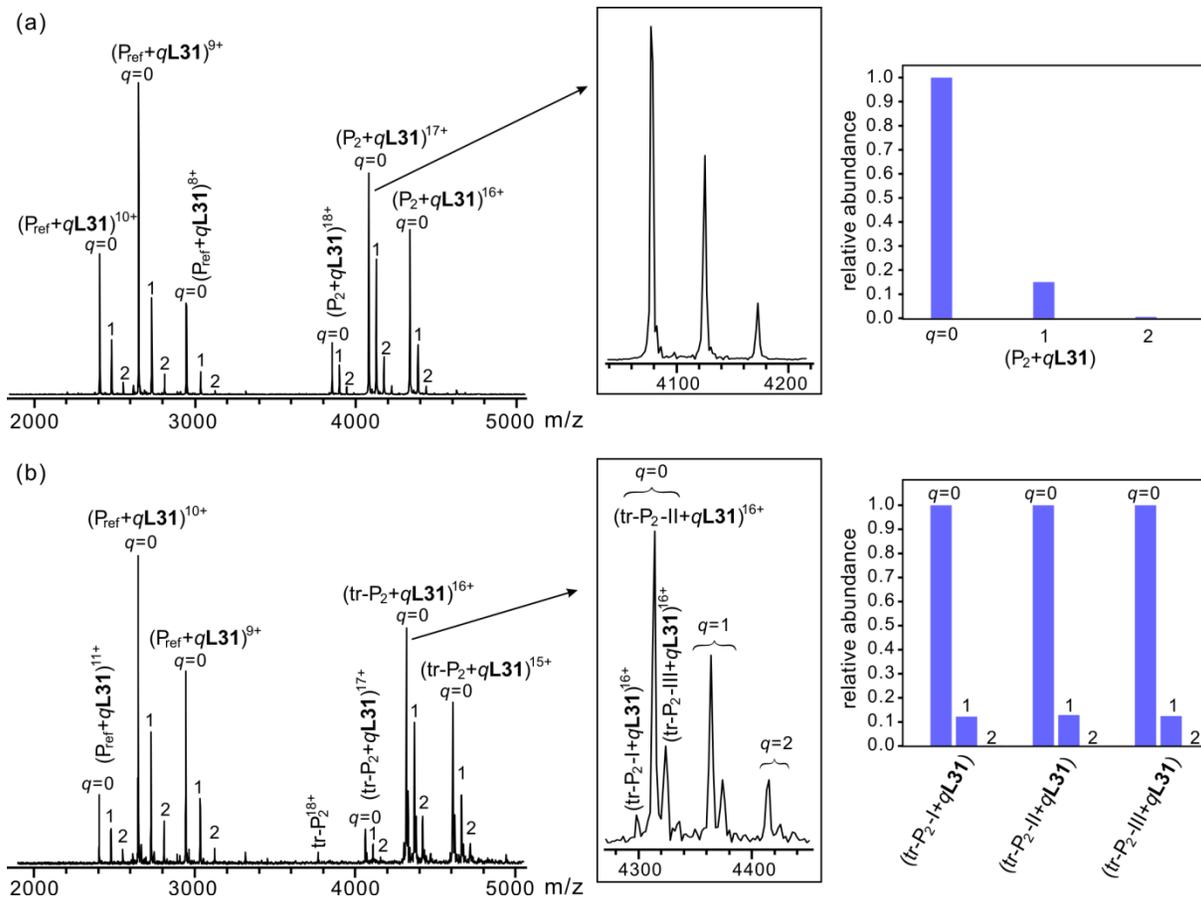
### 2.3.1 HBGA affinities for NoV P dimers

The *direct* ESI-MS assay was carried out to test for specific binding between the P dimer and the truncated P dimer and each of the 47 HBGA oligosaccharides and to quantify their affinities at pH 7 and 25 °C. Shown in Figure 2.1 are ESI mass spectra acquired for an aqueous ammonium acetate (10 mM) solution of the P dimer (12 μM) and of the truncated P dimer (12 μM). From the mass spectra it can be seen that the recombinant NoV VA387 P domain exists predominantly as a dimer (i.e., P<sub>2</sub>) under these solution conditions, with only protonated P<sub>2</sub><sup>n+</sup> ions detected. The measured MW of 69,312 ± 2 Da of the P dimer is in good agreement with the theoretical value of 69,311 Da. The truncated P domain also forms a dimer (i.e., tr-P<sub>2</sub>) under these conditions. However, the mass spectrum reveals evidence of three different protein species. In addition to signal for the protonated ions (tr-P<sub>2</sub><sup>n+</sup>) of the expected tr-P<sub>2</sub> (measured MW of 69,004 ± 2 Da, theoretical MW 69,006 Da), ions corresponding to protonated ions of proteins with MW of 68,763 ± 8 Da and 69,160 ± 10 Da were detected.



**Figure 2.1.** *Direct* ESI-MS analysis of norovirus VA387 P dimer and truncated-P dimer at pH 7 and 25 °C. Representative mass spectra acquired for 10 mM ammonium acetate solution with (a) P dimer (12  $\mu$ M) and (b) truncated P dimer (12  $\mu$ M). Three truncated-P dimer species (tr-P<sub>2</sub>(I), tr-P<sub>2</sub>(II) and tr-P<sub>2</sub>(III)) were resolved.

Shown in Figure 2.2a is a typical ESI mass spectra measured for an aqueous ammonium acetate solution (10 mM, pH 7, 25 °C) containing 12  $\mu$ M P dimer with 50  $\mu$ M B type 3 tetrasaccharide (**L31**). The scFv (10  $\mu$ M), which served as  $P_{\text{ref}}$ , was present in the ESI solutions used for affinity measurement. According to the ESI-MS data, the P dimer binds up to two molecules of **L31**, i.e.,  $(P_2 + q\mathbf{L31})^{n+}$ , where  $q = 0 - 2$  and  $n = 15 - 18$ . Signals corresponding to unbound and bound  $P_{\text{ref}}$  ions were also detected, i.e.,  $(P_{\text{ref}} + q\mathbf{L31})^{n+}$ , where  $q = 0 - 2$  and  $n = 9 - 11$ , which indicates that nonspecific binding of P dimer to **L31** occurred during the ESI process. As seen in the inset, after correction for nonspecific binding, few ions corresponding to specific  $(P_2 + 2\mathbf{L31})$  complex were identified. Therefore, the P dimer binds predominantly to one molecule of **L31** under the solution conditions, with a  $K_{a,\text{int}}$  of  $1500 \pm 150 \text{ M}^{-1}$ . Similarly, ES-MS binding measurement was performed on solution of truncated-P dimer with **L31**. As illustrated in Figure 2.2b, the three isoforms of tr-P<sub>2</sub> exhibit similar affinities for the HBGA, and a  $K_{a,\text{int}}$  of  $1200 \pm 100 \text{ M}^{-1}$  was obtained. Following the same procedure, the affinities of both P dimer and truncated-P dimer binding to each of the 47 HBGA oligosaccharides were quantified. A summary of the  $K_{a,\text{int}}$  values is listed in Table 2.1.



**Figure 2.2.** ESI mass spectra in positive ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing (a) norovirus VA387 P dimer (12  $\mu$ M) with 50  $\mu$ M **L31** (B type 3 tetrasaccharide, MW 801 Da); and (b) truncated-P dimer (12  $\mu$ M) with 60  $\mu$ M **L31**. A  $P_{ref}$  (10  $\mu$ M) was added to each solution to correct the mass spectra for the occurrence nonspecific carbohydrate-protein binding during ESI process. Insets: Normalized distribution of **L31** bound to the P dimer or three truncated-P dimer species (tr-P<sub>2</sub>(I), tr-P<sub>2</sub>(II) and tr-P<sub>2</sub>(III)) after correcting the ESI mass spectra for nonspecific ligand binding.

**Table 2.1.** Intrinsic association constants,  $K_{a,int}$  ( $M^{-1}$ ) for binding of the HBGA oligosaccharides (L1-L47) and HMOs (L48-L49) with norovirus VA387 P dimer and truncated (tr)-P dimer, measured at 25 °C and pH 7 by the *direct* ESI-MS assay.<sup>a</sup>

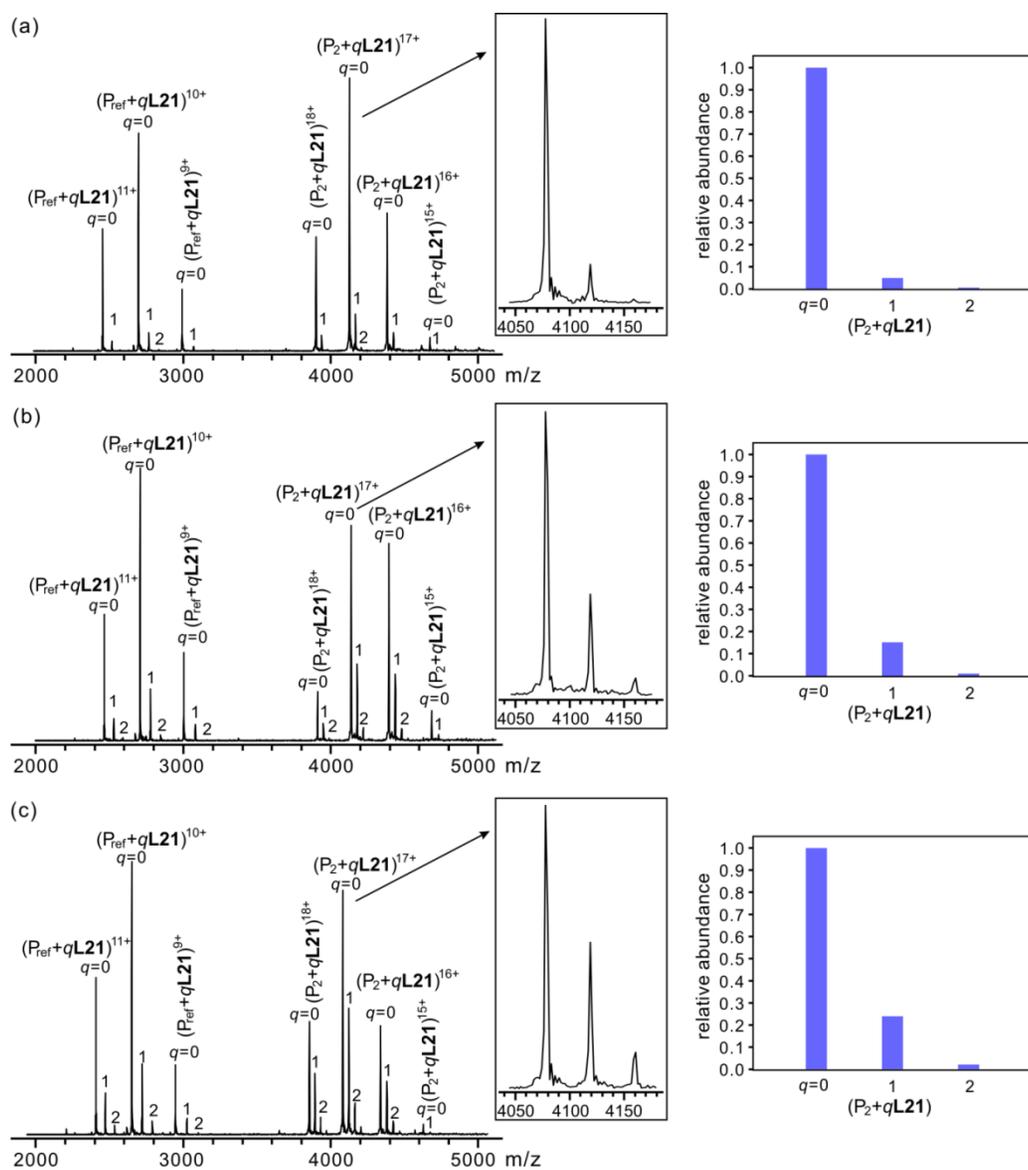
	MW / Da	HBGA	Structure	P dimer	tr-P dimer
L1	326.12	H disaccharide	$\alpha$ -L-Fuc-(1→2)-D-Gal	240 ± 35	330 ± 30
L2	510.27	H disaccharide	$\alpha$ -L-Fuc-(1→2)-D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	240 ± 40	n.d. <sup>d</sup>
L3	639.31	H type 1 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	250 ± 45	230 ± 35
L4	529.20	H type 2 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)-D-GlcNAc	200 ± 25	150 ± 30
L5	639.31	H type 3 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\alpha$ -D-GalNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	<b>650 ± 65</b>	<b>550 ± 150</b>
L6	639.31	H type 4 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	350 ± 70	n.d. <sup>d</sup>
L7	598.28	H type 5 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-Gal- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	350 ± 50	400 ± 45
L8	488.17	H type 6 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)-D-Glc	330 ± 25 [330 ± 10] <sup>b</sup>	250 ± 55
L9	691.25	H type 1 tetrasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc- (1→3)-D-Gal	320 ± 15	430 ± 25
L10	691.25	H type 2 tetrasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc- (1→4)-D-Gal	210 ± 40	250 ± 30
L11	691.25	H type 4 tetrasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc- (1→3)-D-Gal	290 ± 20	300 ± 50
L12	853.31	H type 1 pentasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc- (1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	310 ± 35	360 ± 40
L13	853.31	H type 2 pentasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc- (1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	250 ± 20	270 ± 40
L14	529.20	A trisaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]-D-Gal	430 ± 45	350 ± 15
L15	713.35	A trisaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]-D-Gal- O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	500 ± 50	n.d. <sup>d</sup>
L16	842.39	A type 1 tetrasaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal- (1→3)- $\beta$ -D-GlcNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	600 ± 65	600 ± 100

	MW / Da	HBGA	Structure	P dimer	tr-P dimer
L17	842.39	A type 2 tetrasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	290 $\pm$ 30	340 $\pm$ 40
L18	842.39	A type 3 tetrasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\alpha$ -D-GalNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	800 $\pm$ 50	550 $\pm$ 65
L19	842.39	A type 4 tetrasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc- O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	560 $\pm$ 130	n.d. <sup>d</sup>
L20	801.36	A type 5 tetrasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	560 $\pm$ 40	510 $\pm$ 25
L21	691.25	A type 6 tetrasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	<b>1200 <math>\pm</math> 50</b> <b>[1180 <math>\pm</math> 30]<sup>b</sup></b>	<b>850 <math>\pm</math> 80</b>
L22	894.33	A type 1 pentasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)-D-Gal	650 $\pm$ 50	420 $\pm$ 35
L23	894.33	A type 2 pentasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)-D-Gal	280 $\pm$ 30	220 $\pm$ 45
L24	894.33	A type 4 pentasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)-D-Gal	520 $\pm$ 30	500 $\pm$ 55
L25	1056.39	A type 1 hexasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	600 $\pm$ 60	600 $\pm$ 75
L26	1056.39	A type 2 hexasaccharide	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	310 $\pm$ 40	270 $\pm$ 45
L27	516.21	B trisaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-OC <sub>2</sub> H <sub>5</sub>	620 $\pm$ 45	600 $\pm$ 100
L28	672.32	B trisaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>7</sub> COOC <sub>2</sub> H <sub>5</sub>	800 $\pm$ 100	n.d. <sup>d</sup>
L29	801.36	B type 1 tetrasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	700 $\pm$ 100	550 $\pm$ 100
L30	801.36	B type 2 tetrasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	410 $\pm$ 45	400 $\pm$ 55
L31	801.36	B type 3 tetrasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\alpha$ -D-GalNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	<b>1500 <math>\pm</math> 150</b>	<b>1200 <math>\pm</math> 100</b>
L32	801.36	B type 4 tetrasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	700 $\pm$ 80	n.d. <sup>d</sup>
L33	760.34	B type 5 tetrasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	700 $\pm$ 100	550 $\pm$ 45
L34	650.23	B type 6 tetrasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	600 $\pm$ 45 [620 $\pm$ 30] <sup>b</sup>	550 $\pm$ 50
L35	853.31	B type 1 pentasaccharide	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)-D-Gal	530 $\pm$ 40	610 $\pm$ 30

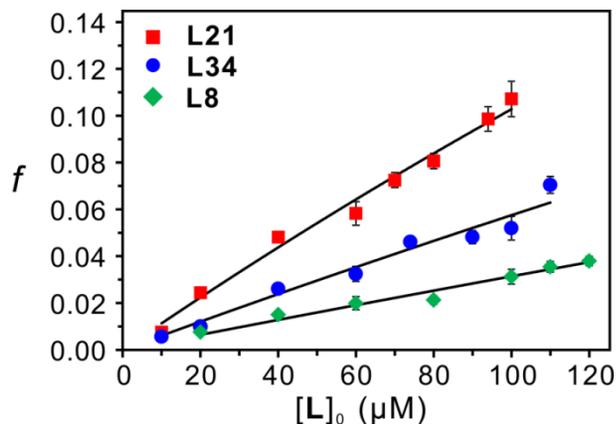
	MW / Da	HBGA	Structure	P dimer	tr-P dimer
L36	853.31	B type 2 pentasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-(1→3)-D-Gal	390 ± 20	370 ± 45
L37	853.31	B type 4 pentasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→3)-D-Gal	660 ± 30	600 ± 50
L38	1015.36	B type 1 hexasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	700 ± 50	650 ± 95
L39	1015.36	B type 2 hexasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	430 ± 30	410 ± 40
L40	529.20	Le <sup>a</sup> trisaccharide	$\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]-D-GlcNAc	180 ± 25	240 ± 35
L41	691.25	Le <sup>a</sup> tetrasaccharide	$\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]- $\beta$ -D-GlcNAc-(1→3)-D-Gal	240 ± 20	300 ± 45
L42	529.20	Le <sup>x</sup> trisaccharide	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]-D-GlcNAc	280 ± 35	350 ± 40
L43	691.25	Le <sup>x</sup> tetrasaccharide	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→3)-D-Gal	310 ± 25	340 ± 35
L44	675.26	Le <sup>b</sup> tetrasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]-D-GlcNAc	320 ± 35	300 ± 45
L45	837.31	Le <sup>b</sup> pentasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]- $\beta$ -D-GlcNAc-(1→3)-D-Gal	250 ± 30	250 ± 30
L46	675.26	Le <sup>y</sup> tetrasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]-D-GlcNAc	340 ± 20	350 ± 45
L47	837.31	Le <sup>y</sup> pentasaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→3)-D-Gal	250 ± 40	200 ± 35
L48	707.25	Lacto-N-tetraose	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	NB <sup>c</sup>	n.d. <sup>d</sup>
L49	707.25	Lacto-N-neotetraose	$\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	NB <sup>c</sup>	n.d. <sup>d</sup>

a. Uncertainties correspond to one standard deviation. b. Values obtained from the ESI-MS titration experiments. c. NB = no binding detected. d. n.d. = not determined.

Because the interactions between the P dimer and the HBGA oligosaccharides are quite weak, it was desirable to establish the reliability of the ESI-MS binding protocol used for the measurements. To this end, ESI-MS titration experiments were carried out on a small number of HBGA oligosaccharides (**L8**, **L21** and **L34**, which represent H, A and B type 6 antigens) using a fixed P dimer concentration (12  $\mu\text{M}$ ) and seven or more different HBGA concentrations (10 – 120  $\mu\text{M}$ ). Figure 2.3 show three typical ESI mass spectra measured for aqueous ammonium acetate solutions (10 mM) of P dimer (12  $\mu\text{M}$ ),  $P_{\text{ref}}$  (10  $\mu\text{M}$ ), and 20, 70 and 100  $\mu\text{M}$  A type 6 tetrasaccharide (**L21**), respectively. The normalized distributions (after correction for nonspecific binding) of  $(P_2 + q\mathbf{L21})$ ,  $q = 0 - 2$ , in the insets reveal an obvious trend that the fraction of **L21**-bound P dimer ( $f$ ) increases with ligand concentration. Figure 2.4 illustrates the plot of  $f$  versus **L21** concentrations and the curve of fitting eq 2.8 to the experimental data. Non-linear fitting yields a  $K_a$  value of  $1180 \pm 30 \text{ M}^{-1}$ . In a similar way,  $K_a$  values of  $330 \pm 10 \text{ M}^{-1}$  and  $620 \pm 30 \text{ M}^{-1}$  were determined for ligands **L8** and **L34**, respectively. Notably, the affinities obtained from the titration experiments are in excellent agreement with the values obtained using a limited number of HBGA oligosaccharide concentrations.



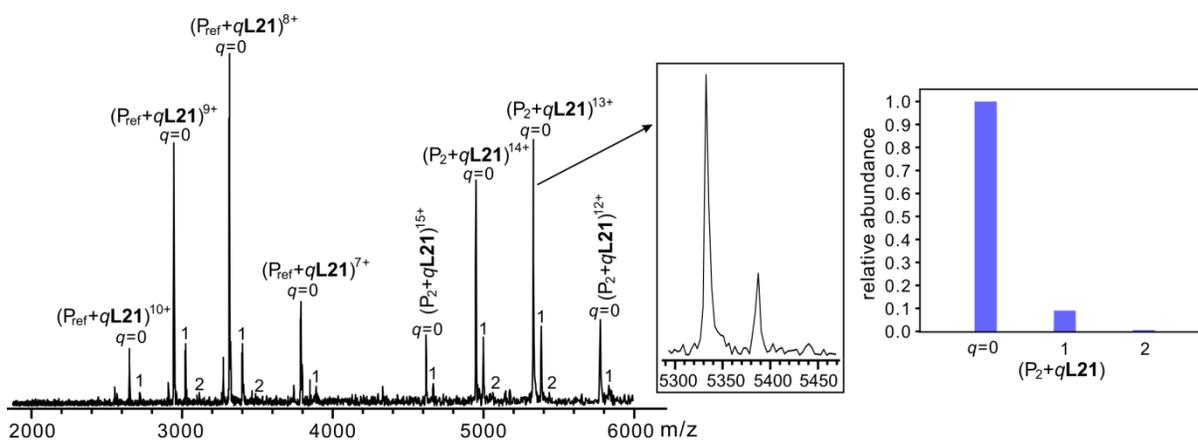
**Figure 2.3.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing norovirus VA387 P dimer (12 μM) with (a) 20 μM L21 (A type 6 tetrasaccharide, MW 691 Da), (b) 70 μM L21, and (c) 100 μM L21. A  $P_{ref}$  (10 μM) was added to each solution to correct the mass spectra for the occurrence nonspecific carbohydrate-protein binding during ESI process. Insets: Normalized distribution of L21 (at concentrations of (a) 20 μM, (b) 70 μM and (c) 100 μM) bound to the P dimer after correcting the ESI mass spectra for nonspecific ligand binding.



**Figure 2.4.** Fraction of ligand-bound P dimer (i.e.,  $f = (R_1 + 2R_2)/(1 + R_1 + R_2)/2$ , after correction for the nonspecific ligand binding) versus ligand concentration measured for **L8**, **L21** and **L34**, which represent H, A and B type 6 antigens, respectively. The titration experiments were carried out on aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing P dimer (12 μM),  $P_{ref}$  (10 μM) and ligand concentrations of between 10 and 120 μM. The solid curves correspond to the best fit of eq 2.8 to the experimental data for each ligand. The errors bars represent one standard deviation.

For a limited number of HBGAs (**L1**, **L5**, **L8**, **L14**, **L21**, **L25**, **L27**, **L31** and **L34**), imidazole (10 mM), which is known to stabilize labile protein-ligand complexes during ESI-MS analysis,<sup>46,47</sup> was added to the solution to test for the occurrence of in-source dissociation. Figure 2.5 shows a representative ESI mass spectrum acquired in positive ion mode for an aqueous ammonium acetate solution (10 mM, pH 7, 25 °C) containing P dimer (12 μM),  $P_{ref}$  (10 μM), **L21** (40 μM) and 10 mM imidazole. Notably, the addition of imidazole did not result in a measurable increase (after correction for nonspecific binding) in the relative abundance of

ligand-bound P dimer (Figure 2.5). P dimer-HBGA oligosaccharide affinities measured in the presence of imidazole are compared with the  $K_{a,int}$  values obtained without using imidazole (Table 2.2). These results established that the affinity measurements were not adversely affected by in-source dissociation.



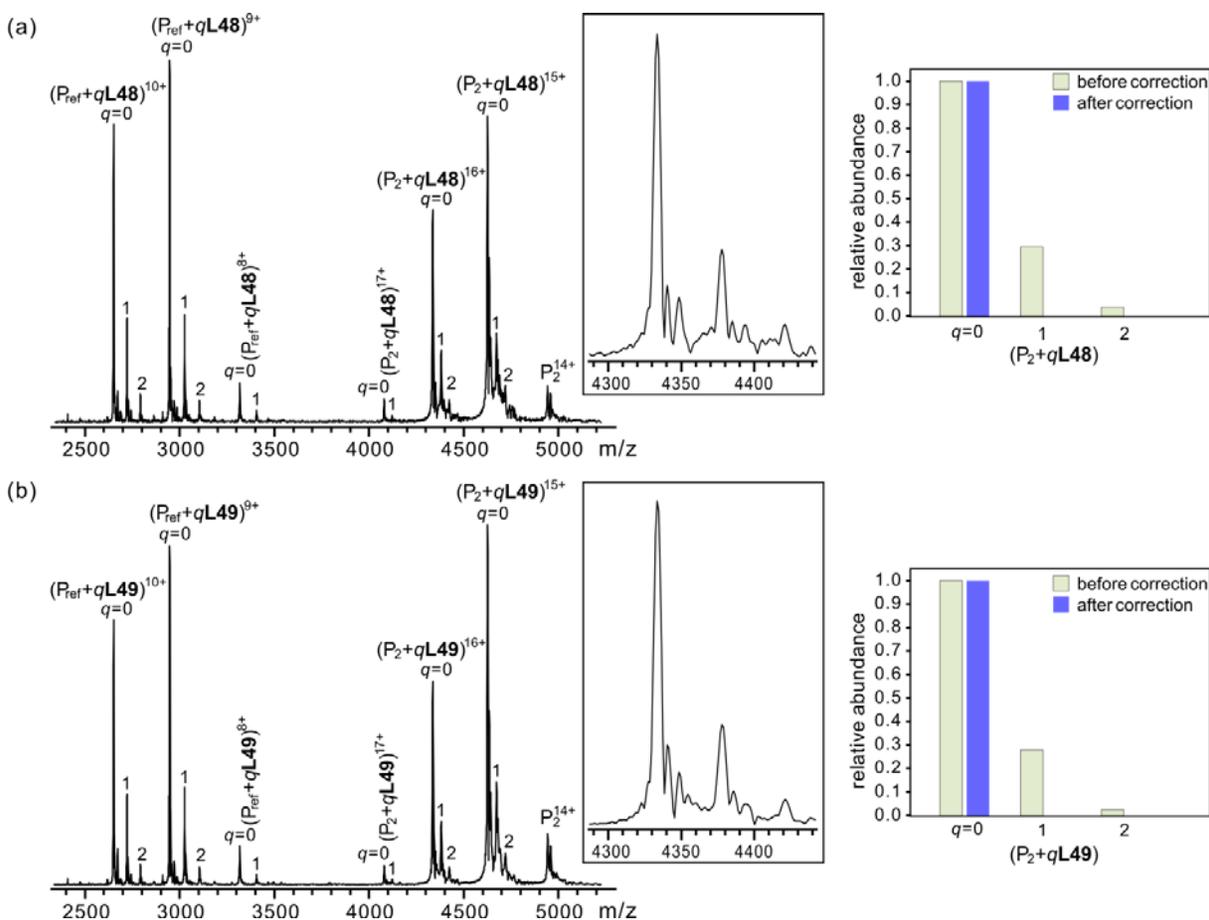
**Figure 2.5.** (a) ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate (10 mM) solution at pH 7 and 25 °C containing norovirus VA387 P dimer (12  $\mu$ M), L21 (40  $\mu$ M),  $P_{ref}$  (10  $\mu$ M) and imidazole (10 mM). Inset: Normalized distribution of L21 bound to the P dimer after correction for nonspecific binding determined from the mass spectra. The calculated  $K_{a,int}$  is  $1100 \pm 150 \text{ M}^{-1}$ .

**Table 2.2.** Intrinsic association constants,  $K_{a,int}$  ( $M^{-1}$ ) for binding of the HBGA oligosaccharides with norovirus VA387 P dimer measured by the *direct* ESI-MS assay.<sup>a</sup> All measurements were performed at 25 °C and pH 7 in 10 mM ammonium acetate buffer with /without 10 mM imidazole.

	MW / Da	HBGA	$K_{a,int}$ ( $M^{-1}$ ) with 10 mM imidazole	$K_{a,int}$ ( $M^{-1}$ ) <sup>b</sup> without imidazole
<b>L1</b>	326.12	H disaccharide	320 ± 45	240 ± 35
<b>L5</b>	639.31	H type 3 trisaccharide	600 ± 70	650 ± 65
<b>L8</b>	488.17	H type 6 trisaccharide	350 ± 70	330 ± 25
<b>L14</b>	529.20	A trisaccharide	450 ± 75	425 ± 45
<b>L21</b>	691.25	A type 6 tetrasaccharide	1100 ± 150	1200 ± 50
<b>L25</b>	1056.39	A type 1 hexasaccharide	650 ± 85	600 ± 60
<b>L27</b>	516.21	B trisaccharide	650 ± 45	620 ± 45
<b>L31</b>	801.36	B type 3 tetrasaccharide	1400 ± 120	1500 ± 150
<b>L34</b>	650.23	B type 6 tetrasaccharide	600 ± 100	600 ± 45

a. Uncertainties correspond to one standard deviation. b. Values duplicated from Table 2.1.

As a further test the reliability of the ESI-MS assay, measurements were performed on two human milk oligosaccharides (HMO), Lacto-N-tetraose (LNT, **L48**) and Lacto-N-neotetraose (LNnT, **L49**), which were reported not to bind to the P dimer.<sup>27</sup> Figure 2.6 shows representative ESI mass spectra acquired for aqueous ammonium acetate solutions (10 mM) with P dimer (12  $\mu$ M) and 80  $\mu$ M of **L48** or **L49** and scFv (10  $\mu$ M). Although there were signals corresponding to binding of each of the oligosaccharides to the P dimer, these were found to be due entirely to nonspecific binding.



**Figure 2.6.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing norovirus VA387 P dimer (12  $\mu$ M) and 80  $\mu$ M (a) **L48** (LNT) or (b) **L49** (LNnT). A  $P_{ref}$  (10  $\mu$ M) was added to each solution to correct the mass spectra for the occurrence nonspecific carbohydrate-protein binding during the ESI process. Insets: Normalized distributions of **L48** and **L49** bound to the P dimer before and after correction for nonspecific binding determined from the corresponding mass spectra.

### 2.3.2 Interpretation of the binding data

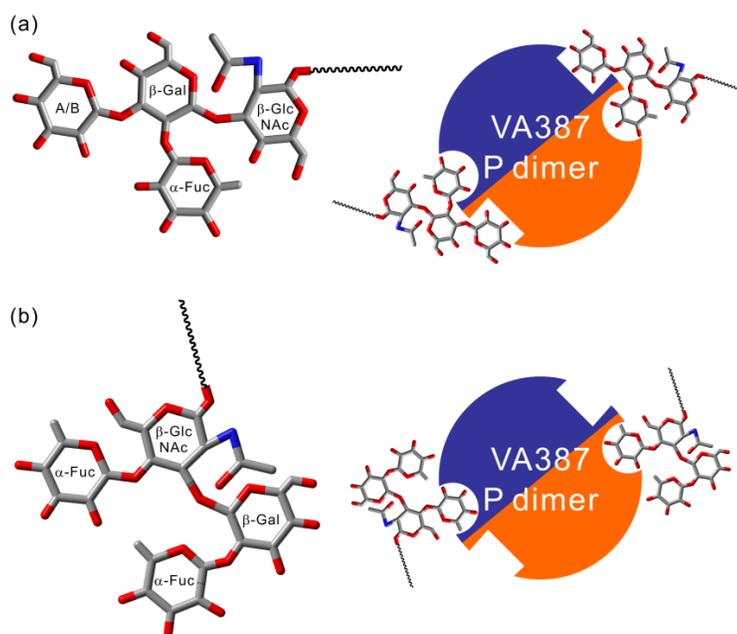
The ESI-MS binding measurements revealed that the P dimer recognizes all HBGA

oligosaccharides investigated. However, the binding ( $K_{a,int}$ ) is uniformly weak, ranging from  $\sim 200$  to  $\sim 1500$   $M^{-1}$ . The highest affinity ligands for P dimer are B tetrasaccharide type 3 (**L31**) and A tetrasaccharide type 6 (**L21**), with  $K_{a,int}$  values of  $1500 \pm 150$   $M^{-1}$  and  $1200 \pm 50$   $M^{-1}$ , respectively. Interestingly, the affinities measured for the P dimer are similar in magnitude to that estimated for the VLP of another GII.4 NoV (Ast6139).<sup>27</sup> Given that the P dimer only possesses two binding sites, while the VLP has 180 sites and assuming that the binding sites of the P dimer resembles those of the VLP, the apparent affinities of the VA387 VLP for these soluble oligosaccharides are expected to range from  $10^4$   $M^{-1}$  to  $10^6$   $M^{-1}$  (eq 1.15).

The binding data in Table 2.1 reveal that, overall, the A and B antigens bind more strongly to the P dimer than do the H and Lewis antigens. This finding appears to be consistent with the previously proposed NoV binding model, in which VA387 can recognize HBGAs through either a  $\alpha$ -D-GalNAc or  $\alpha$ -D-Gal epitope, and a  $\alpha$ -L-Fuc epitope, using two different binding pockets (Figure 2.7).<sup>4</sup> The A and B antigens possess both epitopes, which could be the reason for the stronger binding than those of the H and Lewis antigens, which lack the  $\alpha$ -D-GalNAc/ $\alpha$ -D-Gal epitope. The binding data also indicate that both 1,2-linked  $\alpha$ -L-Fuc (H epitope) and 1,3/4-linked  $\alpha$ -L-Fuc ( $Le^a$  and  $Le^x$  epitope) are recognized by the P dimer with comparable affinities.

The present data also indicate that the precursor chain type influences the strength of HBGA binding. For the H antigens, type 3 displays higher affinity over the remaining types 1, 2, 4, 5 and 6, which exhibit similar affinities ranging from 400 to 700  $M^{-1}$ . The finding that type 2 binds with comparable affinity as 1, 4, 5 and 6 is in agreement with results determined by the STD-NMR spectroscopy reported by Peters and coworkers.<sup>27</sup> However, this is inconsistent with

those measured by ELISA, which suggested that VA387 VLP does not bind to H type 2 trisaccharide.<sup>22</sup> For the A antigens, type 6 exhibits the strongest, type 2 the weakest; while types 1, 3, 4 and 5 exhibit similar affinities ranging from 520 to 800 M<sup>-1</sup>. For the B antigens, type 3 displays the highest affinity, and similar to the A antigens, type 2 the weakest; while types 1, 4, 5 and 6 exhibit similar affinities (in the range of 530 to 700 M<sup>-1</sup>). The K<sub>a,int</sub> values for these three strongest binders of H, A and B antigens, respectively, are highlighted in Table 2.1.



**Figure 2.7.** Interaction model for the norovirus VA387 P dimer binding to HBGA oligosaccharides adapted from the proposed model for norovirus-HBGA binding.<sup>4</sup> (a) Binding model proposed for A/B/H oligosaccharides; illustrated by an A/B/H type 1 antigen. (b) Binding model proposed for Lewis oligosaccharides; illustrated by a Le<sup>b</sup> antigen. This model assumes that VA387 P dimer can accommodate both the H epitope and A/B epitope independently in two nearby binding pockets. In addition, the VA387 P dimer can also accommodate Lewis epitope in the H epitope binding pocket.

The affinities of the truncated P dimer for the listed HBGAs were also measured (Table 2.1). Overall, they are similar in magnitude to those measured for the P dimer. These results indicate that elimination of the C-terminal arginine tail, which is remote from the binding pocket, does not influence substantially the binding of the P domain with the HBGAs. It is worthwhile the mention that although a previous study carried out by Tan and coworkers suggested that the removal of arginine tail eliminated binding to the HBGA,<sup>19</sup> a subsequent investigation using the linear polymer composed of glutathione S-transferase (GST)-truncated P domain fusion protein demonstrated that the truncated P dimer is still capable of binding to HBGAs.<sup>48</sup> A more detailed description of the GST-P domain fusion proteins can be found in reference 48 and Chapter 5.

## 2.4 Conclusion

In summary, the interactions between the NoV VA387 P dimer and a truncated P dimer with a library of 47 HBGA oligosaccharides were quantified for the first time. The results of the binding measurements performed at 25 °C and pH 7 indicate that the P dimer binds to all of the HBGA oligosaccharides tested with intrinsic affinities ranging from 200 to 1500 M<sup>-1</sup>. The affinities of the truncated P dimer, which lacks the Arg-cluster at the C-terminus, for the HBGAs are similar in magnitude to those measured for the P dimer. Based on the intrinsic affinities measured for the P dimer, the apparent affinities of NoV VA387 for the HBGA are estimated 10<sup>4</sup> to 10<sup>6</sup> M<sup>-1</sup>.

Overall, the P dimer exhibits higher affinities for the A and B antigens compared to those of the H and Lewis antigens. This finding is consistent with a proposed norovirus binding model, in which VA387 can recognize HBGAs through both a  $\alpha$ -D-GalNAc-(1→3) / $\alpha$ -D-Gal-(1→3)

epitope, and a  $\alpha$ -L-Fuc(1→2) epitope, using two different binding sites. From the binding measurements, the influence of chain type on the affinities was found to be:  $3 > 1 \approx 2 \approx 4 \approx 5 \approx 6$  for H antigens;  $6 > 1 \approx 3 \approx 4 \approx 5 > 2$  for A antigens;  $3 > 1 \approx 4 \approx 5 \approx 6 > 2$  for B antigens.

## 2.5 References

- (1) Lindesmith, L.; Moe, C.; Marionneau, S.; Ruvoen, N.; Jiang, X.; Lindbland, L.; Stewart, P.; LePendu, J.; Baric, R. *Nat. Med.* **2003**, *9*, 548.
- (2) Tan, M.; Jiang, X. *Expet. Rev. Mol. Med.* **2007**, *9*, 1.
- (3) Patel, M. M.; Widdowson, M.-A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (4) Tan, M.; Jiang, X. *Trends Microbiol.* **2005**, *13*, 285.
- (5) Yang, Y.; Xia, M.; Tan, M.; Huang, P.; Zhong, W.; Pang, X. L.; Lee, B. E.; Meller, J.; Wang, T.; Jiang, X. *J. Virol.* **2010**, *84*, 9595.
- (6) Baert, L.; Mattison, K.; Loisy-Hamon, F.; Harlow, J.; Martyres, A.; Lebeau, B.; Stals, A.; Van Coillie, E.; Herman, L.; Uyttendaele, M. *Int. J. Food Microbiol.* **2011**, *151*, 261.
- (7) Cannon, J. L.; Lindesmith, L. C.; Donaldson, E. F.; Saxe, L.; Baric, R. S.; Vinje, J. *J. Virol.* **2009**, *83*, 5363.
- (8) Johnston, C. P.; Qiu, H.; Ticehurst, J. R.; Dickson, C.; Rosenbaum, P.; Lawson, P.; Stokes, A. B.; Lowenstein, C. J.; Kaminsky, M.; Cosgrove, S. E.; Green, K. Y.; Perl, T. M. *Clin. Infect. Dis.* **2007**, *45*, 534.

- (9) Jiang, X.; Wang, M.; Graham, D. Y.; Estes, M. K. *J. Virol.* **1992**, *66*, 6527.
- (10) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (11) Cao, S.; Lou, Z.; Tan, M.; Chen, Y.; Liu, Y.; Zhang, Z.; Zhang, X. C.; Jiang, X.; Li, X.; Rao, Z. *J. Virol.* **2007**, *81*, 5949.
- (12) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (13) Tan, M.; Huang, P.; Xia, M.; Fang, P.; Zhong, W.; McNeal, M.; Wei, C.; Jiang, W.; Jiang, X. *J. Virol.* **2011**, *85*, 753.
- (14) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (15) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.
- (16) Bereszczak, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R. *J. Struct. Biol.* **2012**, *177*, 273.
- (17) Hardy, M. E.; White, L. J.; Ball, J. M.; Estes, M. K. *J. Virol.* **1995**, *69*, 1693.
- (18) Greenberg, H. B.; Valdesuso, J. R.; Kalica, A. R.; Wyatt, R. G.; McAuliffe, V. J.; Kapikian, A. Z.; Chanock, R. M. *J. Virol.* **1981**, *37*, 994.
- (19) Tan, M.; Meller, J.; Jiang, X. *J. Virol.* **2006**, *80*, 7322.
- (20) Bu, W.; Mamedova, A.; Tan, M.; Xia, M.; Jiang, X.; Hegde, R. S. *J. Virol.* **2008**, *82*, 5340.
- (21) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendou, J.; Jiang, X. *J. Infect. Dis.* **2003**,

188, 19.

- (22) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang, X. *J. Virol.* **2005**, *79*, 6714.
- (23) Oriol, R. *J. Immunogenet.* **1990**, *17*, 235.
- (24) Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1.
- (25) Tan, M.; Xia, M.; Cao, S.; Huang, P.; Farkas, T.; Meller, J.; Hegde, R. S.; Li, X.; Rao, Z.; Jiang, X. *Virology* **2008**, *379*, 324.
- (26) de Rougemont, A.; Ruvoen-Clouet, N.; Simon, B.; Estienney, M.; Elie-Caille, C.; Aho, S.; Pothier, P.; Le Pendu, J.; Boireau, W.; Belliot, G. *J. Virol.* **2011**, *85*, 4057.
- (27) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (28) Chen, Y.; Tan, M.; Xia, M.; Hao, N.; Zhang, X. C.; Huang, P.; Jiang, X.; Li, X.; Rao, Z. *PLoS Pathog.* **2011**, *7*, e1002152.
- (29) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.
- (30) Tan, M.; Zhong, W.; Song, D.; Thornton, S.; Jiang, X. *J. Med. Virol.* **2004**, *74*, 641.
- (31) Tan, M.; Xia, M.; Chen, Y.; Bu, W.; Hegde, R. S.; Meller, J.; Li, X.; Jiang, X. *PLoS One* **2009**, *4*, e5058.
- (32) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.
- (33) Tan, M.; Jiang, X. *Nanomedicine* **2012**, *7*, 1.
- (34) Hansman, G. S.; Shahzad-ul-Hussan, S.; McLellan, J. S.; Chuang, G.-Y.; Georgiev, I.;

- Shimoike, T.; Katayama, K.; Bewley, C. A.; Kwong, P. D. *J. Virol.* **2012**, *86*, 284.
- (35) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945.
- (36) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423.
- (37) Meloncelli, P. J.; Lowary, T. L. *Aust. J. Chem.* **2009**, *62*, 558.
- (38) Meloncelli, P. J.; Lowary, T. L. *Carbohydr. Res.* **2010**, *345*, 2305.
- (39) Meloncelli, P. J.; West, L. J.; Lowary, T. L. *Carbohydr. Res.* **2011**, *346*, 1406.
- (40) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.
- (41) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (42) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (43) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472.
- (44) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2010**, *82*, 2170.
- (45) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.
- (46) Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 416.
- (47) Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801.
- (48) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.

## Chapter 3

# Identifying Carbohydrate Ligands of a Norovirus P Particle using a Catch and Release Electrospray Ionization Mass Spectrometry Assay\*

### 3.1 Introduction

Noroviruses (NoVs), a group of non-enveloped, single-stranded positive-sense RNA viruses in the *Calciviridae* family, are the most common viral pathogens causing acute gastroenteritis.<sup>1</sup> Human histo-blood group antigens (HBGAs) have been shown to be the receptors or attachment factors of human NoVs that control the host susceptibility of NoV infection.<sup>2-9</sup> The HBGAs are either present as free oligosaccharides in bodily fluids, such as blood, saliva,- milk and the intestinal contents, or as complex carbohydrates covalently linked to proteins or lipids on red blood cells or mucosal epithelial cells.<sup>10-11</sup> The structures of the carbohydrates at the non-reducing end determine the type of HBGAs, including A, B, H or Lewis antigens. In addition, each HBGA can be further divided into six subtypes (types 1 – 6), based on their detailed carbohydrate structures at the reducing ends. The binding specificity and affinity of NoVs to HBGAs are strain or genotype dependent and different NoV-HBGA binding patterns are known.<sup>4</sup> For example, VA387, a member of the widely circulated GII.4 NoVs, has been shown to bind to a variety of HBGAs, including all A, B and H antigens and some Lewis antigens.<sup>3,4,12</sup> On the other hand, the MOH, a strain of GII.5 genotype, binds only saliva samples from type A and

---

\*A version of this chapter has been published: Han, L.; Kitova, E.; Tan, M.; Jiang, X.; Klassen, J., *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 111-119.

B individuals, while the VA207 (GII.9) shows a preference for Le<sup>X</sup> and Le<sup>Y</sup> antigens.<sup>3,4</sup>

Currently, there is no *in vitro* cell culture system or a suitable animal model for human NoVs, which has hindered the characterization of NoV receptors. The 7.7 kb RNA genome of NoVs has three open reading frames (ORFs), in which ORF2 encodes the capsid protein (VP1). VP1 possesses two major domains, the N-terminal shell (S) domain and the C-terminal protrusion (P) domain, linked by a short flexible hinge.<sup>13</sup> The S domain forms the interior capsid shell, which maintains the icosahedral structure of the virion, whereas the P domain forms an exterior dimeric structure that is important for the virus-receptor interactions and host immune response.<sup>14,15</sup> Recombinant VP1 can spontaneously assemble into a virus-like particle (VLP) *in vitro*.<sup>16</sup> X-ray crystallography of Norwalk virus (GI.1) VLPs revealed that the icosahedral NoV capsid is composed of 180 VP1s that organize into 90 homodimers.<sup>13</sup> The results of a recent electrospray ionization mass spectrometry (ESI-MS) study of Norwalk virus VLPs<sup>17</sup> support this observation.

The P domains of NoV VP1 can be structurally and functionally independent. Expression of the P domain alone produced P dimers with HBGA binding function.<sup>14,18,19</sup> Crystallography of NoV P dimers in complex with HBGAs demonstrated that the structure of the recombinant P dimers is identical to that of the native NoV capsids and revealed two symmetric HBGA binding sites on the tops of the P dimers,<sup>20-26</sup> which correspond to the outermost surface of the capsid. Two other P domain complexes with authentic HBGA binding function, the 24-mer P particles<sup>27-29</sup> and the 12-mer small P particles,<sup>30</sup> were also made through end-modifications of the P domain. These P domain complexes provide multiple tools for the study of NoV-host interactions,<sup>12,19,31-34</sup>

whereas the P particles have been shown as useful platforms for vaccine development against NoVs and other pathogens.<sup>29, 35–40</sup>

Identification of inhibitors as potential antivirals against NoVs represents an active area of NoV research. Using VLPs as a model, Jiang and coworkers screened a library comprising 5000 drug-like small molecules and identified 14 compounds that efficiently inhibited binding of VA387 VLP to HBGAs in a saliva-binding assay.<sup>41</sup> Screening a library of 340 compounds using saturation transfer difference-nuclear magnetic resonance (STD-NMR) spectroscopy and spin-lock filtered NMR spectroscopy, Peters and coworkers identified 26 compounds that bound to the HBGA binding sites of a GII.4 VLP.<sup>42</sup> In a more recent study using *in silico* screening of a large online library comprising over two million compounds, followed by a validation of blocking assays on P dimer-saliva interaction, five compounds that shared a common structure of cyclopenta [a] dimethyl phenanthren with an  $IC_{50} < 10.0 \mu M$  were identified.<sup>43</sup> Glycan array screening has also provided insights into the carbohydrate binding specificities of some human and murine VLPs. For example, the VLP of VA207 (GII.9) was shown to bind strongly to oligosaccharides with Lewis epitopes (1,3/4 linked  $\alpha$ -L-fucose residue), while the VLP of Norwalk virus (GI.1) was found to bind to a variety of structures not found in the HBGAs (Consortium for Functional Glycomics, <http://www.functionalglycomics.org/>).

Recently, ESI-MS has emerged as a promising tool for identifying and quantifying protein-carbohydrate interactions *in vitro*. In particular, the catch-and-release (CaR) ESI-MS assay enables the rapid screening of carbohydrate libraries against target proteins.<sup>44–46</sup> The assay involves incubating the target protein with a library of compounds, followed by direct ESI-MS

analysis of intact protein-ligand complexes. In principle, the identity of ligands “caught” by the protein can be found from the molecular weight (MW) of the corresponding protein-ligand complex. In cases where MW cannot be accurately determined or when dealing with isomeric species, the ligands are “released” (as ions) from the protein using collision-induced dissociation (CID), followed by accurate mass analysis, alone or in combination with ion mobility separation (IMS) or another stage of CID.<sup>46</sup> Because carbohydrates have relatively low gas-phase acidities and are able to effectively compete with the protein for negative charge, the assay is normally carried out in negative ion mode.<sup>46</sup> It has been shown that moderate-to-high affinity ligands ( $K_a > 10^4 \text{ M}^{-1}$ ) can be identified from libraries containing over 200 carbohydrates in a single CaR-ESI-MS measurement, which is typically completed within 1–2 min.<sup>46</sup>

Here, we report the application of the CaR ESI-MS assay for screening carbohydrate libraries against the NoV P particle. A series of control experiments, including the screening of a 50 compound library containing multiple HBGA oligosaccharides with known affinities for the corresponding P dimer,<sup>12</sup> were performed to confirm the reliability of the assay. The validated assay was then used to screen a carbohydrate library of 146 compounds to identify new carbohydrate ligands and potential inhibitors against NoV–HBGA interaction.

## **3.2 Experimental**

### **3.2.1 Proteins**

The P particle (24-mer, MW 865 036 Da) of NoV strain VA387 (GII.4) was produced from the P domain (residues 222-539) of VP1. A cysteine rich peptide CDCRGDCFC was linked to the C

terminus of the P domain to enhance the stability of the P particle.<sup>27,28</sup> The procedures used for the production and purification of the P particle have been described previously.<sup>27,28</sup> Prior to ESI-MS analysis, the P particle was concentrated (to a final concentration of 20  $\mu$ M) and exchanged into aqueous 200 mM ammonium acetate (pH 7) using Vivaspin 0.5 mL centrifugal filters (Sartorius Stedim Biotech, Göttingen, Germany) with a MW cutoff of 10 kDa and stored at  $-20$  °C until needed. The concentration of the P particle (24-mer) was measured using a Pierce BCA assay kit (Thermo Scientific, Ottawa, Canada) according to the manufacturer's instructions.

### 3.2.2 Carbohydrates

A full list of the carbohydrates used in the present study is shown in Table 3.1, along with their MWs. Compounds **L1–L6**, **L12–L13**, **L55–L63**, **L68**, **L83–L113** were gifts from Prof. T. Lowary (University of Alberta); **L16**, **L37–L38**, **L114–L115** were gifts from Alberta Innovates Technology Futures (Alberta, Canada); **L39–L40**, **L69–L70**, **L72–L82**, **L116–L117** were gifts from Prof. C-C. Ling (University of Calgary); **L14–L15**, **L28–30**, **L50**, **L128–L140** were purchased from Dextra (Reading, UK); **L41–L49** were purchased from Sigma-Aldrich Canada (Oakville, Canada); **L64–L67**, **L141–L143** were purchased from IsoSep AB (Tullinge, Sweden) and **L7–L11**, **L17–L27**, **L31–L36**, **L51–L54**, **L71**, **L118–L127** and **L144–L146** were purchased from Elicityl SA (Crolles, France). The structures of all the oligosaccharides used in this study are listed in Table 3.1. For each compound, an aqueous 2.5 mM stock solution was prepared by dissolving a known mass of the solid sample into ultrafiltered Milli-Q water (Millipore, Billerica, MA) and was stored at  $-20$  °C until needed. To apply the CaR-ESI-MS assay, solutions of P

particle (5  $\mu\text{M}$ ) with one or more carbohydrates (at a concentration of 10  $\mu\text{M}$  each) in 200 mM ammonium acetate (25  $^{\circ}\text{C}$ , pH 7) were prepared.

**Table 3.1.** Structures and molecular weights (MWs) of components of carbohydrate library.<sup>a</sup>

Carbohydrate	Structure	MW (Da)
L1	$\alpha\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\alpha\text{-D-GalNAc-O(CH}_2\text{)}_6\text{CH=CH}_2$	801.36
L2	$\alpha\text{-D-GalNAc-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\alpha\text{-D-GalNAc-O(CH}_2\text{)}_6\text{CH=CH}_2$	842.39
L3	$\alpha\text{-L-Fuc-(1}\rightarrow\text{2)-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\alpha\text{-D-GalNAc-O(CH}_2\text{)}_6\text{CH=CH}_2$	639.31
L4	$\alpha\text{-L-Fuc-(1}\rightarrow\text{2)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-Glc-O(CH}_2\text{)}_6\text{CH=CH}_2$	598.28
L5	$\alpha\text{-D-GalNAc-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-Glc-O(CH}_2\text{)}_6\text{CH=CH}_2$	801.36
L6	$\alpha\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-Glc-O(CH}_2\text{)}_6\text{CH=CH}_2$	760.34
L7	$\alpha\text{-L-Fuc-(1}\rightarrow\text{2)-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\beta\text{-D-GalNAc-(1}\rightarrow\text{3)-D-Gal}$	691.25
L8	$\alpha\text{-D-GalNAc-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\beta\text{-D-GalNAc-(1}\rightarrow\text{3)-D-Gal}$	894.33
L9	$\alpha\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\beta\text{-D-GalNAc-(1}\rightarrow\text{3)-D-Gal}$	853.31
L10	$\alpha\text{-D-GalNAc-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-Glc}$	1056.39
L11	$\alpha\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-Glc}$	1015.36
L12	$\alpha\text{-D-GalNAc-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)]-}\beta\text{-D-Gal-O(CH}_2\text{)}_8\text{COOC}_2\text{H}_5$	713.35
L13	$\alpha\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{2)-}]\beta\text{-D-Gal-O(CH}_2\text{)}_8\text{COOC}_2\text{H}_5$	672.32
L14	$\beta\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{4)]-D-GlcNAc}$	529.20
L15	$\alpha\text{-L-Fuc-(1}\rightarrow\text{2)-}\beta\text{-D-Gal-(1}\rightarrow\text{3)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{4)]-D-GlcNAc}$	675.26
L16	$\beta\text{-D-Gal-(1}\rightarrow\text{4)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{3)]-}\beta\text{-D-GlcNAc-O(CH}_2\text{)}_8\text{CONH(CH}_2\text{)}_2\text{NH}_2$	727.37
L17	$\alpha\text{-L-Fuc-(1}\rightarrow\text{2)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{3)]-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-D-Gal}$	837.31
L18	$\beta\text{-D-Gal-(1}\rightarrow\text{4)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{3)]-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}[\alpha\text{-L-Fuc-(1}\rightarrow\text{3)]-D-Glc}$	999.36
L19	$\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-Glc}$	545.20
L20	$\beta\text{-D-Gal-(1}\rightarrow\text{3)-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-Glc}$	707.25
L21	$\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-Glc}$	707.25
L22	$\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-}\beta\text{-D-GlcNAc-(1}\rightarrow\text{3)-}\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-Glc}$	1072.38

<b>Carbohydrate</b>	<b>Structure</b>	<b>MW (Da)</b>
<b>L23</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	1437.51
<b>L24</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)-D-Gal	545.20
<b>L25</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	707.25
<b>L26</b>	$\beta$ -D-GalNAc-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	545.20
<b>L27</b>	$\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	504.17
<b>L28</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 6)-[ $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	1072.38
<b>L29</b>	$\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-D-GlcNAc	830.33
<b>L30</b>	$\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-D-GlcNAc	1236.49
<b>L31</b>	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	869.30
<b>L32</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	869.30
<b>L33</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	869.30
<b>L34</b>	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	910.33
<b>L35</b>	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)-D-Gal	586.22
<b>L36</b>	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	910.33
<b>L37</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 3)- $\alpha$ -D-Man-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-O(CH <sub>2</sub> ) <sub>8</sub> COOCH <sub>3</sub>	674.30
<b>L38</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 3)- $\alpha$ -D-Man-(1 $\rightarrow$ 2)- $\alpha$ -D-Man-O(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	546.22
<b>L39</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	847.38
<b>L40</b>	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	847.38
<b>L41</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	504.17
<b>L42</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	666.22
<b>L43</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	828.27
<b>L44</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	990.33
<b>L45</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	1152.38
<b>L46</b>	$\alpha$ -D-Glc-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	1314.43
<b>L47</b>	$\beta$ -D-Glc-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	666.22
<b>L48</b>	$\beta$ -D-Glc-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-(1 $\rightarrow$ 4)-D-Glc	990.33
<b>L49</b>	$\alpha$ -D-Man-(1 $\rightarrow$ 6)-[ $\alpha$ -D-Man-(1 $\rightarrow$ 3)]- $\alpha$ -D-Man-(1 $\rightarrow$ 6)-[ $\alpha$ -D-Man-(1 $\rightarrow$ 3)]-D-Man	828.27
<b>L50</b>	$\alpha$ -D-Man-(1 $\rightarrow$ 6)-[ $\alpha$ -D-Man-(1 $\rightarrow$ 3)]- $\alpha$ -D-Man-(1 $\rightarrow$ 6)-D-Man	666.22
<b>L51</b>	$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	1015.36

Carbohydrate	Structure	MW (Da)
L52	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→3)- $\alpha$ -D-Gal-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	1218.44
L53	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→3)- $\alpha$ -D-Gal-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	1177.41
L54	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1144.40
L55	4',6'O-benzylidene- $\alpha$ -D-Glc-(1→4)- $\alpha$ -D-Glc-OCH <sub>3</sub>	444.16
L56	$\alpha$ -D-Araf-(1→3)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-OCH <sub>3</sub>	428.15
L57	$\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)-[ $\alpha$ -D-Araf-(1→3)]- $\alpha$ -D-Araf-OCH <sub>3</sub>	560.20
L58	$\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> N <sub>3</sub>	699.31
L59	$\alpha$ -D-Araf-(1→3)-[ $\alpha$ -D-Araf-(1→5)]- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	769.30
L60	$\alpha$ -L-Rha(1→3)-2-OCH <sub>3</sub> - $\alpha$ -L-Rha-O(p-OCH <sub>3</sub> Ph)	430.18
L61	$\alpha$ -L-Rha-(1→3)- $\alpha$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	479.27
L62	2,4-di-OCH <sub>3</sub> - $\alpha$ -L-Fuc-(1→3)- $\alpha$ -L-Rha-(1→3)-2-OCH <sub>3</sub> - $\alpha$ -L-Rha-O(p-OCH <sub>3</sub> Ph)	604.27
L63	2,3,4-tri-OCH <sub>3</sub> - $\alpha$ -L-Fuc-(1→3)- $\alpha$ -L-Rha-(1→3)-2-OCH <sub>3</sub> - $\alpha$ -L-Rha-O(p-OCH <sub>3</sub> Ph)	618.29
L64	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]-D-Glc	488.17
L65	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]-D-Glc	634.23
L66	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]-D-Glc	779.27
L67	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→6)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1364.50
L68	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]-D-GlcNAc-OCH <sub>2</sub> COOH	733.26
L69	$\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]- $\beta$ -D-GlcNAc-(1→3)-O(CH <sub>2</sub> ) <sub>2</sub> N <sub>3</sub>	963.37
L70	$\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]- $\beta$ -D-GalNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→3)-O(CH <sub>2</sub> ) <sub>2</sub> N <sub>3</sub>	1109.42
L71	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→3)-D-Gal	982.35
L72	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	993.44
L73	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -Fuc-(1→3)]- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	790.36
L74	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -Fuc-(1→3)]- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	628.31
L75	$\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -Fuc-(1→3)]- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	993.44
L76	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	1139.50
L77	$\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	831.38
L78	$\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>2</sub> NH <sub>2</sub>	775.32

Carbohydrate	Structure	MW (Da)
L79	$\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> N <sub>3</sub>	670.29
L80	$\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	806.35
L81	$\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> NH <sub>2</sub>	685.33
L82	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>2</sub> N <sub>3</sub>	817.31
L83	$\alpha$ -D-GlcNAc-(1 $\rightarrow$ 3)-D-Glc	383.14
L84	$\alpha$ -D-Tal-(1 $\rightarrow$ 3)-D-Glc	342.12
L85	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNH <sub>2</sub> x AcOH-OCH <sub>3</sub>	415.17
L86	$\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-OCH <sub>3</sub>	356.13
L87	$\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\alpha$ -D-Glc-OCH <sub>2</sub> COOH	400.12
L88	$\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-OCH <sub>2</sub> CCl <sub>3</sub>	554.08
L89	$\alpha$ -D-Araf-(1 $\rightarrow$ 5)-(2,3)-anhydro- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	376.21
L90	3,6-di-OCH <sub>3</sub> - $\beta$ -D-Glc-(1 $\rightarrow$ 4)-2,3-di-OCH <sub>3</sub> - $\alpha$ -L-Rha-(1 $\rightarrow$ 2)-3-OCH <sub>3</sub> - $\alpha$ -L-Rha-O(p-OCH <sub>3</sub> Ph)	648.30
L91	(6-OCH <sub>3</sub> )- $\beta$ -D-Glc-(1 $\rightarrow$ 4)-(2,3-di-OCH <sub>3</sub> )- $\alpha$ -L-Rha(1 $\rightarrow$ 2)-(3-OCH <sub>3</sub> )- $\alpha$ -L-Rha-O(p-OCH <sub>3</sub> Ph)	634.28
L92	$\beta$ -D-Galf-(1 $\rightarrow$ 5)- $\beta$ -D-Galf-(1 $\rightarrow$ 6)- $\beta$ -D-Galf-O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	616.29
L93	$\beta$ -D-Galf-(1 $\rightarrow$ 6)- $\beta$ -D-Galf-(1 $\rightarrow$ 5)- $\beta$ -D-Galf-O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	616.29
L94	$\alpha$ -D-Araf-(1 $\rightarrow$ 5)-[ $\beta$ -D-Galf-(1 $\rightarrow$ 5)- $\beta$ -D-Galf-(1 $\rightarrow$ 6)]- $\beta$ -D-Galf-O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	748.34
L95	$\beta$ -D-Galf-(1 $\rightarrow$ 6)-[ $\alpha$ -D-Araf-(1 $\rightarrow$ 5)]- $\beta$ -D-Galf-(1 $\rightarrow$ 5)- $\beta$ -D-Galf-O(CH <sub>2</sub> ) <sub>7</sub> CH <sub>3</sub>	748.34
L96	$\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)-[ $\alpha$ -D-Araf-(1 $\rightarrow$ 3)]- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-OCH <sub>3</sub>	692.24
L97	$\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 3)-[ $\alpha$ -D-Araf-(1 $\rightarrow$ 5)]- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-OCH <sub>3</sub>	692.24
L98	$\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 3)-[ $\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)]- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-OCH <sub>3</sub>	824.28
L99	5-SCH <sub>3</sub> - $\alpha$ -D-Xylf-(1 $\rightarrow$ 4)- $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp-(1 $\rightarrow$ 5)- $\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NH <sub>2</sub>	1321.51
L100	5-SCH <sub>3</sub> - $\alpha$ -D-Xylf-(1 $\rightarrow$ 4)- $\alpha$ -D-Manp-(1 $\rightarrow$ 5)- $\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)-[ $\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 3)]- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NH <sub>2</sub>	1261.49
L101	5-SCH <sub>3</sub> - $\alpha$ -D-Xylf-(1 $\rightarrow$ 4)- $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp-(1 $\rightarrow$ 5)- $\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NH <sub>2</sub>	1159.46
L102	$\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> N <sub>3</sub>	1227.48
L103	$\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 3)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> N <sub>3</sub>	1227.48
L104	$\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 3)-[ $\beta$ -D-Araf-(1 $\rightarrow$ 2)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)]- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-(1 $\rightarrow$ 5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> N <sub>3</sub>	1095.43

Carbohydrate	Structure	MW (Da)
L105	$\alpha$ -D-Manp-(1→2)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→3)-[ $\alpha$ -D-Manp-(1→2)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)]- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	1357.49
L106	$\alpha$ -D-Manp-(1→5)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	931.35
L107	$\alpha$ -D-Manp-(1→2)- $\alpha$ -D-Manp-(1→5)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	1093.40
L108	$\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→3)-[ $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)]- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	1033.38
L109	$\alpha$ -D-Manp-(1→2)- $\alpha$ -D-Manp-(1→2)- $\alpha$ -D-Manp-(1→5)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	1255.46
L110	$\alpha$ -D-Manp-(1→2)- $\alpha$ -D-Manp-(1→2)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→3)-[ $\alpha$ -D-Manp-(1→2)- $\alpha$ -D-Manp-(1→2)- $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)]- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> NHCOCF <sub>3</sub>	1681.59
L111	$\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-(1→5)-[ $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→3)]- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> N <sub>3</sub>	1623.60
L112	$\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)-[ $\beta$ -D-Araf-(1→2)- $\alpha$ -D-Araf-(1→5)- $\alpha$ -D-Araf-O(CH <sub>2</sub> ) <sub>8</sub> N <sub>3</sub>	1623.60
L113	$\alpha$ -D-Tal(1→2)-[ $\alpha$ -D-Abe-(1→3)]- $\alpha$ -D-Man-OCH <sub>3</sub>	486.19
L114	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>8</sub> COOCH <sub>3</sub>	715.33
L115	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>8</sub> CONHNH <sub>2</sub>	674.31
L116	$\beta$ -D-GlcNHHis-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> N <sub>3</sub>	765.34
L117	$\beta$ -D-GlcNArg-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> N <sub>3</sub>	784.38
L118	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)-D-Gal	836.29
L119	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc	924.31
L120	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc	1215.40
L121	$\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1127.39
L122	$\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1418.48
L123	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	998.34
L124	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1289.44
L125	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1580.53
L126	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	795.26

Carbohydrate	Structure	MW (Da)
L127	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→3)- $\alpha$ -Gal-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	1160.40
L128	$\beta$ -D-GlcA(2S)-(1→3)- $\beta$ -D-GalNAc(4S,6S)	707.92
L129	$\beta$ -D-GlcA(2S)-(1→3)-D-GalNAc(6S)	605.98
L130	$\beta$ -D-GlcA-(1→3)-D-GalNAc(4S,6S)	605.98
L131	$\beta$ -D-GlcA-(1→3)-D-GalNAc(6S)	504.04
L132	$\beta$ -D-GlcA-(1→3)- $\beta$ -D-GalNAc(4S)-(1→3)- $\beta$ -D-GlcA-(1→3)- $\beta$ -D-GalNAc(4S)-OCH <sub>2</sub> CH=CH <sub>2</sub>	1064.11
L133	$\beta$ -D-Gal-(1→3)-D-GlcNAc	383.14
L134	$\beta$ -D-Gal-(1→4)-D-GlcNAc	383.14
L135	$\beta$ -D-Gal-(1→6)-D-GlcNAc	383.14
L136	$\beta$ -D-Gal-(1→3)-D-GalNAc	383.14
L137	$\alpha$ -D-Gal-(1→3)-D-Gal	342.12
L138	$\beta$ -D-Gal-(1→4)-D-Gal	342.12
L139	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)-D-GlcNAc	545.20
L140	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)- $\alpha$ -D-Gal-(1→3)-D-Gal	666.22
L141	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)-[ $\alpha$ -D-Neu5Ac-(2→6)]- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	1289.44
L142	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	998.34
L143	$\alpha$ -D-Neu5Ac-(2→6)-[ $\beta$ -D-Gal-(1→3)]- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc	998.34
L144	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→3)-D-Gal	836.29
L145	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1289.44
L146	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	1580.53

a. *f* : furanose ring; *p* : pyranose ring. Oligosaccharide residues in pyranose form unless otherwise indicated.

### 3.2.3 Mass spectrometry

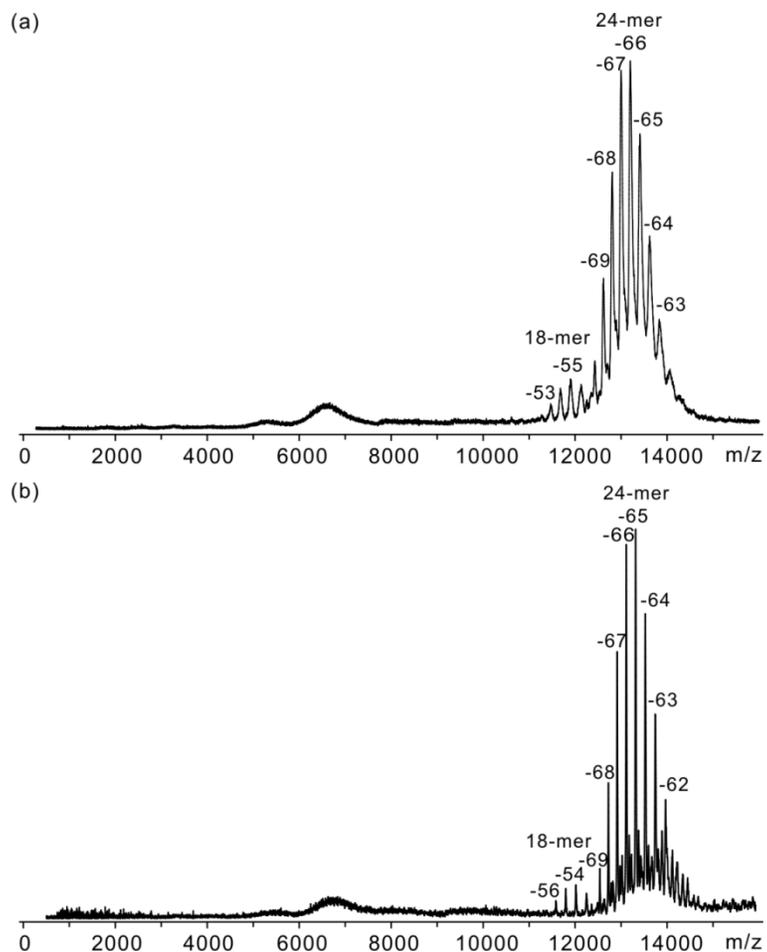
All CaR-ESI-MS measurements were carried out in negative ion mode using a Synapt G2S quadrupole-ion mobility separation-time of flight (Q-IMS-TOF) mass spectrometer (Waters,

Manchester, UK) equipped with a nanoflow ESI (nanoESI) source. NanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.) pulled to  $\sim 5 \mu\text{m}$  using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). A platinum wire was inserted into the nanoESI tip and a capillary voltage of 0.80–1.00 kV was applied. The source parameters were: source temperature 60 °C, cone voltage 50 V, Trap voltage 5 V, and Transfer voltage 2 V. To identify carbohydrate ligands for the P particle, ions corresponding to ligand-bound P particle were isolated using the quadrupole mass filter. The quadrupole was set to transmit a broad  $m/z$  window (approximately 200  $m/z$  units), which allows for the simultaneous passage of free and ligand-bound P particle complexes at a given charge state. Protein-ligand complexes were subjected to CID in the Trap region of the Synapt G2S by increasing the trap voltage from 5 V to 200 V. Argon ( $1.42 \times 10^{-2}$  mbar) was used for CID in the Trap region. In most instances, the deprotonated ligands released from the complexes could be identified from their MWs. Where required, IMS was used to separate the released isomeric ligands. For IMS separation a wave height of 35 V was used while ramping the wave velocity from 2000 to 500  $\text{m s}^{-1}$ . In all cases a helium flow rate of 150  $\text{mL min}^{-1}$  and a nitrogen flow rate of 40  $\text{mL min}^{-1}$  were used. The arrival time distributions (ATDs) for the released ligands were compared to reference ATDs, which were measured for the deprotonated carbohydrates produced directly from solution. Data acquisition and processing were performed using Waters MassLynx software (version 4.1).

### 3.3 Results and Discussion

#### 3.3.1 Validating the CaR-ESI-MS assay

Shown in Figure 3.1a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM, pH 7) solution of P particle (5  $\mu$ M) at 25°C. Under these solution conditions, the recombinant NoV VA387 P particle exists predominantly as a 24-mer, with a charge state distribution from -61 to -70. The 18-mer was also detected, but with much lower abundance, with a charge state distribution ranging from -53 to -56, consistent with observation of a recent ESI-MS study of the P particle.<sup>47</sup> The identity of the broad, unresolved peak centred at  $m/z \sim 7000$  is not known; a similar spectral feature was also reported in the previous ESI-MS study.<sup>47</sup> The peaks corresponding to the different charge states of the P particle are broadened because of the adduct formation during the ESI process. Therefore, in order to more precisely establish the MW of the 24- and 18-mers, modest collisional heating in the Trap region was applied to strip away labile adducts. Shown in Figure 3.1b is a representative ESI mass spectrum obtained using a Trap voltage of 80 V. From the mass spectrum, MWs of  $864\,700 \pm 80$  Da and  $648\,300 \pm 110$  Da are found for the 24-mer and 18-mer, respectively. These values are in reasonable agreement with the theoretical values of 865 036 Da (24-mer) and 648 782 Da (18-mer).<sup>27</sup>



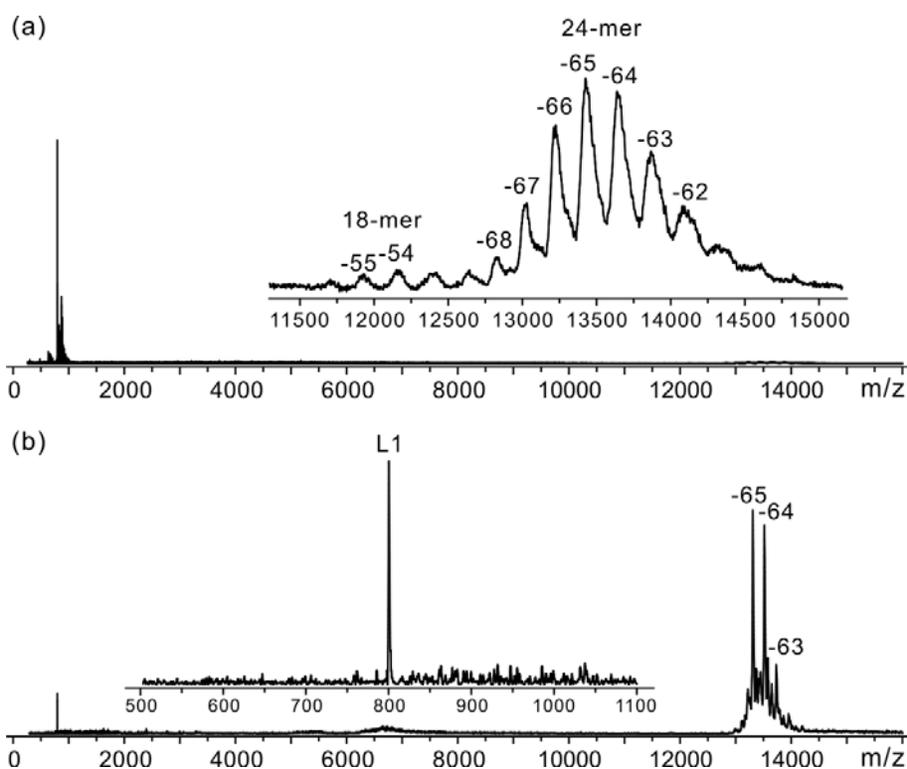
**Figure 3.1.** ESI mass spectra acquired in negative ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) of 5  $\mu$ M P particle using a Trap voltage of (a) 5 V and (b) 80 V (a mild CID condition that facilitates removing solvent adducts but without cause fragmentation of the P particle).

As a starting point for establishing the reliability of the CaR-ESI-MS assay for detecting specific interactions between the P particle and carbohydrate ligands, the assay was applied to a solution of P particle and the B type 3 tetrasaccharide (**L1**), which has been shown to bind to norovirus VA387 P dimer with an intrinsic (per binding site) association constant ( $K_{a,int}$ ) of  $1.5 \times$

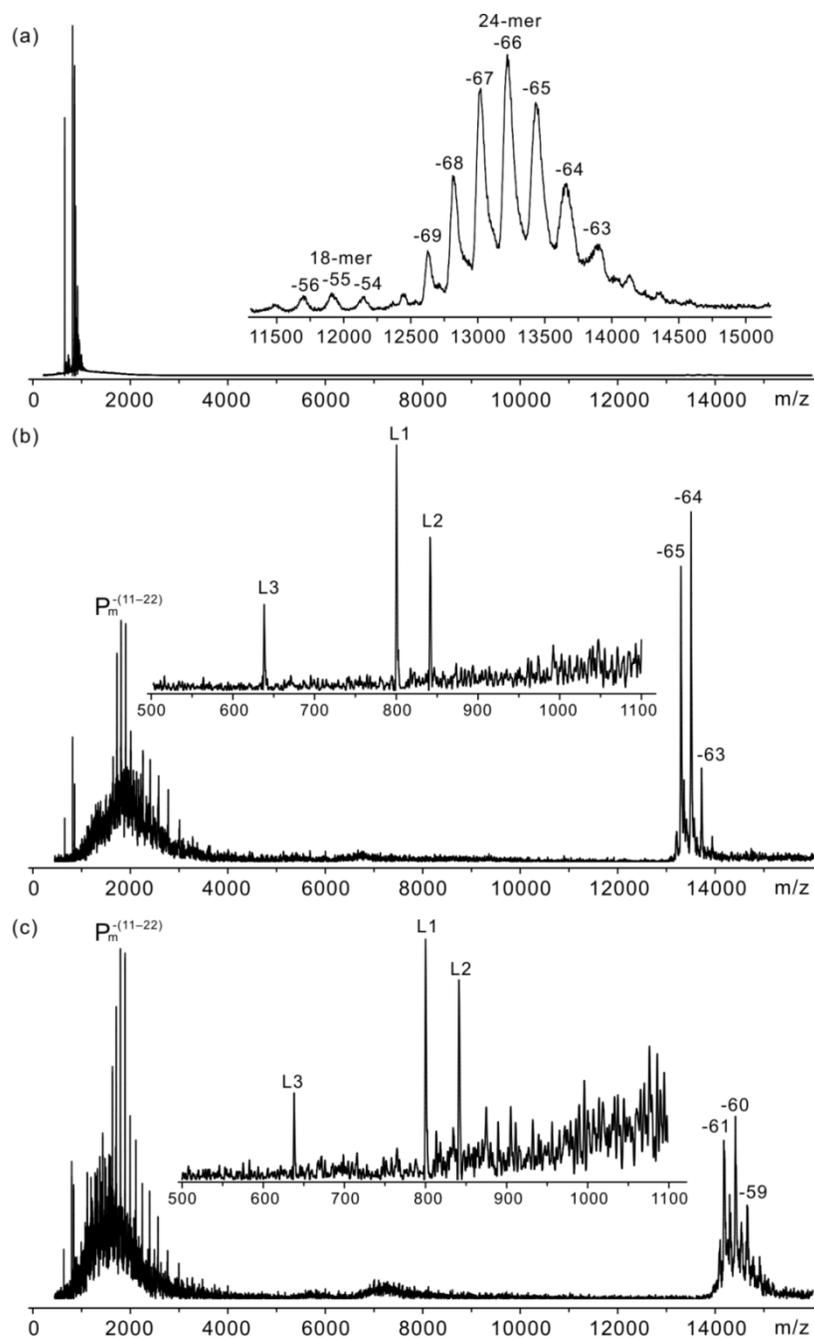
$10^3 \text{ M}^{-1}$ .<sup>12</sup> Shown in Figure 3.2a is a representative ESI mass spectrum measured in negative ion mode for an aqueous ammonium acetate solution of P particle ( $5 \mu\text{M}$ ) and **L1** ( $10 \mu\text{M}$ ). Based on the reported  $K_{a,\text{int}}$  value and assuming all 24 HBGA binding sites of the P particle are equivalent and independent, the P particle is expected to bind a maximum of two molecules of **L1** under these solution conditions. However, it was not possible to resolve the signal corresponding to the individual complexes, i.e., the (P particle +  $q\text{L1}$ ) complexes where  $q = 0 - 2$ , at any of the observed charge states. To establish that **L1** was bound to the P particle in solution, the quadrupole mass filter was set to pass a broad  $m/z$  window centred at  $m/z$  13 420 (which corresponds to the -65 charge state) and the transmitted ions were subjected to CID in order to release bound ligands. The resulting CID mass spectrum reveals strong signal at  $m/z$  800.4, which corresponds to deprotonated **L1** (Figure 3.2b). This result demonstrates that the CaR-ESI-MS assay, as implemented here, can detect specific HBGA ligands for the P particle.

Measurements were also carried out on solutions of P particle with three different HBGA ligands, **L1**, **L2** and **L3** (Figure 3.3a). The corresponding  $K_{a,\text{int}}$  values for **L2** and **L3**, determined from binding measurements performed on the P dimer, are  $8.0 \times 10^2 \text{ M}^{-1}$  and  $6.5 \times 10^2 \text{ M}^{-1}$ , respectively.<sup>12</sup> Collision-induced dissociation of the -65 charge state, at a Trap voltage of 200 V (which was used in order to maximize the release of ligands) produced abundant signal for the deprotonated ions of **L1**, **L2** and **L3** (Figure 3.3b). Under these conditions, the P particle also releases protein monomers. Similar results were obtained when CID was performed on other charge states (Figure 3.3c). These results demonstrate that multiple carbohydrate ligands can be identified, simultaneously, using the CaR-ESI-MS assay. Moreover, the relative abundances of

the three ligands are qualitatively consistent with the trends in  $K_{a,int}$  values determined for the P dimer.<sup>12</sup> This finding suggests that the release efficiencies for the three ligands are similar, despite the differences in their size and structure, and raises the possibility of using the CaR-ESI-MS assay, not only to identify carbohydrate ligands from mixtures, but also to establish their relative affinities.

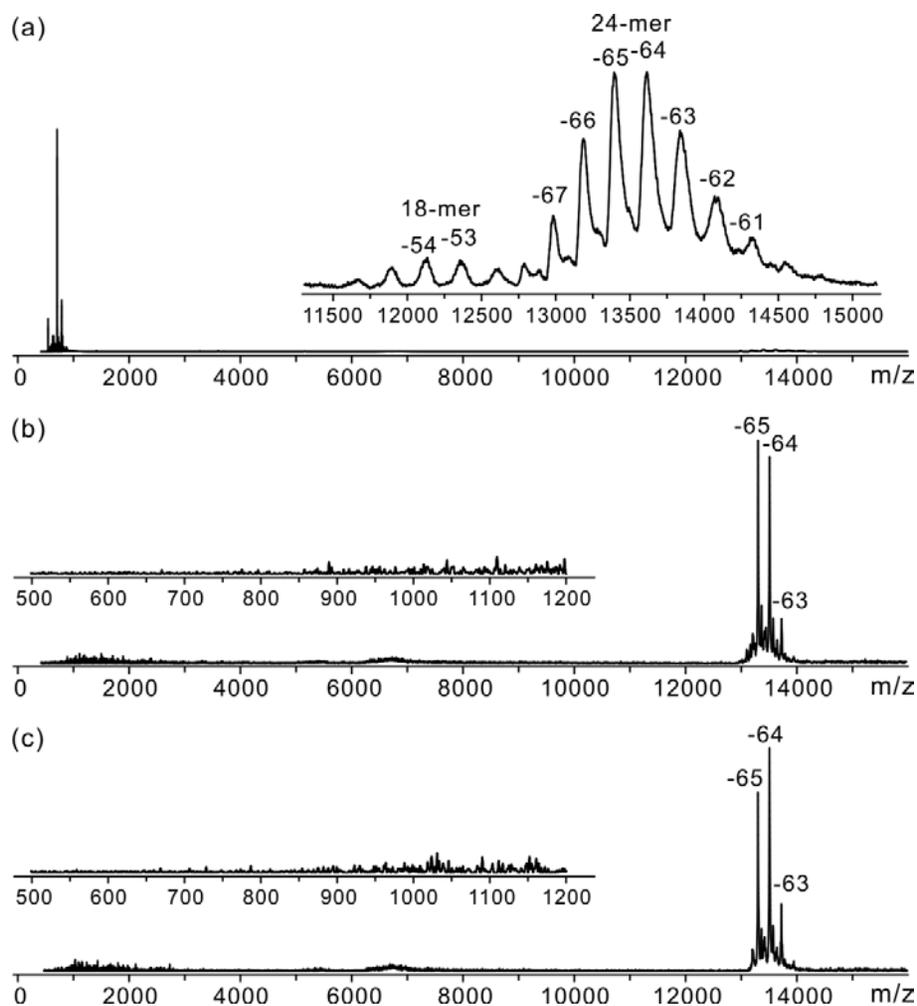


**Figure 3.2.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, at pH 7 and 25 °C) of P particle (5  $\mu$ M) and L1 (B type 3 tetrasaccharide, 10  $\mu$ M), (b) CID mass spectrum measured for the same solution as in part (a) using a broad ( $m/z$  200 units) quadrupole isolation window centered at -65 charge state ( $m/z$  13 420) and a Trap voltage of 150 V.



**Figure 3.3.** (a) ESI mass spectrum obtained in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5  $\mu$ M) and L1, L2 and L3 (10  $\mu$ M each), at pH 7 and 25  $^{\circ}$ C. (b) and (c) CID mass spectra measured with a Trap voltage of 200 V for the P particle and its carbohydrate complexes using a broad ( $m/z$  200 units) quadrupole isolation window centered at the charge state of (b) -65 ( $m/z$  13 420) and (c) -61 ( $m/z$  14 200).

Given the propensity for carbohydrates to bind nonspecifically to proteins during the ESI process, leading to false positives,<sup>48,49</sup> it was important to establish that ligands identified by the CaR-ESI-MS assay originate from specific binding in solution. To test for the occurrence of nonspecific binding, the CaR-ESI-MS assay was applied to solutions of P particle and the type 1 tetrasaccharide **L20**, which lacks the minimal recognition moiety ( $\alpha$ -L-fucose residue) and is not expected to specifically bind to the VA387 P particle in the solution,<sup>6,12</sup> at varying concentrations (10 – 35  $\mu$ M). Shown in Figure 3.4a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM) solution of P particle (5  $\mu$ M) and **L20** (10  $\mu$ M). The CID measurements, which were carried out in a manner analogous to that described above, failed to produce any signal corresponding to the deprotonated **L20** ion (Figure 3.4b). Similar results were also obtained at the higher concentrations investigated (Figure 3.4c). These results suggest that the CaR-ESI-MS assay is not prone to false positives, at least over the range of carbohydrate concentrations used here.



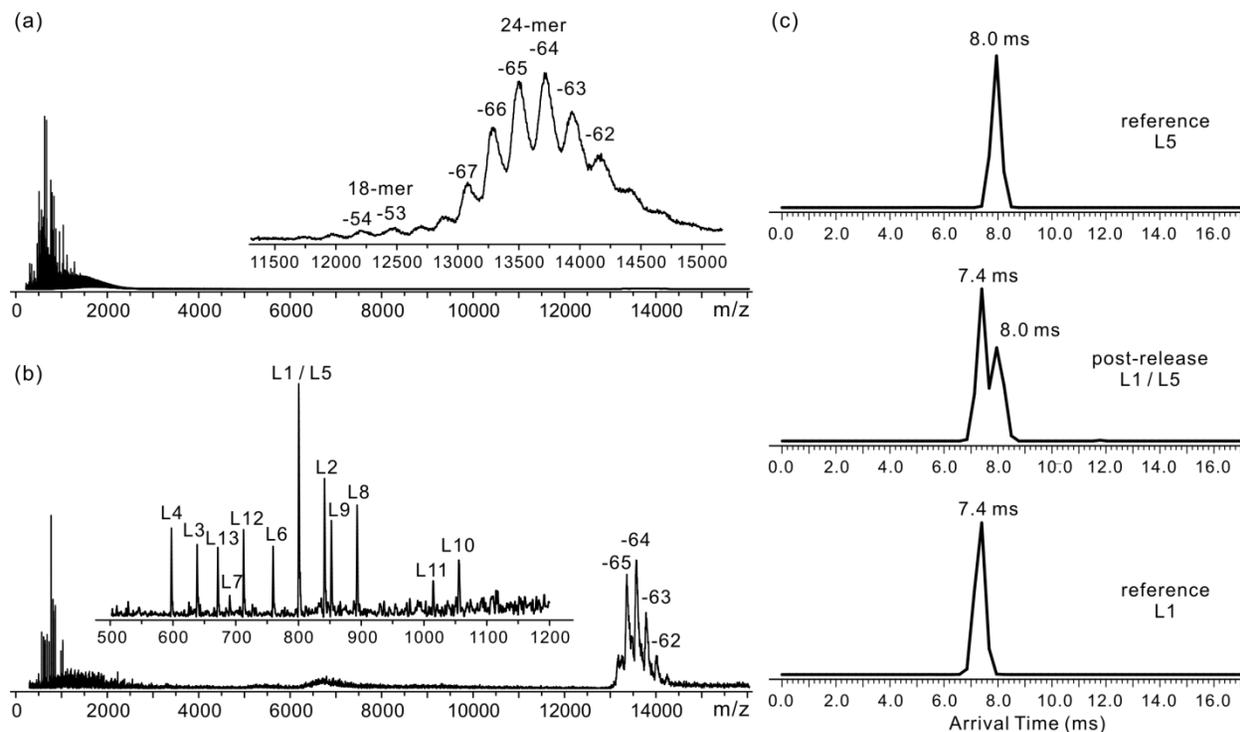
**Figure 3.4.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of P particle (5  $\mu$ M) and **L20** (10  $\mu$ M). (b) CID mass spectrum acquired for solution in part (a) using a broad ( $m/z$  200 units) quadrupole isolation window centered at  $m/z$  13 420 (-65 charge state). (c) CID mass spectrum acquired for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of P particle (5  $\mu$ M) and **L20** (10  $\mu$ M) using a broad ( $m/z$  200 units) quadrupole isolation window centered at  $m/z$  13 420. For (b) and (c), a Trap voltage of 150 V was used.

### 3.3.2 Screening carbohydrate libraries against the NoV P particle

The CaR ESI-MS was used to screen two carbohydrate libraries against the NoV P particle. One of the libraries (*Library1*) was composed of 50 carbohydrates and included 18 HBGA oligosaccharides (**L1 – L18**) known to bind to the NoV VA387 P dimer, with intrinsic affinities ranging from 200 to 1500 M<sup>-1</sup>.<sup>12</sup> The rest of the library was made up from human and plant oligosaccharides (**L19 – L50**) that were not expected to bind to the NoV P particle. The second library (*Library2*) was composed of 146 compounds (**L1 – L146**) and included 24 HBGA oligosaccharides (**L1 – L18, L51 – L53, L69 – L71**), other human, as well as plant and bacterial oligosaccharides.

#### 3.3.2.1 *Library1*

This library served as an additional control to validate the CaR ESI-MS assay for screening carbohydrates against the NoV P particle. Shown in Figure 3.5a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM) solution of P particle (5 μM) and *Library1* (each 10 μM). The -65 charge state was selected for CID (at 150 V) and the resulting mass spectrum is shown in Figure 3.5b. Inspection of the CID mass spectrum reveals signal corresponding to 12 different HBGA ligand MWs. Eleven of the HBGA ligands released from the P particle have unique MWs and can be identified simply from the measured m/z of their deprotonated ions ((**L2** (m/z 841.4), **L3** (m/z 638.3), **L4** (m/z 597.3), **L6** (m/z 759.3), **L7** (m/z 690.3), **L8** (m/z 893.4), **L9** (m/z 852.3), **L10** (m/z 1055.4), **L11** (m/z 1014.5), **L12** (m/z 712.4) and **L13** (m/z 671.3)).

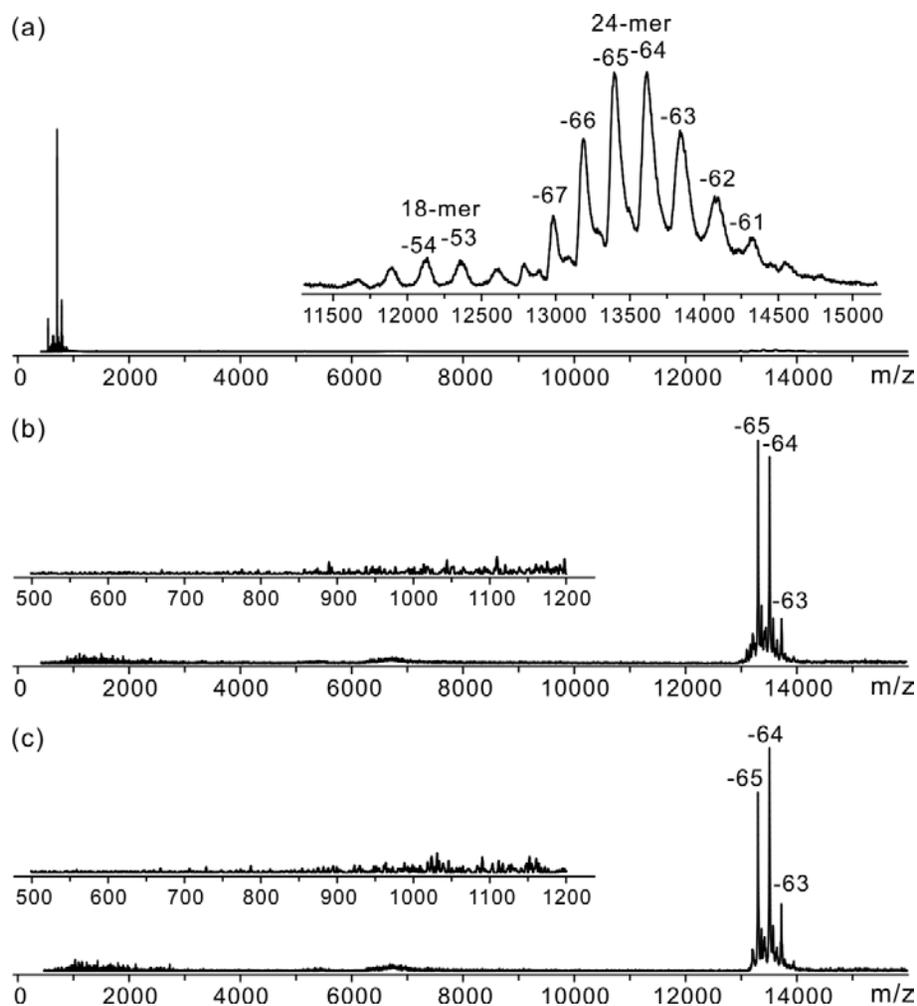


**Figure 3.5.** (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5  $\mu$ M) and *Library1* (10  $\mu$ M each), at pH 7 and 25  $^{\circ}$ C. (b) CID mass spectrum of the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 150 V. (c) Arrival time distributions measured for the deprotonated **L1** and **L5** ions ( $m/z$  800.4) following release from the P particle (*post-release*) and the deprotonated **L1** and **L5** ions obtained directly from solution (*reference*).

Overall, the relative abundances of the released ligands agree qualitatively with their relative affinities.<sup>12</sup> The one notable exception is **L4**, which is as abundant as some of the higher affinity ligands, such as **L9** and **L12**. Although there is no definitive explanation for the unusually high

abundance of **L4** in the CID spectrum, it is possible that the relative HBGA affinities measured for the P particle do not reflect their affinities for the P dimer. Efforts to quantify HBGA ligand binding to the P particle are now underway in our laboratory. *Library1* contained two isomeric HBGA ligands, **L1** and **L5**, and the signal at  $m/z$  800.4 could correspond to either or both of them. To confirm that both **L1** and **L5** bind to the P particle, the released ions were subjected to IMS (Figure 3.5c). Inspection of the ATD measured for the released ions reveals two features (at 7.4 ms and 8.0 ms), which indicates that there are at least two structures present. Comparison of the ATD measured for the released ions with those of the individual **L1** and **L5** ions (Figure 3.5c) confirms that both ligands were released from the P particle. Furthermore, the relative areas of the ATDs for the two ligands are consistent with **L1** having a higher affinity than **L5** for the P particle.<sup>12</sup>

Analysis of the CaR-ESI-MS data obtained for *Library1* reveals that 13 of the 18 HBGA ligands could be identified in a single measurement. Furthermore, all ligands with  $K_{a,int} > 500 \text{ M}^{-1}$  (as determined for the P dimer) were successfully detected. Interestingly, the five HBGA ligands that were not detected (**L14** – **L18**) correspond to Lewis oligosaccharides that have very low affinity ( $K_{a,int} < 300 \text{ M}^{-1}$  for the P dimer). Measurements carried out on solutions of the P particle and **L14** – **L18** at higher concentrations (each 20  $\mu\text{M}$ ) failed to identify binding of any of these ligands (Figures 3.6a and 3.6b). These results suggest that the Lewis antigens have affinities that are lower than expected based on the results obtained for the P dimer.

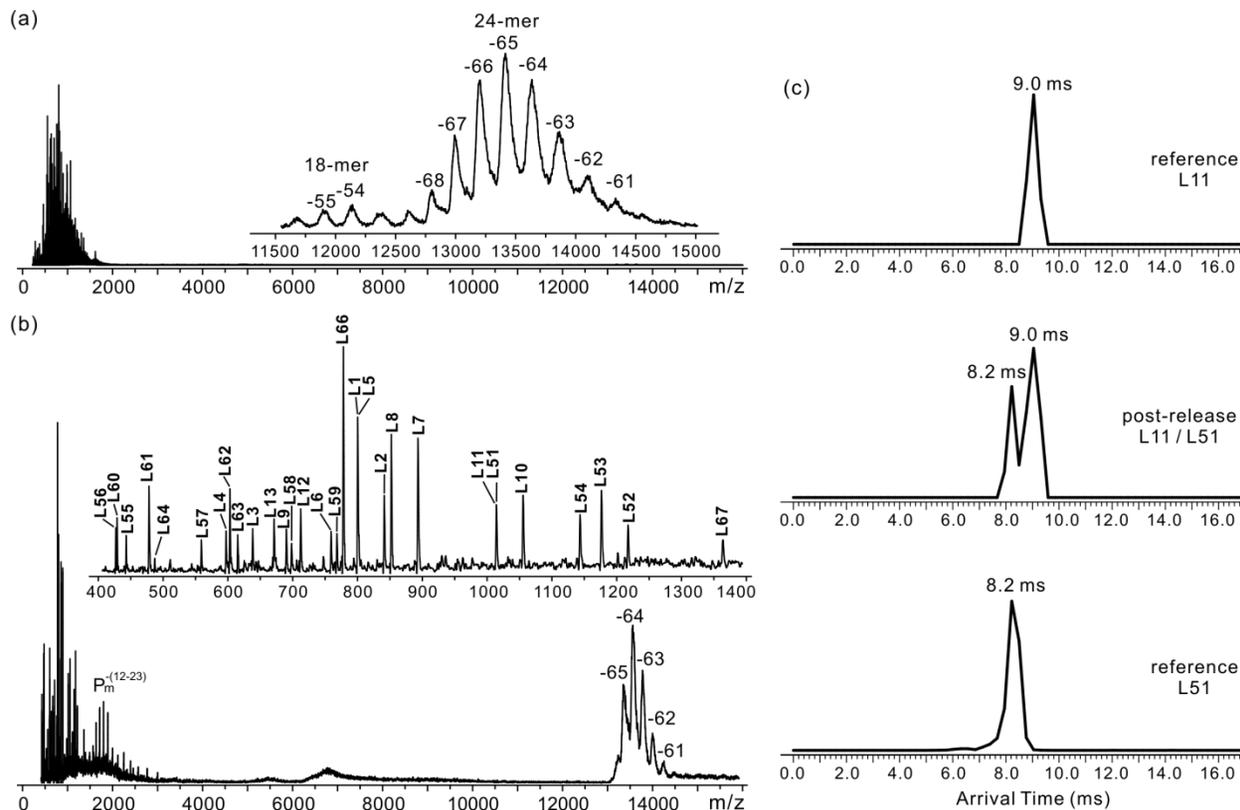


**Figure 3.6.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of P particle (5  $\mu$ M) and **L14**, **L15**, **L16**, **L17** and **L18** (20  $\mu$ M each). (b) CID mass spectrum acquired for solution in part (a) using a broad ( $m/z$  200 units) quadrupole isolation window centered at  $m/z$  13 420 (-65 charge state) and a Trap voltage of 150 V.

### 3.3.2.2 Library2

With the goal of identifying new carbohydrate ligands and potential inhibitors for NoVs, the

CaR-ESI-MS was used to screen a library containing 146 carbohydrates against the VA387 P particle. Shown in Figure 3.7a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM) solution of P particle (5  $\mu$ M) and *Library2* (10  $\mu$ M each component). The -65 charge state was selected for CID (at 180 V) and the resulting mass spectrum is shown in Figure 3.7b. A total of 28 “hits” were identified by CaR-ESI-MS, including the 13 ABH oligosaccharides that were detected from *Library1*. Signals for three other HBGA oligosaccharides not previously tested against the corresponding P dimer, Globo A heptasaccharide (**L52**, m/z 1217.4), Globo B heptasaccharide (**L53**, m/z 1176.4) and Globo H hexasaccharide (**L51**, m/z 1014.5), were detected (Figure 3.7b). The deprotonated **L51** ion is isobaric with deprotonated **L11** ion; however, these two ions can be distinguished by IMS (Figure 3.7c). Moreover, comparison of the relative area of the ATDs for the two compounds indicates that **L11** has higher affinity to P particle than **L51**. The fucosyl GM1 hexasaccharide (**L54**, m/z 1143.5), which has a 1,2-linked  $\alpha$ -L-fucose residue, was also detected and its abundance is similar to that of the H type oligosaccharides (e.g. **L3**, **L4** and **L7**). Of the 17 human milk oligosaccharides (HMOs) present in *Library2* (**L4**, **L18** – **L23**, **L28**, **L64** – **L67**, **L86** – **L87**, **L141** – **L143**), four were found to bind to the P particle, **L4** (m/z 597.3), **L64** (m/z 487.2), **L66** (m/z 778.2) and **L67** (m/z 1363.9). Each of these possesses an  $\alpha$ -L-fucose residue, which has been shown to be an important recognition element for NoV VA387. Of the four HMOs, **L66** ( $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 3)]-D-Glc) exhibited the highest affinity, comparable to the strongest HBGA binders in the library, for the P particle based on the intensity of the released HMOs.



**Figure 3.7.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5  $\mu\text{M}$ ) and *Library2* (10  $\mu\text{M}$  each), at pH 7 and 25  $^{\circ}\text{C}$ . (b) Corresponding CID mass spectrum of the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 180 V. (c) Arrival time distributions measured for the deprotonated **L11** and **L51** ions ( $m/z$  1014.5) following release from the P particle (*post-release*) and the deprotonated **L11** and **L51** ions obtained directly from solution (*reference*).

Surprisingly, a number of bacterial oligosaccharides that are based on structures found in the cell wall of mycobacteria and contain  $\alpha$ -L-rhamnose (**L60** (m/z 429.1), **L61** (m/z 478.2), **L62** (m/z 603.2) and **L63** (m/z 617.2)) or  $\alpha$ -D-arabinofuranose residues (**L56** (m/z 427.0), **L57** (m/z 559.1), **L58** (m/z 698.2) and **L59** (m/z 768.2)) were detected; their abundances are comparable to those of some of the HBGA oligosaccharides tested. Notably, **L56** – **L61** do not possess  $\alpha$ -L-fucose. This finding is consistent with earlier reports of HBGA binding NoVs that also interact with compounds that do not possess  $\alpha$ -L-fucose residues.<sup>34,41,42,50</sup> It has been suggested that these ligands may either mimic the fucose moiety and interact with the NoV in the same binding site, or bind at a different site.<sup>42,50</sup> Both **L60** and **L61** possess an  $\alpha$ -L-rhamnose residue, which may mimic the structure of  $\alpha$ -L-fucose and interact with P particle through the same binding site. To our knowledge, interactions between NoVs and arabinofuranose containing glycans have not been previously reported and the nature of these interactions is not known. Efforts to localize the binding sites of these bacterial oligosaccharides using hydrogen-deuterium exchange MS are currently underway in our laboratory. Curiously, 4',6'-O-benzylidene maltose (**L55**) was also detected. Given that none of the unmodified maltooligosaccharides present in the library (**L41** – **L46**) were found to interact with the P particle, it seems likely that the benzylidene group is responsible for binding.

### 3.4 Conclusions

In summary, a CaR-ESI-MS assay was used to screen carbohydrate libraries against the P particle of NoV VA387 to identify new carbohydrate ligands. To our knowledge this is the first

reported example of the application of a CaR-ESI-MS assay to a large protein assembly. The results of control experiments demonstrated the reliability of the assay for rapidly (1 – 2 min) identifying multiple HBGA ligands present in mixtures carbohydrates. Isomeric ligands could be distinguished by performing IMS on the released ligands (in their deprotonated form). Moreover, the relative abundances of the released ligands provide a qualitative measurement of their relative affinities for the P particle. Application of the CaR-ESI-MS assay to a library of 146 carbohydrates identified all 16 ABH type ligands present. Furthermore, screening revealed interactions with a series of oligosaccharides with structures found in the cell wall of mycobacteria and human milk. The affinities of these newly discovered ligands are comparable to those of the HBGA receptors, as estimated from the relative abundance of released ligand ions.

### 3.5 References

- (1) Patel, M. M.; Widdowson, M.-A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (2) Hutson, A. M.; Atmar, R. L.; Marcus, D. M.; Estes, M. K. *J. Virol.* **2003**, *77*, 405.
- (3) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendu, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.
- (4) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang, X. *J. Virol.* **2005**, *79*, 6714.

- (5) Tan, M.; Jiang, X. *Trends Microbiol.* **2005**, *13*, 285.
- (6) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (7) Tan, M.; Jiang, X. *Expert. Rev. Mol. Med.* **2007**, *9*, 1.
- (8) Tan, M.; Jiang, X. *PLoS pathogens* **2010**, *6*, e1000983.
- (9) Tan, M.; Jiang, X. *Trends Microbiol.* **2011**, *19*, 382.
- (10) Oriol, R. *J. Immunogenet.* **1990**, *17*, 235.
- (11) Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1.
- (12) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. *Glycobiology* **2013**, *23*, 276.
- (13) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (14) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (15) Tan, M.; Huang, P.; Meller, J.; Zhong, W.; Farkas, T.; Jiang, X. *J. Virol.* **2003**, *77*, 12562.
- (16) Jiang, X.; Wang, M.; Graham, D. Y.; Estes, M. K. *J. Virol.* **1992**, *66*, 6527.
- (17) Shoemaker, G. K.; van Duijn, E.; Crawford, S. E.; Utrecht, C.; Baclayon, M.; Roos, W. H.; Wuite, G. J. L.; Estes, M. K.; Prasad, B. V. V.; Heck, A. J. R. *Mol. Cell. Proteomics* **2010**, *9*, 1742.
- (18) Tan, M.; Meller, J.; Jiang, X. *J. Virol.* **2006**, *80*, 7322.
- (19) Tan, M.; Xia, M.; Cao, S.; Huang, P.; Farkas, T.; Meller, J.; Hegde, R. S.; Li, X.; Rao,

- Z.; Jiang, X. *Virology* **2008**, 379, 324.
- (20) Bu, W.; Mamedova, A.; Tan, M.; Xia, M.; Jiang, X.; Hegde, R. S. *J. Virol.* **2008**, 82, 5340.
- (21) Cao, S.; Lou, Z.; Tan, M.; Chen, Y.; Liu, Y.; Zhang, Z.; Zhang, X. C.; Jiang, X.; Li, X.; Rao, Z. *J. Virol.* **2007**, 81, 5949.
- (22) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, 105, 9175.
- (23) Chen, Y.; Tan, M.; Xia, M.; Hao, N.; Zhang, X. C.; Huang, P.; Jiang, X.; Li, X.; Rao, Z. *PLoS Pathog.* **2011**, 7, e1002152.
- (24) Hansman, G. S.; Biertumpfel, C.; Georgiev, I.; McLellan, J. S.; Chen, L.; Zhou, T.; Katayama, K.; Kwong, P. D. *J. Virol.* **2011**, 85, 6687.
- (25) Shanker, S.; Choi, J. M.; Sankaran, B.; Atmar, R. L.; Estes, M. K.; Prasad, B. V. V. *J. Virol.* **2011**, 85, 8635.
- (26) Kubota, T.; Kumagai, A.; Ito, H.; Furukawa, S.; Someya, Y.; Takeda, N.; Ishii, K.; Wakita, T.; Narimatsu, H.; Shirato, H. *J. Virol.* **2012**, 86, 11138.
- (27) Tan, M.; Jiang, X. *J. Virol.* **2005**, 79, 14017.
- (28) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, 382, 115.
- (29) Tan, M.; Huang, P.; Xia, M.; Fang, P.; Zhong, W.; McNeal, M.; Wei, C.; Jiang, W.; Jiang, X. *J. Virol.* **2011**, 85, 753.
- (30) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, 410, 345.

- (31) Tan, M.; Zhong, W.; Song, D.; Thornton, S.; Jiang, X. *J. Med. Virol.* **2004**, *74*, 641.
- (32) Tan, M.; Xia, M.; Chen, Y.; Bu, W.; Hegde, R. S.; Meller, J.; Li, X.; Jiang, X. *PLoS One* **2009**, *4*, e5058.
- (33) Shang, J.; Cheng, F.; Dubey, M.; Kaplan, J. M.; Rawal, M.; Jiang, X.; Newburg, D. S.; Sullivan, P. A.; Andrade, R. B.; Ratner, D. M. *Langmuir* **2012**, *28*, 3338.
- (34) Zhang, X.; Dai, Y.; Zhong, W.; Tan, M.; Lv, Z.; Zhou, Y.; Jiang, X. *Bioorg. Med. Chem.* **2012**, *20*, 1616.
- (35) Tan, M.; Jiang, X. *Nanomedicine* **2012**, *7*, 1.
- (36) Xia, M.; Tan, M.; Wei, C.; Zhong, W.; Wang, L.; McNeal, M.; Jiang, X. *Vaccine* **2011**, *29*, 7670.
- (37) Fang, H.; Tan, M.; Xia, M.; Wang, L.; Jiang, X. *PLoS One* **2013**, *8*, e63269.
- (38) Dai, Y.; Wang, Y.; Zhang, X.; Tan, M.; Xia, M.; Wu, X.; Jiang, X.; Nie, J. *J. Virol. Methods* **2012**, *186*, 126.
- (39) Dai, Y.; Zhang, X.; Tan, M.; Huang, P.; Lei, W.; Fang, H.; Zhong, W.; Jiang, X. *Antivir. Res.* **2012**.
- (40) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.
- (41) Feng, X.; Jiang, X. *Antimicrob. Agents Chemother.* **2007**, *51*, 324.
- (42) Rademacher, C.; Guiard, J.; Kitov, P. I.; Fiege, B.; Dalton, K. P.; Parra, F.; Bundle, D. R.; Peters, T. *Chem. Eur. J.* **2011**, *17*, 7442.
- (43) Zhang, X.; Tan, M.; Chhabra, M.; Dai, Y.; Meller, J.; Jiang, X. *Plos One* **2013**, *8*,

e69379.

- (44) Cederkvist, F.; Zamfir, A. D.; Bahrke, S.; Eijsink, V. G. H.; Sorlie, M.; Peter-Katalinic, J.; Peter, M. G. *Angew. Chem. Int. Ed.* **2006**, *45*, 2429.
- (45) Abzalimov, R. R.; Dubin, P. L.; Kaltashov, I. A. *Anal. Chem.* **2007**, *79*, 6055.
- (46) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50.
- (47) Bereszczyk, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R. *J. Struct. Biol.* **2012**, *177*, 273.
- (48) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (49) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.
- (50) Hansman, G. S.; Shahzad-ul-Hussan, S.; McLellan, J. S.; Chuang, G.-Y.; Georgiev, I.; Shimoike, T.; Katayama, K.; Bewley, C. A.; Kwong, P. D. *J. Virol.* **2012**, *86*, 284.

## Chapter 4

### Affinities of Human Histo-Blood Group Antigens for Norovirus Capsid

#### Protein Complexes\*

#### 4.1 Introduction

Human noroviruses (NoVs), which are the predominant cause of acute gastroenteritis outbreaks worldwide,<sup>1</sup> is a genus of non-enveloped, single-stranded +RNA viruses in the *Caliciviridae* family. As there are no *in vitro* cell culture systems or suitable animal models available for human NoVs, the characterization of their structures and receptor interactions has relied on recombinant forms of their major capsid protein (VP1). For example, recombinant VP1, expressed using the baculovirus system, spontaneously assembles into a virus-like particle (VLP) that is devoid of genomic RNA for infection and replication but is structurally and antigenically indistinguishable from the authentic human NoVs.<sup>2</sup> X-ray crystallography performed on the Norwalk VLP revealed that the intact particle is composed of 180 copies of VP1, which form a  $T = 3$  icosahedral virion.<sup>3</sup> The formation of a smaller  $T = 1$  viral capsid, consisting of 60 copies of VP1, has also been reported.<sup>4</sup>

Production of VLPs using the baculovirus cell system is expensive and time consuming. Consequently, alternative protein complexes that can act as surrogates to VLPs are desirable.

---

\* A version of this chapter has been published: Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinaige, B.; Boraston, A. B.; Klassen, J. S., *Glycobiology* **2015**, *25*, 170-180.

VP1 consists of two major domains bound by a flexible peptide linker, the N-terminal shell (S) domain and the protrusion (P) domain at the C-terminus.<sup>3</sup> The interior S domain is critical to maintaining the icosahedral structure of the virion, whereas the P domain forms a dimeric structure that is located on the outer surface and is implicated in the virus-receptor recognition process and, thus, cell entry. Expression of the P domain in *E. coli* has been shown to produce homodimers, called P dimers.<sup>5</sup> The P dimers can also assemble into larger complexes, a 12-mer small P particle<sup>6</sup> and a 24-mer P particle.<sup>7,8</sup> A recent native electrospray ionization mass spectrometry (ESI-MS) study revealed that, in 100 mM ammonium acetate (pH 7.4), the P particle is made up of approximately 85% of 24-mer and 15% of 18-mer.<sup>9</sup> Importantly, both the sub-viral particles and the P dimer are believed to retain the authentic antigenicity and receptor binding capability of the VLP<sup>5,7,10</sup> and are, therefore, seen as attractive substitutes to VLPs for investigating the nature of human NoV-host cell interactions and discovering potential inhibitors. However, to our knowledge, a quantitative comparison of the receptor binding properties of a NoV VLP and its corresponding P particle and P dimer has yet to be carried out.

It is well established that many human NoVs recognize human histo-blood group antigens (HBGAs), which are found on the surfaces of red blood cells and mucosal epithelial cells in the form of glycoproteins and glycolipids,<sup>11,12</sup> as cellular receptors or attachment factors.<sup>13-16</sup> The HBGAs are divided into four types, namely A, B, H and Lewis, based on the carbohydrate structure at the non-reducing end. Additionally, each HBGA is further divided into six subtypes (type 1 to 6) based on the carbohydrates structure at the reducing end. To date, there have been few quantitative binding studies performed on the capsid proteins of human NoVs. Using

saturation-transfer difference nuclear magnetic resonance (STD-NMR) spectroscopy, Peters and coworkers<sup>17</sup> estimated the apparent association constants ( $K_{a,app}$ ) of VLP from human NoV Ast6139 (GII.4 strain) for a variety of HBGA oligosaccharides (including H disaccharide, A and B trisaccharides, H type 1, 2 and 6 trisaccharides, Lewis a, b, x, y and sialyl-Lewis a and x oligosaccharides) to be in the  $\sim 10^4$   $M^{-1}$  range. Based on these results, the intrinsic (per binding site) association constants ( $K_{a,int}$ ) are predicted to be in the  $\sim 10^2$   $M^{-1}$  range. Using surface plasmon resonance spectroscopy, Belliot and colleagues<sup>18</sup> analyzed BSA-conjugated type H, A and B carbohydrates binding to VLPs isolated from six GII.4 strains of human NoV. However, affinities could not be accurately determined due to the uncertainty in the number of active binding sites in each VLP, as well as the number of immobilized (on the chip) carbohydrates that participated in interactions. Quantitative  $K_{a,app}$  values have also been reported for human NoV P dimers. Using STD-NMR spectroscopy, Kwong and co-workers<sup>19</sup> measured  $K_{a,app}$  values of 2200  $M^{-1}$  and 2600  $M^{-1}$  for fucose and H type 2 trisaccharide, respectively, binding to a GII.10 human NoV P dimer. More recently, the  $K_{a,int}$  values of a library of 42 HBGA oligosaccharides for the human NoV VA387 P dimer were measured using the *direct* ESI-MS assay.<sup>20</sup> The results of this study revealed that the P dimer exhibits broad specificity and binds to A, B, H and Lewis type antigens, although with low affinities ( $\leq 3000$   $M^{-1}$ ).

Here, we report  $K_{a,int}$  values of variety of HBGA oligosaccharides for the P particle and VLP of human NoV VA387. Application of the *direct* ESI-MS assay, which was used for the P dimer measurements, to these large capsid protein complexes is not feasible due to the difficulty in resolving the free and ligand-bound forms of the protein complexes. Consequently, an

adaptation of the *proxy protein* ESI-MS method<sup>21</sup> was used to carry out the measurements. The *proxy protein* ESI-MS method combines *direct* ESI-MS binding measurements and competitive protein binding to evaluate protein-ligand affinities. Specifically, a proxy protein ( $P_{\text{proxy}}$ ), which binds to the ligand of interest with known affinity and can be detected directly by ESI-MS, is used to quantitatively monitor the extent of ligand binding to the protein of interest – the P particle or VLP in the present case. Using this method the  $K_{a,\text{int}}$  values of thirteen HBGA oligosaccharides, containing A, B and H epitopes, with variable sizes (disaccharide to tetrasaccharide) and different precursor chain types (type 1, 2, 3, 5 and 6), for the P particle were measured; affinities for the A and B trisaccharides and A and B type 6 tetrasaccharides for the VLP were also quantified. Comparison of these values with binding data recently reported for the P dimer of human NoV VA387 provides a unique opportunity to assess the similarity of the HBGA binding sites in the P dimer, P particle and VLP.

## **4.2 Experimental Section**

### **4.2.1 Proteins**

For VLP production, the gene encoding the capsid protein (VP1) of human NoV VA387 (GII.4, GenBank accession no. AY038600) was cloned and expressed through the Bac-to-Bac baculovirus expression system (Invitrogen Life Technologies, Grand Island, NY) as described elsewhere.<sup>22,23</sup> VLPs, comprised of 180 copies of VP1 (which has a molecular weight (MW) of 58,887 Da), assembled spontaneously. For P particle and P dimer production, the gene fragments encoding the P domain (residues 222-539) of VP1 with and without a C-terminus-fused peptide

CDCRGDCFC, respectively, were cloned and expressed in bacteria through the GST-Gene Fusion System (GE-Healthcare Life Sciences, Piscataway, NJ) following a protocol described previously.<sup>5,7</sup> The resulting P dimer (MW 69,312 Da) or P particle (24-mer, MW 865,036 Da) that assembled spontaneously were purified using glutathione affinity chromatography, followed by gel filtration chromatography.

The N-terminal family 51 carbohydrate-binding module (CBM, expected MW 20,735 Da) from *Streptococcus pneumonia* SP3-BS71 GH98 was recombinantly produced in *Escherichia coli* and purified by Ni<sup>2+</sup> immobilized metal affinity chromatography (GE-Healthcare Life Sciences) using procedures described elsewhere.<sup>24</sup> A recombinant fragment of the C-terminus of human galectin-3 (Gal-3C, MW 16,330 Da) was a gift from Prof. C. Cairo (University of Alberta). A recombinant soluble fragment of  $\alpha$ -(1→3)N-acetyl galactosaminyltransferase (GTA, homodimer, MW 69,040 Da), which contains a full C-terminal and catalytic domain, as well as a truncated N-terminal domain, was expressed in bacterial cells and purified by ion exchange chromatography using a SP-Sepharose FF resin (GE-Healthcare Life Sciences), followed by affinity purification using a UDP-hexanolamine resin.<sup>25</sup> Bovine ubiquitin (Ubq, MW 8,565 Da), which was purchased from Sigma-Aldrich Canada (Oakville, Canada), and recombinant single chain fragment (scFv, MW 26,539 Da) of monoclonal antibody Se155-4, which was produced using procedures described before,<sup>26</sup> served as reference proteins ( $P_{ref}$ ) for the binding measurements. Each protein was dialyzed and concentrated against 50 mM aqueous ammonium acetate (pH 7) using Vivaspin 0.5 mL centrifugal filters (Sartorius Stedim Biotech, Göttingen, Germany) with a 10 kDa MW cutoff. The concentrations were measured by UV spectroscopy.

Protein stock solutions were stored at  $-80\text{ }^{\circ}\text{C}$  until used.

#### **4.2.2 Carbohydrates**

The oligosaccharides (**L1** – **L14**) were a gift from Prof. T. Lowary (University of Alberta).<sup>27-29</sup>

Their structures are shown in Table 4.1. To prepare stock solutions, solid sample of each compound was weighed and dissolved in a known volume of ultrafiltered water (Milli-Q, Millipore, Billerica, MA) to yield a final concentration of 1 mM. These solutions were stored at  $-20\text{ }^{\circ}\text{C}$  until needed.

#### **4.2.3 Mass spectrometry**

The ESI-MS measurements were carried out in positive ion mode using either a Synapt G2S quadrupole-ion mobility separation-time of flight mass spectrometer (Waters, Manchester, UK) or a 9.4T ApexQe Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker-Daltonics, Billerica, MA). The Synapt G2S mass spectrometer, with its high mass capabilities, was used for the ESI-MS analysis of the human NoV P particle and VLP. Both the G2S and the ApexQe mass spectrometers were used to carry out the *direct* and *proxy protein* ESI-MS measurements. Both mass spectrometers were equipped with nanoflow ESI (nanoESI) sources. To carry out ESI, the sample solution was loaded into a nanoESI tip pulled from a borosilicate capillary (1.0 mm o.d., 0.68 mm i.d.) using a micropipette puller (P-1000, Sutter Instruments, Novato, CA). To initiate ESI, a voltage of 1.0 – 1.4 kV was applied to a platinum wire inserted into the nanoESI tip. A detailed description of the instrumental conditions used to

implement the *direct* and *proxy protein* ESI-MS binding measurements can be found elsewhere.<sup>20,21</sup> Unless otherwise indicated, the ESI-MS measurements performed on the ApexQe mass spectrometer were carried out using 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing protein and ligand of interest, while those on the Synapt G2S mass spectrometer were carried out using 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing protein and ligand of interest. To carry out the *proxy protein* ESI-MS measurements, a  $P_{\text{proxy}}$  was also added to the solutions. In all cases, the reference protein method, which involves the addition of a  $P_{\text{ref}}$  to the solution, was used to correct the mass spectra for the occurrence of nonspecific protein-carbohydrate binding during the ESI process. A complete description of the correction method can be found elsewhere.<sup>30,31</sup>

#### **4.2.4 Gel filtration chromatography**

Gel filtration chromatography was performed using a Superdex 200 size exclusion column (HiLoad 16/60, 120 mL bed volume, GE Healthcare Life Sciences) powered by an AKTA fast-performance liquid chromatography (FPLC) system (model 920, GE Healthcare Life Sciences). The column was equilibrated and run in 1X phosphate-buffered saline (PBS, pH 7.4) at a flow rate of 1.0 mL min<sup>-1</sup>; 1.0 mL of the VLP sample (5 mg mL<sup>-1</sup>) was loaded onto the column using a manual injector. The MW of the proteins in each elution volume was calibrated with the Gel Filtration Calibration kit (GE Healthcare Life Sciences) and purified P particle as described elsewhere.<sup>32,33</sup>

## 4.2.5 Determination of $K_a$ values

### 4.2.5.1 Direct ESI-MS assay.

The *direct* ESI-MS assay was used to quantify oligosaccharide affinities for CBM and Gal-3C, two of four  $P_{\text{proxy}}$ 's used in this study, and the affinity of **L14** for the P dimer. For a 1:1 protein-ligand complex the association constant ( $K_a$ ) is calculated from the abundance ( $Ab$ ) ratio ( $R$ , eq 1.8) of the ligand-bound (PL) to free protein (P) ions measured by ESI-MS and the initial concentrations of protein ( $[P]_0$ ) and ligand ( $[L]_0$ ), using eq 1.9.<sup>31</sup> The abundances of free and ligand-bound proteins were calculated as the sum of the peak areas for all of the charge states detected for each species. In cases where ligand binding was weak ( $K_a < 10^4 \text{ M}^{-1}$ ), a titration approach was employed, whereby the protein concentration was kept constant and the ligand concentration was varied.<sup>34</sup> Nonlinear regression analysis of the concentration-dependence of the fraction of ligand-bound protein,  $[R/(R+1)]$  was used to determine  $K_a$ , as given by eq 1.10.<sup>31</sup>

### 4.2.5.2 Proxy protein ESI-MS assay.

The *proxy protein* ESI-MS assay was recently developed to quantify the affinities of protein–ligand complexes that could not be directly measured by ESI-MS.<sup>21</sup> Briefly, the method involves the use of the *direct* ESI-MS assay to monitor the extent of ligand binding to a proxy protein ( $P_{\text{proxy}}$ ) in the presence of the target protein (P).<sup>21</sup> The general expression relating the intrinsic association constant of P ( $K_{a,\text{int,P}}$ ), with  $h$  identical and independent binding sites, to  $R_{\text{proxy}}$ , the abundance ratio of ligand-bound to free  $P_{\text{proxy}}$  ions, for a  $P_{\text{proxy}}$  with a single binding site is given by eq 4.1a:<sup>21</sup>

$$K_{a,int,P} = \frac{K_{a,P_{proxy}} / R_{proxy}}{\frac{h[P]_0}{[L]_0 - \frac{[P_{proxy}]_0 R_{proxy}}{R_{proxy} + 1} - \frac{R_{proxy}}{K_{a,P_{proxy}}}} - 1} \quad (4.1a)$$

As described in more detail below, the human NoV P particle and VLP investigated in the present study exist in multiple forms in solution. The P particle exists predominately as a 24-mer, however, both the 18-mer and 36-mer have also been detected.<sup>9</sup> Similarly, although VLP exists predominantly as a 180-mer, the 60-mer and 80-mer, as well as dimer, are also present.<sup>35</sup> Consequently, it is more appropriate to rewrite eq 4.1a in terms of the total concentration of ligand binding sites ( $[P]_{m,0} = \sum h_i [P_i]_0$ ), which is equal to the number of protomers that make up the assembly, eq 4.1b:

$$K_{a,P,int} = \frac{K_{a,P_{proxy}} / R_{proxy}}{\frac{[P]_{m,0}}{[L]_0 - \frac{[P_{proxy}]_0 R_{proxy}}{R_{proxy} + 1} - \frac{R_{proxy}}{K_{a,P_{proxy}}}} - 1} \quad (4.1b)$$

It must be stressed that the  $K_{a,int,P}$  values determined in this way represent the weighted average of the affinities of the different assemblies present in solution.

#### 4.2.5.3 Application of the proxy protein ESI-MS method using $P_{proxy}$ that possesses multiple ligand binding sites.

The application of the *proxy protein* method was previously demonstrated using a  $P_{proxy}$  with a single binding site. Two of the  $P_{proxy}$  used in the present study possess multiple (two) ligand binding sites. To address this challenge, a special extension of the *proxy protein* ESI-MS method, in which both the target protein (P) and proxy protein ( $P_{proxy}$ ) possesses multiple ligand (L)

binding sites, was implemented. The relevant equations of mass balance for the situation where P and P<sub>proxy</sub> possess *h* and *g* identical L binding sites, respectively, are given by eqs 4.2a – c:

$$[P]_0 = [P] + \sum_{i=1}^h [PL_i] \quad (4.2a)$$

$$[P_{\text{proxy}}]_0 = [P_{\text{proxy}}] + \sum_{j=1}^g [P_{\text{proxy}}L_j] \quad (4.2b)$$

$$[L]_0 = [L] + \sum_{i=1}^h i[PL_i] + \sum_{j=1}^g j[P_{\text{proxy}}L_j] \quad (4.2c)$$

where  $[P]_0$ ,  $[P_{\text{proxy}}]_0$  and  $[L]_0$  are the initial concentrations of P, P<sub>proxy</sub> and L, respectively,  $[P]$ ,  $[P_{\text{proxy}}]$  and  $[L]$  are the equilibrium concentrations of P, P<sub>proxy</sub> and L, respectively, and  $[PL_i]$  and  $[P_{\text{proxy}}L_j]$  are the equilibrium concentrations of P and P<sub>proxy</sub> bound to *i* and *j* molecules of L, respectively.

The values of  $[P_{\text{proxy}}]$  and  $[P_{\text{proxy}}L_j]$  can be calculated from eqs 4.3a and 4.3b, respectively:

$$[P_{\text{proxy}}] = \frac{[P_{\text{proxy}}]_0}{1 + \sum_{j=1}^g [P_{\text{proxy}}L_j]/[P_{\text{proxy}}]} = \frac{[P_{\text{proxy}}]_0}{1 + \sum_{j=1}^g R_{\text{proxy},j}} \quad (4.3a)$$

$$[P_{\text{proxy}}L_j] = [P_{\text{proxy}}]R_{\text{proxy},j} = \frac{[P_{\text{proxy}}]_0 R_{\text{proxy},j}}{1 + \sum_{j=1}^g R_{\text{proxy},j}} \quad (4.3b)$$

where  $R_{\text{proxy},j}$  corresponds to the concentration ratio of ligand-bound (bound to *j* molecules of L) to free P<sub>proxy</sub> and is taken to be equal to the abundance (*Ab*) ratio of ligand-bound (bound to *j* molecules of L) to free P<sub>proxy</sub> gas phase ions, eq 4.4:

$$R_{\text{proxy},j} = \frac{\sum Ab(P_{\text{proxy}}L_j)}{\sum Ab(P_{\text{proxy}})} = \frac{[P_{\text{proxy}}L_j]}{[P_{\text{proxy}}]} \quad (4.4)$$

The ratio of occupied-to-free binding sites for P<sub>proxy</sub> ( $\theta$ ) can be calculated using eq 4.5:

$$\theta = \frac{[\text{occupied binding sites}]_{\text{p}_{\text{proxy}}}}{[\text{free binding sites}]_{\text{p}_{\text{proxy}}}} = \frac{\sum_{j=1}^g j[\text{P}_{\text{proxy}}\text{L}_j]}{g[\text{P}_{\text{proxy}}]_0 - \sum_{j=1}^g j[\text{P}_{\text{proxy}}\text{L}_j]} = \frac{\sum_{j=1}^g jR_{\text{proxy}j}}{g(1 + \sum_{j=1}^g R_{\text{proxy}j}) - \sum_{j=1}^g jR_{\text{proxy}j}} \quad (4.5)$$

and the intrinsic association constant for L binding to P<sub>proxy</sub> (K<sub>a,P<sub>proxy</sub>,int</sub>) can be expressed by eq 4.6:

$$K_{\text{a,P}_{\text{proxy}},\text{int}} = \frac{[\text{occupied binding sites}]_{\text{p}_{\text{proxy}}}}{[\text{free binding sites}]_{\text{p}_{\text{proxy}}}[L]} = \frac{\theta}{[L]} \quad (4.6)$$

It follows that [L] can be found using eq 4.7:

$$[L] = \frac{\theta}{K_{\text{a,P}_{\text{proxy}},\text{int}}} \quad (4.7)$$

The intrinsic association constant of P (K<sub>a,P,int</sub>) can be expressed by eq 4.8:

$$K_{\text{a,P},\text{int}} = \frac{[\text{occupied binding sites}]_{\text{p}}}{[\text{free binding sites}]_{\text{p}}[L]} \quad (4.8)$$

and the concentrations of occupied and free binding sites are given by eqs 4.9a and 4.9b:

$$[\text{occupied binding sites}]_{\text{p}} = \sum_{i=1}^h i[\text{PL}_i] \quad (4.9a)$$

$$[\text{free binding sites}]_{\text{p}} = [\text{P}]_{\text{m},0} - \sum_{i=1}^h i[\text{PL}_i] \quad (4.9b)$$

where the total concentration of binding sites can be written as eq 4.9c:

$$[\text{total binding sites}]_{\text{p}} \equiv [\text{P}]_{\text{m},0} = h[\text{P}]_0 \quad (4.9c)$$

From eqs 4.2c, 4.5 and 4.7, the following expression can be derived, eq 4.10:

$$\sum_{i=1}^h i[\text{PL}_i] = [L]_0 - \frac{\theta}{K_{\text{a,P}_{\text{proxy}},\text{int}}} - \frac{g[\text{P}_{\text{proxy}}]_0 \theta}{\theta + 1} \quad (4.10)$$

Finally, the value of K<sub>a,P,int</sub> can be calculated from eq 4.11:

$$K_{a,P,int} = \frac{K_{a,P_{proxy},int} / \theta}{\frac{[P]_{m,0}}{[L]_0 - \frac{g[P_{proxy}]_0 \theta}{\theta + 1} - \frac{\theta}{K_{a,P_{proxy},int}}} - 1} \quad (4.11)$$

It can be shown that, in the case where  $P_{proxy}$  possesses a single binding site (i.e.,  $g = 1$ ), eq 4.11 reduces to eq 4.1b. Moreover, in the case of a  $P_{proxy}$  with two ligand binding sites (i.e.,  $g = 2$ ),  $\theta$  is given by eq 4.12:

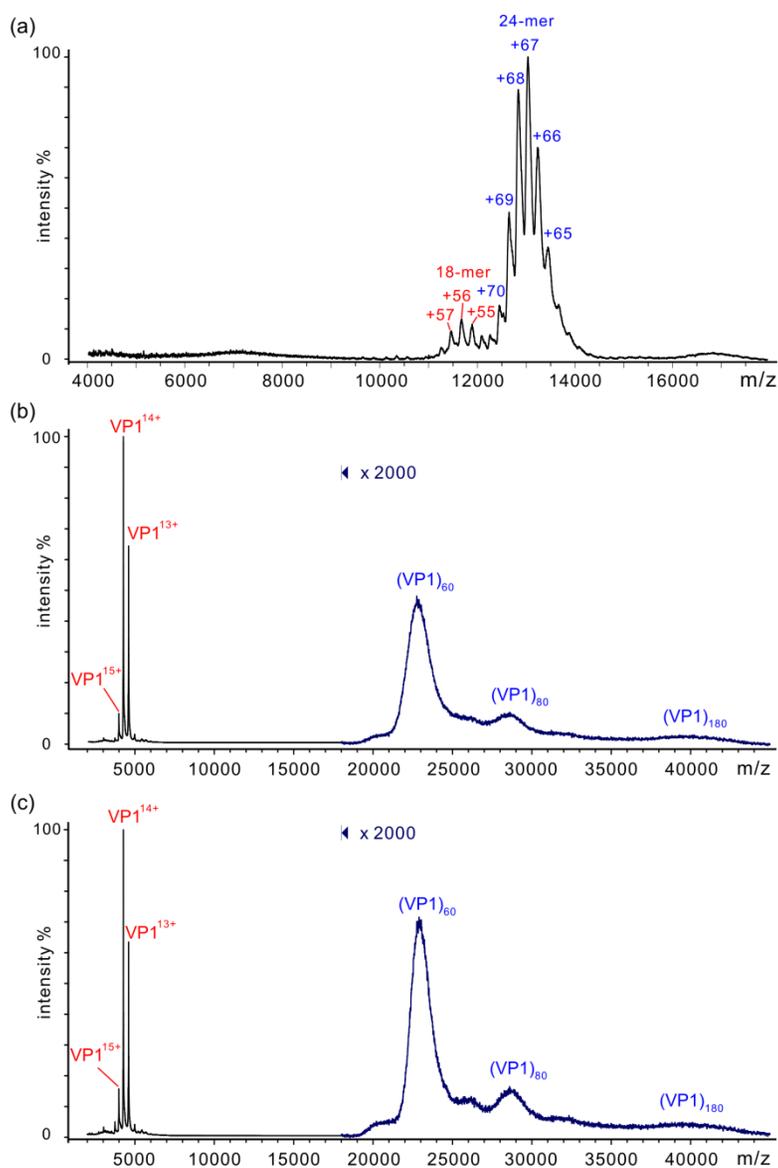
$$\theta = \frac{R_{proxy,1} + 2R_{proxy,2}}{2 + R_{proxy,1}} \quad (4.12)$$

## 4.3 Results and Discussion

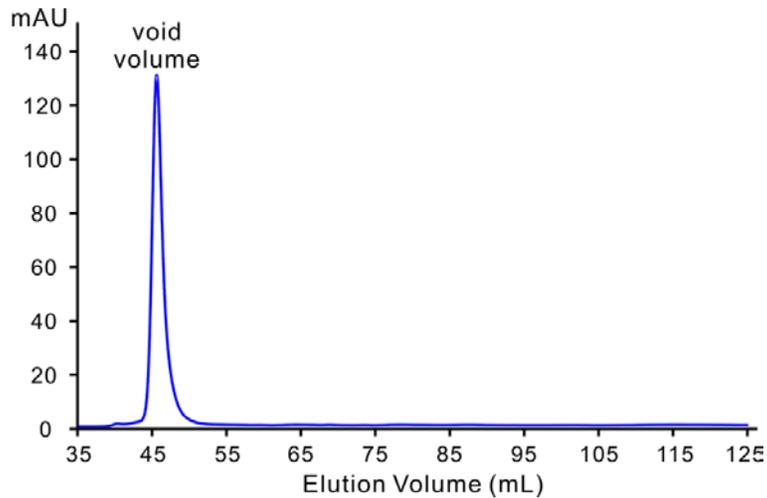
### 4.3.1 ESI-MS analysis of human NoV VA387 P particle and VLP

Representative ESI mass spectra measured for a 200 mM ammonium acetate aqueous solution (pH 7 and 25 °C) containing 3  $\mu$ M of P particle (which corresponds to 72  $\mu$ M of monomer) or 0.2  $\mu$ M VLP (corresponding to 36  $\mu$ M VP1) are shown in Figure 4.1. As seen in Figure 4.1a, the P particle is present predominantly as a 24-mer (with a charge state distribution centred around +67), along with the 18-mer (charge state distribution centred around +56) at lower abundance. These observations are consistent with those reported previously.<sup>9</sup> The MWs of the 24-mer and 18-mer,  $865,000 \pm 540$  Da and  $648,900 \pm 400$  Da, respectively, are in reasonably good agreement with the expected values calculated from the protein sequence, 865,036 Da and 648,782 Da, respectively.<sup>7</sup> As seen in Figure 4.1b, the VP1 monomer (MW 58,887 Da), 60-mer (~3.5 MDa), 80-mer (~4.7 MDa) and 180-mer (~10 MDa) are all present in solution. Notably, the distribution of VP1 species is reproducible over a period of weeks, Figure 4.1c. The smaller oligomers,

which are believed to exist in dynamic equilibrium with the 180-mer, were previously observed by native ESI-MS, ion mobility separation MS, as well as atomic force microscopy.<sup>35,36</sup> Due to the high MWs of the oligomers, it was not possible to resolve individual charge states and, thus, their identification was based on previously reported ESI-MS results.<sup>35</sup> Moreover, since the ESI-MS ionization/detection efficiencies of high MW oligomers are expected to be significantly different than those of VP1 monomer, the relative abundances of the monomer and oligomers measured by ESI-MS likely do not accurately reflect their relative concentrations in solution. In fact, gel-filtration chromatography performed on 5 mg mL<sup>-1</sup> human NoV VA387 VLP samples (in 1x PBS, pH 7.4, 25 °C) produced a single peak (Figure 4.2), which corresponds to the void volume, suggesting that the capsid protein assembles predominantly into large complexes (>800 kDa) in solution.



**Figure 4.1.** (a) ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate (200 mM, pH 7 and 25 °C) solution with (a) 3  $\mu$ M P particle (corresponding to 72  $\mu$ M monomer) and (b) 0.2  $\mu$ M VLP (corresponding to 36  $\mu$ M VP1) of human NoV VA387. (b) and (c) ESI mass spectra acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 0.2  $\mu$ M (corresponding to 36  $\mu$ M VP1) NoV VA387 VLP. The measurements in (b) and (c) were carried out on two different days. The measurements were acquired using a Waters Synapt G2S mass spectrometer.

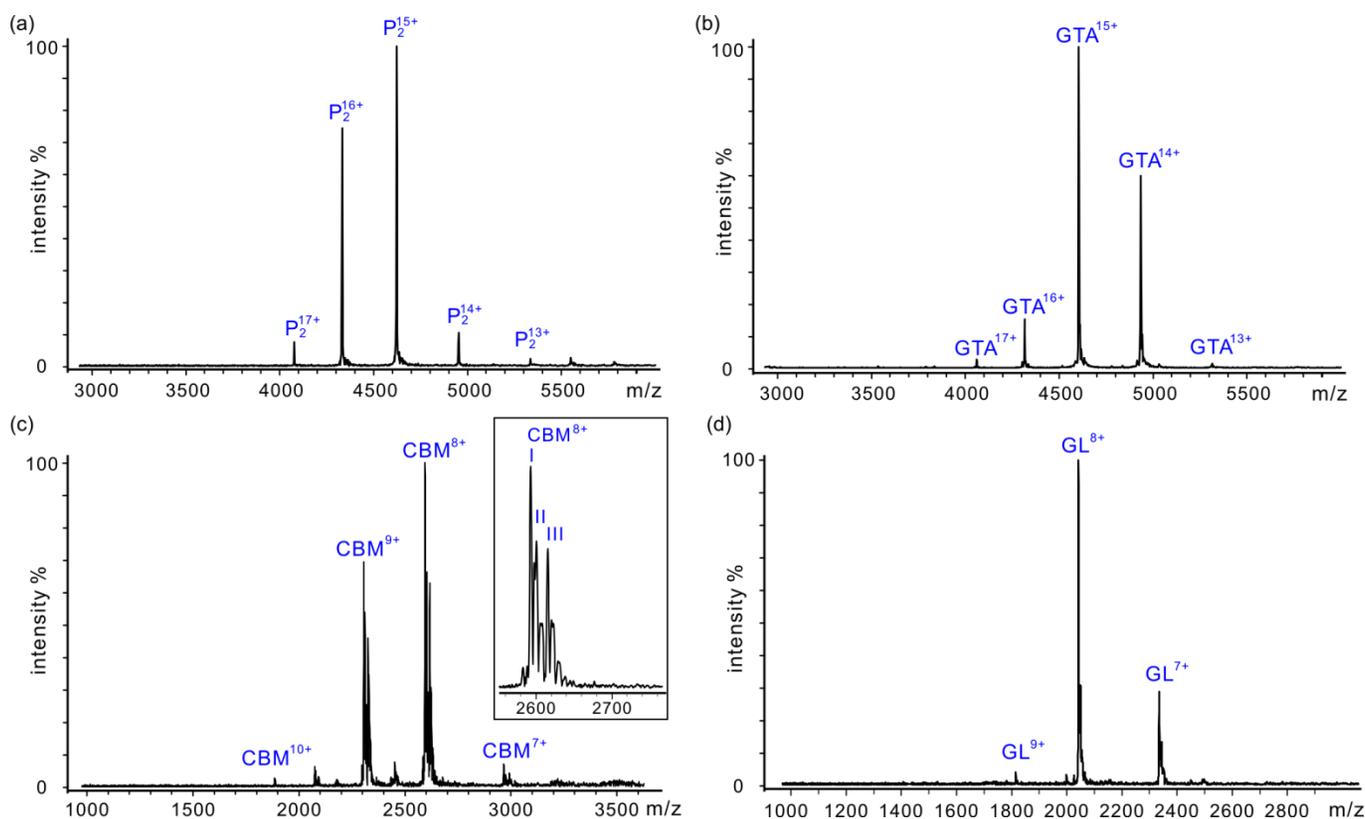


**Figure 4.2.** Chromatograph of gel filtration of the human NoV VA387 VLP ( $5 \text{ mg mL}^{-1}$ ) in 1x PBS (pH 7.4,  $25^\circ\text{C}$ ) buffer. The observation of only a single peak (corresponding to the void volume at  $\sim 45 \text{ mL}$ ) indicates the capsid protein of NoV VA387 predominantly assembles into large complexes, with molecular weights  $\geq 800 \text{ kDa}$ . The Superdex 200 gel filtration column was calibrated as described previously.<sup>7</sup>

#### 4.3.2 Affinities of HBGA oligosaccharides for the human NoV VA387 P particle and VLP

The *proxy protein* ESI-MS assay was used to evaluate the affinities of the thirteen HBGA oligosaccharides, **L1** – **L13**, for the human NoV VA387 P particle. Due to the limited availability of human NoV VA387 VLP, measurements were restricted to four oligosaccharides, **L1**, **L2**, **L7** and **L8**. In order to implement the assay, a suitable  $P_{\text{proxy}}$  (one that exhibits moderate/high affinity for the ligand) was required for each oligosaccharide tested. Four different  $P_{\text{proxy}}$  were used for these measurements, the P dimer of human NoV VA387, a truncated recombinant form of the

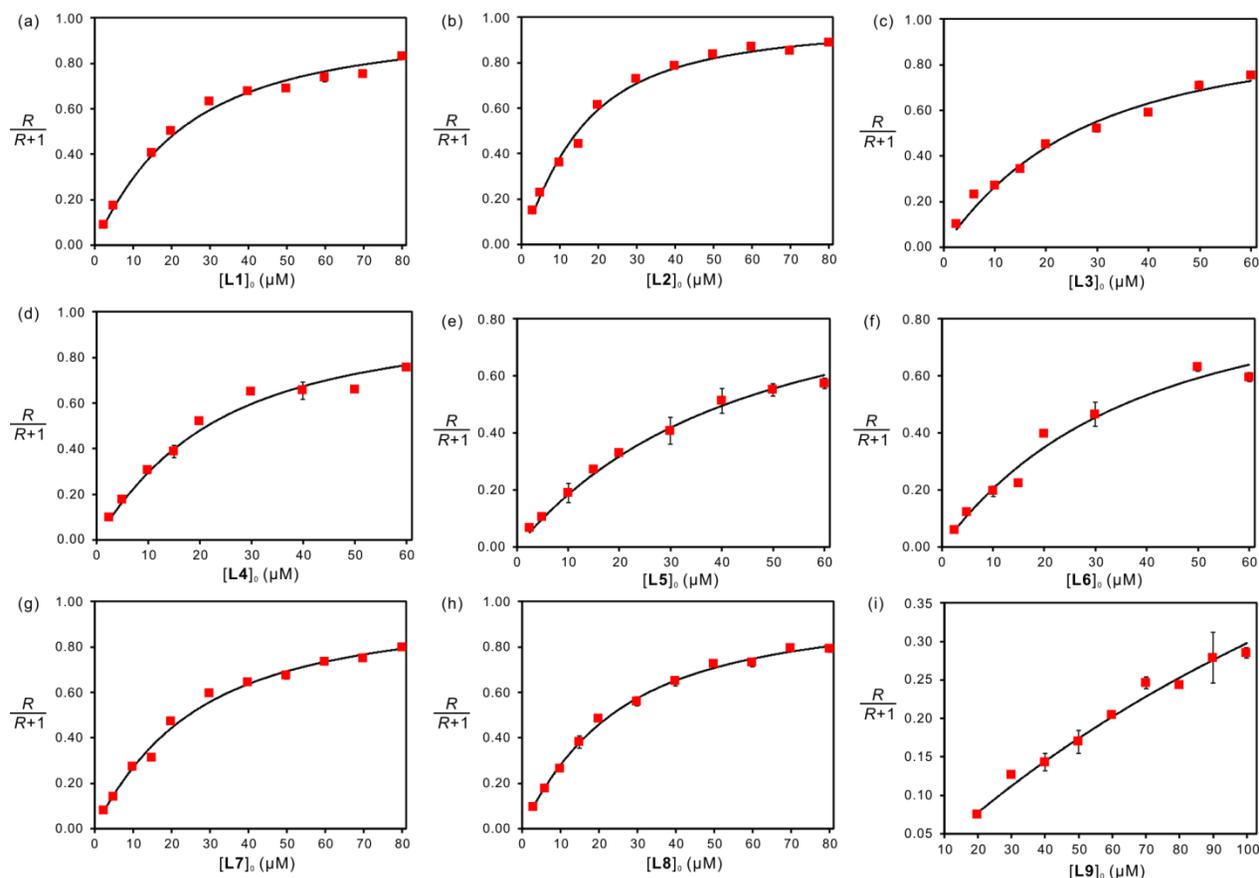
human blood group glycosyltransferase  $\alpha$ -(1 $\rightarrow$ 3)-N-acetylgalactosaminyltransferase (GTA), the family 51 carbohydrate-binding module (CBM) and a recombinant fragment of the C-terminus of human galectin-3 (Gal-3C). Representative ESI mass spectra acquired for each of the P<sub>proxy</sub> are shown in Figure 4.3.



**Figure 4.3.** ESI mass spectra obtained in positive ion mode for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing P dimer (P<sub>2</sub>, 12  $\mu$ M), (b) GTA (10  $\mu$ M), (c) CBM (12  $\mu$ M) and (d) Gal-3C (GL, 5  $\mu$ M). The measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.

Although the VA387 P dimer binds to a broad range of HBGAs, the interactions are uniformly weak.<sup>5,20</sup> Consequently, in the present study the P dimer was only used to quantify the interaction between the P particle and B type 3 tetrasaccharide (**L12**), which is the highest affinity HBGA ligand ( $K_{a,int} = 1500 \text{ M}^{-1}$ ) identified for the P dimer.<sup>20</sup> Truncated recombinant GTA, which forms a homodimer in aqueous solution at neutral pH, possesses two thermodynamically equivalent and independent acceptor substrate binding sites.<sup>37,38</sup> Recent ESI-MS measurements revealed that GTA exhibits modest intrinsic affinities for B trisaccharide (**L1**,  $1.6 \times 10^4 \text{ M}^{-1}$ ) and H disaccharide (**L13**,  $3.2 \times 10^4 \text{ M}^{-1}$ ),<sup>38,39</sup> which enabled the use of GTA as a  $P_{\text{proxy}}$  to quantify the interactions of **L1** and **L13** with the P particle.

Three distinct CBM species (labeled as I, II and III) with MWs of 20,738 Da, 20,798 Da and 20,916 Da, respectively, are evident from the mass spectrum shown in Figure 4.3c. The MW measured for CBM I agrees well with the theoretical MW of 20,735, the nature of the modifications giving rise to the other two forms of CBM were not established. However, all three forms bind to HBGA oligosaccharides with identical affinities (data not shown). Consequently, for the *direct* and *proxy protein* ESI-MS binding measurements, the abundances of all three forms were summed together to calculate the ratio of ligand-bound to free CBM. Glycan array screening carried out on CBM revealed binding to the A and B trisaccharides, as well as A and B type 2 and type 6 tetrasaccharides.<sup>24</sup> According to isothermal titration calorimetry (ITC), binding between CBM and the A and B type 2 tetrasaccharides is quite strong, with  $K_a$  values in the range of  $10^4 - 10^5 \text{ M}^{-1}$ .<sup>24</sup> To extend the utility of CBM as a  $P_{\text{proxy}}$  for the current study, the *direct* ESI-MS assay was used to measure the affinities of **L1** – **L14** for CBM (Table 4.1).



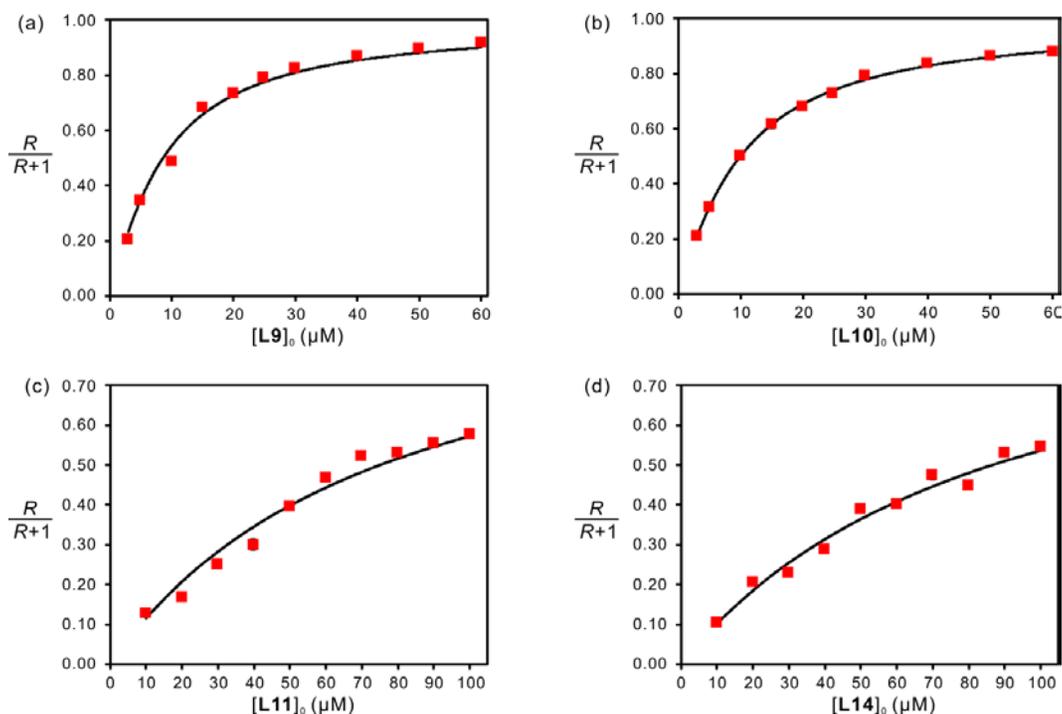
**Figure 4.4.** Plot of fraction of ligand-bound CBM versus ligand concentration measured for (a) **L1** (B trisaccharide), (b) **L2** (A trisaccharide), (c) **L3** (B type 2 tetrasaccharide), (d) **L4** (A type 2 tetrasaccharide), (e) **L5** (B type 5 tetrasaccharide), (f) **L6** (A type 5 tetrasaccharide), (g) **L7** (B type 6 tetrasaccharide), (h) **L8** (A type 6 tetrasaccharide), and (i) **L12** (B type 3 tetrasaccharide).

The ESI-MS binding measurements were carried out on 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing CBM (12 μM), Ubq ( $P_{ref}$ , 8 μM) and each ligand at a minimum of eight different concentrations ranging from 2.5 to 100 μM. The solid curves correspond to the best fit of eq 1.10 to the experimental data and the error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.

For a majority of the ligands (**L1 – L8** and **L12**), ESI-MS titration experiments were performed and the  $K_a$  values obtained by fitting eq 1.10 to the fraction of ligand bound protein measured experimentally (Figure 4.4). The  $K_a$  values for the other ligands (**L9 – L11** and **L13** and **L14**) were determined from ESI-MS measurements carried out at  $\geq 3$  different ligand concentrations. The results show that CBM only binds to A and B oligosaccharides, with the following trend in affinities: A/B trisaccharides > type 2 ~ type 6 > type 5 > type 3 > type 1 tetrasaccharides. It should be noted that the binding data for **L3** ( $(5.3 \pm 0.3) \times 10^4 \text{ M}^{-1}$ ) and **L4** ( $(7.4 \pm 0.3) \times 10^4 \text{ M}^{-1}$ ) measured by ESI-MS are in reasonable agreement with the  $K_a$  values obtained using ITC,  $7.8 \times 10^4 \text{ M}^{-1}$  and  $6.6 \times 10^4 \text{ M}^{-1}$ , respectively.<sup>24</sup>

Recombinant Gal-3C contains a carbohydrate recognition domain that interacts with the  $\beta$ -galactoside motif.<sup>40,41</sup> The affinities of the HBGA oligosaccharides (**L1 – L14**) for Gal-3C were measured by the *direct* ESI-MS assay (Table 4.1). For **L9 – L11** and **L14** the  $K_a$  values were obtained from ESI-MS titration experiments (Figure 4.5). For the remaining oligosaccharides the  $K_a$  values were determined from ESI-MS measurements carried out at  $\geq 3$  different ligand concentrations. Notably, Gal-3C binds strongly ( $K_a \sim 10^5 \text{ M}^{-1}$ ) to A/B type 1, 2 and 6 tetrasaccharides; exhibits moderately strong binding ( $K_a \sim 10^4 \text{ M}^{-1}$ ) to B type 3, A type 5, B type 5 tetrasaccharides, H type 6 and type 2 trisaccharides, and weak binding ( $K_a < 10^3 \text{ M}^{-1}$ ) to A/B trisaccharides and H disaccharide. It can also be seen from these data that Gal-3C binding to B type oligosaccharides is consistently stronger than to the corresponding A type oligosaccharides. It should also be pointed out that the  $K_a$  values for **L7** [ $(1.02 \pm 0.04) \times 10^5 \text{ M}^{-1}$ ], **L8** [ $(6.4 \pm 0.6) \times 10^4 \text{ M}^{-1}$ ], **L10** [ $(1.34 \pm 0.02) \times 10^5 \text{ M}^{-1}$ ] and **L11** [ $(1.38 \pm 0.06) \times 10^4 \text{ M}^{-1}$ ] agree

reasonably well with values measured using frontal affinity chromatography for B type 6 tetrasaccharide ( $1.7 \times 10^5 \text{ M}^{-1}$ ), A type 6 tetrasaccharide ( $7.1 \times 10^4 \text{ M}^{-1}$ ), A type 1 hexasaccharide ( $5.6 \times 10^5 \text{ M}^{-1}$ ) and H type 6 trisaccharide ( $1.1 \times 10^4 \text{ M}^{-1}$ ).<sup>40</sup>



**Figure 4.5.** Plot of fraction of ligand-bound Gal-3C versus ligand concentration measured for (a) **L9** (B type 1 tetrasaccharide), (b) **L10** (A type 1 tetrasaccharide), (c) **L11** (H type 6 trisaccharide), and (d) **L14** (type 2 trisaccharide). The ESI-MS binding measurements were carried out on 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing Gal-3C (5  $\mu\text{M}$ ), Ubq ( $P_{\text{ref}}$ , 3  $\mu\text{M}$ ) and each ligand at a minimum of nine different concentrations ranging from 2.5 to 100  $\mu\text{M}$ . The solid curves correspond to the best fit of eq 1.10 to the experimental data and the error bars correspond to one standard derivation. These measurements were performed on Waters Synapt G2S mass spectrometer.

**Table 4.1.** Association constants ( $K_a$ ) for binding of the HBGA oligosaccharides (L1 – L14) with CBM, Gal-3C and GTA measured at 25 °C and pH 7 by the ESI-MS assay.<sup>a</sup>

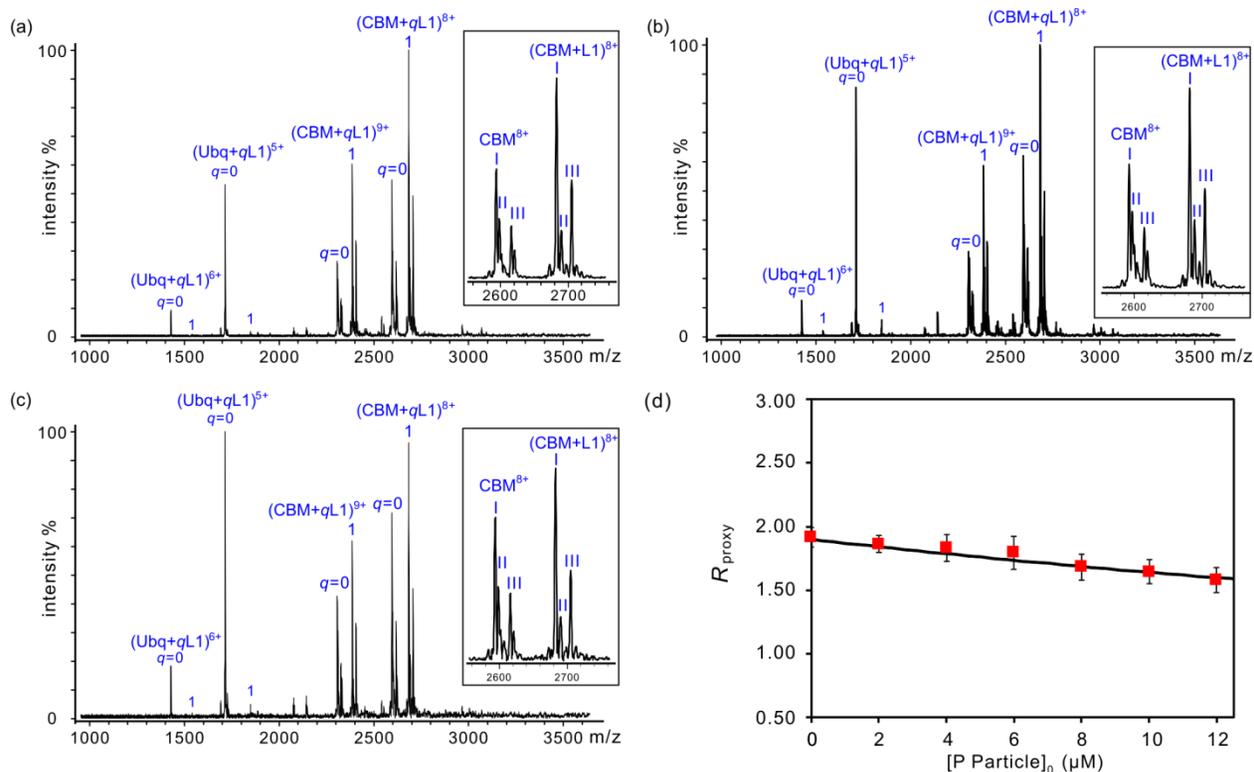
HBGA	Structure	$K_a$	$K_a$	$K_{a,int}^b$
		( $\times 10^4 M^{-1}$ ) CBM	( $\times 10^4 M^{-1}$ ) Gal-3C	( $\times 10^4 M^{-1}$ ) GTA
L1	B trisaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	(7.3 $\pm$ 0.4)	(0.08 $\pm$ 0.01)	1.6 <sup>c</sup>
L2	A trisaccharide $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	(11.3 $\pm$ 0.7)	(0.056 $\pm$ 0.002)	n.d. <sup>d</sup>
L3	B type 2 tetrasaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(5.3 $\pm$ 0.3)	(22.3 $\pm$ 1.7)	n.d. <sup>d</sup>
L4	A type 2 tetrasaccharide $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(7.4 $\pm$ 0.3)	(17.4 $\pm$ 1.4)	n.d. <sup>d</sup>
L5	B type 5 tetrasaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(3.1 $\pm$ 0.1)	(3.02 $\pm$ 0.06)	n.d. <sup>d</sup>
L6	A type 5 tetrasaccharide $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(3.3 $\pm$ 0.6)	(1.05 $\pm$ 0.04)	n.d. <sup>d</sup>
L7	B type 6 tetrasaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(5.6 $\pm$ 0.3)	(10.2 $\pm$ 0.4)	n.d. <sup>d</sup>
L8	A type 6 tetrasaccharide $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(5.8 $\pm$ 0.2)	(6.4 $\pm$ 0.6)	n.d. <sup>d</sup>
L9	B type 1 tetrasaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(0.36 $\pm$ 0.05)	(16.4 $\pm$ 0.9)	n.d. <sup>d</sup>
L10	A type 1 tetrasaccharide $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(0.31 $\pm$ 0.05)	(13.4 $\pm$ 0.2)	n.d. <sup>d</sup>
L11	H type 6 trisaccharide $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	< 0.02	(1.38 $\pm$ 0.06)	n.d. <sup>d</sup>
L12	B type 3 tetrasaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\alpha$ -D-GalNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	(0.44 $\pm$ 0.03)	(0.70 $\pm$ 0.05)	n.d. <sup>d</sup>
L13	H disaccharide $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	< 0.02	NB <sup>e</sup>	3.2 <sup>c</sup>
L14	type 2 trisaccharide $\alpha$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	< 0.02	(1.20 $\pm$ 0.04)	n.d. <sup>d</sup>

a. Uncertainties correspond to one standard deviation. b.  $K_{a,int}$  values corresponds to the intrinsic (per binding site) association constants. c. Values are adapted from references 38 and 39. d. n.d.  $\equiv$  not determined. e. NB  $\equiv$  no binding detected.

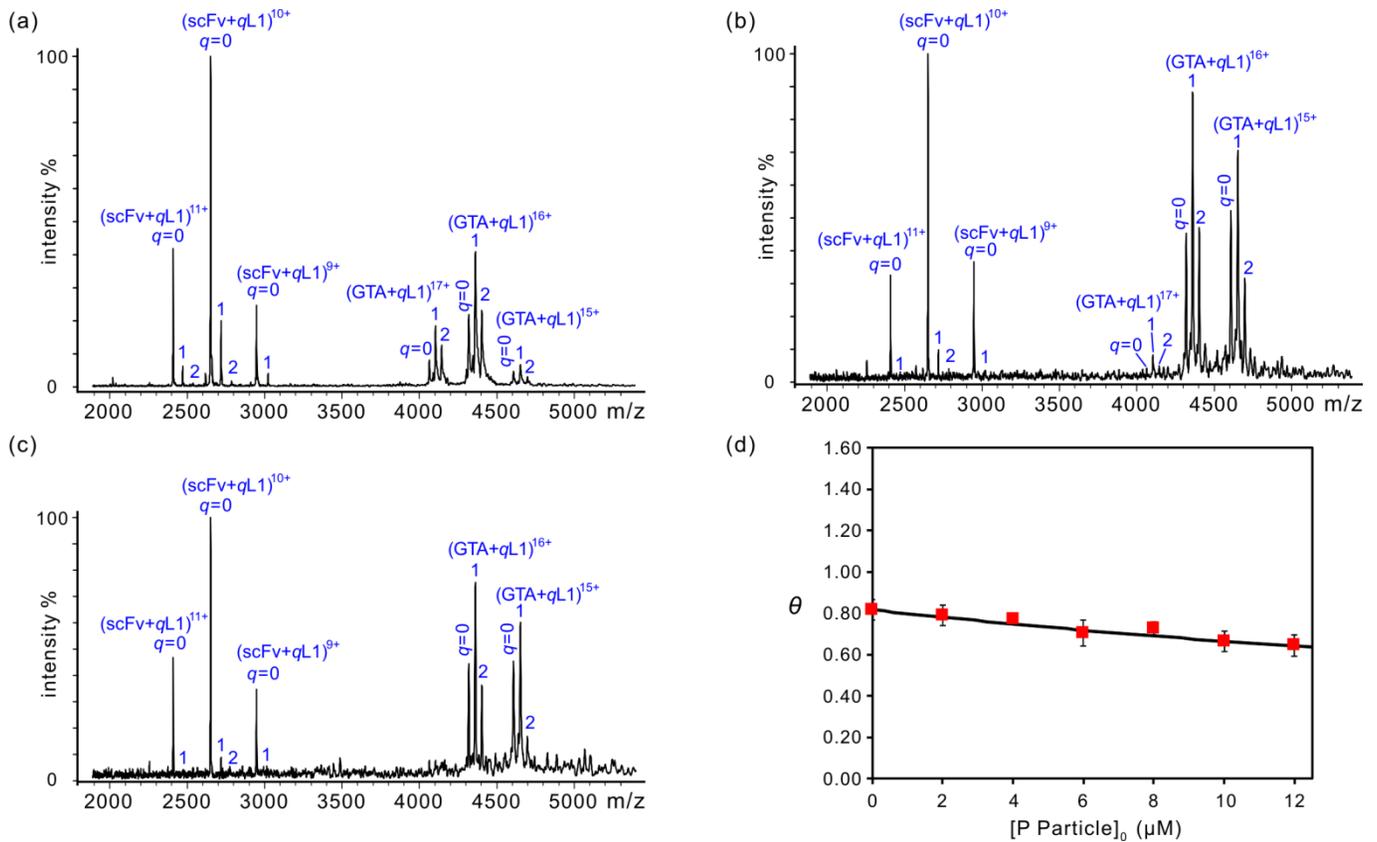
To quantify the affinities of **L1** – **L13** for the P particle, the *proxy protein* ESI-MS method was implemented using a titration format, whereby the concentrations of  $P_{\text{proxy}}$ , HBGA oligosaccharide and  $P_{\text{ref}}$  were fixed, while the concentration of the target protein (i.e. P particle or VLP) was varied. From the dependence of  $R_{\text{proxy}}$  (which corresponds to the abundance ratio of ligand-bound to free  $P_{\text{proxy}}$ ) on target protein concentration,  $K_{a,\text{int}}$  for the target protein could be determined. Shown in Figure 4.6 are representative ESI mas spectra acquired for aqueous 50 mM ammonium acetate solutions (pH 7, 25 °C) containing CBM (12  $\mu\text{M}$ ), B trisaccharide (**L1**, 35  $\mu\text{M}$ ), Ubq (8  $\mu\text{M}$ ) with 0, 6 and 12  $\mu\text{M}$  P particle (corresponding to 0, 144 and 288  $\mu\text{M}$  of monomer, respectively). Visual inspection reveals that the relative abundance of **L1**-bound  $P_{\text{proxy}}$  decreased with increasing P particle concentration, indicating that **L1** binds to the P particle. In Figure 4.6d the measured  $R_{\text{proxy}}$  values are plotted versus P particle concentration. Fitting eq 4.1b to these data gives a  $K_{a,\text{int}}$  of  $940 \pm 90 \text{ M}^{-1}$ .

Analogous measurements were carried out using GTA as  $P_{\text{proxy}}$ . Representative mass spectra acquired for aqueous 50 mM ammonium acetate solutions (pH 7, 25 °C) containing GTA (10  $\mu\text{M}$ ), **L1** (60  $\mu\text{M}$ ), scFv (8  $\mu\text{M}$ ,  $P_{\text{ref}}$ ) with 0, 6 and 12  $\mu\text{M}$  P particle are shown in Figure 4.7. Clearly, the relative abundance of **L1**-bound  $P_{\text{proxy}}$  decreased with the addition of P particle to the solution, consistent with the occurrence of binding of **L1** to the P particle. From the ratios  $R_{\text{proxy},1}$  and  $R_{\text{proxy},2}$ , which correspond to abundance ratios of **L1**-bound GTA (to one or two **L1**, respectively) to free GTA, the magnitude  $\theta$  was calculated (eq 4.12). Shown in Figure 4.7d is a plot of the calculated values of  $\theta$  versus P particle concentration. Fitting eq 4.11 to these data gives a  $K_{a,\text{int}}$  of  $1100 \pm 100 \text{ M}^{-1}$ . Importantly, these two values of  $K_{a,\text{int}}$  (which were determined

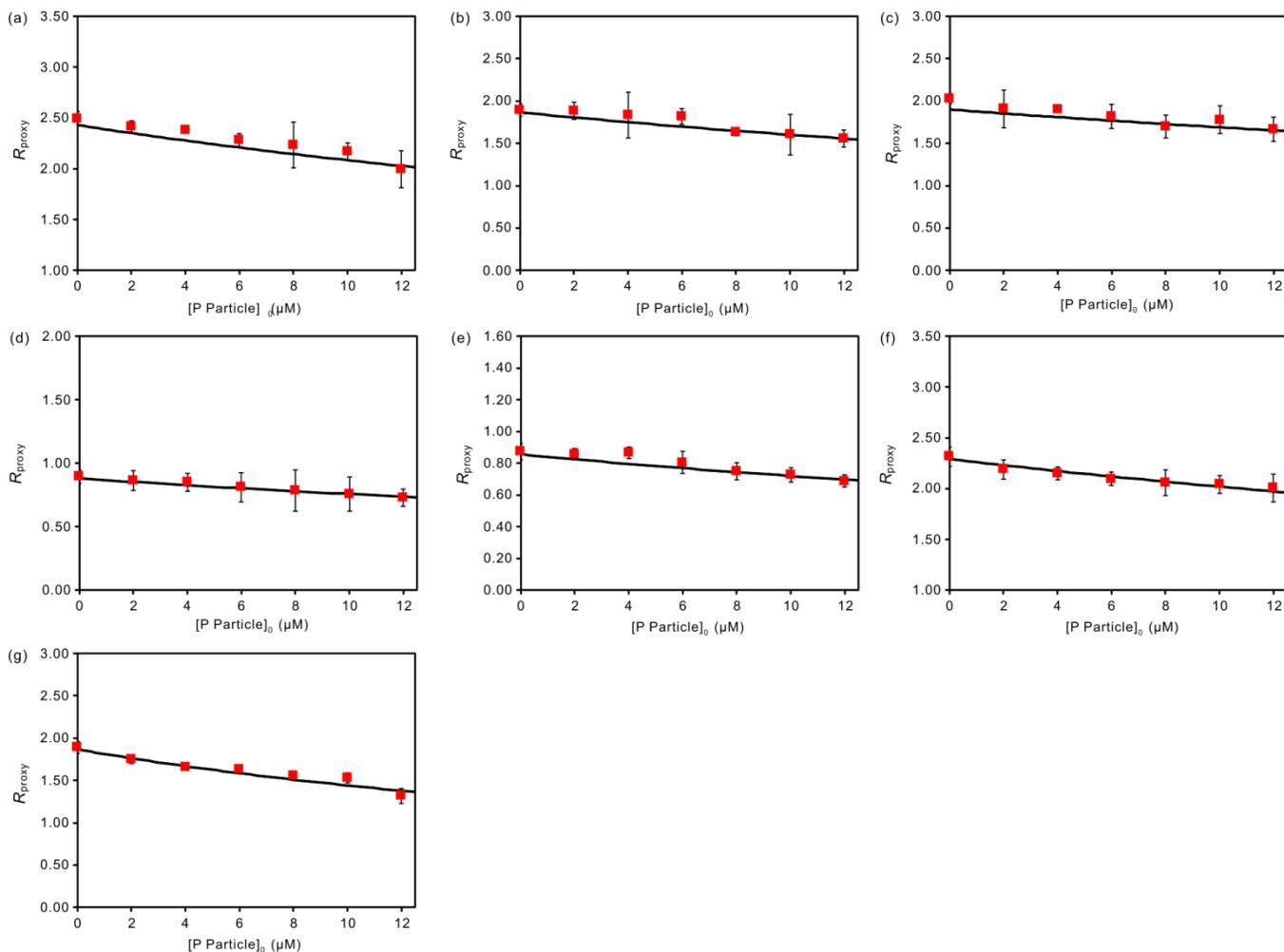
using different  $P_{\text{proxy}}$ ) are indistinguishable, within experimental error. Using an analogous strategy (Figures 4.8 – 4.12),  $K_{a,\text{int}}$  values for the interactions of **L2** – **L13** with the P particle were determined, Table 4.2.



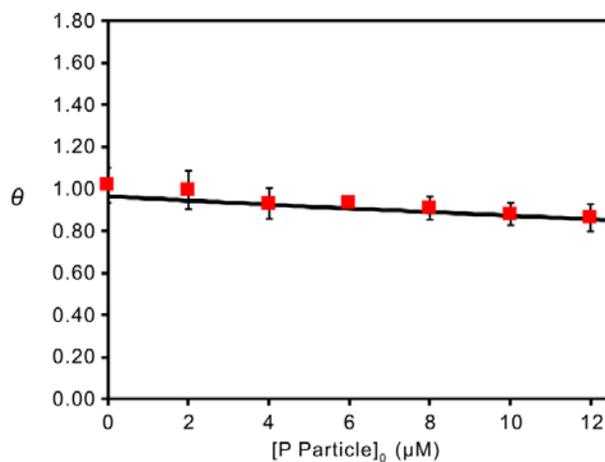
**Figure 4.6.** Representative ESI mass spectra obtained in positive ion mode for a 50 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12 μM CBM, 8 μM Ubq ( $P_{\text{ref}}$ ), 35 μM **L1** (B trisaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. (d) Plot of  $R_{\text{proxy}}$  versus concentration of monomer in the P particle. The solution condition for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to the best fit of eq 4.1b for the experimental data. The error bars correspond to one standard deviation. The measurements were carried out using a Bruker ApexQe FT-ICR mass spectrometer.



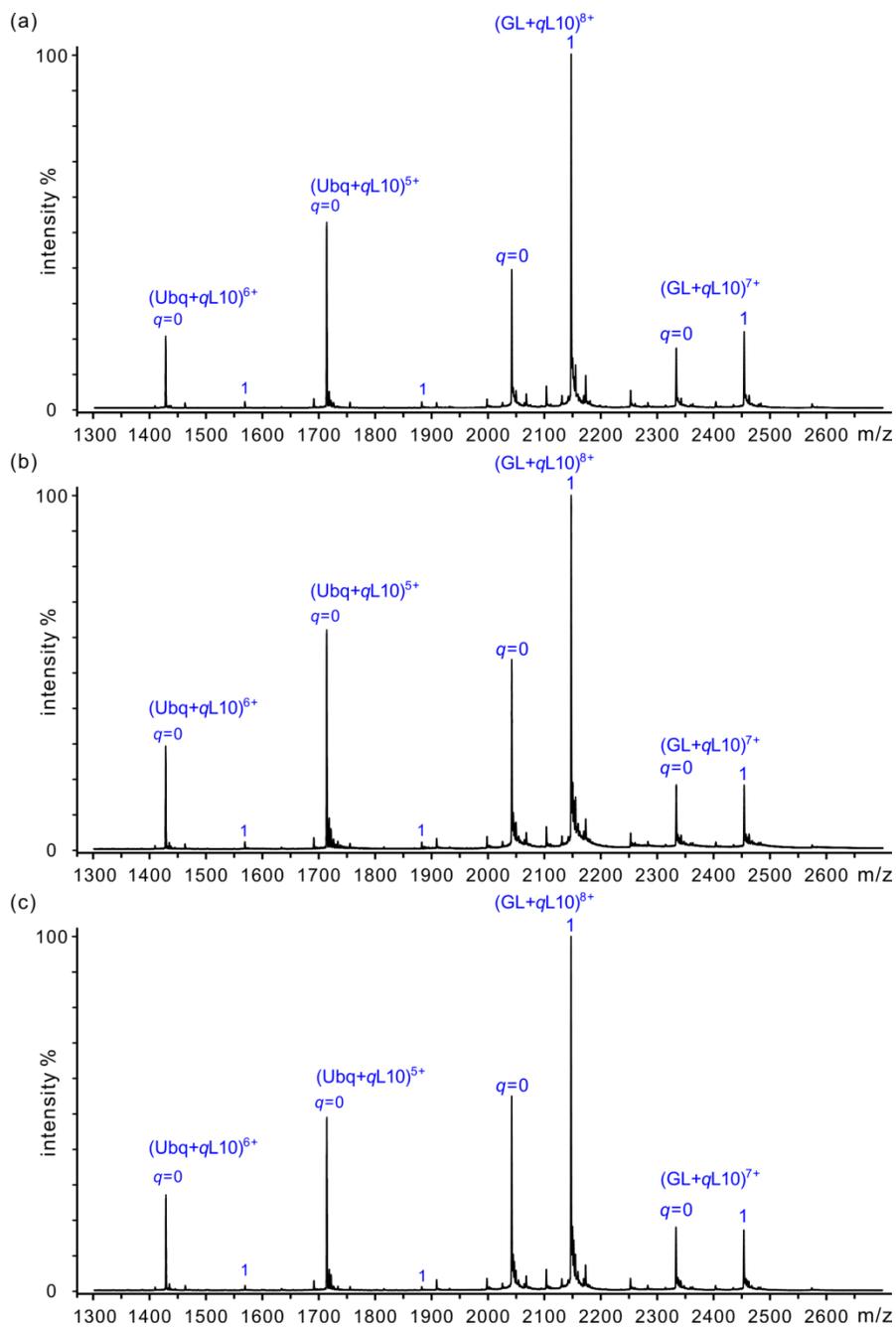
**Figure 4.7.** Representative ESI mass spectra obtained in positive ion mode for a 50 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 10 μM GTA, 10 μM scFv (P<sub>ref</sub>), 60 μM L1 (B trisaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. (d) Plot of  $\theta$  versus concentration of monomer in the P particle. The solution condition for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to the best fit of eq 4.11 to the experimental data. The error bars correspond to one standard deviation. The measurements were carried out using a Bruker ApexQe FT-ICR mass spectrometer.



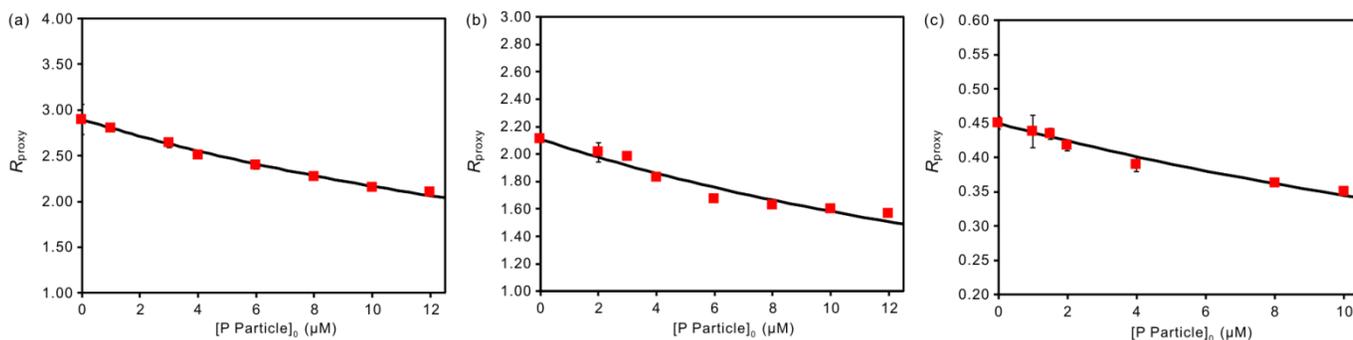
**Figure 4.8.** Plots of  $R_{\text{proxy}}$  versus P particle concentration measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C ) containing CBM (12  $\mu\text{M}$ ), Ubq ( $P_{\text{ref}}$ , 8  $\mu\text{M}$ ), (a) **L2** (A trisaccharide, 30  $\mu\text{M}$ ), (b) **L3** (B type 2 tetrasaccharide, 40  $\mu\text{M}$ ), (c) **L4** (A type 2 tetrasaccharide, 40  $\mu\text{M}$ ), (d) **L5** (B type 5 tetrasaccharide, 35  $\mu\text{M}$ ), (e) **L6** (A type 5 tetrasaccharide, 30  $\mu\text{M}$ ), (f) **L7** (B type 6 tetrasaccharide, 50  $\mu\text{M}$ ) and (g) **L8** (A type 6 tetrasaccharide, 40  $\mu\text{M}$ ), and P particle (0 – 16  $\mu\text{M}$ , which corresponds to 0 – 384  $\mu\text{M}$  of monomer). The solid curves correspond to the best fit of eq 4.1b to the experimental data for each ligand. The error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.



**Figure 4.9.** Plot of the  $\theta$  versus P particle concentration measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing GTA (10  $\mu\text{M}$ ), scFv ( $P_{\text{ref}}$ , 10  $\mu\text{M}$ ), **L13** (H disaccharide, 40  $\mu\text{M}$ ) and human NoV VA387 P particle (0 – 12  $\mu\text{M}$ , which corresponds to 0 – 288  $\mu\text{M}$  of monomer). The solid curve corresponds to the best fit of eq 4.11 to the experimental data. The error bars correspond to one standard derivation. These measurements were performed on Bruker ApexQe FT-ICR mass spectrometer.



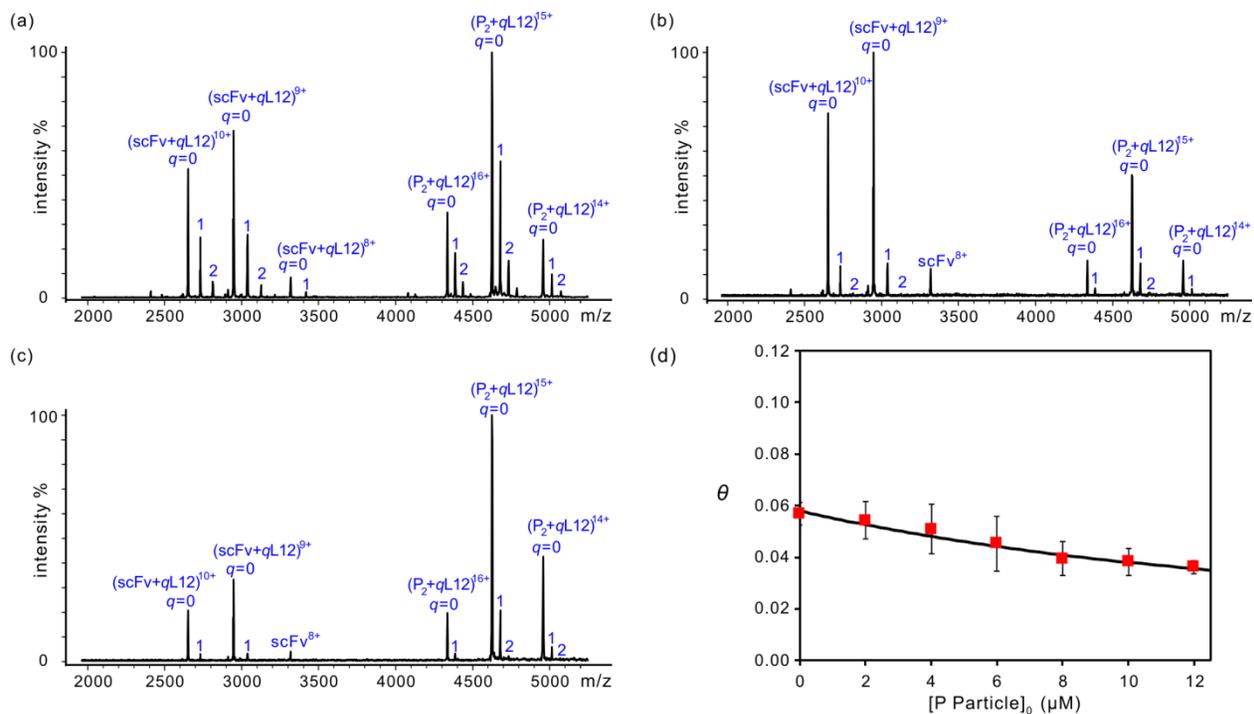
**Figure 4.10.** Representative ESI mass spectra measured for 200 mM ammonium acetate solutions (pH 7 and 25 °C) containing 5 μM Gal-3C (GL), 3 μM Ubiquitin (Ubq, P<sub>ref</sub>), 20 μM **L10** (A type 1 tetrasaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of human NoV VA387. The measurements were performed on a Waters Synapt G2S mass spectrometer.



**Figure 4.11.** Plots of  $R_{\text{proxy}}$  versus P particle concentration measured for 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C ) containing Gal-3C (5  $\mu\text{M}$ ), Ubq ( $P_{\text{ref}}$ , 3  $\mu\text{M}$ ), (a) **L9** (B type 1 tetrasaccharide, 20  $\mu\text{M}$ ), (b) **L10** (A type 1 tetrasaccharide, 20  $\mu\text{M}$ ), and (c) **L11** (H type 6 trisaccharide, 40  $\mu\text{M}$ ), and P particle (0 – 12  $\mu\text{M}$ , which corresponds to 0 – 288  $\mu\text{M}$  of monomer). The solid curves correspond to the best fit of eq 4.1b to the experimental data for each ligand. The error bars correspond to one standard derivation. Measurements were performed on a Waters Synapt G2S mass spectrometer.

The aforementioned measurements were carried out using 50 mM aqueous ammonium acetate solutions. To ensure that the affinity measurements were not sensitive to ionic strength, ESI-MS measurements were repeated using substantially higher concentrations of ammonium acetate. Representative mass spectra acquired for solutions of CBM (12  $\mu\text{M}$ ), **L1** (35  $\mu\text{M}$ ), Ubq (8  $\mu\text{M}$ ) and 6  $\mu\text{M}$  P particle in either 200 mM or 800 mM aqueous ammonium acetate solution (pH 7, 25 °C) are shown in (Figures 4.13a and 4.13b, respectively). Notably, the  $K_{a,\text{int}}$  values of  $940 \pm 60 \text{ M}^{-1}$  and  $900 \pm 110 \text{ M}^{-1}$ , respectively, are indistinguishable from the value measured in

50 mM aqueous ammonium acetate ( $940 \pm 90 \text{ M}^{-1}$ ). From these results it is concluded that the affinities of the HBGA ligands for the P particle and, presumably, VLP are relatively insensitive to the ionic strength of the solution.

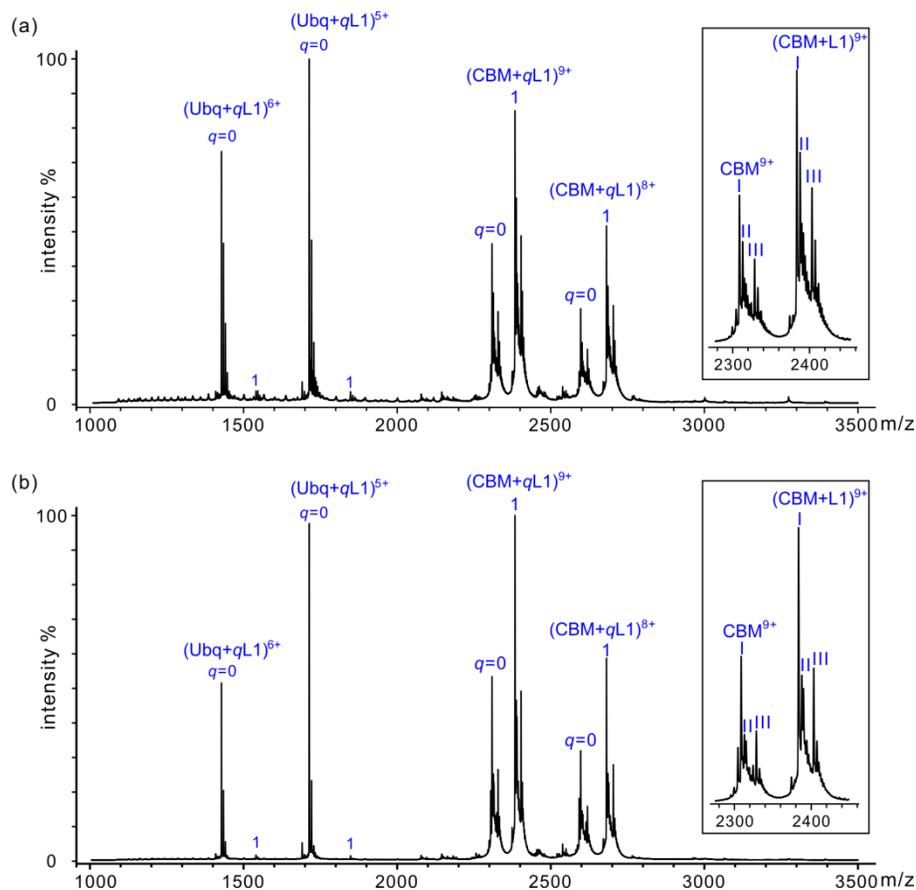


**Figure 4.12.** ESI mass spectra measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 12  $\mu\text{M}$  P dimer ( $\text{P}_2$ ), 10  $\mu\text{M}$  scFv ( $\text{P}_{\text{ref}}$ ), 40  $\mu\text{M}$  L12 (B type 3 tetrasaccharide) with (a) 0  $\mu\text{M}$ , (b) 6  $\mu\text{M}$  and (c) 12  $\mu\text{M}$  P particle (24-mer) of human NoV VA387. (d) Plot of the occupancy function  $\theta$  versus P particle concentration. The solution conditions for each measurement was same as (a), but with the addition of 0 to 12  $\mu\text{M}$  P particle. The solid curve corresponds to the best fit of eq 4.11 to the experimental data. The error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.

**Table 4.2.** Intrinsic association constants ( $K_{a,int}$ ) for HBGA oligosaccharides (**L1 – L14**) binding to human NoV VA387 P dimer, P particle and VLP, measured at 25 °C and pH 7 using the ESI-MS *proxy protein* assay.<sup>a</sup>

HBGA		Structure	$K_{a,int}$ (M <sup>-1</sup> ) P dimer <sup>b</sup>	$K_{a,int}$ (M <sup>-1</sup> ) P particle	$K_{a,int}$ (M <sup>-1</sup> ) VLP
<b>L1</b>	B trisaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	800 ± 100	940 ± 90 <sup>c</sup> 1100 ± 100 <sup>d</sup>	2300 ± 250 <sup>c</sup>
<b>L2</b>	A trisaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	500 ± 50	840 ± 90 <sup>c</sup>	1400 ± 150 <sup>c</sup>
<b>L3</b>	B type 2 tetrasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	410 ± 45	870 ± 60 <sup>c</sup>	n.d. <sup>h</sup>
<b>L4</b>	A type 2 tetrasaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	290 ± 30	710 ± 90 <sup>c</sup>	n.d. <sup>h</sup>
<b>L5</b>	B type 5 tetrasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	700 ± 100	930 ± 80 <sup>c</sup>	n.d. <sup>h</sup>
<b>L6</b>	A type 5 tetrasaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	560 ± 40	900 ± 100 <sup>c</sup>	n.d. <sup>h</sup>
<b>L7</b>	B type 6 tetrasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	600 ± 45	650 ± 30 <sup>c</sup>	1000 ± 160 <sup>c</sup>
<b>L8</b>	A type 6 tetrasaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	1200 ± 50	1400 ± 120 <sup>c</sup>	3900 ± 260 <sup>c</sup>
<b>L9</b>	B type 1 tetrasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	700 ± 100	1530 ± 40 <sup>e</sup>	n.d. <sup>h</sup>
<b>L10</b>	A type 1 tetrasaccharide	$\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	600 ± 65	1500 ± 100 <sup>e</sup>	n.d. <sup>h</sup>
<b>L11</b>	H type 6 trisaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-Glc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	330 ± 25	880 ± 50 <sup>c</sup>	n.d. <sup>h</sup>
<b>L12</b>	B type 3 tetrasaccharide	$\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\alpha$ -D-GalNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	1500 ± 150	2300 ± 110 <sup>f</sup>	n.d. <sup>h</sup>
<b>L13</b>	H disaccharide	$\alpha$ -L-Fuc-(1→2)- $\beta$ -D-Gal-O(CH <sub>2</sub> ) <sub>8</sub> COOC <sub>2</sub> H <sub>5</sub>	240 ± 40	520 ± 100 <sup>d</sup>	n.d. <sup>h</sup>
<b>L14</b>	type 2 trisaccharide	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-O(CH <sub>2</sub> ) <sub>6</sub> CH=CH <sub>2</sub>	NB <sup>g</sup>	NB <sup>g</sup>	n.d. <sup>h</sup>

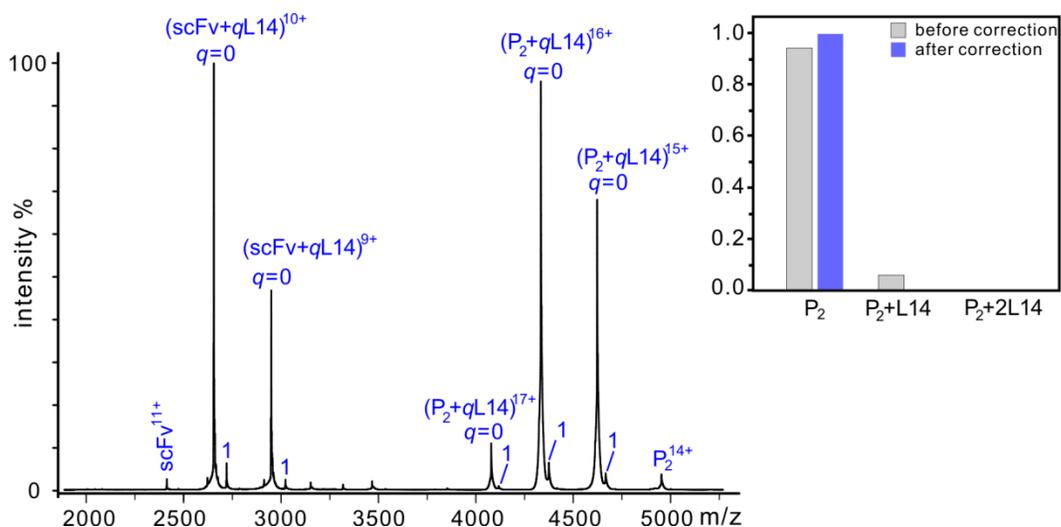
a. Uncertainties correspond to one standard deviation. b. Values are adapted from reference 20. c, d, e, f. Values are measured using CBM, GTA, Gal-3C and P dimer as P<sub>proxy</sub>, respectively. g. NB ≡ no binding detected. h. n.d. ≡ not determined.



**Figure 4.13.** ESI mass spectra obtained in positive ion mode for aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12  $\mu$ M CBM, 8  $\mu$ M Ubq ( $P_{ref}$ ), 35  $\mu$ M L1 (B trisaccharide) with 6  $\mu$ M P particle (24-mer) of human NoV VA387. The concentration of ammonium acetate in (a) was 200 mM and in (b) 800 mM. These measurements were performed on a Waters Synapt G2S mass spectrometer.

Control experiments were also carried out to rule out the possibility that the aglycone of the HBGA oligosaccharides used in the study,  $-(CH_2)_8COOC_2H_5$  for L1, L2 and L3 and –

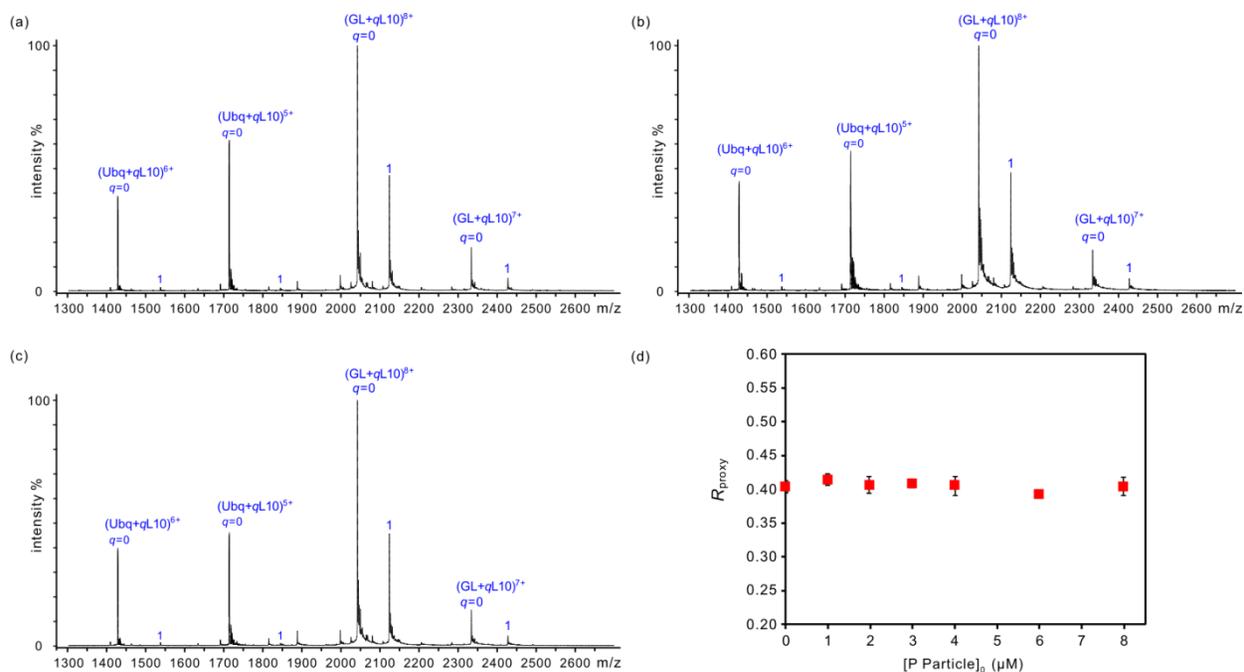
$(\text{CH}_2)_6\text{CH}=\text{CH}_2$  for **L3** – **L12**, promotes nonspecific binding to the P particle or VLP. The affinities of a type 2 trisaccharide (**L14**), which also has  $-(\text{CH}_2)_6\text{CH}=\text{CH}_2$  at the reducing end but does not bind to GII.4 human NoVs<sup>17</sup> for both the VA387 P dimer and the P particle were measured using *direct* and the *proxy protein* ESI-MS assay, respectively. Shown in Figure 4.14 is a representative ESI mass spectrum acquired for an aqueous 200 mM ammonium acetate solutions (pH 7, 25 °C) of P dimer (12  $\mu\text{M}$ ), **L14** (80  $\mu\text{M}$ ) and scFv (8  $\mu\text{M}$ ). It can be seen from the normalized distributions of **L14** bound P dimer (after correction for nonspecific ligand binding) that **L14** does not exhibit any detectable binding to the P dimer.



**Figure 4.14.** A representative ESI mass spectrum measured in positive ion mode for 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12  $\mu\text{M}$  VA387 P dimer ( $\text{P}_2$ ), 8  $\mu\text{M}$  scFv ( $\text{P}_{\text{ref}}$ ) and 80  $\mu\text{M}$  **L14** (type 2 trisaccharide). Inset, normalized distribution of **L14** bound P dimer before and after correction for nonspecific ligand binding. The measurement was performed on a Waters Synapt G2S mass spectrometer.

Shown in Figures 4.15a–c are representative ESI mass spectra acquired for aqueous 200 mM ammonium acetate solutions (pH 7, 25 °C) containing Gal-3C (5  $\mu$ M), **L14** (40  $\mu$ M), Ubq (3  $\mu$ M) with 0, 4 and 8  $\mu$ M P particle (corresponding to 0, 96 and 192  $\mu$ M of monomer, respectively). Notably, it can be seen from Figure 4.15d that  $R_{\text{proxy}}$  is independent of P particle concentration. This result confirms that the  $-(\text{CH}_2)_6\text{CH}=\text{CH}_2$  aglycone does not bind nonspecifically to the P dimer, the P particle or, presumably, the VLP.

According to the results of the *proxy protein* ESI-MS measurements, the thirteen HBGA oligosaccharides investigated (**L1** – **L13**) exhibit binding to the P particle, with  $K_{\text{a,int}}$  values ranging from 500  $\text{M}^{-1}$  to 2300  $\text{M}^{-1}$ . From these data, the apparent affinities of these HBGA oligosaccharides for the P particle (24-mer) can be estimated to be between  $1 \times 10^4$  and  $6 \times 10^4$   $\text{M}^{-1}$ . Notably, the trend in measured affinities of **L1** – **L13** for the P particle mirrors that found for the P dimer, with the B type 3 tetrasaccharide and H disaccharide being the strongest and weakest binders, respectively. Additionally, the H type 6 trisaccharide (**L11**, which is 2'-fucosyllactose), which is abundant in human milk and was recently shown to inhibit VA387 P particles from binding to other HBGAs,<sup>42</sup> exhibits a  $K_{\text{a,int}}$  of  $880 \pm 50$  for the P particle. Overall, the  $K_{\text{a,int}}$  values for the P dimer and P particle are within a factor of two. This finding suggests that the binding sites of the P dimer and P particle are structurally identical, or nearly so.

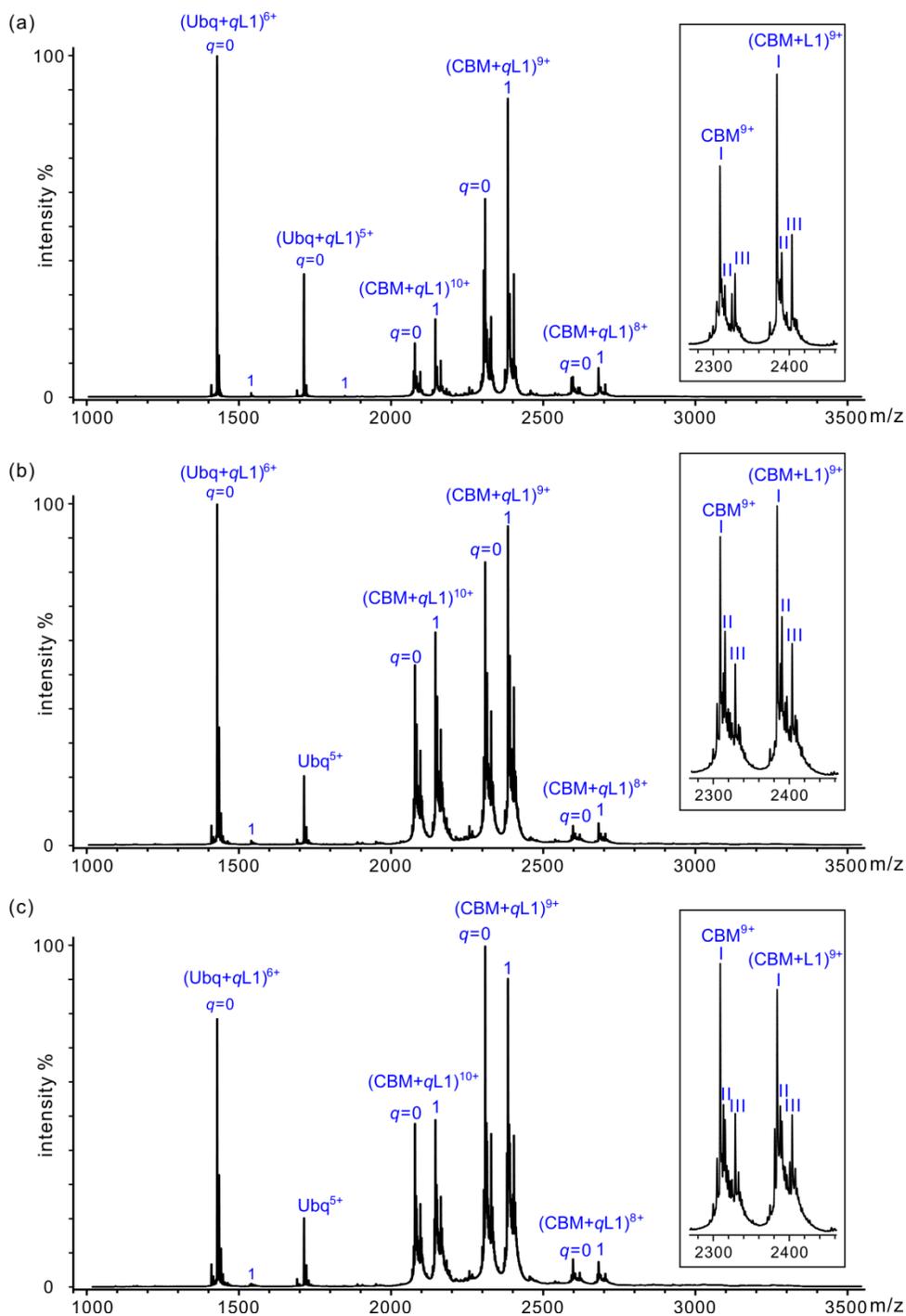


**Figure 4.15.** ESI mass spectra measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 5 μM Gal-3C (GL), 10 μM Ubiquitin (Ubq, P<sub>ref</sub>), 40 μM L14 (type 2 trisaccharide) with (a) 0 μM, (b) 4 μM and (c) 8 μM P particle (24-mer) of human NoV VA387. (d) Plot of  $R_{\text{prox}}$  versus P particle concentration. The solution conditions for each measurement was same as (a), but with the addition of 0 to 8 μM P particle. The error bars correspond to one standard derivation. The measurements were performed on a Waters Synapt G2S mass spectrometer.

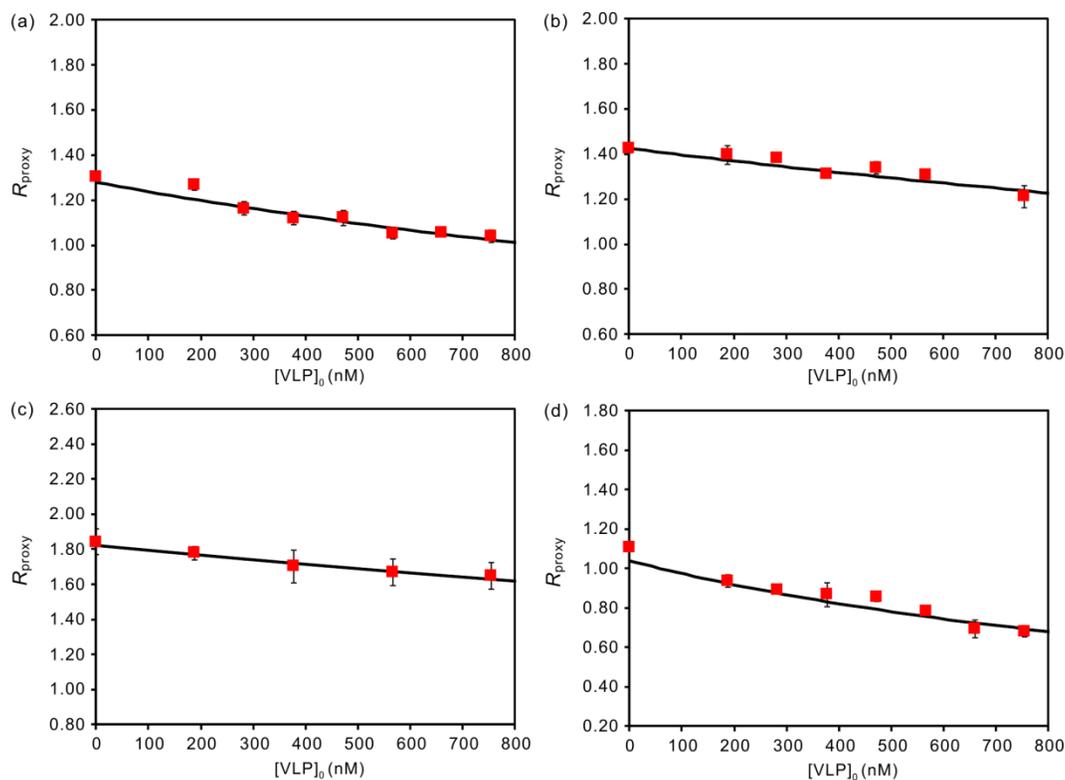
The affinities of four HBGA oligosaccharides (L1, L2, L7 and L8) for the VLP were also measured using CBM as the P<sub>prox</sub>. Shown in Figures 4.16a – 4.16c are representative ESI mass spectra acquired for aqueous 200 mM ammonium acetate solutions (pH 7, 25 °C) containing CBM (12 μM), L1 (25 μM) and Ubq (4 μM) with 0, 380 and 760 nM VLP (corresponding to 0,

68 and 136  $\mu\text{M}$  of monomer, respectively). From visual inspection it is seen that the relative abundance of **L1**-bound  $\text{P}_{\text{proxy}}$  decreased with the addition of the VLP to the solution, consistent with binding between **L1** and the VLP. Shown in Figure 4.17 are plots of  $R_{\text{proxy}}$  versus VLP concentrations determined for **L1**, **L2**, **L7** and **L8**. By fitting eq 4.1b to these data, the corresponding  $K_{a,\text{int}}$  values were determined, Table 4.2. The  $K_{a,\text{int}}$  values, which range from 1000  $\text{M}^{-1}$  to 4000  $\text{M}^{-1}$ , are consistently larger than those measured for the P particle, but are within a factor of three. From the measured  $K_{a,\text{int}}$  values, the apparent affinities of these ligands for the VLP (180-mer form) are estimated to be between  $2 \times 10^5$  and  $7 \times 10^5 \text{ M}^{-1}$ .

Taken together, the binding data measured in the present study and those reported previously for the P dimer indicate that the interactions between the HBGA ligands and the P dimer, P particle and VLP of human NoV VA387, while similar, are not identical and exhibit the following trend: P dimer  $\approx$  P particle < VLP. This finding suggests that there exist subtle differences in the structure of the carbohydrate binding sites presented by the P dimer and P particle and those of the VLP. It is interesting to note, however, that the crystal structures of the VLP<sup>3</sup> and P dimer<sup>43,44</sup> of Norwalk virus (GI.1) do not reveal a distinct structural difference between the P dimer of VLP and the one formed from the isolated P domain. Therefore, our data may imply that such subtle structural difference may not be easily recognized from crystal structures or, alternatively, that such subtle difference may occur between the VLP and the P dimers/P particle of VA387 (GII.4) but not between those of Norwalk virus. Nevertheless, the differences in  $K_{a,\text{int}}$  values for the VA387 P dimer, P particle and VLP are small and support the use of P dimers and P particles as surrogates to the VLP.



**Figure 4.16.** Representative ESI mass spectra obtained in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12  $\mu\text{M}$  CBM, 4  $\mu\text{M}$  Ubq ( $P_{\text{ref}}$ ), 25  $\mu\text{M}$  L1 (B trisaccharide) with (a) 0 nM, (b) 380 nM and (c) 760 nM VLP (180-mer) of human NoV VA387. The measurements were carried out using a Waters Synapt G2S mass spectrometer.



**Figure 4.17.** Plots of  $R_{\text{proxy}}$  versus concentration of monomer in the VLP measured for aqueous ammonium acetate solutions (200 mM, pH 7 and 25 °C) containing CBM (12  $\mu\text{M}$ ), Ubq ( $P_{\text{ref}}$ , 4  $\mu\text{M}$ ), VLP (180-mer) of human NoV VA387 (0 – 760 nM) with (a) **L1** (B trisaccharide, 25  $\mu\text{M}$ ), (b) **L2** (A trisaccharide, 20  $\mu\text{M}$ ), (c) **L7** (B type 6 tetrasaccharide, 40  $\mu\text{M}$ ) and (d) **L8** (A type 6 tetrasaccharide, 25  $\mu\text{M}$ ). The solid curves correspond to the best fit of eq 4.1b to the experimental data for each ligand. The error bars correspond to one standard deviation. The measurements were carried out using a Waters Synapt G2S mass spectrometer.

## 4.4 Conclusion

In summary, the interactions between the human NoV VA387 P particle and VLP and a series of HBGA oligosaccharides were quantified for the first time. The measured  $K_{a,int}$  values of the 13 HBGA oligosaccharides for the P particle range from 500 to 2300  $M^{-1}$ ; those of the A and B trisaccharides and the A and B type 6 tetrasaccharides for the VLP range from 1000  $M^{-1}$  to 4000  $M^{-1}$ . Comparison of the binding data with those measured previously for the P dimer reveals that the HBGA oligosaccharides tested exhibit similar intrinsic affinities for the P dimer and P particle. The intrinsic affinities exhibited for the VLP are consistently higher than those measured for the P particle, by a factor of three. Based on these data, the apparent affinities of the HBGA oligosaccharides tested for the P particle and VLP were estimated to be in the  $10^4 - 10^5 M^{-1}$  and  $10^5 - 10^6 M^{-1}$  range, respectively. While the cause of the subtle differences in HBGA oligosaccharide affinities for the P dimer and P particle and those for the VLP remains unknown, the present data support the use of P dimers and P particles as substitutes to the VLP for NoV-receptor binding studies.

## 4.5 References

- (1) Patel, M. M.; Widdowson, M.-A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (2) Jiang, X.; Wang, M.; Graham, D. Y.; Estes, M. K. *J. Virol.* **1992**, *66*, 6527.
- (3) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.

- (4) White, L. J.; Hardy, M. E.; Estes, H. K. *J. Virol.* **1997**, *71*, 8066.
- (5) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (6) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.
- (7) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (8) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.
- (9) Bereszcak, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R. *J. Struct. Biol.* **2012**, *177*, 273.
- (10) Tamminen, K.; Huhti, L.; Koho, T.; Lappalainen, S.; Hytonen, V. P.; Vesikari, T.; Blazevic, V. *Immunology* **2012**, *135*, 89.
- (11) Oriol, R. *J. Immunogenet.* **1990**, *17*, 235.
- (12) Ravn, V.; Dabelsteen, E. *APMIS* **2000**, *108*, 1.
- (13) Hutson, A. M.; Atmar, R. L.; Graham, D. Y.; Estes, M. K. *J. Infect. Dis.* **2002**, *185*, 1335.
- (14) Hutson, A. M.; Atmar, R. L.; Marcus, D. M.; Estes, M. K. *J. Virol.* **2003**, *77*, 405.
- (15) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang, X. *J. Virol.* **2005**, *79*, 6714.
- (16) Tan, M.; Jiang, X. *Trends Microbiol.* **2005**, *13*, 285.
- (17) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (18) de Rougemont, A.; Ruvoen-Clouet, N.; Simon, B.; Estienney, M.; Elie-Caille, C.; Aho,

- S.; Pothier, P.; Le Pendu, J.; Boireau, W.; Belliot, G. *J. Virol.* **2011**, *85*, 4057.
- (19) Hansman, G. S.; Shahzad-ul-Hussan, S.; McLellan, J. S.; Chuang, G.-Y.; Georgiev, I.; Shimoike, T.; Katayama, K.; Bewley, C. A.; Kwong, P. D. *J. Virol.* **2012**, *86*, 284.
- (20) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. *Glycobiology* **2013**, *23*, 276.
- (21) El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 3867.
- (22) Jiang, X.; Zhong, W.; Farkas, T.; Huang, P.; Wilton, N.; Barrett, E.; Fulton, D.; Morrow, R.; Matson, D. O. *Arch. Virol* **2002**, *147*, 119.
- (23) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendu, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.
- (24) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.
- (25) Seto, N. O. L.; Palcic, M. M.; Compston, C. A.; Li, H.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1997**, *272*, 14133.
- (26) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423.
- (27) Meloncelli, P. J.; Lowary, T. L. *Aust. J. Chem.* **2009**, *62*, 558.
- (28) Meloncelli, P. J.; Lowary, T. L. *Carbohydr. Res.* **2010**, *345*, 2305.
- (29) Meloncelli, P. J.; West, L. J.; Lowary, T. L. *Carbohydr. Res.* **2011**, *346*, 1406.

- (30) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (31) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.
- (32) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.
- (33) Wang, L.; Xia, M.; Huang, P.; Fang, H.; Cao, D.; Meng, X.-J.; McNeal, M.; Jiang, X.; Tan, M. *Biomaterials* **2014**, *35*, 8427.
- (34) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.
- (35) Shoemaker, G. K.; van Duijn, E.; Crawford, S. E.; Uetrecht, C.; Baclayon, M.; Roos, W. H.; Wuite, G. J. L.; Estes, M. K.; Prasad, B. V. V.; Heck, A. J. R. *Mol. Cell. Proteomics* **2010**, *9*, 1742.
- (36) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3*, 126.
- (37) Patenaude, S. I.; Seto, N. O. L.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol.* **2002**, *9*, 685.
- (38) Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587.
- (39) Soya, N.; Shoemaker, G. K.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2009**, *19*, 1224.
- (40) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E. G.; Yagi, F.; Kasai, K. *Biochim. Biophys. Acta Gen. Subj.* **2002**, *1572*, 232.

- (41) Collins, P. M.; Bum-Erdene, K.; Yu, X.; Blanchard, H. *J. Mol. Biol.* **2014**, *426*, 1439.
- (42) Shang, J.; Piskarev, V. E.; Xia, M.; Huang, P.; Jiang, X.; Likhoshesterov, L. M.; Novikova, O. S.; Newburg, D. S.; Ratner, D. M. *Glycobiology* **2013**, *23*, 1491.
- (43) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.
- (44) Bu, W.; Mamedova, A.; Tan, M.; Xia, M.; Jiang, X.; Hegde, R. S. *J. Virol.* **2008**, *82*, 5340.

## Chapter 5

### Gangliosides are Ligands for Human Noroviruses\*

#### 5.1 Introduction

Noroviruses (NoVs), a group of small, round-structured RNA viruses constituting the *Norovirus* genus in the family *Caliciviridae*, infect both humans and animals. Human NoVs cause epidemic acute gastroenteritis, affecting millions of people and claiming over 200,000 lives annually worldwide.<sup>1,2</sup> At present, there is no effective vaccine or antiviral against human NoVs. Structurally, NoVs are nonenveloped, containing an outer protein capsid that encapsulates the single-stranded, positive sense RNA genome of ~7.7 kb. The NoV capsid is made from a single major structural viral protein, VP1. Crystallography of recombinant NoV-like particles (VLPs) reveals a  $T = 3$  icosahedral symmetry consisting of 180 copies of VP1 organized into 90 dimers.<sup>3</sup> VP1 is divided into two major domains, the shell (S) and the protruding (P) domains. The S domain forms the interior, icosahedral shell; while the P domain forms the dimeric protrusions extending outward from the shell.<sup>3</sup> The P domain can be further divided into P1 and P2 subdomains, corresponding to the legs and the head of the arch-like protrusion, respectively. The P2 subdomain forms the outermost surface of the capsid with highly variable sequence, responsible for the virus-host interactions and immune recognitions of NoVs.<sup>2,4-6</sup>

---

\* A version of this chapter has been published: Han, L.; Tan, M.; Xia, M.; Kitova, E. N.; Jiang, X.; Klassen, J. S., *J. Am. Chem. Soc.* **2014**, *136*, 12631-12637.

Human NoVs are difficult to study due to the lack of an efficient cell culture system and a small animal model. Currently, research into NoV-host interactions relies on various NoV subviral particles. Expression of full-length VP1 results in VLPs that are structurally similar to an authentic virus.<sup>3</sup> Furthermore, expression of various subdomains results in smaller subviral particles or complexes. For example, production of the S domain forms S particles,<sup>7,8</sup> corresponding to the interior shell of the capsid, while expressions of the P domains with or without modifications can form P dimers,<sup>8-12</sup> 12-mer small P particles<sup>13</sup> or 24-mer P particles.<sup>14,15</sup> In addition, various glutathione S-transferase (GST)-P domain fusion proteins have been shown to form polyvalent complexes owing to the dimeric and oligomeric features of the GST and the P domain.<sup>16,17</sup> These VLPs, P particles and P complexes retain the basic structures of the capsid or P dimer, recognize host ligands and, thus, have been used as tools or models for the study of NoV-host interactions.

Human NoVs recognize histo-blood group antigens (HBGAs) as attachment factors or receptors, which play an important role in the host susceptibility of NoV infection, as shown by both human challenge studies and outbreak investigations.<sup>18-20</sup> HBGAs are oligosaccharides linked to membrane proteins or lipids as glycoprotein or glycolipid that are distributed extensively on the surfaces of red blood cells and mucosal epithelia.<sup>21</sup> They are also present as free oligosaccharides in biological fluids, such as saliva or milk.<sup>21</sup> Human NoVs interact with HBGAs in a strain-specific manner, whereby a number of NoV-HBGA binding patterns involved in all ABO, Lewis and secretor/nonsecretor types have been identified.<sup>22,23</sup> The structural basis of these interactions have been elucidated by X-ray crystallography of NoV P dimers in complex

with HBGA oligosaccharides.<sup>9-12,24</sup> However, it has been observed that some human NoVs, such as VA115 (GI.3),<sup>23</sup> Desert Shield virus (GI.3)<sup>23</sup> and Noda485 (GII.1),<sup>25</sup> do not bind any HBGAs. A human challenge study of Snow Mountain virus (SMV, GII.2) did not reveal a dependence of host susceptibility on HBGA type, despite the fact that the SMV VLP recognizes only the B antigen.<sup>26</sup> In addition, a recent study showed that NoV VLPs of Ueno 7k (GII.6) and Noda485 binds Caco-2 cells and human small intestinal epithelium biopsy in a HBGA independent manner.<sup>27</sup> These data suggest that HBGAs may not be the only receptors for human NoVs.

Recent studies have implicated glycosphingolipids and acidic oligosaccharides as human NoV ligands. For example, using thin-layer chromatography and quartz crystal microbalance with dissipation monitoring, Larson *et al.* reported binding of GII.4 VLPs to galactosylceramide and HBGA glycosphingolipids that were purified from human meconium samples.<sup>28,29</sup> Takeda and coworkers demonstrated that VLPs of GII NoVs bound heparan sulfate on the cell surface,<sup>30</sup> while Belliot and coworkers showed that GII.4 VLPs recognized sialic acid-containing carbohydrates, such as sialyl Lewis X (Le<sup>X</sup>), sialyl-lacto-N-fucopentaose, sialyl-lacto-N-tetraose and sialyl-lacto-N-neotetraose, with affinities comparable to those of HBGA ligands.<sup>31</sup> Using saturation transfer difference nuclear magnetic resonance spectroscopy, Peters and coworkers detected the interactions between GII.4 VLPs and the sialic acid moiety of sialyl Le<sup>X</sup> and sialyl Le<sup>a</sup>.<sup>32</sup> However, they also found that carbohydrates containing sialic acid, but not fucose, e.g. 3'-sialyllactose and 6'-sialyllactose, do not exhibit detectable binding with the VLP.<sup>32</sup> The results of these studies, taken together, imply that sialic acid-containing oligosaccharides could also be ligands of human NoVs. In fact, sialic acid-containing oligosaccharides have been shown to be

ligands or receptors for some animal caliciviruses (CVs), including murine NoV (MNV1),<sup>33</sup> feline calicivirus (FCV)<sup>34</sup> and a porcine sapovirus (PSaV, Cowden strain).<sup>35</sup> However, solid evidence to establish the ligand status of sialic acid for human CVs (human NoVs and human sapoviruses) is lacking.

Here, we report the first experimental evidence that human NoVs recognize sialic acid-containing glycosphingolipids (gangliosides). The catch-and-release electrospray ionization mass spectrometry (CaR-ESI-MS) assay<sup>36</sup> was used to screen a library of gangliosides against the P particle of human NoV VA387 (GII.4). The affinities of thirteen gangliosides for the P dimer of VA387 and of a second human NoV strain, VA115 (GI.3), were measured using the *direct* ESI-MS assay.<sup>37</sup> Using a competitive ESI-MS assay, the *proxy protein* method,<sup>38</sup> the highest affinity ligand, GM3, was subjected to additional binding measurements and the affinities for both the VA387 P particle and VLP were determined. Notably, the ganglioside affinities measured for NoV VA387 are comparable to those of known HBGA oligosaccharide receptors.<sup>39</sup> Enzyme-linked immunosorbent assays (ELISA) provided additional evidence that both strains of NoVs exhibit binding to sialic acid-containing oligosaccharides.

## 5.2 Experimental Section

### 5.2.1 Proteins

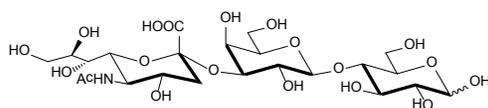
The VLPs of VA387 (GII.4) were produced in insect cells (SF9) through a recombinant baculovirus containing the gene encoding VA387 VP1 (GenBank accession number AY038600, molecular weight (*MW*) of monomer 58,887 Da) as described previously.<sup>22</sup> The resulting VLPs

were purified by sucrose gradient. VA387 P particles (24-mer, *MW* 865,036 Da), P dimers (*MW* 69,312 Da)<sup>8</sup> and GST-P domain fusion proteins were produced based on the P domain sequences (residues 222-539) of VP1 via *E. coli* as reported in our previous studies.<sup>16,17</sup> The GST– Gene Fusion System (GE Healthcare Life Sciences, Piscataway, NJ) with plasmid vector pGEX-4T-1 was used for the P proteins expression. Preparations of VA115 (GI.3) VLPs and P particles were attempted based on the VP1 sequences (GenBank accession number AY038598) and the established procedure described above, but the yields for both particles were found to be very low. The P dimers (*MW* 67,712 Da) and the GST-P fusion proteins of VA115 were produced in high yield (>20 mg L<sup>-1</sup> bacteria) through the same procedure as used for the production of the P proteins of VA387. Formations of the 24-mer P particles, P dimers, and the GST-P polymers were analyzed by gel-filtration chromatography via a Superdex 200 size exclusion column (GE Healthcare Life Sciences) controlled by an Akta Fast Performance Liquid Chromatography system (FPLC, Model 920, GE Healthcare Life Sciences).

A single chain fragment (scFv, *MW* 26,539 Da) of the monoclonal antibody Se155–4 was produced using recombinant technology as described elsewhere.<sup>40</sup> A recombinant fragment of the C-terminus of human galectin-3 (Gal-3C, *MW* 16,330 Da) was generously provided by Prof. C. Cairo (University of Alberta). Bovine ubiquitin (Ubq, *MW* 8,565 Da) was purchased from Sigma-Aldrich Canada (Oakville, Canada). The proteins were concentrated and exchanged into an aqueous 200 mM ammonium acetate (pH 7) using Vivaspin 0.5 mL centrifugal filters (Sartorius Stedim Biotech, Göttingen, Germany) with a *MW* cutoff of 10 kDa and stored at –80 °C until use. The concentrations of protein stock solutions were estimated by UV absorption.

## 5.2.2 Ligands

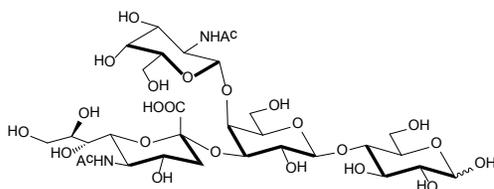
The structures of the oligosaccharides and glycoconjugates used in this study are shown in Figure 5.1. The seventeen ganglioside and globoside oligosaccharides (GM3, GM2, GM1a, GM1b, GD3, GD2, GD1a, GD1b, GT3, GT2, GT1a, GT1c, fucosyl-GM1, asialo-GM2, asialo-GM1, Gb3 and Gb4) were purchased from Elicityl SA (Crolles, France). H type 3 trisaccharide, A type 3 tetrasaccharide and B type 3 tetrasaccharide were a gift from Prof. T. Lowary (University of Alberta).<sup>41</sup> Each solid compound was dissolved in ultrafiltered Milli-Q water (Millipore, MA) to give a 1 mM stock solution. The stock solutions were stored at  $-20\text{ }^{\circ}\text{C}$  until needed. Polyacrylamide (PAA)-conjugated Neu5Ac, 6'-sialylacNAc and GM3 trisaccharide were purchased from Vector Lab (Burlingame, CA). They were stored at  $-20\text{ }^{\circ}\text{C}$  until used.



### GM3

MW 633.21 Da

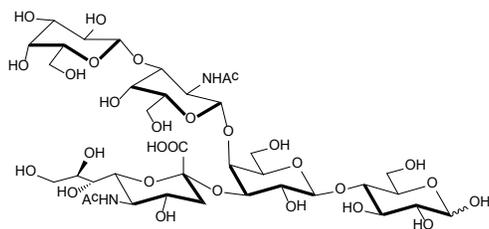
$\alpha$ -D-Neu5Ac-(2,3)- $\beta$ -D-Gal-(1,4)-D-Glc



### GM2

MW 836.29 Da

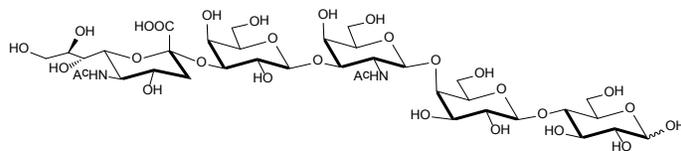
$\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**GM1a**

MW 998.34 Da

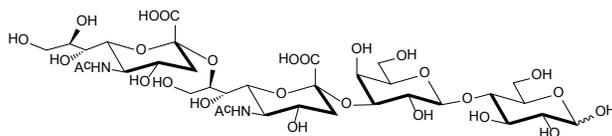
$\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**GM1b**

MW 998.34 Da

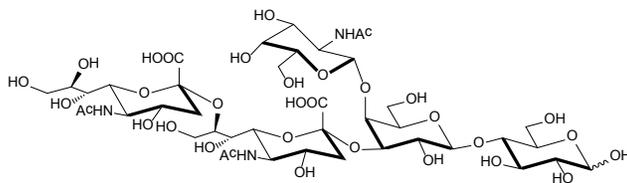
$\alpha$ -D-Neu5Ac-(2,3)- $\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)- $\beta$ -D-Gal-(1,4)-D-Glc



**GD3**

MW 924.31 Da

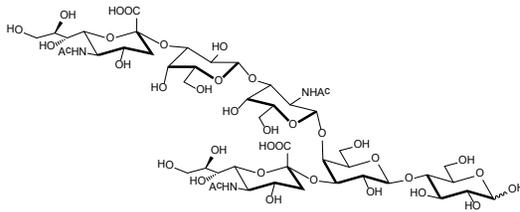
$\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,3)- $\beta$ -D-Gal-(1,4)-D-Glc



**GD2**

MW 1127.39 Da

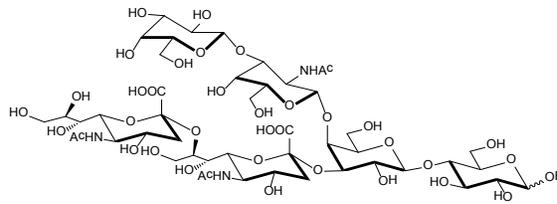
$\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**GD1a**

MW 1289.44 Da

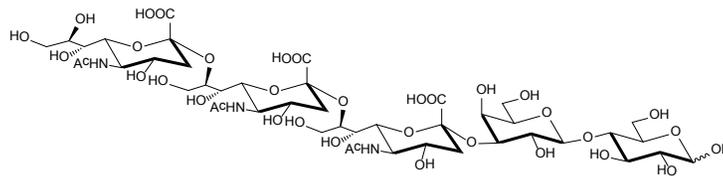
$\alpha$ -D-Neu5Ac-(2,3)- $\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**GD1b**

MW 1289.44 Da

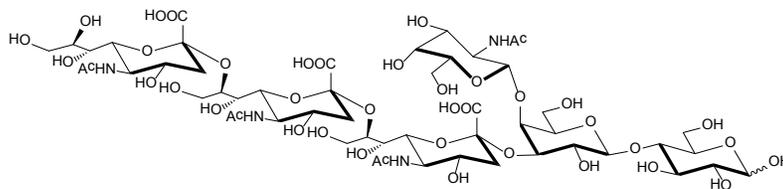
$\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**GT3**

MW 1215.40 Da

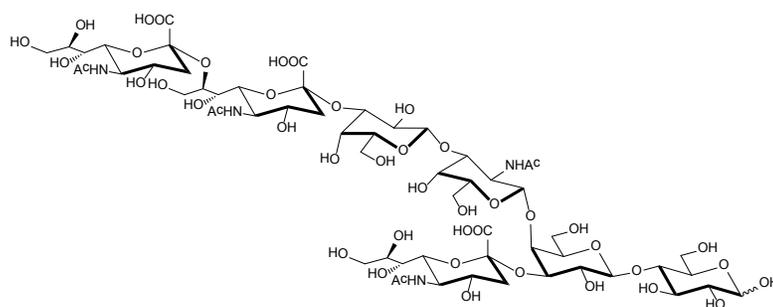
$\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,3)- $\beta$ -D-Gal-(1,4)-D-Glc



**GT2**

MW 1418.48 Da

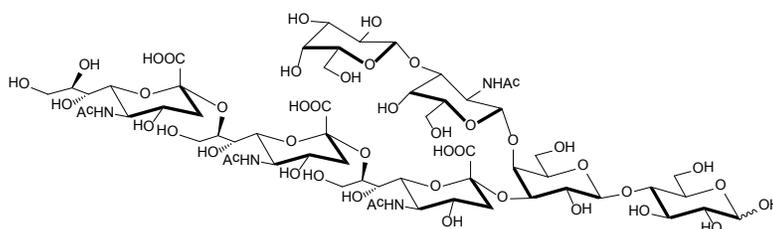
$\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,3)]-  
 $\beta$ -D-Gal-(1,4)-D-Glc



**GT1a**

MW 1580.53 Da

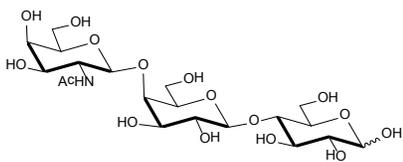
$\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,3)- $\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)-  
[ $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**GT1c**

MW 1580.53 Da

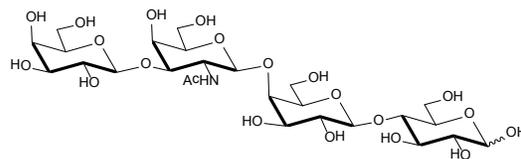
$\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,8)- $\alpha$ -D-Neu5Ac-(2,8)-  
 $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**asialo-GM2**

MW 545.20 Da

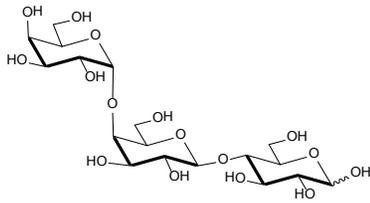
$\beta$ -D-GalNAc-(1,4)- $\beta$ -D-Gal-(1,4)-D-Glc



**asialo-GM1**

MW 707.25 Da

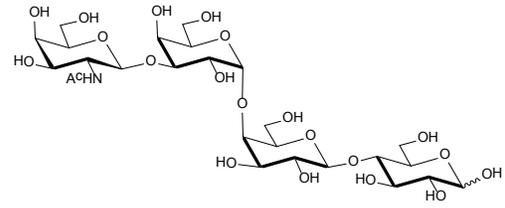
$\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)- $\beta$ -D-Gal-  
(1,4)-D-Glc



**Gb3**

MW 545.20 Da

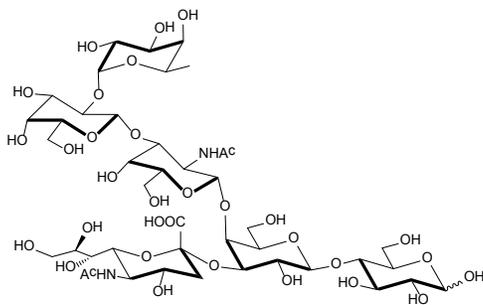
$\alpha$ -D-Gal-(1,4)- $\beta$ -D-Gal-(1,4)-D-Glc



**Gb4**

MW 707.25 Da

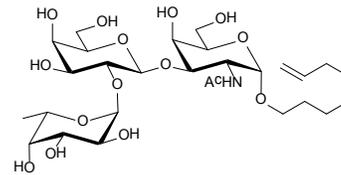
$\beta$ -D-GalNAc-(1,3)- $\alpha$ -D-Gal-(1,4)- $\beta$ -D-Gal-(1,4)-D-Glc



**fucosyl-GM1 (Fuc-GM1)**

MW 1144.40 Da

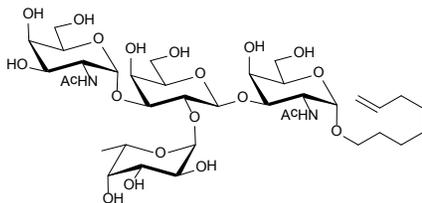
$\alpha$ -L-Fuc-(1,2)- $\beta$ -D-Gal-(1,3)- $\beta$ -D-GalNAc-(1,4)-[ $\alpha$ -D-Neu5Ac-(2,3)]- $\beta$ -D-Gal-(1,4)-D-Glc



**H type 3 trisaccharide (H3)**

MW 639.31 Da

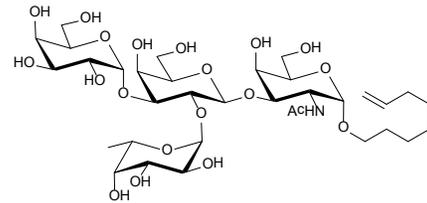
$\alpha$ -L-Fuc-(1,2)- $\beta$ -D-Gal-(1,3)- $\alpha$ -D-GalNAc-O(CH<sub>2</sub>)<sub>6</sub>CH=CH<sub>2</sub>



**A type 3 tetrasaccharide (A3)**

MW 842.39 Da

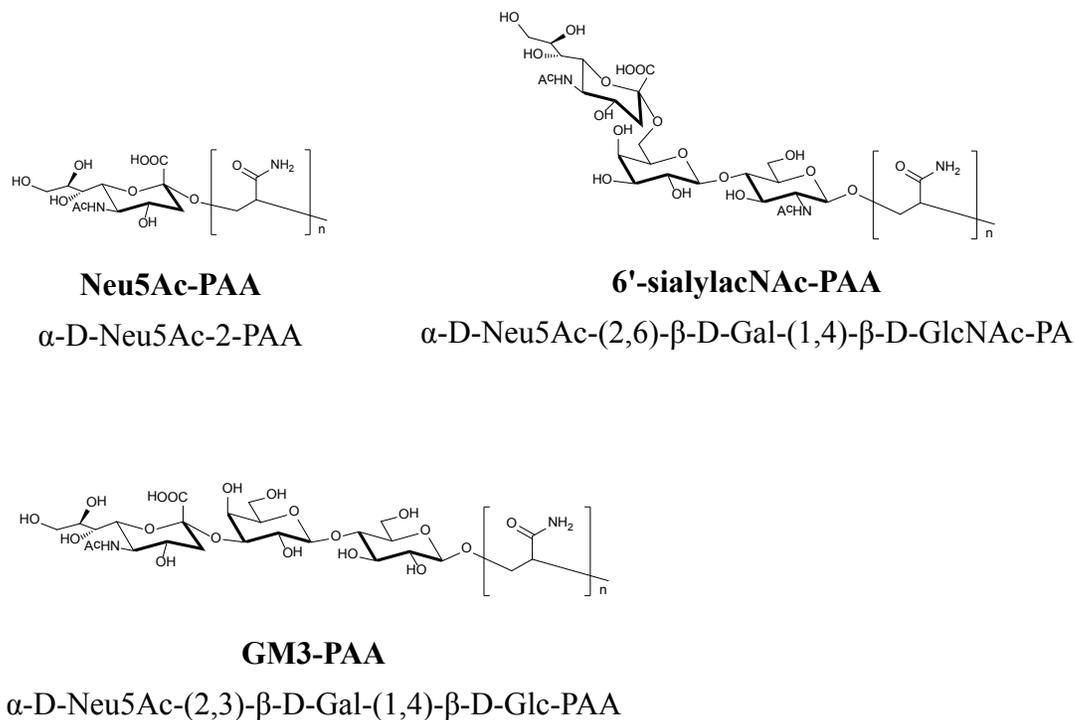
$\alpha$ -D-GalNAc-(1,3)-[ $\alpha$ -L-Fuc-(1,2)]- $\beta$ -D-Gal-(1,3)- $\alpha$ -D-GalNAc-O(CH<sub>2</sub>)<sub>6</sub>CH=CH<sub>2</sub>



**B type 3 tetrasaccharide (B3)**

MW 801.36 Da

$\alpha$ -D-Gal-(1,3)-[ $\alpha$ -L-Fuc-(1,2)]- $\beta$ -D-Gal-(1,3)- $\alpha$ -D-GalNAc-O(CH<sub>2</sub>)<sub>6</sub>CH=CH<sub>2</sub>



**Figure 5.1.** Structures of the twenty-component carbohydrate library consisting of the oligosaccharides of gangliosides, globosides and HBGAs and the structures of polyacrylamide (PAA)-conjugated sialic acid-containing oligosaccharides used for the ELISA binding measurements.

### 5.2.3 Mass spectrometry

All of the ESI-MS assays were carried out on 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. The CaR-ESI-MS and *direct* ESI-MS assays were performed in negative ion mode, whereas the *proxy protein* ESI-MS assay was implemented in positive ion mode. NanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.68

mm i.d.) pulled to  $\sim 5 \mu\text{m}$  using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). A platinum wire was inserted into the nanoESI tip and a capillary voltage was applied to carry out ESI. The source parameters for both negative and positive ion modes were: capillary voltage  $-0.8 \text{ kV}$  (negative ion mode) or  $1.0 \text{ kV}$  (positive ion mode), source temperature  $60 \text{ }^\circ\text{C}$ , cone voltage  $60 \text{ V}$  (negative ion mode) or  $35 \text{ V}$  (positive ion mode), Trap voltage  $5 \text{ V}$ , and Transfer voltage  $2 \text{ V}$ . Data acquisition and processing were performed using MassLynx software (v. 4.1).

#### ***5.2.3.1 Catch-and-Release ESI-MS assay***

The CaR-ESI-MS assay was performed to identify carbohydrate ligands of the NoV VA387 P particle. Ions corresponding to ligand-bound P particle were isolated using the quadrupole mass filter. The quadrupole was set to transmit a broad mass-to-charge-ratio ( $m/z$ ) window (approximately  $200 \text{ } m/z$  units), which allows for the simultaneous passage of free and ligand-bound P particle complexes at a given charge state. Protein-ligand complexes were subjected to collision-induced dissociation (CID) in the Trap region of the Synapt G2S by increasing the Trap voltage from  $5 \text{ V}$  to  $200 \text{ V}$ . Argon ( $1.42 \times 10^{-2} \text{ mbar}$ ) was used to carry out CID in the Trap region. In most instances, the deprotonated ligands released from the complexes could be identified from their MWs. Where required, IMS was used to separate the released isomeric ligands. For IMS separation a wave height of  $35 \text{ V}$  was used and the wave velocity was ramped from  $2000$  to  $500 \text{ m s}^{-1}$ . In all cases a helium flow rate of  $150 \text{ mL min}^{-1}$  and a nitrogen flow rate of  $40 \text{ mL min}^{-1}$  were used. The arrival time distributions (ATDs) for the released ligands were compared to reference ATDs, which were measured for the deprotonated

carbohydrates produced directly from solution.

### 5.2.3.2 Direct ESI-MS assay

The *direct* ESI-MS assay was used to quantify the affinities of the carbohydrate ligands for the NoV P dimers of VA387 and VA115. At least four different initial ligand concentrations were used for each oligosaccharide tested and the binding measurements were carried out in triplicate. A complete description of the data analysis method employed to calculate the intrinsic association constants ( $K_{a,int}$ ) can be found elsewhere.<sup>37,42</sup> Briefly, the abundance ratio ( $R_q$ ) of the ligand-bound protein ( $PL_q$ ), bound to  $q$  molecules of L, to free protein (P) measured by ESI-MS (after correction for nonspecific ligand-protein binding) is taken to be equal to the equilibrium concentration ratio in solution, eq 5.1:

$$R_q = \frac{\sum Ab(PL_q)}{\sum Ab(P)} = \frac{[PL_q]}{[P]} \quad (5.1)$$

Assuming the protein has  $h$  independent and identical binding sites,  $K_{a,int}$  can be expressed by eq 5.2:

$$K_{a,int} = \frac{1}{\left(\frac{1}{f} - 1\right)([L]_0 - h[P]_0 f)} \quad (5.2)$$

where  $[P]_0$  and  $[L]_0$  are the initial concentrations of the protein and ligand, respectively, and  $f$  is the fraction of occupied binding sites, eq 5.3:

$$f = \frac{\sum iR_i}{h(1 + \sum R_i)} \quad (5.3)$$

In the case of the P dimer, which has two equivalent binding sites,  $K_{a,int}$  can be found using eq 4:

$$K_{a,int} = \frac{R_1 + 2R_2}{(R_1 + 2) \left( [L]_0 - \frac{[P]_0 (R_1 + 2R_2)}{1 + R_1 + R_2} \right)} \quad (5.4)$$

Additional details on the derivation of these equations can be found in Chapter 2.

### 5.2.3.3 Proxy protein ESI-MS method

The *proxy protein* ESI-MS assay was used to quantify the affinities of GM3 trisaccharide for NoV VA387 P particle and VLP. A complete description of the data analysis method employed to calculate  $K_{a,int}$  can be found elsewhere.<sup>38</sup> Briefly, a proxy protein ( $P_{proxy}$ ), which binds to L with a known affinity, is used to monitor the extent of L binding to P. Specifically, in the presence of P, the abundance ratio  $R_{proxy}$  ( $= [P_{proxy}L]/[P_{proxy}]$ ) will quantitatively reflect the concentration of L bound to P. If the initial concentrations of target protein ( $[P]_0$ ), proxy protein ( $[P_{proxy}]_0$ ) and ligand ( $[L]_0$ ), as well as the association constant for binding of  $P_{proxy}$  to the ligand ( $K_{a,P_{proxy}}$ ) are known,  $K_{a,int}$  can be evaluated from eq 5.5.

$$K_{a,P,int} = \frac{K_{a,P_{proxy}} / R_{proxy}}{\frac{[P]_{m,0}}{[L]_0 - \frac{[P_{proxy}]_0 R_{proxy}}{R_{proxy} + 1} - \frac{R_{proxy}}{K_{a,P_{proxy}}}} - 1} \quad (5.5)$$

where the initial concentrations of target protein ( $[P]_0$ ), proxy protein ( $[P_{proxy}]_0$ ) and ligand ( $[L]_0$ ), as well as the association constant for binding of  $P_{proxy}$  to the ligand ( $K_{a,P_{proxy}}$ ) are known;  $[P]_{m,0}$  is the initial concentration of binding sites in the target protein, i.e.,  $[P]_{m,0} = h \times [P]_0$ . Detailed descriptions of the proxy protein ESI-MS assay can be found in Chapter 4.

#### **5.2.4 Enzyme-linked immunosorbent assay (ELISA)**

PAA-conjugated Neu5Ac, 6'-sialylacNAc and GM3 trisaccharide were dissolved in 1x PBS (pH 7.4). They were diluted and coated on a 96-well microtiter plate at concentration of 2  $\mu\text{g mL}^{-1}$  and stored at 4 °C overnight. After blocking with 5% non-fat dry milk, NoV VLP, P particle or GST-P domain fusion proteins, as well as GST (negative control), at 50  $\text{ng } \mu\text{L}^{-1}$  were added and incubated for 2 hours at 37 °C. The ligand-bound NoV VLP and P proteins were detected by homemade guinea pig hyperimmune serum against VA387 VLP and VA115 P protein (1:3000), respectively, followed by horseradish peroxidase (HRP)-conjugated goat anti-guinea pig immunoglobulin G (IgG, 1:3000; ICN, Aurora, OH). Bound GST was detected by a homemade GST antibody. The signals were displayed using a TMB kit (Thermo Fisher Scientific, Rockford, IL).

### **5.3 Results and Discussion**

#### **5.3.1 Ganglioside binding to NoV VA387 P particle**

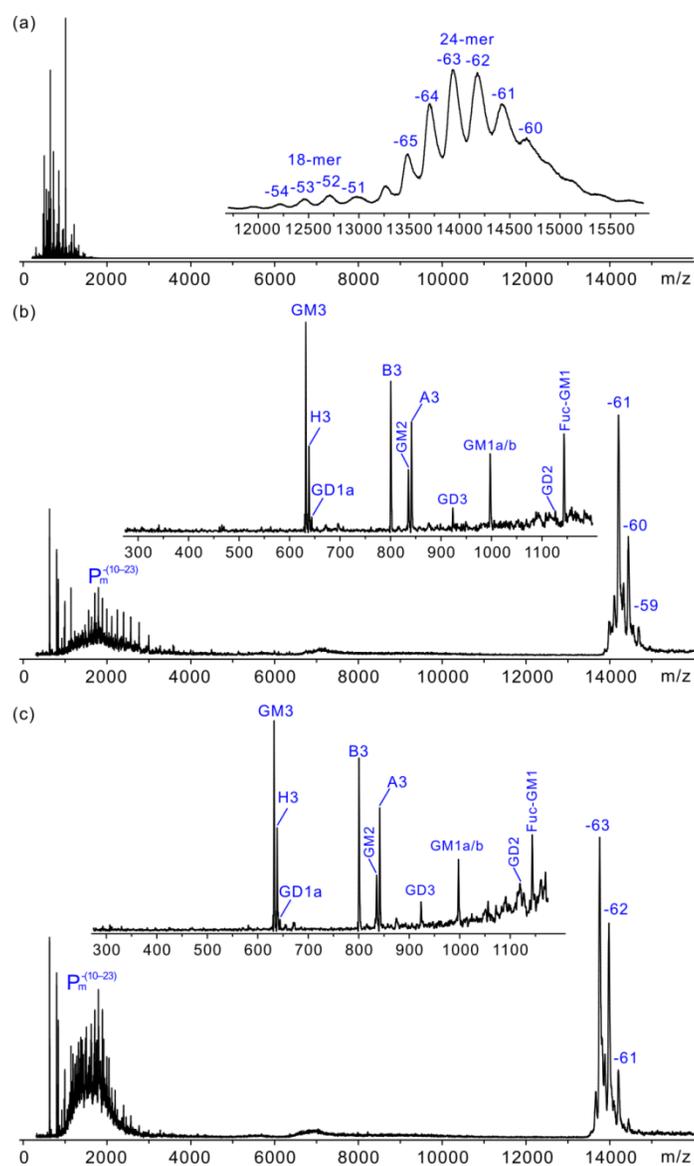
Evidence of ganglioside binding to NoVs was initially revealed through the screening of a small (20 components) carbohydrate library against the P particle (24-mer, *MW* 865,036 Da) of NoV VA387 (GII.4) using the CaR-ESI-MS assay.<sup>36</sup> The library consisted of the oligosaccharides of seventeen glycosphingolipids, GM1a, GM1b, GM2, GM3, GD1a, GD1b, GD2, GD3, GT1a, GT1c, GT2, GT3, fucosyl-GM1 (referred to as Fuc-GM1), asialo GM1, asialo GM2, Gb3 and Gb4, as well as three known HBGA oligosaccharide ligands, H type 3 trisaccharide (referred to as H3), A type 3 tetrasaccharide (A3) and B type 3 tetrasaccharide (B3). The intrinsic affinities of

the HBGA ligands range from 700 to 1500  $M^{-1}$ .<sup>39</sup> The CaR-ESI-MS assay was carried out by first incubating the P particle with the carbohydrate library, followed by direct ESI-MS analysis of the mixture. Because of the high  $MW$  of the P particle, the identity of the bound ligands could not be established directly from the mass spectrum. Instead, using a quadrupole mass filter set to pass a range of mass-to-charge-ratio ( $m/z$ ) ions, all of the ligand-bound P particle ions at a given charge state were isolated and then activated (heated) using CID to release the ligands (as ions) from the complex. Given that carbohydrates have relatively low gas-phase acidities and are able to effectively compete with proteins for negative charge, the CaR-ESI-MS assay was carried out in negative ion mode.<sup>36</sup> Accurate mass analysis, alone or in combination with ion mobility separation (IMS), which separates ions based on size and shape, allowed for positive ligand identification.

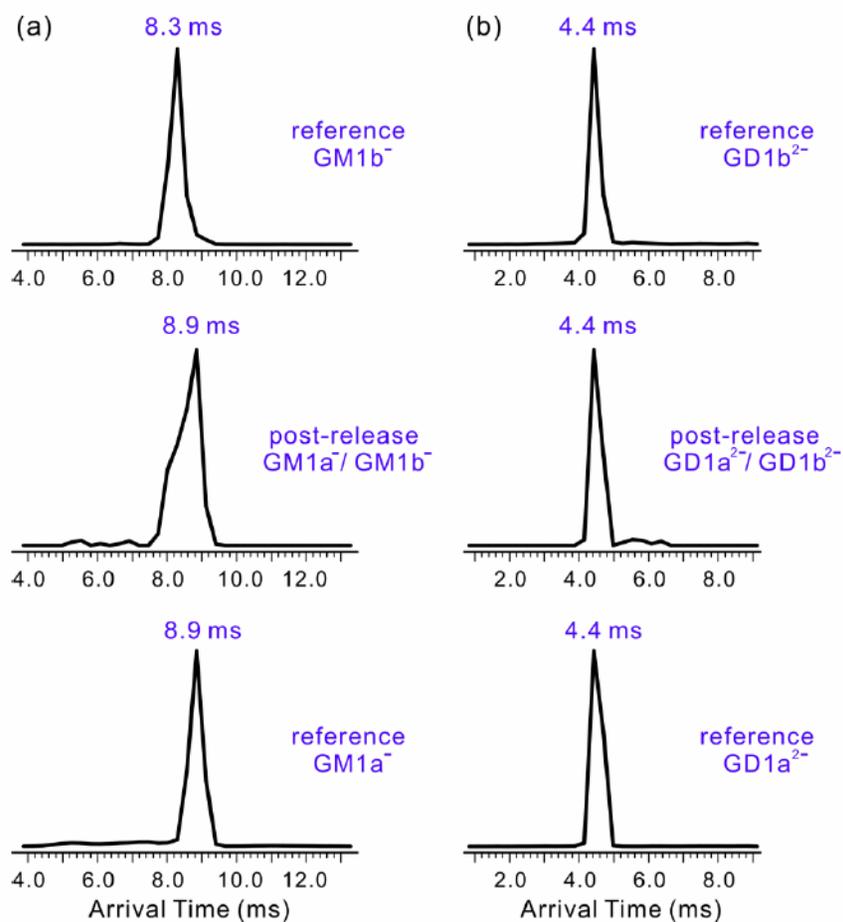
Shown in Figure 5.2a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM, pH 7, 25 °C) solution of P particle (3  $\mu$ M) and the carbohydrate library (10  $\mu$ M each). From the mass spectrum it can be seen that the P particle exists predominately as a 24-mer, with a charge state distribution ranging from -60 to -65. Signal corresponding to an 18-mer is also present, although at lower abundance, with a charge state distribution of -51 to -54. Due to the high  $MW$  of the P particle and the formation of adducts during the ESI process, it was impossible to resolve the ions corresponding to free P particle and its complexes with one or more oligosaccharide ligands. However, CID, performed using a 200  $m/z$  wide isolation window centered at 14,350 to pass ions corresponding to the -61 charge state of the P particle, led to the appearance of singly deprotonated ions of the three HBGA

oligosaccharides, as well as GM3 ( $m/z$  632.2), GM2 ( $m/z$  835.3), GD3 ( $m/z$  923.3), GM1a and/or GM1b ( $m/z$  997.3) and Fuc-GM1 ( $m/z$  1143.4) (Figure 5.2b). Ions corresponding to the singly deprotonated GD2 ( $m/z$  1126.4) and the doubly deprotonated ions of GD1a and/or GD1b ( $m/z$  644.1) were also detected, although at low abundance (Figure 5.2b). Abundant multiply charged protein monomer ions,  $P_m^{n-}$  at  $n = 10 - 23$ , were also evident (Figure 5.2b). Implementation of the CaR-ESI-MS assay using other charge states of the P particle complexes produced similar results (Figure 5.2c).

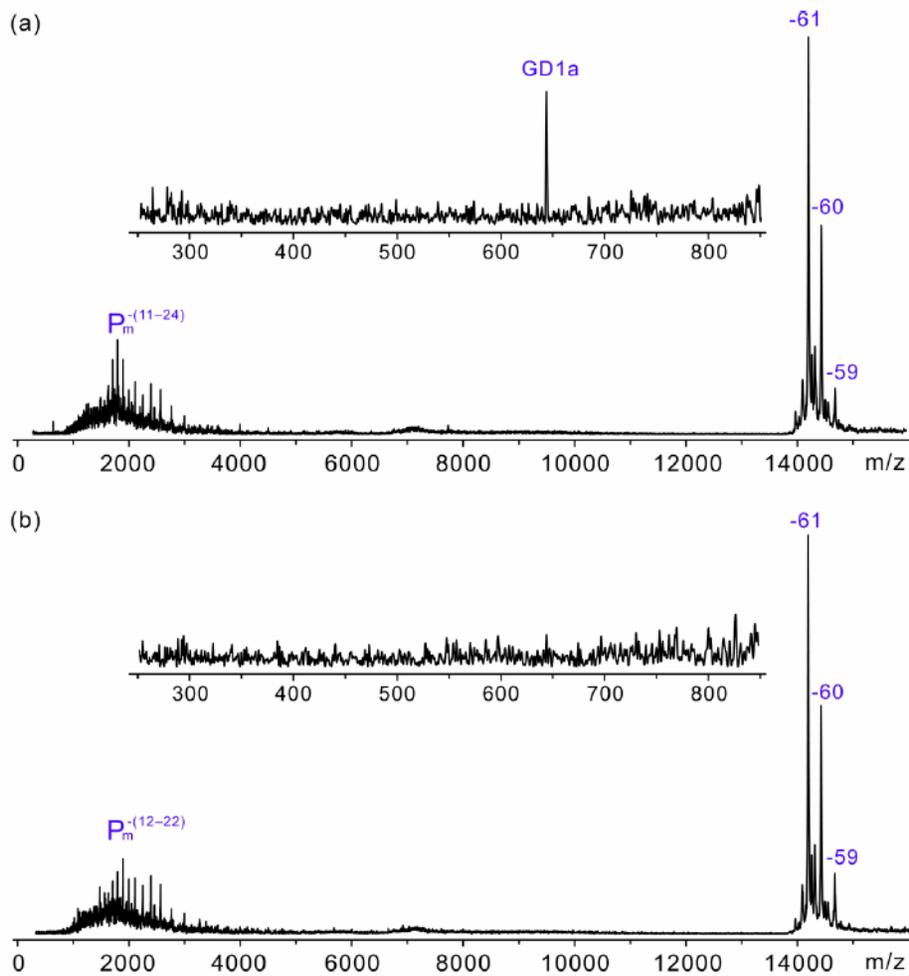
Ion mobility separation of the released ligands revealed evidence that both GM1a and GM1b are released from the P particle, with GM1a being more abundant (Figure 5.3a). The doubly deprotonated ions of GD1a and GD1b could not be differentiated using optimized IMS conditions (Figure 5.3b) and, therefore, it was not possible to establish whether one or both oligosaccharides bind to the P particle directly from these measurements. Instead, the CaR-ESI-MS assay was applied to solutions containing P particle (3  $\mu$ M) and 10  $\mu$ M of GD1a or GD1b. These data revealed that only GD1a binds to the P particle under these solution conditions (Figure 5.4). The CaR-ESI-MS results provide compelling evidence that the P particle of VA387 exhibits a broad specificity for mono- and di-sialylated gangliosides. However, there is a clear preference for GM3 and the addition of saccharides to Gal (e.g. GM1 or GM2) or Sia (e.g. GD3, GD2 or GD1b) decreases binding, compared to GM3. These data, combined with affinities measured for ganglioside oligosaccharides, *vide infra*, suggest that the Sia-Gal-Glc moiety represents the dominant recognition epitope for this NoV.



**Figure 5.2.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of NoV VA387 P particle (3  $\mu$ M) and a twenty-component (10  $\mu$ M each) carbohydrate library consisting of the oligosaccharides of GM1a, GM1b, GM2, GM3, GD1a, GD1b, GD2, GD3, GT1a, GT1c, GT2, GT3, Fuc-GM1, asialo GM1, asialo GM2, Gb3 and Gb4, as well as the H3, B3 and A3 oligosaccharides. (b) and (c) CID mass spectrum measured for the -61 and -63 charge states, respectively, of the free and ligand-bound P particle. A Trap voltage of 200 V was used.



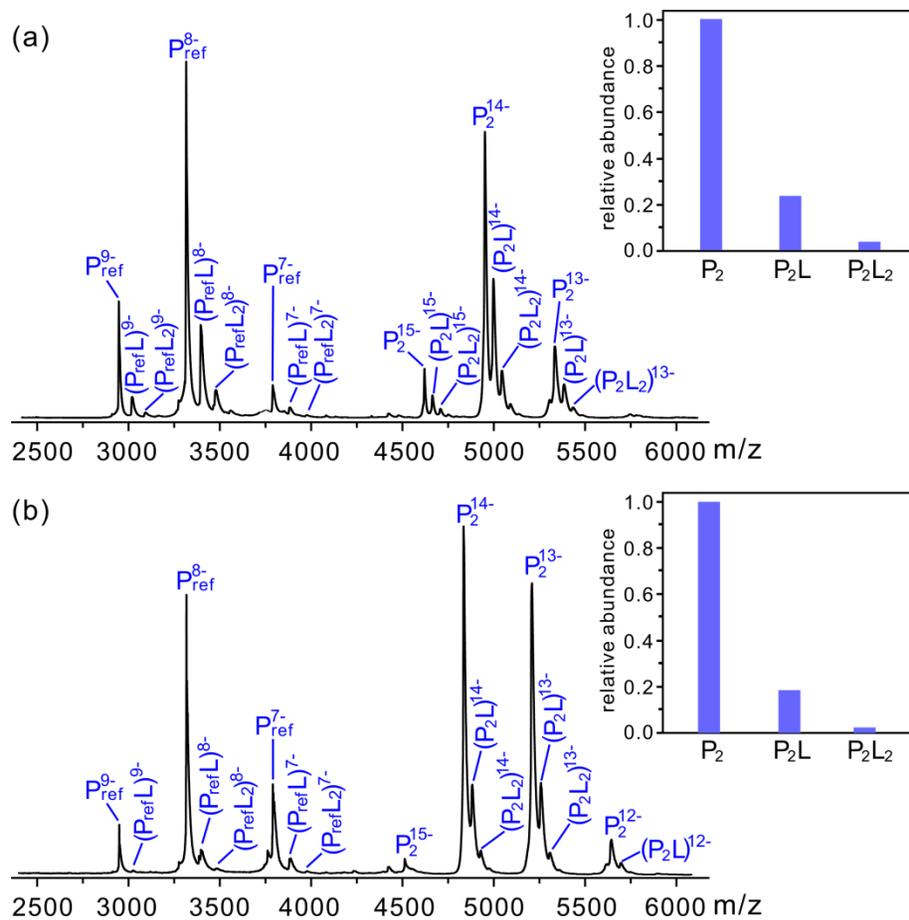
**Figure 5.3.** (a) Arrival time distributions measured for the (a) deprotonated ions of GM1a/GM1b ( $m/z$  997.3) following release from the P particle (post-release) and the deprotonated GM1a and GM1b ions obtained directly from solution (reference) and (b) double deprotonated GD1a/GD1b ions ( $m/z$  644.1) following release from the P particle (post-release) and the GD1a and GD1b ions obtained directly from solution (reference).



**Figure 5.4.** CID mass spectra acquired in negative ion mode for an aqueous ammonium acetate solutions (200 mM, pH 7 and 25 °C) of P particle (3  $\mu$ M) and 10  $\mu$ M of (a) GD1a and (b) GD1b using a broad (200  $m/z$ ) quadrupole isolation window centered at  $m/z$  14,350. A Trap voltage of 200 V was used.

### 5.3.2 Ganglioside affinities for NoV VA387 capsid proteins

Based on the relative abundances of the released oligosaccharide ligands measured by CaR-ESI-MS (Figure 5.2) it would appear that the affinities of the ganglioside ligands are similar to those of the highest affinity HBGA oligosaccharides.<sup>39</sup> However, this conclusion is predicated on the assumption that the release efficiency of the bound-ligands is essentially independent of structure. Because of the presence of the sialic acid, it is possible that gangliosides (which are likely deprotonated in the gaseous complexes) are preferentially released from the P particle due to a lower activation energy resulting from Coulombic repulsion.<sup>43,44</sup> Therefore, it was important to measure the affinities directly. In order to do this, the corresponding P dimer of the NoV was used. The affinities were measured using the *direct* ESI-MS assay, which has been shown to provide reliable  $K_a$  values for many protein-carbohydrate interactions.<sup>37</sup> Affinities were measured for the oligosaccharides of thirteen gangliosides (GM3, GM2, GM1a, GM1b, GD3, GD2, GD1a, GD1b, GT3, GT2, GT1a, GT1c, and Fuc-GM1) for the VA387 P dimer (MW 69,312 Da). A reference protein ( $P_{ref}$ ) was used in all cases to correct the mass spectra for the occurrence of nonspecific carbohydrate-protein interactions during the ESI process.<sup>45,46</sup> A representative ESI mass spectrum acquired for an aqueous ammonium acetate solution (200 mM, pH 7, 25 °C) of VA387 P dimer (12  $\mu$ M) and GM3 trisaccharide (80  $\mu$ M) is shown in Figure 5.5a, as well as the distribution of ligand-bound P dimer after correction for nonspecific binding. From the ESI-MS data,  $K_{a,int}$  values were calculated for each oligosaccharide (Table 5.1). Affinities were also measured for A3, B3 and H3 and shown to agree well with the reported values (Table 5.2).<sup>39</sup>



**Figure 5.5.** ESI mass spectra acquired in negative ion mode for aqueous ammonium acetate solution (200 mM, pH 7 and 25 °C) of (a) NoV VA387 P dimer ( $P_2$ , 12  $\mu\text{M}$ ), GM3 trisaccharide (80  $\mu\text{M}$ ) and  $P_{\text{ref}}$  (4  $\mu\text{M}$ ) and (b) NoV VA115 P dimer ( $P_2$ , 12  $\mu\text{M}$ , MW 67,712 Da), GM3 trisaccharide (80  $\mu\text{M}$ ) and  $P_{\text{ref}}$  (4  $\mu\text{M}$ ). Insets, normalized distribution of GM3 bound to  $P_2$  of (a) VA387 and (b) VA115) after correction for nonspecific ligand binding.

**Table 5.1.** Intrinsic (per binding site) association constants ( $K_{a,int}$ ) for P dimer and the oligosaccharides of thirteen gangliosides measured in aqueous ammonium acetate (200 mM) at pH 7 and 25 °C using the *direct* ESI-MS assay.<sup>a</sup>

L	$K_{a,int}$ ( $M^{-1}$ )	
	P dimer VA387	P dimer VA115
GM3	1500 ± 150	1300 ± 130
GM2	360 ± 90	700 ± 200
GM1a	350 ± 80	400 ± 140
GM1b	180 ± 60	480 ± 110
GD3	340 ± 60	420 ± 140
GD2	150 ± 60	700 ± 140
GD1a	<100	320 ± 90
GD1b	NB <sup>b</sup>	340 ± 50
GT3	NB <sup>b</sup>	310 ± 150
GT2	<100	230 ± 50
GT1a	NB <sup>b</sup>	210 ± 70
GT1c	<100	260 ± 110
fucosyl-GM1	460 ± 150	600 ± 130

a. The reported errors are one standard deviation. b. NB ≡ No binding detected.

**Table 5.2.** Intrinsic association constants ( $K_{a,int}$ ) for the NoV VA387 P dimer and the histo-blood group type 3 oligosaccharides (H3, A3 and B3) in an aqueous ammonium acetate solution (200 mM) at pH 7 and 25 °C measured using the *direct* ESI-MS assay performed in negative ion mode.<sup>a</sup>

L	$K_{a,int}$ ( $M^{-1}$ )	$K_{a,int}$ ( $M^{-1}$ ) <sup>b</sup>
H3	$720 \pm 80$	$650 \pm 65$
A3	$760 \pm 60$	$800 \pm 50$
B3	$1500 \pm 230$	$1500 \pm 150$

a. The reported errors are one standard deviation. b. Values taken from reference 39.

Inspection of the  $K_{a,int}$  values reveals that, of the tested gangliosides, GM3 exhibits the highest affinity for the VA387 P dimer, which is consistent with the results of the CaR ESI-MS measurements, *vide supra*. Moreover, the  $K_{a,int}$  ( $1500 M^{-1}$ ) is identical, within experimental error, to that of B3 ( $1500 \pm 150 M^{-1}$ ).<sup>39</sup> Of the twelve other gangliosides investigated, nine bind weakly ( $K_{a,int} < 500 M^{-1}$ ) and three (GD1b, GT3 and GT1a) do not show any detectable binding. Notably, the quantitative binding data obtained for the P dimer agree qualitatively with the relative affinities inferred from the CaR-ESI-MS measurements performed on the P particle. Moreover, all ligands with affinities  $> 100 M^{-1}$  were detected in the CaR-ESI-MS measurements (Table 5.3).

**Table 5.3.** Summary of results from screening of ganglioside oligosaccharide against the P particle of human NoV VA387 in an aqueous ammonium acetate solution (200 mM) at pH 7 and 25 °C using the CaR-ESI-MS assay.

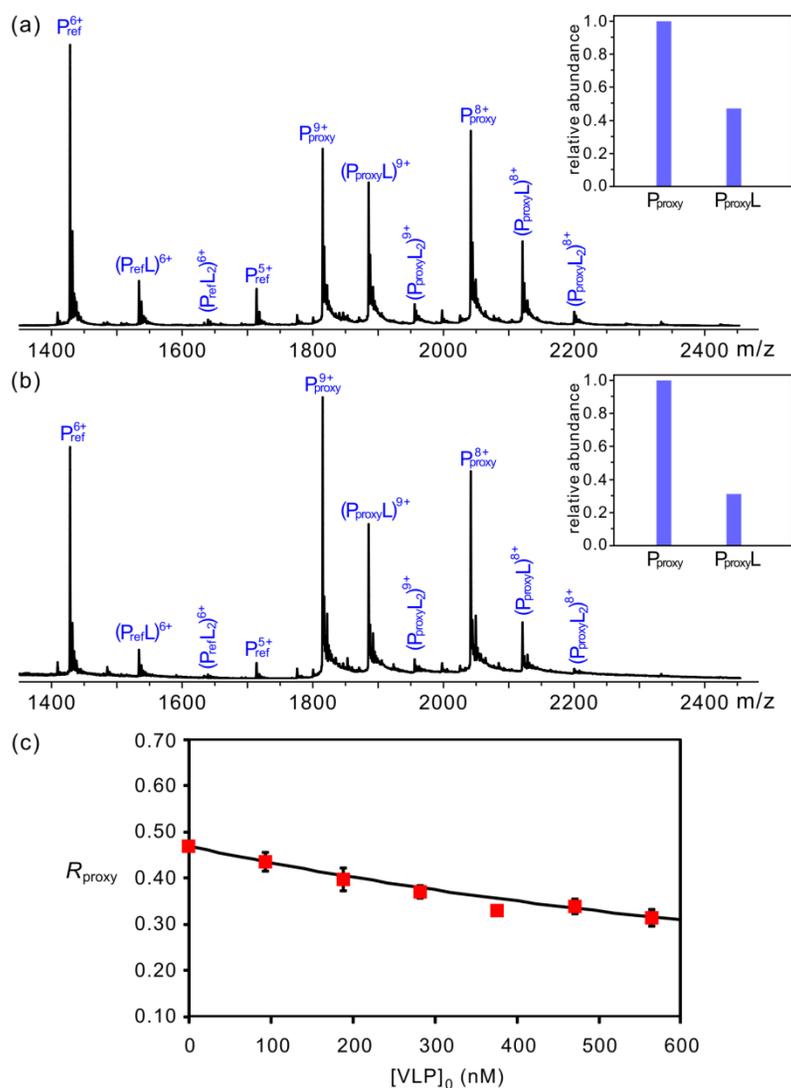
Oligosaccharide	Binder (+) / Non-binder (-)
GM3	+
GM2	+
GM1a	+
GM1b	+
GD3	+
GD2	+
GD1a	-
GD1b	-
GT3	-
GT2	-
GT1a	-
GT1c	-
fucosyl-GM1	+
asialo-GM1	-
asialo-GM2	-

To demonstrate the relevance of the affinity data acquired for the P dimer, affinity measurements were also carried out for GM3 trisaccharide binding to the VA387 P particle and VLP (180-mer,  $MW \sim 10.5$  MDa). An adaptation of the *proxy protein* ESI-MS method, which combines *direct* ESI-MS binding measurements and competitive protein binding, was used to evaluate the affinities.<sup>38</sup> A recombinant fragment of the C-terminus of human galectin-3 (Gal-3C,  $MW$  16,330 Da), which contains a carbohydrate recognition domain and interacts with a  $\beta$ -galactoside moiety,<sup>47,48</sup> served as the proxy protein ( $P_{\text{proxy}}$ ). Importantly, Gal-3C binds to GM3 trisaccharide with an affinity of  $(1.20 \pm 0.02) \times 10^4 \text{ M}^{-1}$ . The extent of binding of GM3 trisaccharide to Gal-3C, as determined by ESI-MS, in the presence of known concentrations of the target protein (P particle or VLP) allowed for a quantitative measure of GM3 binding to the target.

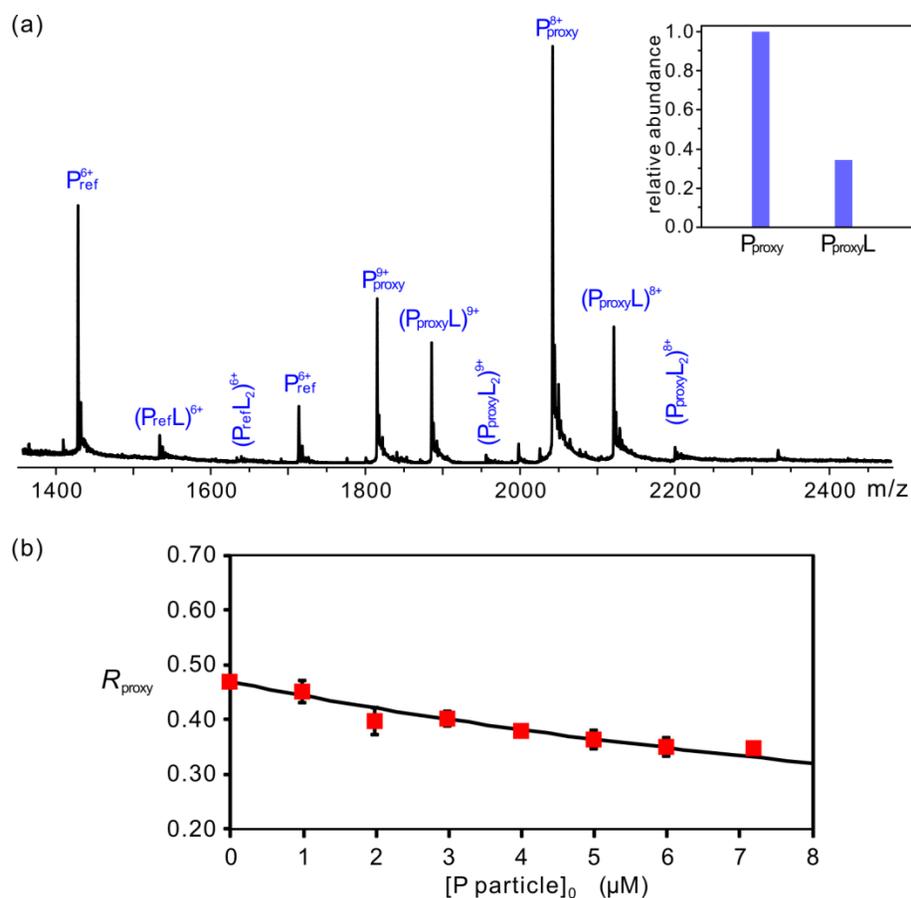
ESI-MS measurements were performed on aqueous ammonium acetate solutions (160 mM, pH 7 and 25 °C) of  $P_{\text{proxy}}$  (3.0  $\mu\text{M}$ ),  $P_{\text{ref}}$  (1.0  $\mu\text{M}$ ), GM3 trisaccharide (40  $\mu\text{M}$ ) and either P particle, at concentrations ranging from 0 to 7.2  $\mu\text{M}$  (corresponds to monomer concentration of 0 – 172.8  $\mu\text{M}$ ), or VLP, at concentrations ranging from 0 to 570 nM (monomer concentration of 0 – 102.6  $\mu\text{M}$ ). Representative ESI mass spectra acquired in positive ion mode in the absence and presence of NoV VLP (570 nM) are shown in Figures 5.6a and 5.6b, respectively. The distributions of ligand-bound  $P_{\text{proxy}}$ , following correction for nonspecific ligand binding, are also given. Inspection of the distributions reveals a measurable decrease in the extent of GM3 trisaccharide binding to Gal-3C upon addition of VLP. This observation confirms that the VLP binds the trisaccharide. The dependence of the extent of GM3 trisaccharide binding to  $P_{\text{proxy}}$  on

VLP concentration is shown in Figure 5.6c. Binding measurements performed on solutions containing P particle yielded qualitatively similar results (Figure 5.7). Analysis of the P<sub>proxy</sub> binding data acquired in the presence of VLP or P particle using the procedure outlined in Experimental Section yields GM3 affinities of  $2600 \pm 200 \text{ M}^{-1}$  and  $5500 \pm 600 \text{ M}^{-1}$  for the P particle and VLP, respectively. The slight differences in the magnitude of the affinities measured for the binding of a common carbohydrate ligand to the P dimer, P particle and VLP of a NoV (the first such data set to be reported), likely reflect subtle differences in the structure of the carbohydrate binding site presented by these related protein complexes.<sup>3</sup> These differences notwithstanding, the present results suggest that the P dimer can serve as a surrogate of the VLP for carbohydrate binding studies.

It has been proposed that NoV VA387 has a binding interface that recognizes HBGAs through the  $\alpha$ -L-Fuc epitope as the major binding interaction and either the  $\alpha$ -D-GalNAc or  $\alpha$ -D-Gal epitope as a minor binding interaction.<sup>5,9,23</sup> However, these core recognition elements are missing in the ganglioside ligands identified in the present study. Therefore, it is of interest to establish whether the ganglioside ligands interact with the NoV through the HBGA binding site or through a distinct ganglioside binding site. It is not possible to answer this question through competitive binding measurements carried out using a ganglioside oligosaccharide (e.g. GM3 trisaccharide) and VA387 P dimer in the presence of varying concentrations of a HBGA oligosaccharide ligand due to the low affinities of these ligands. Instead, future efforts will rely on X-ray crystallography to establish whether VA387 NoV has distinct binding sites for HBGA and ganglioside ligands.



**Figure 5.6.** Representative ESI mass spectra measured in positive ion mode for aqueous ammonium acetate solutions (160 mM, pH 7 and 25 °C) of  $P_{proxy}$  (Gal-3C, 3.0  $\mu$ M),  $P_{ref}$  (Ubq, 1.0  $\mu$ M), GM3 trisaccharide (40  $\mu$ M) without (a) or with (b) NoV VA387 VLP (570 nM, 180-mer). Insets show the normalized distributions of free and GM3-bound  $P_{proxy}$ , after correction for nonspecific ligand binding. (c) Plot of the abundance ratio of GM3-bound  $P_{proxy}$  to free  $P_{proxy}$  ( $R_{proxy}$ ) versus VLP concentration. The solution conditions for each measurement were the same as in (a), but with the addition of VLP. The curve represents the best fit of eq 5.5 to the experimental data.



**Figure 5.7.** (a) Representative ESI mass spectrum measured in positive ion mode for aqueous ammonium acetate solution (160 mM, pH 7 and 25 °C) of  $P_{\text{proxy}}$  (Galectin-3, 3.0 μM),  $P_{\text{ref}}$  (ubiquitin, 1.0 μM), GM3 trisaccharide (40 μM) and NoV P particle (6.0 μM, 24-mer). Inset corresponds to normalized distribution of GM3 bound to  $P_{\text{proxy}}$  after correction for nonspecific ligand binding. (b) Plot of abundance ratio of GM3 trisaccharide-bound  $P_{\text{proxy}}$  to free  $P_{\text{proxy}}$  ( $R_{\text{proxy}}$ ) versus P particle concentration. The solution conditions for each measurement were the same as in (a), but with the addition of P particle. The curve represents the best fit of eq 5.5 to the experimental data.

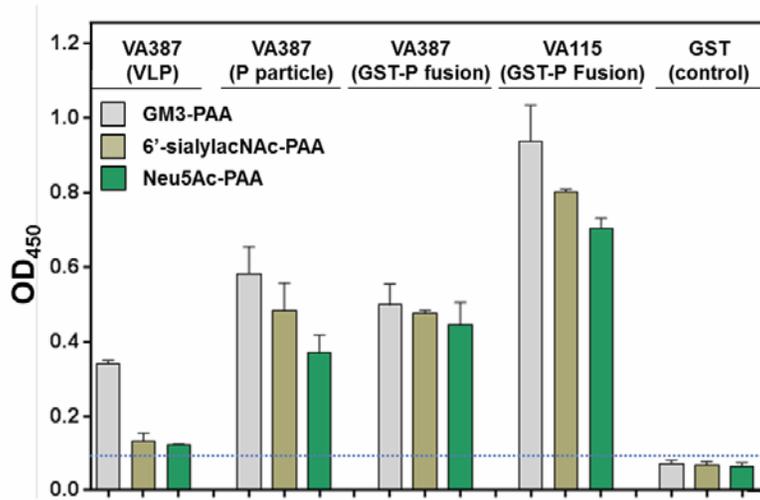
### **5.3.3 Ganglioside affinities for NoV VA115 P dimer**

The aforementioned binding data reveal that NoV VA387 binds to mono- and di-sialylated gangliosides, with affinities comparable to those of the highest affinity HBGA oligosaccharide ligands. To demonstrate that this is not an isolated example of a human NoV that recognizes gangliosides, the affinities of the thirteen ganglioside oligosaccharides for the P dimer of NoV VA115 (GI.3 genotype), which does not bind to human HBGAs,<sup>23</sup> were also measured (Table 5.1). A representative ESI mass spectrum acquired for an aqueous ammonium acetate solution (200 mM, pH 7, 25 °C) of NoV VA115 P dimer (12 μM) and GM3 trisaccharide (80 μM) is shown in Figure 5.5b, as well as the distribution of ligand-bound P dimer after correction for nonspecific binding. Notably, the VA115 P dimer binds to all thirteen oligosaccharides tested and, overall, the affinities are slightly higher than those for VA387. These results suggest that human NoVs generally recognize gangliosides as ligands.

### **5.3.4 Binding of sialic acid-containing glycoconjugates to NoV VA387 and VA115**

Additional evidence for the recognition of sialic acid by human NoVs comes from ELISA measurements carried out on the VA387 VLP, P particle and GST-P fusion protein,<sup>16,17</sup> as well as VA115 GST-P fusion protein, with PAA-conjugated Neu5Ac, 6'-sialylacNAc and GM3 trisaccharide. As shown in Figure 5.8, the two NoV capsid proteins bind all three sialic acid-containing glycoconjugates. It is curious that the VA387 VLP exhibited weaker binding than that of the P particle to the three glycoconjugates and the cause of the weaker binding is, at this time, unknown. Nevertheless, the fact that all three assemblies of NoV capsid protein exhibit a

similar binding pattern to the three glycoconjugates (GM3 > 6'-sialylacNAc > Neu5Ac) validate their applications as models for NoV-ligand interaction. Moreover, comparing the binding of GST-P fusion protein of VA387 to that of VA115 indicates that the sialic acid-containing glycoconjugates have slightly higher affinities for VA115, consistent with the ESI-MS data.



**Figure 5.8.** Binding of NoV VLP, P particles and GST-P fusion protein of VA387, as well as GST-P fusion protein of VA115 to PAA-conjugated GM3 trisaccharide, 6'-sialylacNAc and Neu5Ac in 1x PBS (pH 7.4). GST, which does not show binding to any of the three glycoconjugates, served as a negative control.

These results, together with those from ESI-MS, suggest both  $\alpha$ -(2,3)- and  $\alpha$ -(2,6)- linked sialic acids as critical motifs in VA387 and VA115 binding, similar to what has been reported for MNV1<sup>33</sup> and PSaV.<sup>35</sup> It is important to point out that, although sialic acid-containing oligosaccharides have been identified as receptors for an animal NoV (MNV1)<sup>33,49</sup> and two other animal CVs (FCV and PSaV),<sup>34,35</sup> human NoVs generally recognize gangliosides in addition to

HBGAs. Furthermore, human NoVs differ greatly from MNVs in many other important aspects, including host tropism (human vs. mouse), clinical manifestation (with vs. without diarrhea/vomiting), and pathogenesis.<sup>2</sup>

## 5.4 Conclusions

Taken together, the results of ESI-MS and ELISA measurements performed on two human NoVs representing two different genogroups (GI and GII) provide the first experimental evidence of interactions between human NoVs and gangliosides and sialic acid-containing glycoconjugates. Notably, the affinities measured for the oligosaccharides of the ganglioside ligands by ESI-MS are comparable in magnitude to those reported for the oligosaccharides of known HBGA receptors. These experimental data demonstrate sialic acid-containing oligosaccharides as alternative (to HBGAs) ligands for human NoVs and suggest a new mechanism of human NoV-host interaction, one that involves HBGA and sialic acid-containing oligosaccharide receptors and co-receptors for attachment and penetration into host cells, and opens a new direction in human NoV research. Further studies to characterize the role of cell surface sialic acids/gangliosides in the early stage of viral infection and its potential coordination with HBGAs for viral attachment and/or entry are needed.

## 5.5 References

- (1) Patel, M. M.; Widdowson, M. A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.

- (2) Tan, M.; Jiang, X. *PLoS Pathog.* **2010**, *6*, e1000983.
- (3) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (4) Tan, M.; Jiang, X. *Expert Rev. Mol. Med.* **2007**, *9*, 1.
- (5) Tan, M.; Jiang, X. *Trends Microbiol.* **2011**, *19*, 382.
- (6) Tan, M.; Jiang, X. *Expert Rev. Mol. Med.* **2014**, *16*, e5.
- (7) Bertolotti-Ciarlet, A.; White, L. J.; Chen, R.; Prasad, B. V. V.; Estes, M. K. *J. Virol.* **2002**, *76*, 4044.
- (8) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (9) Cao, S.; Lou, Z.; Tan, M.; Chen, Y.; Liu, Y.; Zhang, Z.; Zhang, X. C.; Jiang, X.; Li, X.; Rao, Z. *J. Virol.* **2007**, *81*, 5949.
- (10) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.
- (11) Hansman, G. S.; Biertumpfel, C.; Georgiev, I.; McLellan, J. S.; Chen, L.; Zhou, T.; Katayama, K.; Kwong, P. D. *J. Virol.* **2011**, *85*, 6687.
- (12) Chen, Y.; Tan, M.; Xia, M.; Hao, N.; Zhang, X. C.; Huang, P.; Jiang, X.; Li, X.; Rao, Z. *PLoS Pathog.* **2011**, *7*, e1002152.
- (13) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.
- (14) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (15) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.

- (16) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.
- (17) Wang, L.; Xia, M.; Huang, P.; Fang, H.; Cao, D.; Meng, X.; McNeal, M.; Jiang, X.; Tan, M. *Biomaterials* **2014**, *35*, 8427.
- (18) Lindesmith, L.; Moe, C.; Marionneau, S.; Ruvoen, N.; Jiang, X.; Lindbland, L.; Stewart, P.; LePendu, J.; Baric, R. *Nat. Med.* **2003**, *9*, 548.
- (19) Hutson, A. M.; Atmar, R. L.; Graham, D. Y.; Estes, M. K. *J. Infect. Dis.* **2002**, *185*, 1335.
- (20) Tan, M.; Jin, M.; Xie, H.; Duan, Z.; Jiang, X.; Fang, Z. *J. Med. Virol.* **2008**, *80*, 1296.
- (21) Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1.
- (22) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendu, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.
- (23) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang X. *J. Virol.* **2005**, *79*, 6714.
- (24) Shanker, S.; Choi, J.-M.; Sankaran, B.; Atmar, R. L.; Estes, M. K.; Prasad, B. V. V. *J. Virol.* **2011**, *85*, 8635.
- (25) Shirato, H.; Ogawa, S.; Ito, H.; Sato, T.; Kameyama, A.; Narimatsu, H.; Zheng, X.; Miyamura, T.; Wakita, T.; Ishii, K.; Takeda, N. *J. Virol.* **2008**, *82*, 10756.
- (26) Lindesmith, L.; Moe, C.; LePendu, J.; Frelinger, J. A.; Treanor, J.; Baric, R. S. *J. Virol.* **2005**, *79*, 2900.

- (27) Murakami, K.; Kurihara, C.; Oka, T.; Shimoike, T.; Fujii, Y.; Takai-Todaka, R.; Park, Y.; Wakita, T.; Matsuda, T.; Hokari, R.; Miura, S.; Katayama, K. *PLoS One* **2013**, *8*, e66534.
- (28) Rydell, G. E.; Dahlin, A. B.; Hook, F.; Larson, G. *Glycobiology* **2009**, *19*, 1176.
- (29) Bally, M.; Rydell, G. E.; Zahn, R.; Nasir, W.; Eggeling, C.; Breimer, M. E.; Svensson, L.; Hook, F.; Larson, G. *Angew. Chem. Int. Ed.* **2012**, *51*, 12020.
- (30) Tamura, M.; Natori, K.; Kobayashi, M.; Miyamura, T.; Takeda, N. *J. Virol.* **2004**, *78*, 3817.
- (31) de Rougemont, A.; Ruvoen-Clouet, N.; Simon, B.; Estienney, M.; Elie-Caille, C.; Aho, S.; Pothier, P.; Le Pendu, J.; Boireau, W.; Belliot, G. *J. Virol.* **2011**, *85*, 4057.
- (32) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (33) Taube, S.; Perry, J. W.; Yetming, K.; Patel, S. P.; Auble, H.; Shu, L. M.; Nawar, H. F.; Lee, C. H.; Connell, T. D.; Shayman, J. A.; Wobus, C. E. *J. Virol.* **2009**, *83*, 4092.
- (34) Stuart, A. D.; Brown, T. D. K. *J. Gen. Virol.* **2007**, *88*, 177.
- (35) Kim, D. S.; Hosmillo, M.; Alfajaro, M. M.; Kim, J. Y.; Park, J. G.; Son, K. Y.; Ryu, E. H.; Sorgeloos, F.; Kwon, H. J.; Park, S. J.; Lee, W. S.; Cho, D.; Kwon, J.; Choi, J. S.; Kang, M. I.; Goodfellow, I.; Cho, K. O. *PLoS Pathog.* **2014**, *10*, e1004172.
- (36) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50.
- (37) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.*

- 2012**, *23*, 431.
- (38) El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 3867.
- (39) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. *Glycobiology* **2013**, *23*, 276.
- (40) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423.
- (41) Meloncelli, P. J.; West, L. J.; Lowary, T. L. *Carbohydr. Res.* **2011**, *346*, 1406.
- (42) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (43) Jiang, Y.; Cole, R. B. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 60.
- (44) Zhang, Y.; Deng, L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1573.
- (45) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (46) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472.
- (47) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E. G.; Yagi, F.; Kasai, K. *Biochim. Biophys. Acta* **2002**, *1572*, 232.
- (48) Collins, P. M.; Bum-Erdene, K.; Yu, X.; Blanchard, H. *J. Mol. Biol.* **2014**, *426*, 1439.
- (49) Taube, S.; Jiang, M.; Wobus, C. E. *Viruses* **2010**, *2*, 1011.

## Chapter 6

### Protein-Glycolipid Interactions Studied *in vitro* using ESI-MS and Nanodiscs.

#### Insights into the Mechanisms and Energetics of Binding\*

### 6.1 Introduction

Glycolipids (GLs) on the surfaces of cells serve a number of important roles. They function as receptors in signaling, pathogen recognition and cellular adhesion processes and convey immunological identity.<sup>1-3</sup> Due to the poor solubility of GL receptors, together with the low affinities that are typical of individual protein-carbohydrate interactions ( $K_a < 10^4 \text{ M}^{-1}$ ),<sup>4,5</sup> the direct quantification of interactions between water-soluble proteins and GL ligands *in vitro* is generally not possible using conventional binding assays, such as isothermal titration calorimetry. Moreover, the structural and functional properties of the receptors may be significantly altered upon removal from a membrane environment.<sup>6,7</sup> Indeed, it is increasingly recognized that protein-GL binding is context dependent (e.g. cell versus model membrane and membrane composition) and is sensitive to GL concentration and fatty acid/ceramide content.<sup>8-10</sup> At present, quantitative binding data are typically obtained using spectroscopy- or microscopy-based measurements and GL that are solubilized by model membranes (e.g. supported lipid bilayer and tethered bilayer lipid membranes and vesicles).<sup>11-15</sup> However, the heterogeneous nature and

---

\* A version of this chapter has been published: Han, L.; Kitova, E. N.; Li, J.; Nikjah, S.; Lin, H.; Pluvinaige, B.; Boraston, A. B.; Klassen, J. S., *Anal. Chem.* **2015**, *87*, 4888-4896.

limited stability of these model membranes make protein-GL interactions difficult to study experimentally and the interpretation of the binding data is not always straightforward.

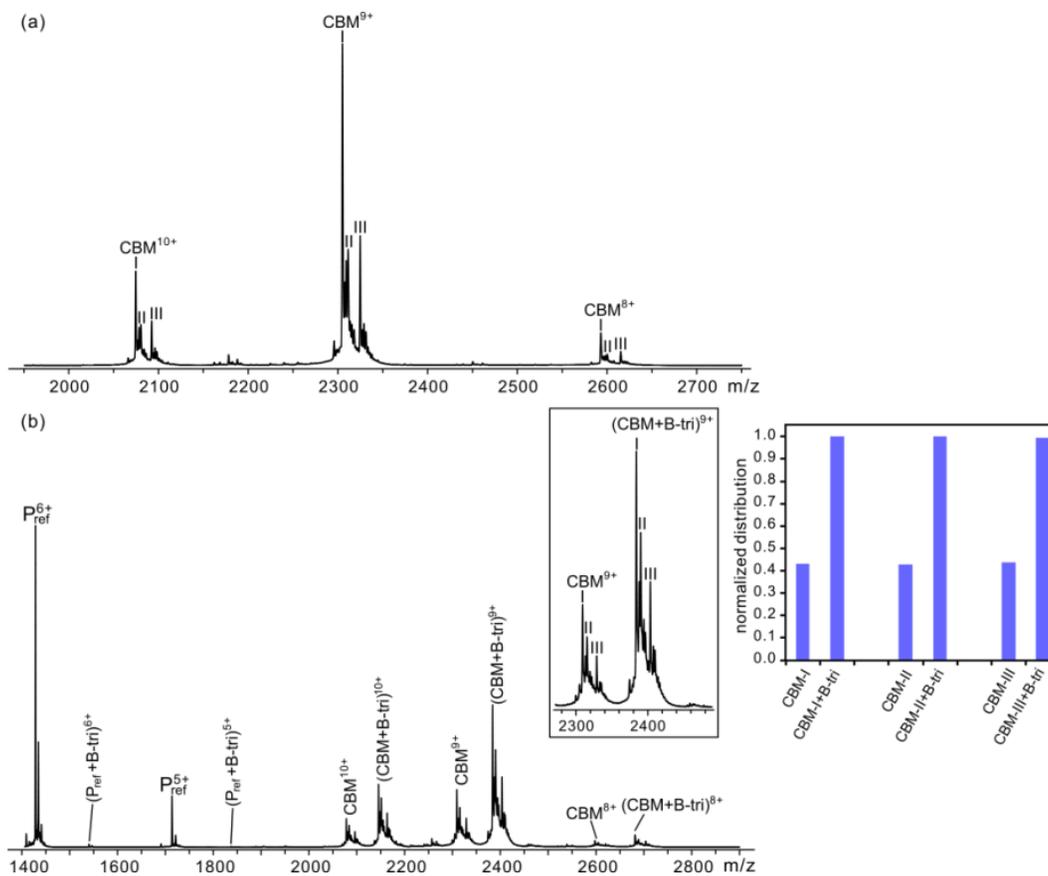
Recently, the use of nanodiscs (NDs), which are water soluble discoidal phospholipid bilayers, has emerged as a promising method for studying protein interactions with GLs in a lipid environment.<sup>16-18</sup> Glycolipids are readily incorporated into NDs allowing their interactions with water-soluble proteins to be investigated in aqueous solutions using a variety of biophysical methods, including surface plasmon resonance (SPR) spectroscopy,<sup>19</sup> electrospray ionization-mass spectrometry (ESI-MS)<sup>20,21</sup> and silicon photonic sensors.<sup>22</sup> However, while it is possible to detect protein binding to GLs in NDs, interpretation of the binding data is challenging owing to a lack of mechanistic insights into the association processes. The goal of the present study was to probe, primarily through the use of ESI-MS measurements, the mechanism(s) of protein binding to GLs contained in NDs and to quantify the thermodynamic stabilities of the resulting protein-GL complexes. The interactions between the cholera toxin B subunit homopentamer (CTB<sub>5</sub>) and its native ganglioside receptor,  $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-ceramide (GM1),<sup>23,24</sup> and between a recombinant family 51 carbohydrate binding module (CBM) originating from *S. pneumoniae*, a gram-positive bacterium responsible for a variety of life-threatening diseases including pneumonia, meningitis, and septicemia,<sup>25</sup> with a synthetic B type 2 neoglycolipid,  $\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-1,2-di-O-dodecyl-sn-glycero (B2<sub>NGL</sub>), served as model protein-GL complexes for this study.

## 6.2 Experimental Section

### 6.2.1 Proteins

Cholera toxin B subunit homopentamer (CTB<sub>5</sub>, molecular weight (MW) 58,040 Da) from *Vibrio cholerae* was purchased from Sigma-Aldrich Canada (Oakville, Canada). A gene fragment encoding a family 51 carbohydrate-binding module (CBM, MW 20,735 Da) was recombinantly produced in *Escherichia coli* and purified as described elsewhere.<sup>26</sup> The ESI-MS analysis of an aqueous solution of CBM (Figure 6.1a) revealed the presence of three isoforms (referred to as CBM-I (MW 20,738 ± 2 Da), CBM-II (MW 20,798 ± 5 Da) and CBM-III (MW 20,916 ± 5 Da)). The origin of the structural heterogeneity is unknown, but the MW of the major form of CBM detected (CBM-I) is consistent with the theoretical value (MW 20,735 Da) obtained from the amino acid sequence (Figure 6.2). Notably, the three CBM forms exhibit similar affinities for A and B blood group oligosaccharides (Figure 6.1b). Bovine ubiquitin (Ubq, MW 8,565 Da) purchased from Sigma-Aldrich Canada (Oakville, Canada) was used as reference protein (P<sub>ref</sub>) for the binding measurements.<sup>27</sup> The recombinant membrane scaffold protein (MSP) MSP1E1 (MW 27,494 Da) used for ND preparation was expressed from the plasmid pMSP1E1 (Addgene, Cambridge, MA) and purified using a reported protocol.<sup>28</sup> Saposin A, used for the preparation of the lipoprotein discs (picodiscs), was a gift from Prof. G. Privé (University of Toronto).<sup>29</sup> Stock solutions of CTB<sub>5</sub> and CBM were concentrated and dialyzed into an aqueous 200 mM ammonium acetate solution (pH 6.8) using Amicon 0.5 mL microconcentrator (Millipore, Billerica, MA) with a MW cutoff of 10 kDa. The concentrations of CTB<sub>5</sub> and CBM stock solutions were determined using a Pierce BCA assay kit (Thermo Scientific, Ottawa, Canada)

following the manufacturer's instructions, whereas the concentration of Ubq, MSP1E1 and saposin A stock solutions were estimated by UV absorption at 280 nm. All the protein stock solutions were stored at  $-80\text{ }^{\circ}\text{C}$  until used.



**Figure 6.1.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25  $^{\circ}\text{C}$  and pH 6.8) of (a) CBM (12  $\mu\text{M}$ ) alone or (b) CBM (12  $\mu\text{M}$ ) with B trisaccharide (B-tri, 40  $\mu\text{M}$ ). Inset shows the normalized distributions of free and B-tri-bound CBM measured for the three isoforms (CBM-I, -II and -III).

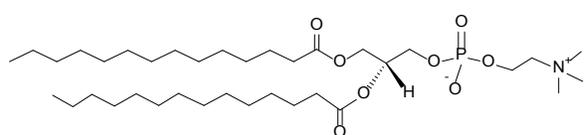
<u>10</u>	<u>20</u>	<u>30</u>	<u>40</u>	<u>50</u>
GSSHHHHHHS	SGLVPRGSHM	ASTYLSMDMW	SSATHGDIDK	TKTVQKDAPF
<u>60</u>	<u>70</u>	<u>80</u>	<u>90</u>	<u>100</u>
TTGNKGEHTK	ISLLTSDDKV	KYFDKGIGTV	ADSPSVISYD	ISGQGFEEKFE
<u>110</u>	<u>120</u>	<u>130</u>	<u>140</u>	<u>150</u>
TYIGIDQSAN	SSRSDHAVVD	RIEIEIDGKV	VYSSSVTNPE	GFRYNTQAQF
<u>160</u>	<u>170</u>	<u>180</u>	<u>190</u>	
ISVTIPQNAK	KISLKSFAGE	HTWGDEVVFA	DAKLIKTVST	

**Figure 6.2.** Amino acid sequence of the recombinant fragment of the family 51 carbohydrate binding module (CBM).

### 6.2.2 Phospholipids, glycolipids and oligosaccharides

1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC, MW 677.9 Da) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC, MW 760.1 Da) were purchased from Avanti Polar Lipids (Alabaster, AL). The ganglioside GM1, purified from bovine brain, was purchased from Axxora LLC (Farmingdale, NY). Two isoforms of GM1, i.e., d18:1-18:0 (MW 1545.9 Da) and d20:1-18:0 (MW 1573.9 Da), were identified in the GM1 sample. Blood group B type 2 tetrasaccharide neoglycolipid (B<sub>2</sub>NGL, MW 1101.7 Da) and A type 2 tetrasaccharide neoglycolipid (A<sub>2</sub>NGL, MW 1142.7 Da) were purchased from Dextra (Reading, UK). The structures of these phospholipids and GLs are shown in Figure 6.3. The GM1 pentasaccharide (GM1<sub>os</sub>, MW 998.34 Da) was purchased from Elicityl SA (Crolles, France). The blood group B trisaccharide (B-tri) was a gift from Prof T. Lowary (University of Alberta). The structures of

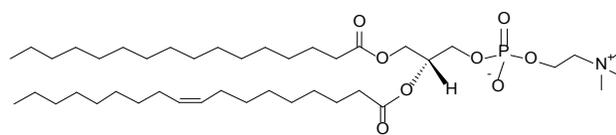
GM1<sub>os</sub> and B-tri are also included in Figure 6.3. DMPC, POPC, GM1, B2<sub>NGL</sub> and A2<sub>NGL</sub> samples were dissolved in HPLC grade methanol/chloroform (1:1 v/v, Thermo Fisher, Ottawa, Canada) to prepare stock solutions of known concentrations. The GM1<sub>os</sub> and B-tri solid samples were weighed and dissolved in ultrafiltered Milli-Q water (Millipore) to yield a stock solution at 1 mM concentration. All the stock solutions were stored at -20 °C until needed.



**DMPC**

MW 677.5 Da

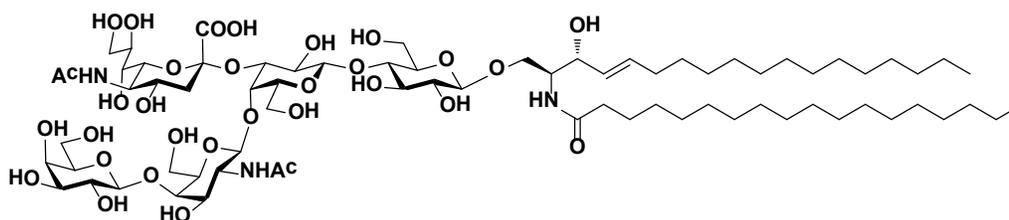
1,2-dimyristoyl-sn-glycero-3-phosphocholine



**POPC**

MW 760.1 Da

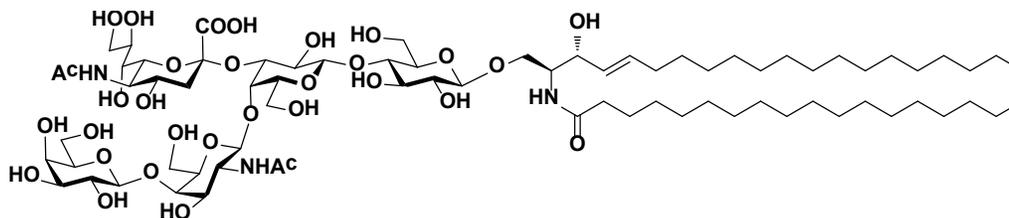
1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine



**GM1** (*d18:1-18:0*)

MW 1545.8 Da

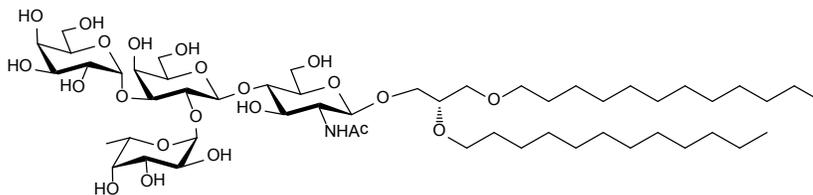
$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-ceramide



**GM1** (*d20:1-18:0*)

MW 1573.9 Da

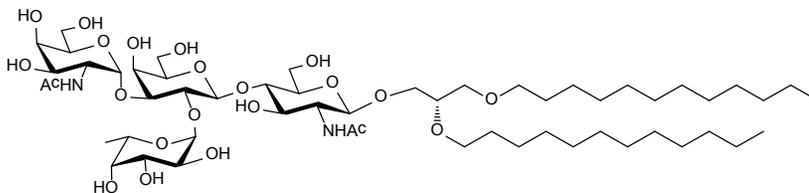
$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-ceramide



**B type 2 neoglycolipid (B2<sub>NGL</sub>)**

MW 1101.7 Da

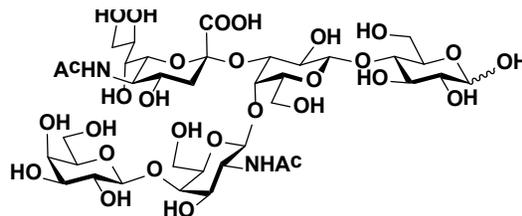
$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-1,2-di-O-dodecyl-sn-glycero



**A type 2 neoglycolipid (A2<sub>NGL</sub>)**

MW 1142.7 Da

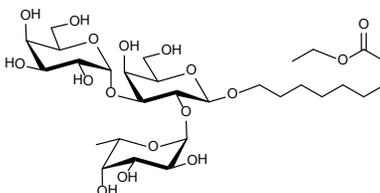
$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-1,2-di-O-dodecyl-sn-glycero



**GM1 pentasaccharide (GM1<sub>os</sub>)**

MW 998.3 Da

$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc



**B trisaccharide (B-tri)**

MW 672.3 Da

$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-O(CH<sub>2</sub>)<sub>8</sub>COOCH<sub>2</sub>CH<sub>3</sub>

**Figure 6.3.** Structures of the phospholipids, glycolipids and oligosaccharides used for the study.

### **6.2.3 Preparation of nanodiscs**

Nanodiscs containing GL (GM1, B2<sub>NGL</sub>, or A2<sub>NGL</sub>) were prepared using a protocol developed by Sligar and coworkers<sup>16,17</sup> and only a brief description is given here. DMPC was mixed with GM1, B2<sub>NGL</sub> or A2<sub>NGL</sub> at the desired ratios. The lipids were dried under a gentle stream of nitrogen overnight at room temperature and re-dissolved in a Tris buffer containing 20 mM sodium cholate (Sigma-Aldrich Canada, Oakville, Canada) at neutral pH. The recombinant membrane scaffold protein MSP1E1 was added to the mixture to yield the an MSP1E1:lipid molar ratio of 1:100. To initiate the ND self-assembly process, an equal volume of pre-washed biobeads (Bio-Rad, Mississauga, Canada) were added and incubated with the mixture for 4 h at room temperature. The supernatant was recovered and then loaded onto the Superdex 200 10/300 size exclusion column (GE-Healthcare Life Sciences, Piscataway, NJ). Finally, the ND fraction was collected, concentrated and dialyzed against 200 mM ammonium acetate (pH 6.8) using an Amicon microconcentrator (Millipore) with a 30 kDa MW cut off . The ND stock solutions were stored at -80 °C before use and the concentration was estimated based on the UV absorption of MSP1E1 at 280 nm. As the nominal molar ratio of MSP to total lipid is 1:100 and each ND possesses two MSPs, the number of GLs per ND is estimated to be two times the percentage of GL.

### **6.2.4 Preparation of picodiscs**

Picodiscs, containing SapA and POPC, alone or with GL, were prepared following a protocol described elsewhere.<sup>29,30</sup> Briefly, GM1 and POPC (dissolved in 1:1 methanol: chloroform) were

mixed in a 1:4 ratio and dried under flowing nitrogen overnight to form a lipid film. The lipid film was re-suspended in 50 mM sodium acetate and 150 mM NaCl (pH 4.8) followed by sonication and thaw cycles to form liposomes. Saposin A protein was then added into the liposomes at 1:10 molar ratio of SapA:(GM1+POPC) to initiate the picodiscs formation and the mixture was incubated at 37 °C for 45 min. Purification of the picodiscs was performed on a Superdex 75 10/300 size exclusion column (GE-Healthcare Life Sciences) equilibrated in 200 mM ammonium acetate (pH 4.8). Finally, picodiscs were concentrated and exchanged into 200 mM ammonium acetate (pH 6.8) and stored at room temperature for a maximum of 1 week. The concentration of SapA in the discs was determined by the UV absorption at 280 nm and the concentration of GM1 was estimated by assuming a 1:1 ratio of GM1:SapA.

### **6.2.5 Mass spectrometry**

All ESI-MS binding measurements were carried out in positive ion mode (unless otherwise indicated) 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. Nanoflow ESI was performed by inserting a platinum wire into a nanoESI tip, which was produced from borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.) pulled to ~5 µm using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). The typical voltage applied to the platinum wire was 1.0 kV. The source conditions for the ESI-MS measurements were: source temperature 60 °C, cone voltage 35 V, Trap voltage 5 V, and Transfer voltage 2 V. For each acquisition at least 60 scans (at 2 s scan<sup>-1</sup>) were measured. Data acquisition and processing were

performed using Waters MassLynx software (version 4.1).

ESI solutions were prepared using 200 mM aqueous ammonium acetate buffer (pH 6.8, 25°C). For the *direct* ESI-MS measurements, solutions of target protein and GL ND were prepared at the desired concentrations. For the *proxy ligand* ESI-MS assays, solutions containing fixed concentrations of target protein and ligand and varying concentrations of GL ND were prepared. All solutions were allowed to equilibrate for 15 min at 25 °C prior to ESI-MS analysis, unless otherwise indicated.

#### **6.2.5.1 Direct ESI-MS assay.**

The *direct* ESI-MS assay was used to measure the extent of ligand (oligosaccharide or GL) binding to CTB<sub>5</sub> and CBM and to quantify the interactions. As described in detail elsewhere,<sup>31</sup> the association constant ( $K_a$ ) for a 1:1 protein-ligand complex can be determined from the abundance ( $Ab$ ) ratio ( $R$ ) of the ligand-bound (PL) to free protein (P) ions measured from ESI-MS, as given by eqs 1.8 and 1.9.

For a protein with  $h$  ligand binding sites, the apparent association constant ( $K_{a,q}$ ) for the addition of a  $q^{\text{th}}$  L to P bound  $(q-1)$  L can be expressed by eq 1.12.<sup>32</sup> Also of interest in the present study was the fraction of occupied ligand binding sites ( $f$ ) in P at a given concentration. A general expression for  $f$ , in terms of abundance or concentration, is given by eq 1.14.

#### **6.2.5.2 Proxy ligand ESI-MS assay.**

It must be stressed that, in the case of P binding to the GL ligands (L) incorporated into NDs, the

$PL_q$  ions detected by ESI-MS are assumed to be associated with NDs in solution and are stripped out of the NDs during the ESI process, *vide infra*,<sup>20,21</sup> while the P ions originate from free P in solution. Differences in ionization efficiencies and others effects, such as incomplete extraction of the  $PL_q$  complexes from the ND or in-source dissociation of the  $PL_q$  ions, could introduce errors to the direct ESI-MS affinity measurements. Consequently, indirect binding measurements were also carried out using the newly developed *proxy ligand* ESI-MS method.

The *proxy ligand* ESI-MS method relies on a proxy ligand ( $L_{\text{proxy}}$ ), which binds to P with known affinity ( $K_{a,\text{proxy}}$ ) and competes with the GL ligand (L). The binding of P to L reduces the concentration of free P in solution, resulting in an increase in the concentration of  $PL_{\text{proxy}}$  complex, relative to P. Consequently, the extent of PL binding in solution can be deduced by monitoring the relative abundance of  $PL_{\text{proxy}}$  by ESI-MS. For the competitive binding of L and  $L_{\text{proxy}}$  to a P possessing a single binding site, the relevant equilibrium expressions are given by eqs 6.1a and 6.1b:

$$K_{a,\text{proxy}} = \frac{[PL_{\text{proxy}}]}{[P][L_{\text{proxy}}]} = \frac{R_{\text{proxy}}}{[L_{\text{proxy}}]} \quad (6.1a)$$

$$K_a = \frac{[PL]}{[P][L]} = \frac{R}{[L]} \quad (6.1b)$$

where  $R_{\text{proxy}}$  corresponds to the abundance ratio of the  $L_{\text{proxy}}$ -bound P ( $PL_{\text{proxy}}$ ) to free P ions, which is taken to be equal to the corresponding concentration ratio in solution, eq 6.2a:

$$R_{\text{proxy}} = \frac{\sum Ab(PL_{\text{proxy}})}{\sum Ab(P)} = \frac{[PL_{\text{proxy}}]}{[P]} \quad (6.2a)$$

and  $R$  is the concentration ratio of L-bound P (PL) to free P in solution, eq 6.2b:

$$R = \frac{[PL]}{[P]} \quad (6.2b)$$

The value of  $R$  can be found from the experimentally determined  $R_{\text{proxy}}$  and the following equations of mass balance, eqs 6.3a – c:

$$[P]_0 = [P] + [PL_{\text{proxy}}] + [PL] \quad (6.3a)$$

$$[L_{\text{proxy}}]_0 = [L_{\text{proxy}}] + [PL_{\text{proxy}}] \quad (6.3b)$$

$$[L]_0 = [L] + [PL] \quad (6.3c)$$

Substituting  $[PL_{\text{proxy}}]$  and  $[PL]$  (in eq 6.3a) with  $R_{\text{proxy}}[P]$  and  $R[P]$ , respectively, gives eq 6.4a:

$$[P] = \frac{[P]_0}{1 + R_{\text{proxy}} + R} \quad (6.4a)$$

It follows that  $[PL_{\text{proxy}}]$  and  $[PL]$  can be expressed as eqs 6.4b and 6.4c, respectively:

$$[PL_{\text{proxy}}] = \frac{R_{\text{proxy}}[P]_0}{1 + R_{\text{proxy}} + R} \quad (6.4b)$$

$$[PL] = \frac{R[P]_0}{1 + R_{\text{proxy}} + R} \quad (6.4c)$$

and  $[L_{\text{proxy}}]$  and  $[L]$  can be expressed as eqs 6.5a and 6.5b, respectively:

$$[L_{\text{proxy}}] = \frac{R_{\text{proxy}}}{K_{a,\text{proxy}}} = [L_{\text{proxy}}]_0 - \frac{R_{\text{proxy}}[P]_0}{1 + R_{\text{proxy}} + R} \quad (6.5a)$$

$$[L] = [L]_0 - \frac{R[P]_0}{1 + R_{\text{proxy}} + R} \quad (6.5b)$$

Rearranging eq 6.5a allows  $R$  to be expressed in terms of  $[P]_0$ ,  $[L]_0$ ,  $R_{\text{proxy}}$  and  $K_{a,\text{proxy}}$ , eq 6.6:

$$R = \frac{R_{\text{proxy}}[P]_0}{[L_{\text{proxy}}]_0 - \frac{R_{\text{proxy}}}{K_{a,\text{proxy}}}} - (R_{\text{proxy}} + 1) \quad (6.6)$$

and  $K_a$  can be calculated from eq 6.7:

$$\begin{aligned}
K_a &= \frac{R}{[L]_0 - \frac{R[P]_0}{1 + R_{\text{proxy}} + R}} = \frac{1}{\frac{[L]_0}{R} - \frac{[P]_0}{1 + R_{\text{proxy}} + R}} \\
&= \frac{1}{\left([L]_{\text{proxy}}]_0 - \frac{R_{\text{proxy}}}{K_{a,\text{proxy}}}\right) \left(\frac{[L]_0}{R_{\text{proxy}} [P]_0 - \left([L]_{\text{proxy}}]_0 - \frac{R_{\text{proxy}}}{K_{a,\text{proxy}}}\right) (R_{\text{proxy}} + 1)} - \frac{1}{R_{\text{proxy}}}\right)} \quad (6.7)
\end{aligned}$$

Where necessary, the *proxy ligand* ESI-MS assay was implemented in conjunction with the reference protein method, which was used to quantitatively correct the mass spectra for the occurrence of nonspecific protein-carbohydrate interactions during the ESI process.<sup>27</sup> This method involves adding a non-interacting reference protein ( $P_{\text{ref}}$ ) to the solution and the extent of nonspecific binding of L to  $P_{\text{ref}}$  was used to subtract the contribution of nonspecific binding of L to P from the mass spectrum. A complete description of the correction method can be found elsewhere.<sup>27</sup>

### 6.2.5.3 Application of proxy ligand ESI-MS method to quantify CTB<sub>5</sub>-GM1 ND interactions.

The binding model (Scheme 6.1), which is an extension of the Homans cooperative binding model established for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>,<sup>33</sup> treats both GM1<sub>os</sub> ( $L_{\text{proxy}}$ ) and GM1 in NDs (L) as monovalent ligands capable of interacting at any of the five binding sites of CTB<sub>5</sub> (P). Based on experimental observations, the model is restricted to the case where, at most, a single GM1 binds but up to five GM1<sub>os</sub> can bind. Notably, the binding GM1<sub>os</sub> and the GM1 are described by three association constants ( $K_{a,\text{proxy},1}$ ,  $K_{a,\text{proxy},2}$ , and  $K_{a,\text{proxy},3}$  and  $K_{a,1}$ ,  $K_{a,2}$  and  $K_{a,3}$ , respectively), which reflect the dependence of the affinity on the number of neighbouring

subunits that are bound to ligand. According to a recent ESI-MS study,<sup>34</sup> the binding of GM1<sub>os</sub> is increased by a factor of 1.7 when one neighbouring binding site is occupied and by a factor of 2.9 when both neighbouring sites are occupied:

$$K_{a,proxy,2} = 1.7K_{a,proxy,1} \quad (6.8a)$$

$$K_{a,proxy,3} = 2.9K_{a,proxy,1} \quad (6.8b)$$

A similar enhancement was assumed for GM1 ND binding:

$$K_{a,2} = 1.7K_{a,1} \quad (6.9a)$$

$$K_{a,3} = 2.9K_{a,1} \quad (6.9b)$$

A summary of all the possible binding interactions, along with the corresponding statistical factors, is given in Scheme 6.1. Based on this model, the equations of mass balance are given by eqs. 6.10 – 6.12:

$$\begin{aligned} [P]_0 = & [P] + [PL_{proxy}] + [P(L_{proxy})_2\alpha] + [P(L_{proxy})_2\beta] + [P(L_{proxy})_3\alpha] + [P(L_{proxy})_3\beta] + [P(L_{proxy})_4] + \\ & [P(L_{proxy})_5] + [PL] + [PL_{proxy}L\alpha] + [PL_{proxy}L\beta] + [P(L_{proxy})_2L\alpha] + [P(L_{proxy})_2L\beta] + [P(L_{proxy})_2L\gamma] + \\ & [P(L_{proxy})_2L\delta] + [P(L_{proxy})_3L\alpha] + [P(L_{proxy})_3L\beta] + [P(L_{proxy})_4L] \end{aligned} \quad (6.10)$$

$$\begin{aligned} [L_{proxy}]_0 = & [L_{proxy}] + [PL_{proxy}] + 2 \times [P(L_{proxy})_2\alpha] + 2 \times [P(L_{proxy})_2\beta] + 3 \times [P(L_{proxy})_3\alpha] + \\ & 3 \times [P(L_{proxy})_3\beta] + 4 \times [P(L_{proxy})_4] + 5 \times [P(L_{proxy})_5] + [PL_{proxy}L\alpha] + [PL_{proxy}L\beta] + 2 \times [P(L_{proxy})_2L\alpha] + \\ & 2 \times [P(L_{proxy})_2L\beta] + 2 \times [P(L_{proxy})_2L\gamma] + 2 \times [P(L_{proxy})_2L\delta] + 3 \times [P(L_{proxy})_3L\alpha] + 3 \times [P(L_{proxy})_3L\beta] + \\ & 4 \times [P(L_{proxy})_4L] \end{aligned} \quad (6.11)$$

$$\begin{aligned} [L]_0 = & [L] + [PL] + [PL_{proxy}L\alpha] + [PL_{proxy}L\beta] + [P(L_{proxy})_2L\alpha] + [P(L_{proxy})_2L\beta] + [P(L_{proxy})_2L\gamma] + \\ & [P(L_{proxy})_2L\delta] + [P(L_{proxy})_3L\alpha] + [P(L_{proxy})_3L\beta] + [P(L_{proxy})_4L] \end{aligned} \quad (6.12)$$

where  $[P]_0$ ,  $[L_{\text{proxy}}]_0$  and  $[L]_0$  are the initial concentrations of P,  $L_{\text{proxy}}$  and L in solution, respectively. The relevant equilibrium expressions for the binding interactions are given by eqs

6.13 – 6.29:

$$\frac{1}{5K_{a,\text{proxy},1}} = \frac{[P][L_{\text{proxy}}]}{[PL_{\text{proxy}}]} \quad (6.13)$$

$$\frac{1}{K_{a,\text{proxy},2}} = \frac{[PL_{\text{proxy}}][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2\alpha]} \quad (6.14)$$

$$\frac{1}{K_{a,\text{proxy},1}} = \frac{[PL_{\text{proxy}}][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2\beta]} \quad (6.15)$$

$$\frac{1}{K_{a,\text{proxy},2}} = \frac{[P(L_{\text{proxy}})_2\alpha][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_3\alpha]} \quad (6.16a)$$

$$\frac{1}{K_{a,\text{proxy},3}} = \frac{[P(L_{\text{proxy}})_2\beta][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_3\alpha]} \quad (6.16b)$$

$$\frac{1}{K_{a,\text{proxy},1}} = \frac{[P(L_{\text{proxy}})_2\alpha][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_3\beta]} \quad (6.17a)$$

$$\frac{1}{K_{a,\text{proxy},2}} = \frac{[P(L_{\text{proxy}})_2\beta][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_3\beta]} \quad (6.17b)$$

$$\frac{1}{K_{a,\text{proxy},2}} = \frac{[P(L_{\text{proxy}})_3\alpha][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_4]} \quad (6.18a)$$

$$\frac{1}{K_{a,\text{proxy},3}} = \frac{[P(L_{\text{proxy}})_3\beta][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_4]} \quad (6.18b)$$

$$\frac{5}{K_{a,\text{proxy},3}} = \frac{[P(L_{\text{proxy}})_4][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_5]} \quad (6.19)$$

$$\frac{1}{5K_{a,1}} = \frac{[P][L]}{[PL]} \quad (6.20)$$

$$\frac{1}{2K_{a,2}} = \frac{[PL_{\text{proxy}}][L]}{[PL_{\text{proxy}}L\alpha]} \quad (6.21a)$$

$$\frac{1}{2K_{a,\text{proxy},2}} = \frac{[PL][L_{\text{proxy}}]}{[PL_{\text{proxy}}L\alpha]} \quad (6.21b)$$

$$\frac{1}{2K_{a,1}} = \frac{[PL_{\text{proxy}}][L]}{[PL_{\text{proxy}}L\beta]} \quad (6.22a)$$

$$\frac{1}{2K_{a,\text{proxy},1}} = \frac{[PL][L_{\text{proxy}}]}{[PL_{\text{proxy}}L\beta]} \quad (6.22b)$$

$$\frac{1}{K_{a,1}} = \frac{[P(L_{\text{proxy}})_2\alpha][L]}{[P(L_{\text{proxy}})_2L\alpha]} \quad (6.23a)$$

$$\frac{2}{K_{a,\text{proxy},2}} = \frac{[PL_{\text{proxy}}L\beta][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2L\alpha]} \quad (6.23b)$$

$$\frac{1}{2K_{a,2}} = \frac{[P(L_{\text{proxy}})_2\alpha][L]}{[P(L_{\text{proxy}})_2L\beta]} \quad (6.24a)$$

$$\frac{1}{K_{a,\text{proxy},3}} = \frac{[PL_{\text{proxy}}L\beta][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2L\beta]} \quad (6.24b)$$

$$\frac{1}{K_{a,\text{proxy},2}} = \frac{[PL_{\text{proxy}}L\alpha][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2L\beta]} \quad (6.24c)$$

$$\frac{1}{2K_{a,2}} = \frac{[P(L_{\text{proxy}})_2\beta][L]}{[P(L_{\text{proxy}})_2L\gamma]} \quad (6.25a)$$

$$\frac{1}{K_{a,\text{proxy},2}} = \frac{[PL_{\text{proxy}}L\beta][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2L\gamma]} \quad (6.25b)$$

$$\frac{1}{K_{a,\text{proxy},1}} = \frac{[PL_{\text{proxy}}L\alpha][L_{\text{proxy}}]}{[P(L_{\text{proxy}})_2L\gamma]} \quad (6.25c)$$

$$\frac{1}{K_{a,3}} = \frac{[P(L_{\text{proxy}})_2\beta][L]}{[P(L_{\text{proxy}})_2L\delta]} \quad (6.26a)$$

$$\frac{2}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2 L \alpha][L_{proxy}]}{[P(L_{proxy})_2 L \delta]} \quad (6.26b)$$

$$\frac{1}{2K_{a,2}} = \frac{[P(L_{proxy})_3 \alpha][L]}{[P(L_{proxy})_3 L \alpha]} \quad (6.27a)$$

$$\frac{1}{2K_{a,proxy,3}} = \frac{[P(L_{proxy})_2 L \alpha][L_{proxy}]}{[P(L_{proxy})_3 L \alpha]} \quad (6.27b)$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2 L \beta][L_{proxy}]}{[P(L_{proxy})_3 L \alpha]} \quad (6.27c)$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_2 L \gamma][L_{proxy}]}{[P(L_{proxy})_3 L \alpha]} \quad (6.27d)$$

$$\frac{1}{2K_{a,3}} = \frac{[P(L_{proxy})_3 \beta][L]}{[P(L_{proxy})_3 L \beta]} \quad (6.28a)$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2 L \beta][L_{proxy}]}{[P(L_{proxy})_3 L \beta]} \quad (6.28b)$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_2 L \gamma][L_{proxy}]}{[P(L_{proxy})_3 L \beta]} \quad (6.28c)$$

$$\frac{1}{2K_{a,proxy,2}} = \frac{[P(L_{proxy})_2 L \delta][L_{proxy}]}{[P(L_{proxy})_3 L \beta]} \quad (6.28d)$$

$$\frac{1}{K_{a,3}} = \frac{[P(L_{proxy})_4][L]}{[P(L_{proxy})_4 L]} \quad (6.29a)$$

$$\frac{2}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_3 L \alpha][L_{proxy}]}{[P(L_{proxy})_4 L]} \quad (6.29b)$$

$$\frac{2}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_3 L \beta][L_{proxy}]}{[P(L_{proxy})_4 L]} \quad (6.29c)$$

Given the initial concentrations ( $[P]_0$ ,  $[L_{proxy}]_0$  and  $[L]_0$ ) and  $K_{a,proxy,1}$ ,  $K_{a,proxy,2}$ , and  $K_{a,proxy,3}$  values, the equilibrium concentrations of all CTB<sub>5</sub> species can be calculated for a given set of

$K_{a,1}$ ,  $K_{a,2}$ , and  $K_{a,3}$  values. The theoretical distribution can then be compared to the experimental distribution. Because of differences in the ESI-MS response factors for the (CTB<sub>5</sub> +  $q$ GM1<sub>os</sub>) and (CTB<sub>5</sub> +  $q$ GM1<sub>os</sub> + GM1) complexes, the  $K_{a,1}$ ,  $K_{a,2}$ , and  $K_{a,3}$  values were found by considering only the abundance ratio  $Ab(P(L_{proxy})_5)/Ab(P(L_{proxy})_4)$  ( $\equiv R_{proxy,5}$ ). Optimum  $K_{a,1}$ ,  $K_{a,2}$ , and  $K_{a,3}$  values were found using a least square analysis, where the sum of squares of residuals ( $SSR$ ) between the experimental to theoretical  $R_{proxy,5}$  values were minimized, eq 6.30:

$$SSR = \sum_n \left( R_{proxy,5}(\text{experimental}) - R_{proxy,5}(\text{theoretical}) \right)^2 \quad (6.30)$$

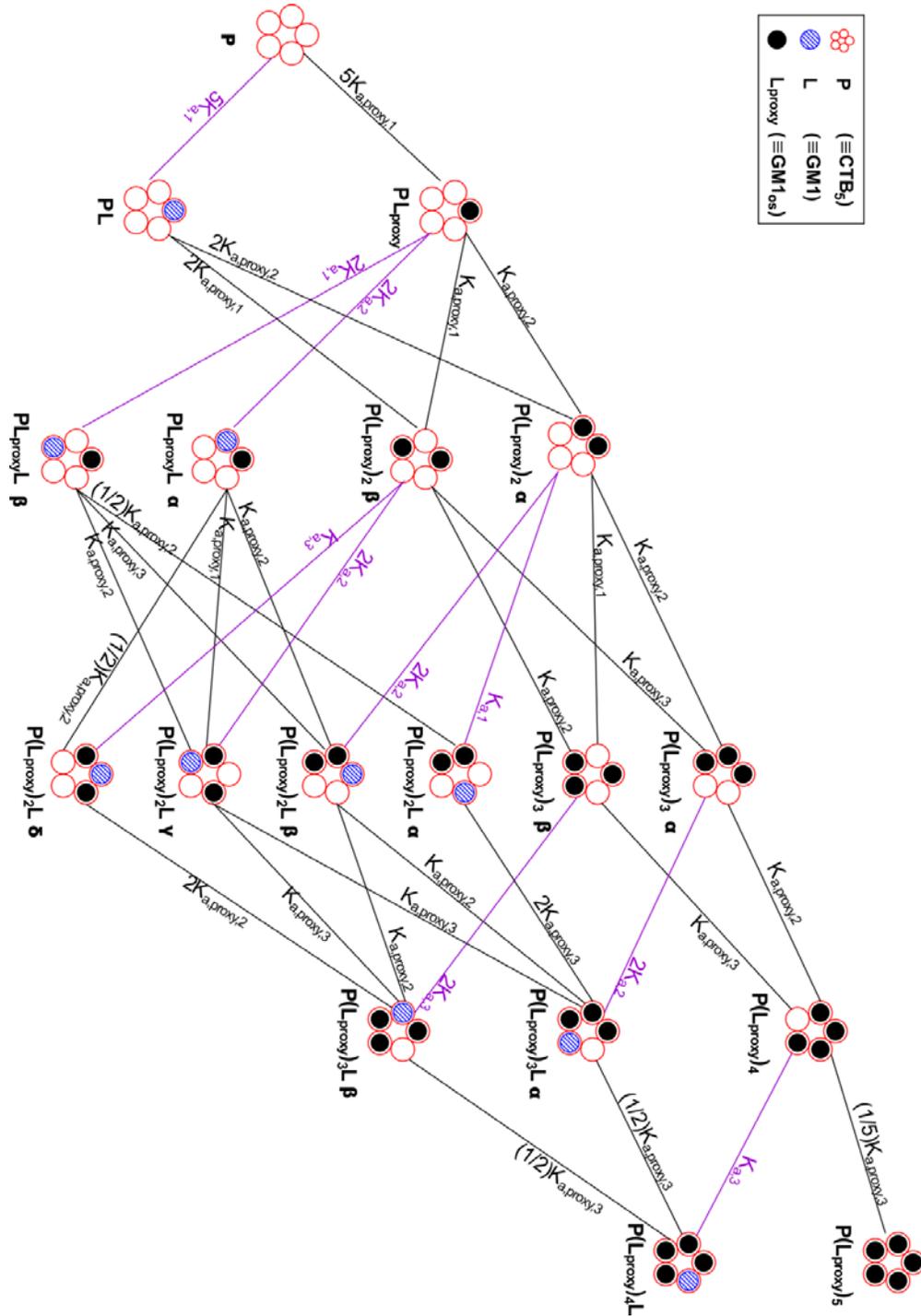
in which  $n$  is the total number of data points.

### 6.2.6 Ultracentrifugation and SDS-PAGE

Ultracentrifugation and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were used to analyze the species present in solutions containing proteins and GL NDs. Briefly, CTB<sub>5</sub> and GM1 ND or CBM and B2<sub>NGL</sub> ND were incubated in a 200 mM ammonium acetate solution (pH 6.8, 25 °C) and placed in a microconcentrator (Millipore) with a MW cutoff of 100 kDa and subjected to ultracentrifugation three times. Each time, 200 mM ammonium acetate buffer was added to the concentrated supernatant solution to maintain the same initial volume. Proteins and protein-ligand complexes with MW  $\geq$ 100 kDa remained in the supernatant while those with MW <100 kDa passed through the membrane to the filtrate. The supernatant and filtrate were further analyzed by ESI-MS and sodium dodecyl sulfate polyacrylamide (15%) gel electrophoresis (SDS-PAGE). To carry out SDS-PAGE, solutions were diluted with an equal volume of 2 $\times$  loading buffer (125 mM TrisHCl pH 6.8, 4% (w/v) SDS, 0.01% (w/v)

bromophenol blue, 20% glycerol and 200 mM dithiothreitol). The solutions were preheated to ~90 °C for 5 min and then allowed to cool to room temperature prior to loading the samples.

Coomassie stain was used to visualize proteins on the gel.



**Scheme 6.1.** Graphical representation of possible protein-ligand interactions involving a monovalent ligand (L), present in a ND, and a monovalent proxy ligand ( $L_{\text{proxy}}$ ) that binding competitively with a protein (P) with five binding sites. Based on the experimental data described in Figures 6.22b and 6.23b, *vide infra*, only one L is bound to P under the experimental conditions used in this study. The binding model is based on the cooperative model for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub> proposed by Homans,<sup>33</sup> wherein ligand affinity is enhanced when a neighbouring binding site is occupied. The  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  labels are used to distinguish the positional isomers. The intrinsic association constants ( $K_{a,\text{proxy},1}$ ,  $K_{a,\text{proxy},2}$ , and  $K_{a,\text{proxy},3}$  and  $K_{a,1}$ ,  $K_{a,2}$  and  $K_{a,3}$ ), along with the statistic coefficients, are given for each interaction.

## 6.3 Results and Discussion

### 6.3.1 CTB<sub>5</sub> binding to GM1 nanodiscs

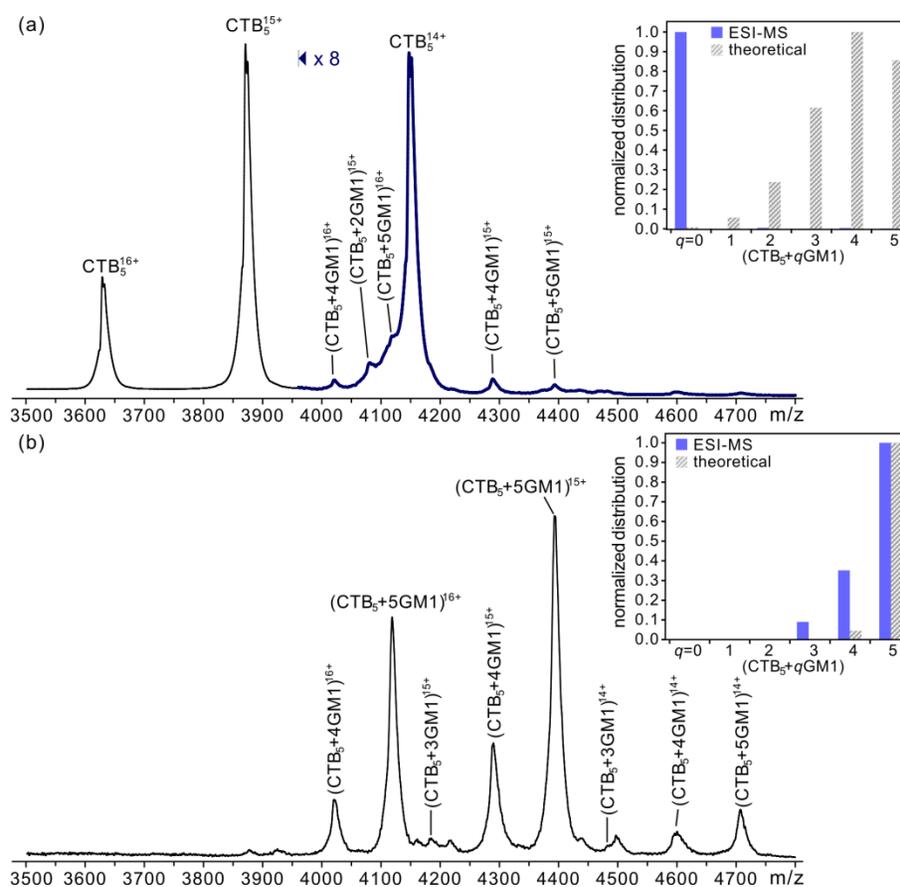
The binding of GM1 to CTB<sub>5</sub> is one of the most extensively studied protein-glycosphingolipid interactions. CTB<sub>5</sub> can bind up to five molecules of GM1 and, according to crystal structures reported for the complex of CTB<sub>5</sub> with the water-soluble GM1 pentasaccharide (GM1<sub>os</sub>), the  $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc and  $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3) motifs in each GM1<sub>os</sub> interact primarily with a single B subunit of CTB<sub>5</sub> through eighteen direct or water mediated H-bonds.<sup>24</sup> The stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub> at neutral pH exhibits positive cooperativity, with intrinsic (per binding site)  $K_a$  values ranging from  $10^6$  to  $10^7$  M<sup>-1</sup>.<sup>33,34</sup> The CTB<sub>5</sub>-GM1 interaction serves as a useful model system for probing various aspects of protein binding to GLs incorporated into NDs. The measured distribution of GM1 bound to the five available CTB<sub>5</sub> binding sites can

provide insights into the nature of the binding processes, such as the reversibility of the individual protein-GL interactions and, relatedly, the ability of CTB<sub>5</sub> to sample GM1 ligands from multiple NDs, as well as the diffusion of GM1 both within and between NDs. Moreover, because of the relatively high affinity of the interactions, the extent of GM1 binding can be used to quantify the concentration of available GM1 and, consequently, establish the efficiency of incorporation of GM1 into NDs.

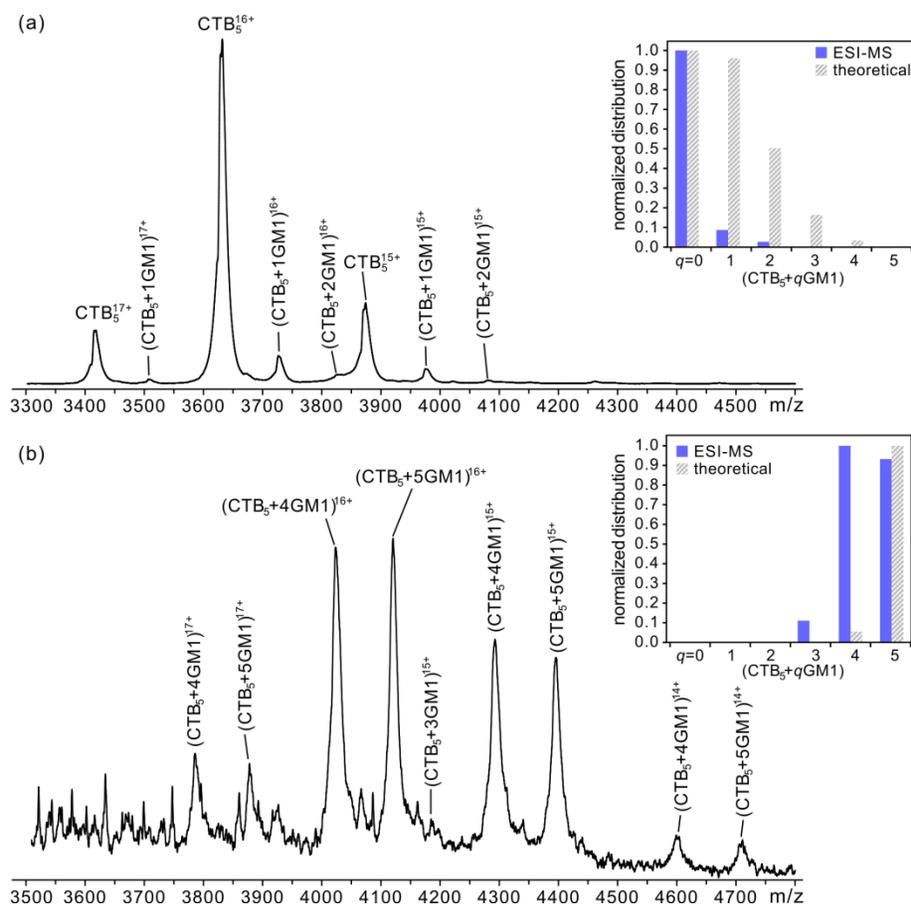
### **6.3.1.1 CTB<sub>5</sub>-GM1 nanodisc interactions revealed by ESI-MS.**

ESI-MS binding measurements were performed on solutions of CTB<sub>5</sub> and ND containing GM1 at percentages ranging from 0.5% to 10%; the corresponding average number of GM1 molecules per ND was estimated to be 1 (0.5%) to 20 (10%). Shown in Figure 6.4 are illustrative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, pH 6.8, 25 °C) containing CTB<sub>5</sub> (3 μM) with 0.6 μM and 1.4 μM 10% GM1 ND. Inspection of the mass spectra reveals signal corresponding to the protonated ions of free and GM1-bound CTB<sub>5</sub>, i.e.,  $(\text{CTB}_5 + q\text{GM1})^{n+}$  with  $q = 0 - 5$  at  $n = 14 - 17$ . Also shown in Figures 6.4 are the normalized distributions of  $(\text{CTB}_5 + q\text{GM1})$  species calculated from the corresponding mass spectra. Illustrative ESI mass spectra and distributions of bound GM1 measured for the 0.5%, 1%, 2.5% and 5% GM1 NDs are given in Figures 6.5 – 6.8. According to the ESI-MS data, the number of GM1 ligands bound to CTB<sub>5</sub> is sensitive to both the ND concentration, as well as the percentage of GM1 in the ND. For example, at low concentrations (e.g. 0.6 μM) of the 10% GM1 ND, CTB<sub>5</sub> exists predominantly as free protein with trace amounts of CTB<sub>5</sub> bound to between two

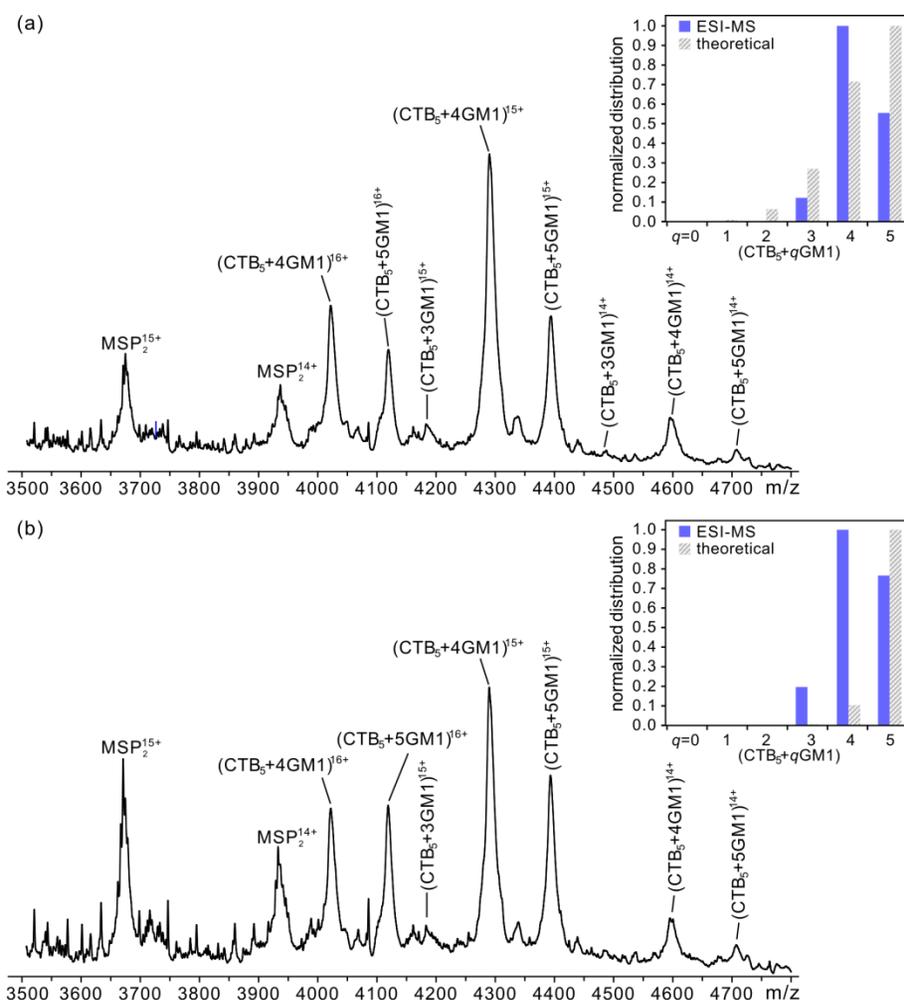
and five GM1 (Figure 6.4a), whereas at higher concentrations (e.g. 1.4  $\mu\text{M}$ ), CTB<sub>5</sub> is bound predominantly to four and five GM1 (Figure 6.4b). Similarly, at low concentrations (e.g. 3  $\mu\text{M}$ ) for the 0.5% GM1 ND, the unbound form CTB<sub>5</sub> is the most abundant species (Figure 6.5a); at higher concentrations of ND (e.g. 24  $\mu\text{M}$ ), the distribution shifts to higher ligand occupancy, with the majority of CTB<sub>5</sub> bound to four GM1 (Figure 6.5b).



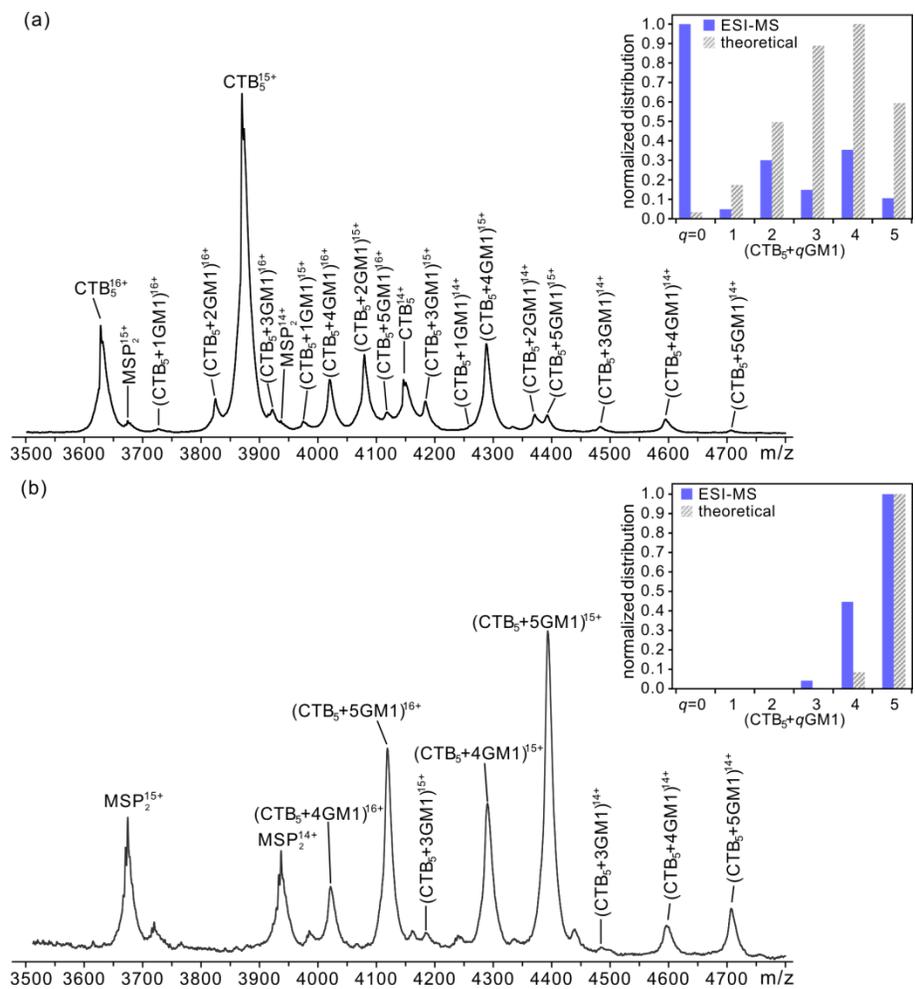
**Figure 6.4.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3  $\mu\text{M}$  CTB<sub>5</sub> with (a) 0.6  $\mu\text{M}$  and (b) 1.4  $\mu\text{M}$  10% GM1 ND (corresponding to 12 and 28  $\mu\text{M}$  GM1, respectively). Insets show normalized distributions of free and GM1-bound CTB<sub>5</sub>; theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>.



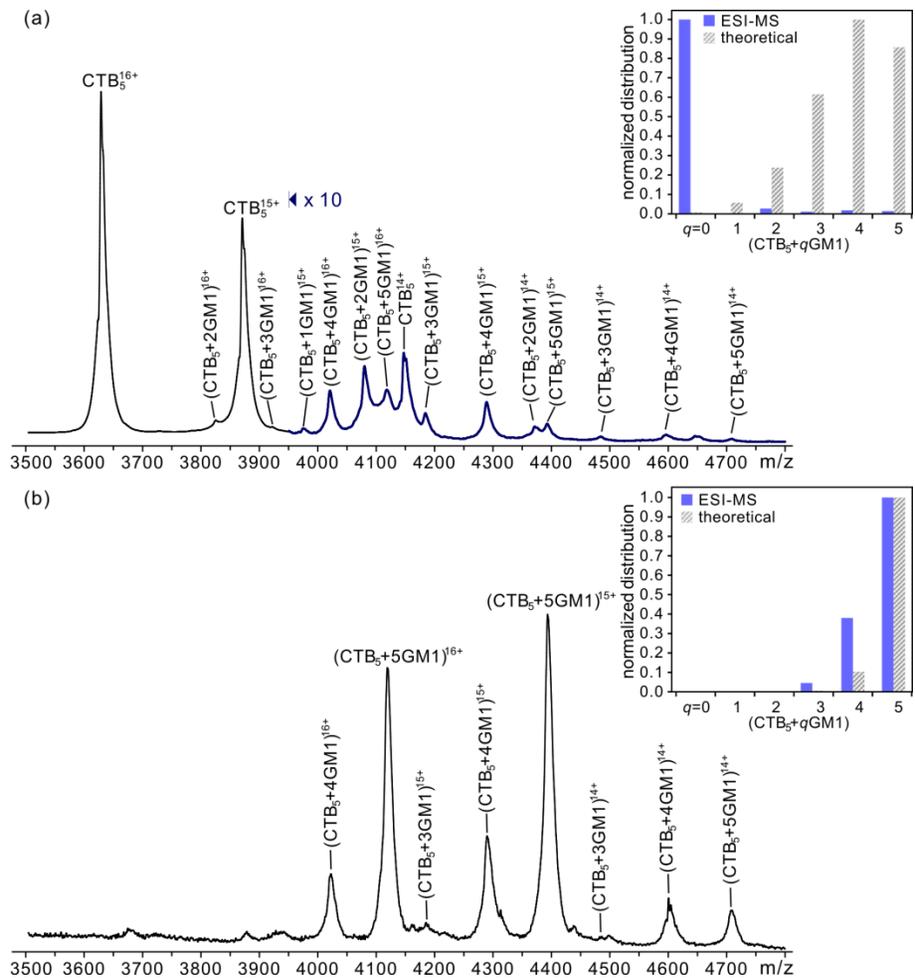
**Figure 6.5.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3 μM CTB<sub>5</sub> with (a) 3 μM and (b) 24.4 μM 0.5% GM1 ND (corresponding to 3 and 24.4 μM GM1, respectively). Insets show normalized distributions of free and GM1-bound CTB<sub>5</sub>; theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>.



**Figure 6.6.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) with (a) 6.8 μM and (b) 10.2 μM 1% GM1 ND (corresponding to 13.6 and 20.4 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB<sub>5</sub>. The theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>.



**Figure 6.7.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) with (a) 2.1 μM and (b) 4.3 μM 2.5% GM1 ND (corresponding to 10.5 and 21.5 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB<sub>5</sub>. The theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>.



**Figure 6.8.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) with (a) 1.2 μM and (b) 2.0 μM 5% GM1 ND (corresponding to 12 and 20 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB<sub>5</sub>. The theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub>.

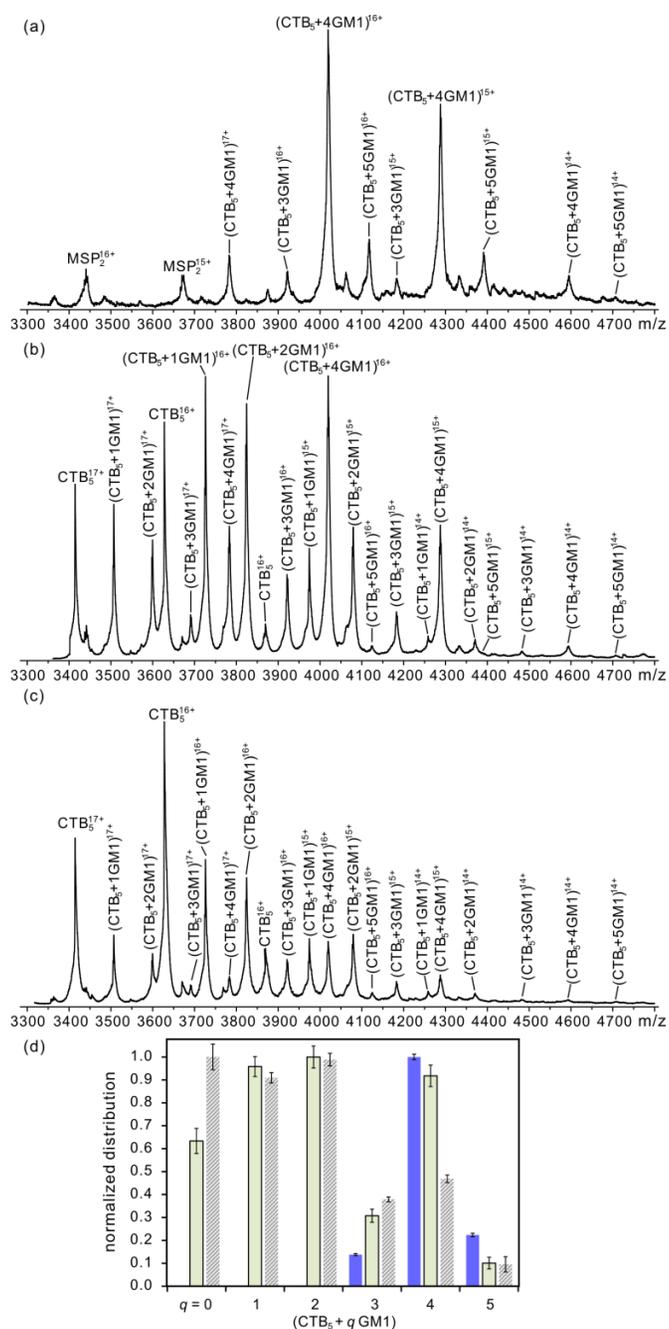
Notably, the distributions of bound GM1 measured using NDs with different percentages but with the same total concentration of GM1 are, in some cases, substantially different. As an example, CTB<sub>5</sub> is found to be bound predominantly to between three and five GM1 for solutions of 0.5% GM1 (12  $\mu$ M) and 1% GM1 (6  $\mu$ M) NDs (Figures 6.9a and 6.6a). In contrast, for solutions of higher percentage GM1 NDs, which also contain a total GM1 concentration of 12  $\mu$ M, CTB<sub>5</sub> is found to be primarily in its free form and the fraction of GM1-bound CTB<sub>5</sub> decreases with the increase of GM1 percentage (Figures 6.4a, 6.7a and 6.8a). The observed differences in the measured distributions are less noticeable at higher GM1 concentrations. For example, for solutions containing GM1 NDs of different GM1 percentages but all with  $\sim$ 20  $\mu$ M GM1, CTB<sub>5</sub> is found bound to between three to five GM1 in all cases (Figures 6.4b, 6.5b, 6.6b, 6.7b and 6.8b).

Comparing the measured distributions of bound GM1 to those expected based on the reported equilibrium constants for stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub> reveals that, under solution conditions that promote extensive GM1 binding (up to four or five GM1), the measured and theoretical distributions are similar, although the extent of GM1 binding measured by ESI-MS is generally less than expected (Figures 6.4b, 6.5b, 6.6b, 6.7b and 6.8b). In contrast, for solutions containing low concentrations of GM1 NDs, there are marked differences between the measured and theoretical distributions. For example, for solutions of 3  $\mu$ M CTB<sub>5</sub> with 2.1  $\mu$ M 2.5% GM1 ND, 1.2  $\mu$ M 5% GM1 ND or 0.6  $\mu$ M 10% GM1 ND, where the total GM1 concentration is 12  $\mu$ M, free CTB<sub>5</sub> dominates the ESI mass spectra. However, based on the concentration of GM1 present in solution and the affinities reported for GM1<sub>os</sub>, CTB<sub>5</sub> is expected

to be nearly fully bound (Figure 6.4a, 6.7a and 6.8a). As described in more detail below, the apparent disagreement between the measured and expected distributions for solutions containing low concentrations of GM1 NDs can be explained in terms of differential ESI-MS response factors for free CTB<sub>5</sub> and the (CTB<sub>5</sub> + *q*GM1) complexes, which are produced by gas-phase dissociation of ND-(CTB<sub>5</sub> + *q*GM1) complexes originating from solution.

### **6.3.1.2 Reversibility of CTB<sub>5</sub>-GM1 nanodisc interactions.**

To test the reversibility of the CTB<sub>5</sub> interactions with GM1 contained in the NDs, the influence of adding free CTB<sub>5</sub> to a solution containing CTB<sub>5</sub> and GM1 ND was investigated. Shown in Figure 6.9a is an ESI mass spectrum acquired for a 200 mM ammonium acetate aqueous solution containing 3 μM CTB<sub>5</sub> and 12 μM 0.5% GM1 ND (incubated for 15 min). Under these conditions, CTB<sub>5</sub> is predominantly bound to between three and five GM1. However, upon addition of 3 μM CTB<sub>5</sub> to this solution, free CTB<sub>5</sub>, as well as CTB<sub>5</sub> bound to between one and five GM1 are detected (Figure 6.9b). This distribution is nearly identical to that observed for a solution initially containing 6 μM CTB<sub>5</sub> and 12 μM 0.5% GM1 ND (Figures 6.9c and 6.9d). These results confirm that the CTB<sub>5</sub> interactions with GM1 (in NDs) in solution are reversible and that GM1 can be readily redistributed among the CTB<sub>5</sub> binding sites.

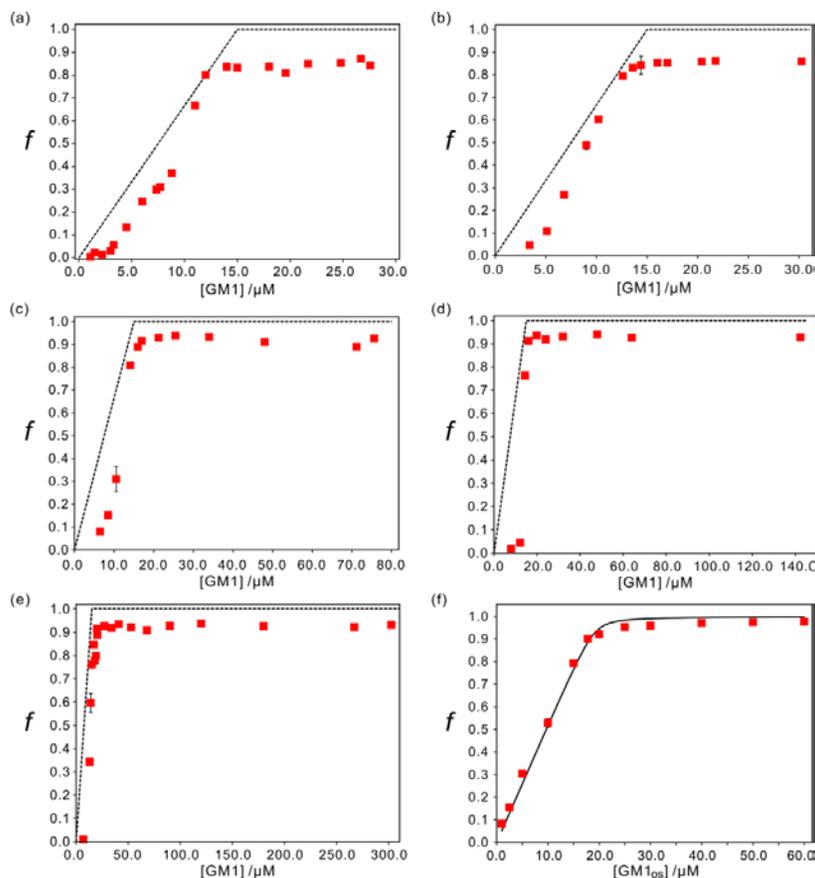


**Figure 6.9.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 12 μM 0.5% GM1 ND (corresponding to 12 μM GM1) with (a) 3 μM and (c) 6 μM CTB<sub>5</sub>. (b) ESI mass spectrum acquired upon addition of another 3 μM CTB<sub>5</sub> to the solution in (a). (d) Normalized distributions of free and GM1-bound CTB<sub>5</sub> measured from mass spectra in (a) ■, (b) ■ and (c) ▨.

### 6.3.1.3 Release of CTB<sub>5</sub>-GM1 complexes from nanodiscs in the gas phase.

From the ESI-MS data acquired for the solutions of 0.5%, 1%, 2.5%, 5% and 10% GM1 ND, plots of the fraction ( $f$ ) of occupied CTB<sub>5</sub> binding sites versus GM1 concentration were calculated (Figures 6.10a–6.10e). Although most noticeable for the low % GM1 ND data, all of the plots are sigmoidal in appearance, which, on its own, is suggestive of positive cooperativity, and reach a maximum  $f$  of between 85% and 94%. Also plotted is the dependence of  $f$  expected assuming complete (stoichiometric) binding. Notably, the experimental values approach the theoretical values, at least at certain concentrations, indicating that the amount of GM1 incorporated into the NDs does not differ significantly from the value expected based on the molar ratios of GM1 to DMPC used to prepare the NDs. To our knowledge, this is the first experimental evidence that the incorporation efficiency of GLs, such as GM1, into NDs is close to 100%. For comparison purposes the corresponding plot of  $f$  versus GM1<sub>os</sub> concentration measured by ESI-MS for solutions of CTB<sub>5</sub> (3.8  $\mu$ M) and GM1<sub>os</sub> (1 – 60  $\mu$ M) is also shown (Figure 6.10f). Notably, the experimental data for GM1<sub>os</sub> binding are well described by the theoretical curve, which was calculated using the Homans' binding model<sup>33</sup> and the reported affinities.<sup>34</sup> Moreover, although GM1<sub>os</sub> binding to CTB<sub>5</sub> exhibits slight positive cooperativity<sup>33,34</sup> the binding isotherm increases nearly linearly with GM1<sub>os</sub> concentration until the binding sites are saturated, i.e.,  $f$  (>99%). This latter result indicates that all five binding sites of CTB<sub>5</sub> are accessible for binding and that the  $f$  values <95% observed for GM1 binding are not due to structural effects related to ligand binding sites. Instead, it is proposed that a fraction of GM1 is retained by the ND upon release of the (CTB<sub>5</sub> +  $q$ GM1) ions in the gas phase, *vide infra*. The

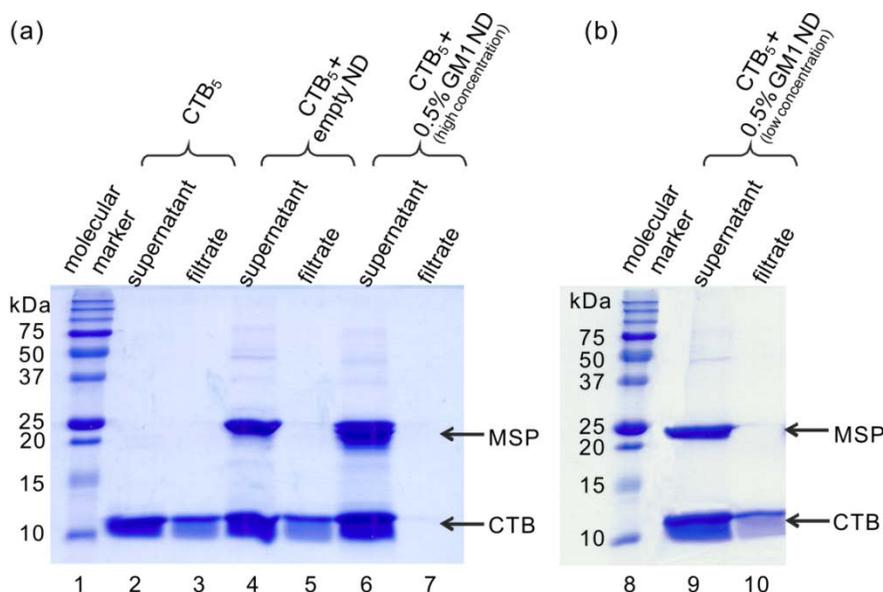
former result, the differences in the binding isotherms measured for the GM1 NDs and GM1<sub>os</sub>, suggests that the origin of apparent cooperative binding is different in the two cases, *vide infra*.



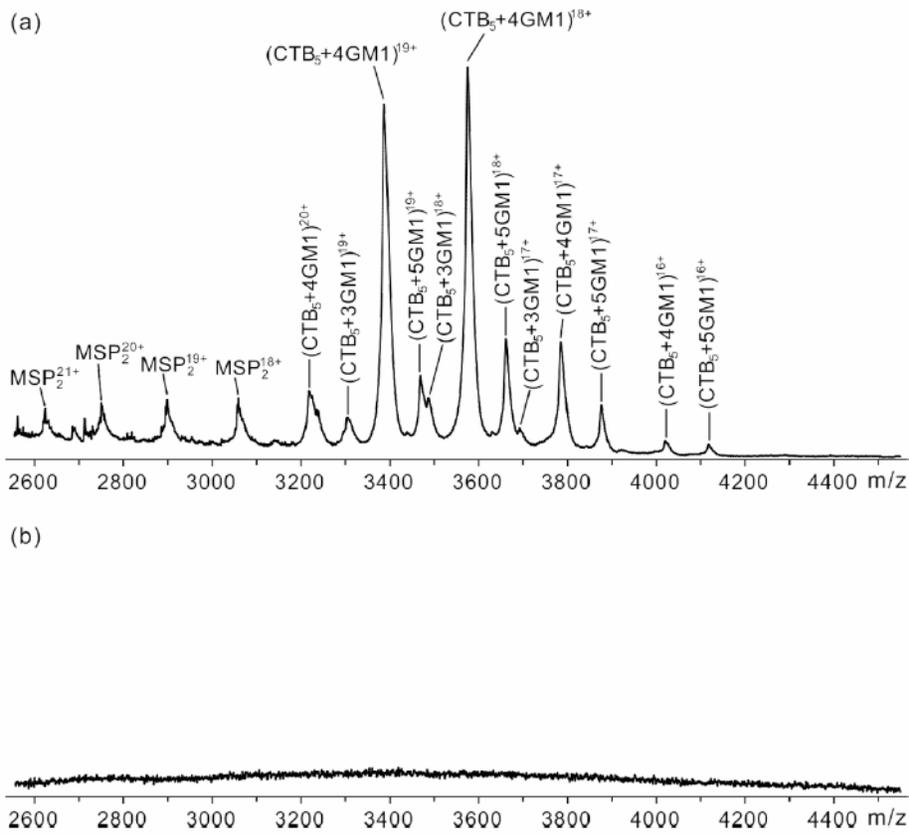
**Figure 6.10.** Plot of fraction of occupied ligand binding sites in CTB<sub>5</sub> ( $f$ ) versus GM1 concentration measured by ESI-MS in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3  $\mu\text{M}$  CTB<sub>5</sub> and varying concentrations of (a) 0.5%, (b) 1%, (c) 2.5%, (d) 5% and (e) 10% GM1 ND. Dashed lines represent the molar ratio of GM1 to the CTB<sub>5</sub> binding sites. (f) Plot of  $f$  versus GM1<sub>os</sub> concentration measured by ESI-MS in positive ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 3.8  $\mu\text{M}$  CTB<sub>5</sub> with GM1<sub>os</sub> (1.0  $\mu\text{M}$  – 60  $\mu\text{M}$ ). The solid curve corresponds to the theoretical plot calculated using affinities for the stepwise binding of GM1<sub>os</sub> to CTB<sub>5</sub> reported in reference 34. The error bars correspond to one standard deviation.

Previously, it was shown that, for solution of CTB<sub>5</sub> (5 μM) and 10% GM1 ND (10 μM), no free CTB<sub>5</sub> could be detected.<sup>21</sup> This finding led to the suggestion that the (CTB<sub>5</sub> + *q*GM1) ions measured by ESI-MS (under gentle sampling conditions) were the result of the kinetically facile dissociation of the (CTB<sub>5</sub> + *q*GM1) complexes from the NDs in the gas phase.<sup>21</sup> Analogous experiments were carried out in the present study to establish whether the (CTB<sub>5</sub> + *q*GM1) complexes present in solutions containing high and low concentrations of low % GM1 ND were associated with the NDs. For the high concentration case, an ammonium acetate solution (200 mM, pH 6.8, 25 °C) of CTB<sub>5</sub> (5 μM) and 0.5% GM1 ND (24 μM) was subjected to ultracentrifugation using a membrane with a 100 kDa MWCO and the filtrate and supernatant solutions analyzed by SDS-PAGE (Figure 6.11a). The results of this analysis failed to reveal the presence of CTB subunit in the filtrate, suggesting that the protein is predominantly bound to ND in solution. ESI-MS measurements were also carried out to identify the species present in the supernatant and filtrate. Notably, ions corresponding to CTB<sub>5</sub> bound to between three and five GM1, as well as MSP dimer, were detected in the supernatant (Figure 6.12a). In contrast, no free or GM1-bound CTB<sub>5</sub> ions were detected in the filtrate (Figure 6.12b). At lower concentration of GM1 ND (e.g. 3 μM), SDS-PAGE revealed bands corresponding to CTB subunit in both the supernatant and filtrate, similar to the results obtained for solutions of CTB<sub>5</sub> (5 μM) alone or with a ND containing no GM1 (Figure 6.11). However, while free CTB<sub>5</sub> and (CTB<sub>5</sub> + *q*GM1) complexes were present in the supernatant (Figure 6.13a), only free CTB<sub>5</sub> was identified in the filtrate (Figure 6.13b). To further confirm that no GM1-bound CTB<sub>5</sub> was present in the filtrate, CID was performed in negative ion mode on all ions with *m/z* >2500. The CID mass spectrum

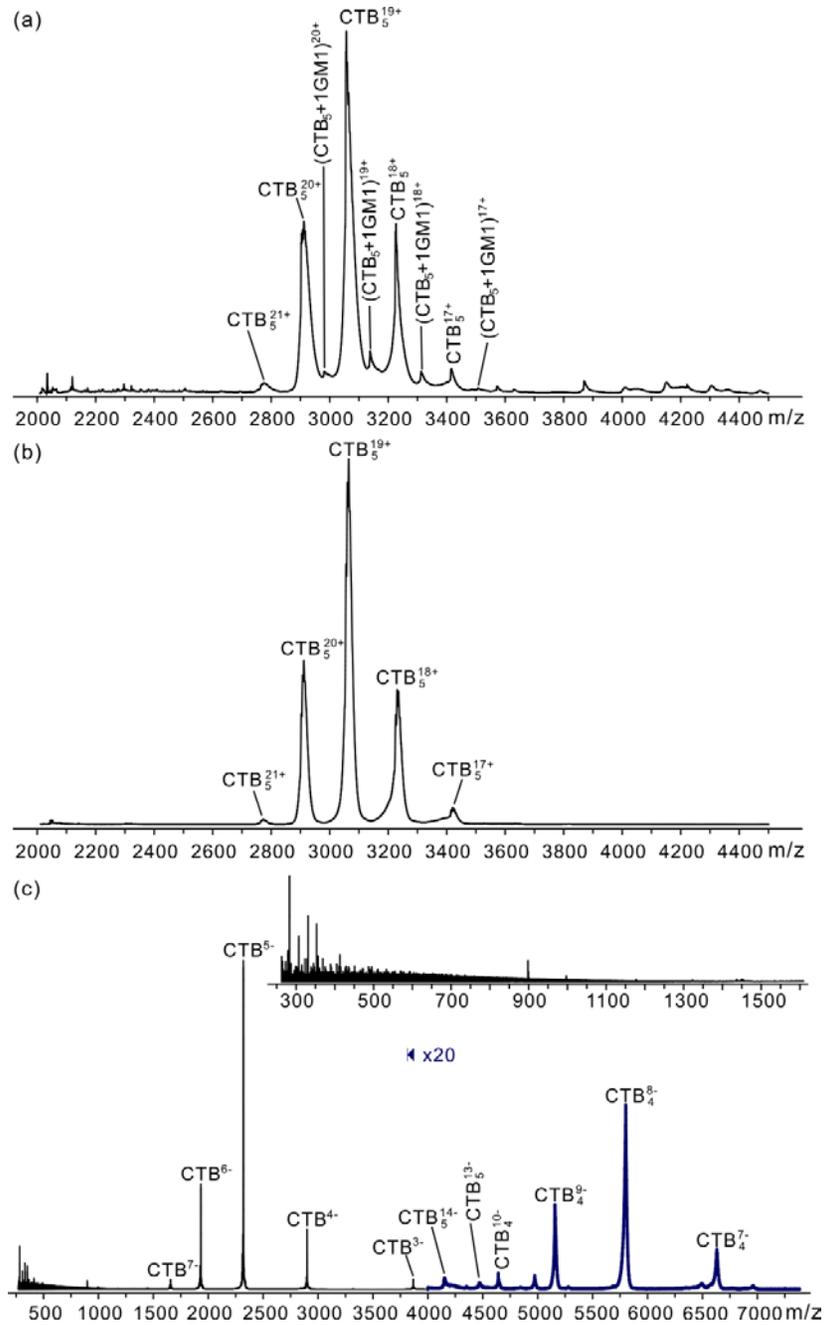
reveals signal corresponding to CTB subunit monomer and tetramer ions; no ions corresponding deprotonated GM1 were detected (Figure 6.13c). Taken together, these results provide compelling evidence that the (CTB<sub>5</sub> + *q*GM1) ions detected by ESI-MS are the results of gas-phase dissociation of the ND complexes, which results in the release of intact (CTB<sub>5</sub> + *q*GM1) complexes.



**Figure 6.11.** (a) and (b) SDS-PAGE analysis of aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> alone, CTB<sub>5</sub> with empty ND and CTB<sub>5</sub> with 0.5% GM1 NDs subjected to ultracentrifugation using a filter with a MWCO of 100 kDa.; supernatant (MW ≥100 kDa) and filtrate (MW ≤100 kDa). Molecular weight markers (lane 1 and 8); supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) alone (lanes 2 and 3, respectively); supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) with ND (11 μM) containing no GM1 (lanes 4 and 5, respectively); supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) with 0.5% GM1 ND (24 μM) (lanes 6 and 7, respectively); and supernatant and filtrate for solution of CTB<sub>5</sub> (5 μM) with 3 μM 0.5% GM1 ND (lanes 9 and 10, respectively).



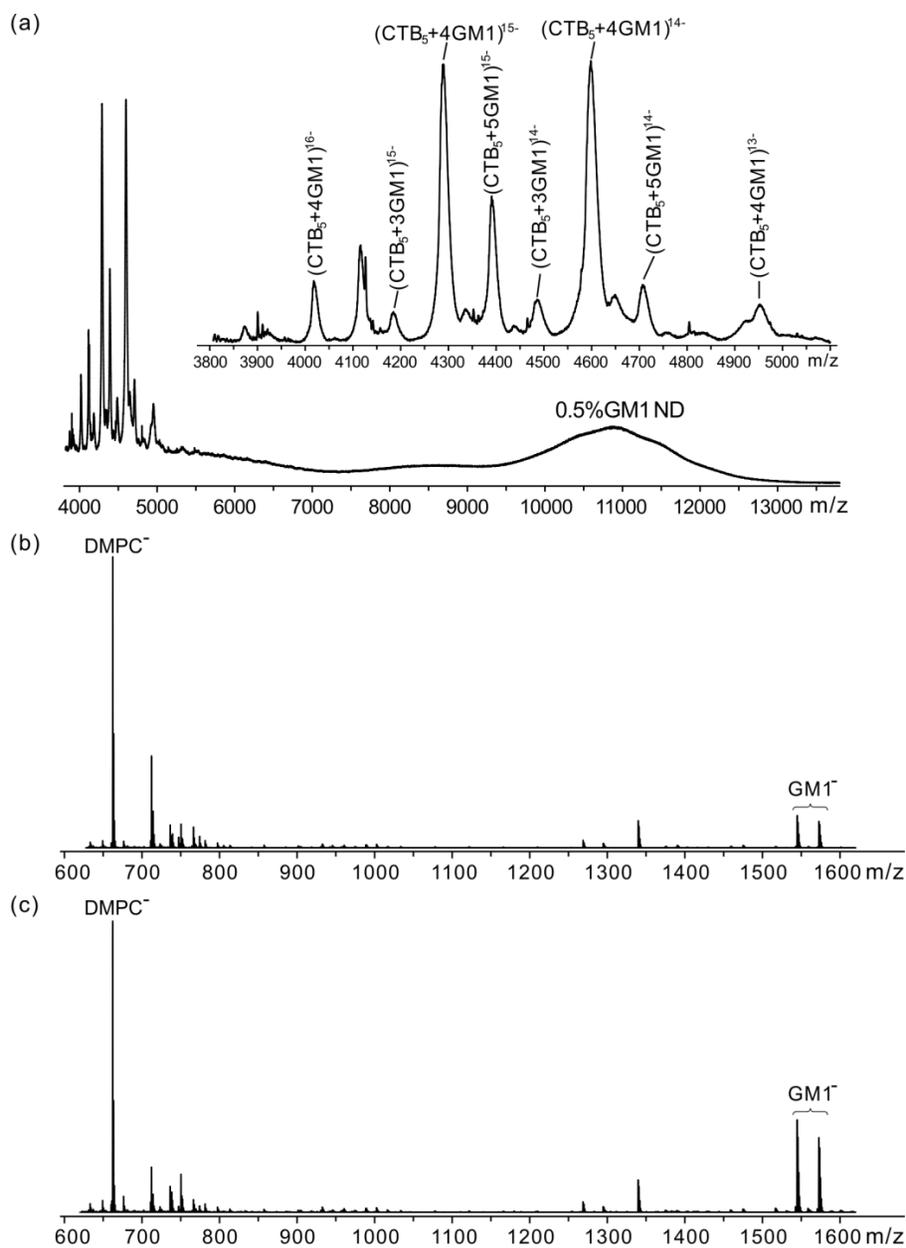
**Figure 6.12.** ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (5 μM) and 0.5% GM1 ND (24 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa.; (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa).



**Figure 6.13.** ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (5 μM) and 0.5% GM1 ND (3 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa: (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa). (c) CID mass spectrum acquired in negative ion mode for ions of m/z >2,500 produced in (b) using a collision energy of 120 V in the Trap.

Experimental support for the incomplete release of CTB<sub>5</sub>-bound GM1 from the NDs in the gas phase can be found in the results of CID experiments performed on the ND ions. Shown in Figures 6.14b and 6.14c are CID mass spectra acquired in negative ion mode for ND ions produced from a 200 mM ammonium acetate solution (pH 6.8, 25 °C) containing 14 μM 0.5% GM1 ND with and without 3 μM CTB<sub>5</sub>, respectively. CID was carried out using an isolation window centered at  $m/z$  11,000, which corresponds to the ND ions. A comparison of the CID mass spectra shows that the abundance ratio of GM1 to DMPC ions decreases after addition of CTB<sub>5</sub>, which is consistent with a fraction of GM1 is extracted from ND, forming (CTB<sub>5</sub> +  $q$ GM1) complex ions. However, deprotonated GM1 ions were found to be released from the ND even in the presence of excess CTB<sub>5</sub>.

The present binding data measured for solutions of CTB<sub>5</sub> and GM1 NDs reveal that the distributions of (CTB<sub>5</sub> +  $q$ GM1) complexes acquired by ESI-MS are sensitive to gas-phase processes. Two key conclusions are: the (CTB<sub>5</sub> +  $q$ GM1) ions detected by ESI-MS are produced by dissociation of the ND complexes in the gas phase and the dissociation process is not 100% efficient, with a small fraction of GM1 left behind in the NDs. Based on these findings, the apparent cooperative nature of CTB<sub>5</sub> binding to GM1 NDs, as suggested from the curvature in the plots of  $f$  versus GM1 concentration (Figure 6.10), can be attributed to a higher ESI-MS response factor for free CTB<sub>5</sub>, compared to the ND-associated (CTB<sub>5</sub> +  $q$ GM1) complexes, *vide supra*. Furthermore, the apparent inability to saturate the CTB<sub>5</sub> binding sites (i.e.,  $f$  reaches a limiting value of <0.95) is attributed to the incomplete extraction of GM1 from the NDs by CTB<sub>5</sub> in the gas phase.

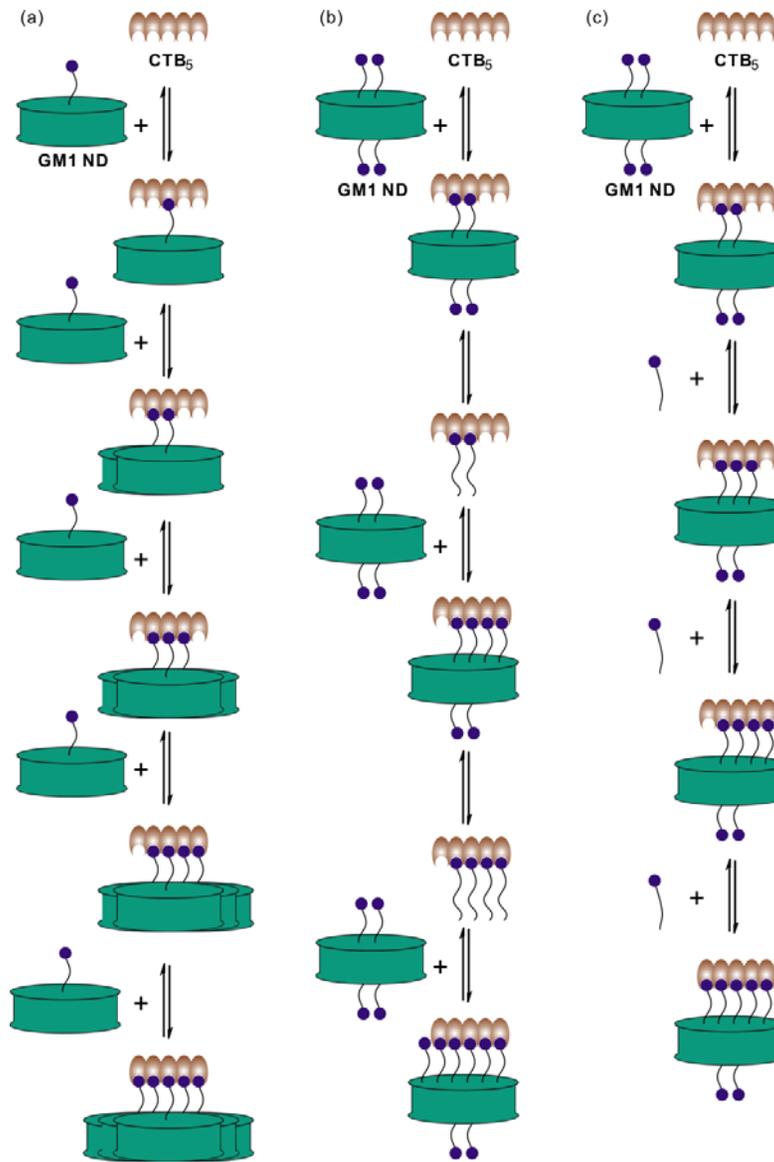


**Figure 6.14.** (a) ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (3 μM) and 0.5% GM1 ND (14 μM). (b) CID mass spectrum for ions produced in (a) and centred at m/z 11,000 (which correspond to GM1 ND). (c) CID mass spectrum for ions centred at m/z 11,000 produced from aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 0.5% GM1 ND (14 μM). A collision energy of 200 V in Trap was used for the CID experiments.

#### 6.3.1.4 Mechanism of CTB<sub>5</sub>-GM1 nanodisc binding.

Although the distributions of (CTB<sub>5</sub> + *q*GM1) complexes measured by ESI-MS are sensitive to gas-phase reactions, the binding data provide new insight into how CTB<sub>5</sub> associates with GM1 in the NDs. Notably, the detection of (CTB<sub>5</sub> + 4GM1) and (CTB<sub>5</sub> + 5GM1) complexes in solutions with low percentage GM1 NDs (i.e., 0.5% and 1% GM1 NDs, which contain an average of 1 and 2 GM1, respectively) is consistent with a stepwise binding model, in which CTB<sub>5</sub> sequentially binds to GM1 originating from multiple NDs. There are three possible mechanisms that could account for this observation.

*i) ND recruitment mechanism.* One possible mechanism would see CTB<sub>5</sub> binding irreversibly to GM1 from multiple NDs (Figure 6.15a). However, by overlaying the relative positions of the five ligand binding sites of CTB<sub>5</sub><sup>24</sup> onto NDs with diameters of ~11 nm,<sup>35,36</sup> it can be concluded that one CTB<sub>5</sub> could bind simultaneously to at most two NDs. Even then, unfavourable steric effects are likely to be significant. Consequently, based on structural considerations it is unlikely that the simultaneous binding of CTB<sub>5</sub> to multiple NDs is responsible for the measured distributions of (CTB<sub>5</sub> + *q*GM1) complexes.



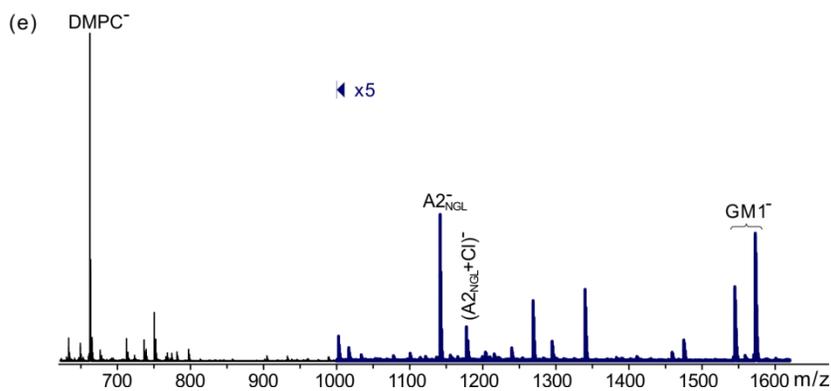
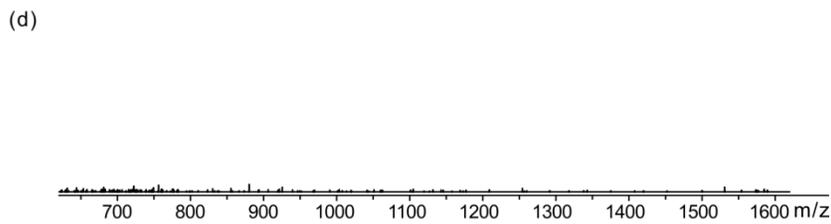
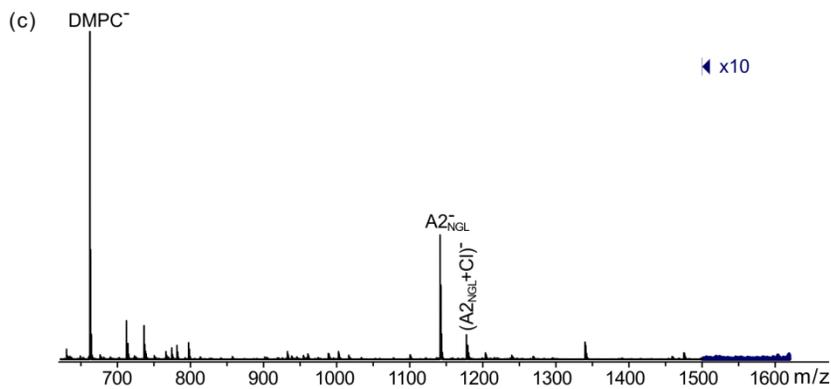
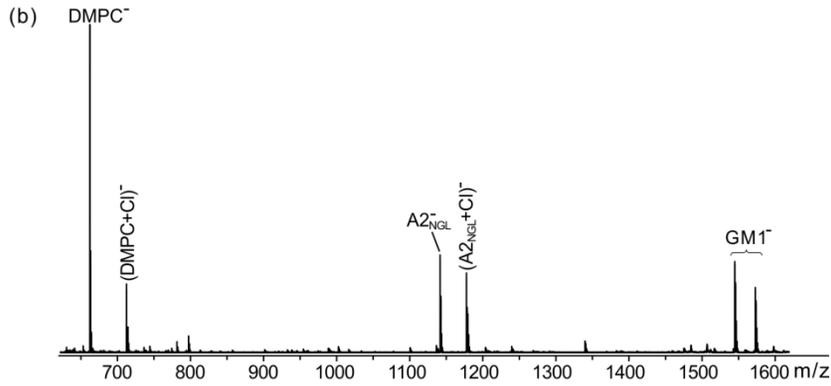
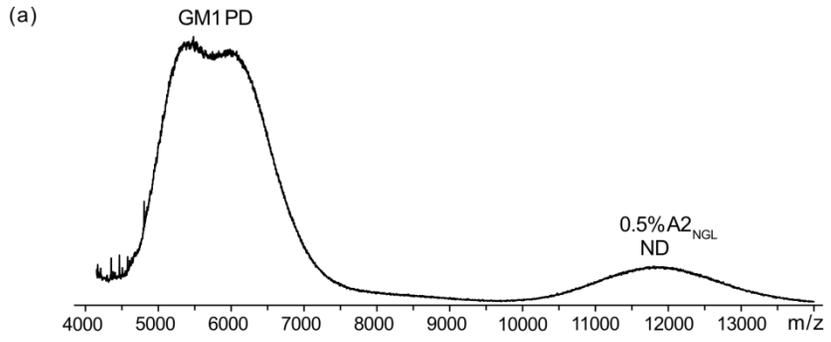
**Figure 6.15.** Possible mechanisms for the stepwise binding of CTB<sub>5</sub> to GM1 ND. (a) *Nanodisc recruitment mechanism* - CTB<sub>5</sub> binds irreversibly to GM1 ligands from multiple NDs. (b) *Glycolipid extraction mechanism* - CTB<sub>5</sub> interacts with GM1 in one ND, followed by dissociation of the resulting (CTB<sub>5</sub> +  $q$ GM1) complex from the original ND and rapid re-binding to GM1 in another ND. (c) *Glycolipid diffusion mechanism* - GM1 rapidly redistribute between NDs and can be recruited by CTB<sub>5</sub>. Note: to facilitate visualizing multivalent binding, a linear arrangement of subunits is used to represent the CTB<sub>5</sub> homopentamer.

ii) *GL extraction mechanism*. A second possible mechanism would involve CTB<sub>5</sub> interacting with GM1 molecules in one ND, followed by dissociation of an intact (CTB<sub>5</sub> + *q*GM1) complex from the ND and rapid re-binding to GM1 in another ND (Figure 6.15b). The number of binding steps would depend on the number of GM1 per ND and, in the case of NDs containing a high numbers of GM1 (e.g.  $\geq 5$  per ND), CTB<sub>5</sub> would be expected to interact with a single ND. An argument against this mechanism comes from kinetic data measured by SPR spectroscopy for the dissociation of CTB<sub>5</sub> from immobilized NDs containing on average one or two GM1 at 25 °C in HEPES-buffered saline (pH 7.4).<sup>19</sup> Based on the measured rate constant,  $0.028 \text{ min}^{-1}$ , the lifetime of ND-bound (CTB<sub>5</sub> + *q*GM1) complexes will be  $>35 \text{ min}$ , which is significantly longer than the time scale of the ESI-MS measurements. Moreover, the rate of dissociation from immobilized NDs containing  $>12$  GM1 was too slow to be accurately measured.<sup>19</sup> Although it was not clear from these measurements whether free CTB<sub>5</sub> or (CTB<sub>5</sub> + *q*GM1) complexes were released from the NDs, the kinetic data suggest that the stepwise binding of CTB<sub>5</sub> to different NDs is too slow to account for the measured distributions of (CTB<sub>5</sub> + *q*GM1) complexes. The absence of detectable amounts of (CTB<sub>5</sub> + *q*GM1) complexes in the filtrate from the ultracentrifugation experiments described above provides additional, although indirect, support for this conclusion.

iii) *GL diffusion mechanism*. A third possible mechanism would proceed through a rapid redistribution of GM1 between NDs such that, upon binding to one ND, CTB<sub>5</sub> can recruit additional GM1 from other NDs (Figure 6.15c). The exchange kinetics for DMPC between NDs have been quantified using small-angle neutron scattering and fluorescence methods.<sup>37</sup> These

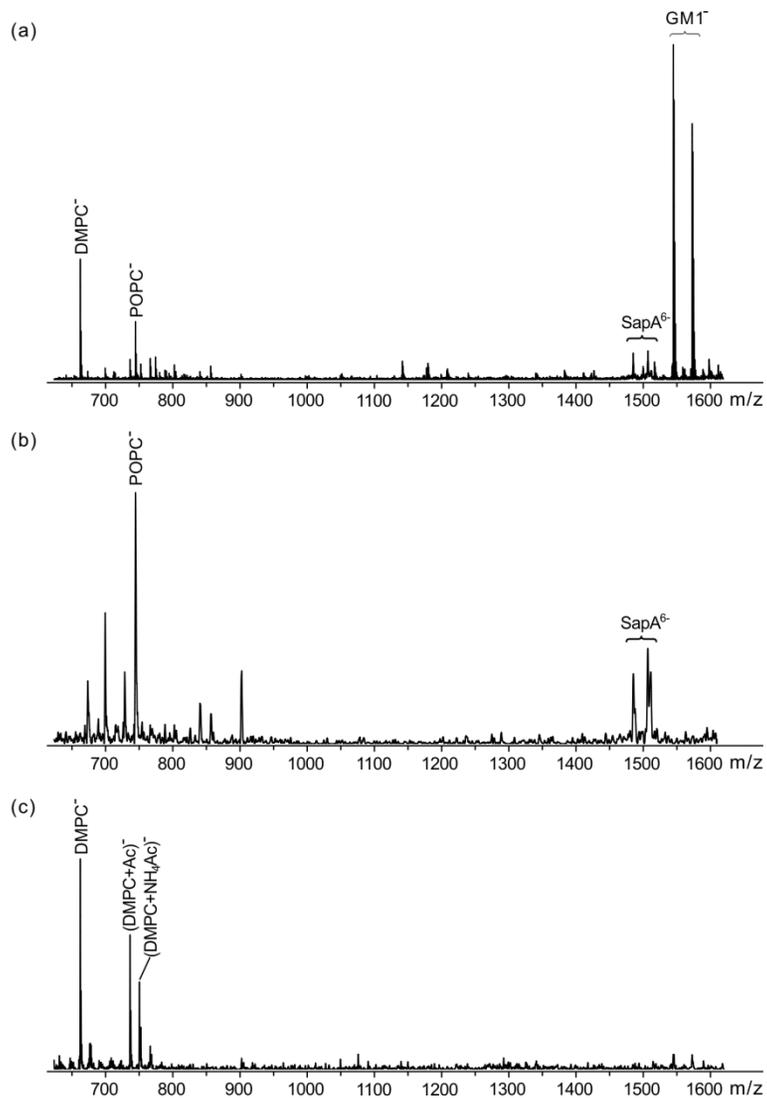
measurements, which support a monomeric lipid diffusion mechanism, yielded exchange rate constants ( $k_{ex}$ ) of 0.0328 min<sup>-1</sup> and 0.0378 min<sup>-1</sup> for DMPC exchange at 27 °C and an activation Gibbs energy of 91.8 kJ mol<sup>-1</sup>.<sup>37</sup> Using an average value of  $k_{ex}$  of 0.035 min<sup>-1</sup>, the lifetime of DMPC in the ND is estimated to be ~29 min at 27 °C.

In an effort to evaluate the rate of exchange of GM1 between NDs, CID measurements were performed in negative ion mode on ions with a narrow range of  $m/z$  centred at 11,500 produced by ESI from four different solutions, one with 0.5% A2<sub>NGL</sub> ND and picodiscs<sup>29,30</sup> containing GM1 and POPC (in a 1:1:4 SapA:GM1:POPC ratio), one with GM1 picodiscs alone (1:1:4 SapA:GM1:POPC ratio), one with 0.5% A2<sub>NGL</sub> ND alone, and one with 0.5% A2<sub>NGL</sub> ND and 0.5% GM1 ND (Figure 6.16). CID performed on solution of 12 μM 0.5% A2<sub>NGL</sub> ND and 54 μM GM1 picodisc produced negatively charged DMPC, A2<sub>NGL</sub> and GM1 ions (Figure 6.16b). In contrast, in the absence of the GM1 picodiscs in solution, CID produced only DMPC and A2<sub>NGL</sub> ions (Figure 6.16c). To rule out the possibility that the GM1 detected in the CID mass spectrum shown in Figure 6.16b originated from picodisc ions, analogous CID measurement was performed on ions with  $m/z$  centred at 11,500 produced from solution of GM1 picodisc. Notably, no GM1 ions were detected (Figures 6.16d). Interestingly, the relative abundances of GM1 and A2<sub>NGL</sub> ions detected in Figure 6.16b are similar to those measured by CID performed on ions ( $m/z$  ~11,500) produced from an equimolar mixture of 0.5% A2<sub>NGL</sub> ND and 0.5% GM1 ND (Figure 6.16e). Taken together, these data suggest that GM1 readily transfers from the picodisc to the ND (on the min timescale), leading to NDs that have ~0.5% GM1.



**Figure 6.16.** (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 0.5% A<sub>2</sub>NGL ND (12 μM) and picodisc (PD, 54 μM) containing POPC and GM1. (b) – (e) CID mass spectra acquired for ions centred at *m/z* 11,500 produced from: (b) the same solution as in (a); aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of (c) 0.5% A<sub>2</sub>NGL ND (12 μM); (d) PD (54 μM) containing POPC and GM1; and (e) 0.5% A<sub>2</sub>NGL ND (12 μM) and 0.5% GM1 ND (12 μM). For all measurements a collision energy of 200 V in the Trap was used.

The rapid transfer of GM1 from NDs to picodiscs was also demonstrated. As shown in Figure 6.17a, GM1 ions were observed in the CID mass spectrum acquired for ions with *m/z* ~5,500 produced from an ammonium acetate solution (pH 6.8, 25 °C) of 16 μM 1% GM1 ND and 60 μM picodiscs (containing only POPC). CID was also performed on ions with *m/z* ~5,500 produced from solutions of either POPC-containing picodiscs (Figure 6.17b) or 1% GM1 ND (Figure 6.17c). In neither case were deprotonated GM1 ions detected; this finding suggests that the GM1 ions detected in Figure 6.17a arise from the transfer of GM1 from the NDs to the picodiscs. Taken together, these results establish the rapid exchange of GM1 between picodiscs and NDs and lend support to the hypothesis that GM1 diffusion between NDs influences, at least to some extent, the measured distributions of (CTB<sub>5</sub> + *q*GM1) complexes.



**Figure 6.17.** CID mass spectra acquired in negative ion mode for ions centred at  $m/z$  5,500 produced by ESI from an aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of (a) 1% GM1 ND (16  $\mu$ M) with picodisc (PD, 60  $\mu$ M) containing POPC; (b) PD (60  $\mu$ M) containing POPC; and (c) 1% GM1 ND (16  $\mu$ M). For all measurements, a collisional energy of 120 V in the Trap was used.

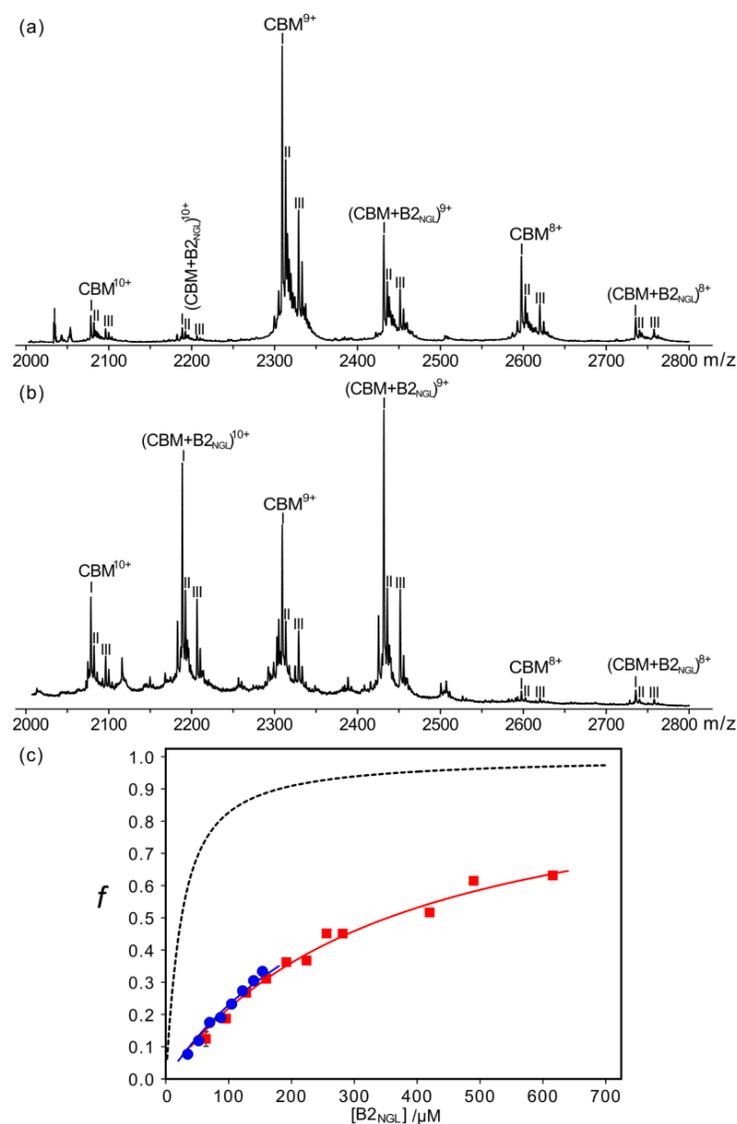
### 6.3.2 CBM binding to B2<sub>NGL</sub> in nanodiscs

The CBM–B2<sub>NGL</sub> interaction served as a second model system for investigating protein binding to GLs contained in NDs. CBM recognizes type A and B blood group oligosaccharides. Recent studies employing glycan array screening (Consortium for Functional Glycomics, <http://www.functionalglycomics.org/>), isothermal titration calorimetry (ITC),<sup>26</sup> as well as ESI-MS<sup>38</sup> revealed that CBM exhibits relatively strong binding for A/B trisaccharides and A/B type 2, 5 and 6 oligosaccharides ( $10^4$  to  $10^5$  M<sup>-1</sup>). Additionally, the X-ray crystal structure of CBM bound to the B type 2 tetrasaccharide (B2<sub>os</sub>) indicates that CBM possesses a single ligand binding site and forms a network of H-bonds with the  $\alpha$ -L-Fuc,  $\alpha$ -D-Gal and  $\beta$ -D-Gal residues.<sup>26</sup> The  $K_a$  of the B2<sub>os</sub> binding to CBM is reported to be  $5 \times 10^4 - 8 \times 10^4$  M<sup>-1</sup>.<sup>26,38</sup>

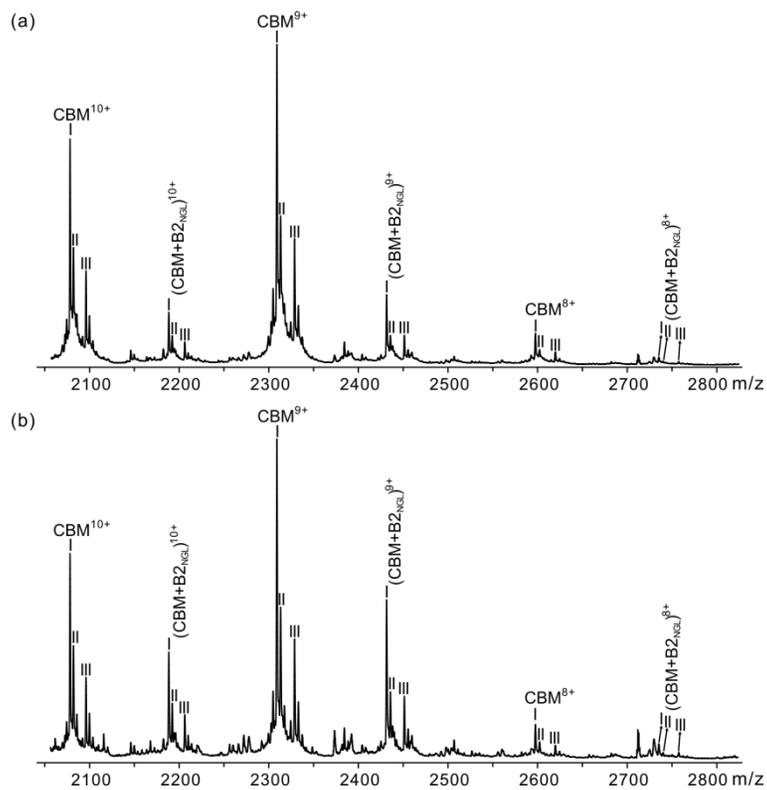
ESI-MS measurements were performed on solutions of CBM and 2.5% and 10% B2<sub>NGL</sub> NDs. Shown in Figures 6.18a and 6.18b are representative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate (200 mM, pH 6.8, 25 °C) solutions containing CBM (12  $\mu$ M) with 8  $\mu$ M and 30.8  $\mu$ M 10% B2<sub>NGL</sub> ND, respectively. Notably, signal corresponding to both free and B2<sub>NGL</sub>-bound CBM (all three CBM species) was detected, i.e., (CBM + B2<sub>NGL</sub>)<sup>n+</sup> at  $n = 8 - 10$ . Representative mass spectra acquired for solutions of CBM (12  $\mu$ M) with 2.5% B2<sub>NGL</sub> NDs are shown in Figure 6.19. Plots of the fraction of ligand-bound CBM versus B2<sub>NGL</sub> concentration are shown in Figure 6.18c, along with the expected curve for B2<sub>os</sub> binding, based on the reported affinity.<sup>38</sup> Fitting eq 1.10 to the experimental data yields similar affinities,  $3200 \pm 100$  M<sup>-1</sup> (2.5% B2<sub>NGL</sub> ND) and  $2900 \pm 100$  M<sup>-1</sup> (10% B2<sub>NGL</sub> ND). These values are significantly smaller (by factor of 17 – 18) than the  $K_a$  reported for B2<sub>os</sub>.<sup>26,38</sup> While this finding is, on its own,

consistent with the reduced protein affinities reported for some surface immobilized glycans,<sup>39</sup> it is likely that measured affinities for B2<sub>NGL</sub> are influenced by non-uniform ESI response factors for the bound and unbound CBM species, *vide infra*.

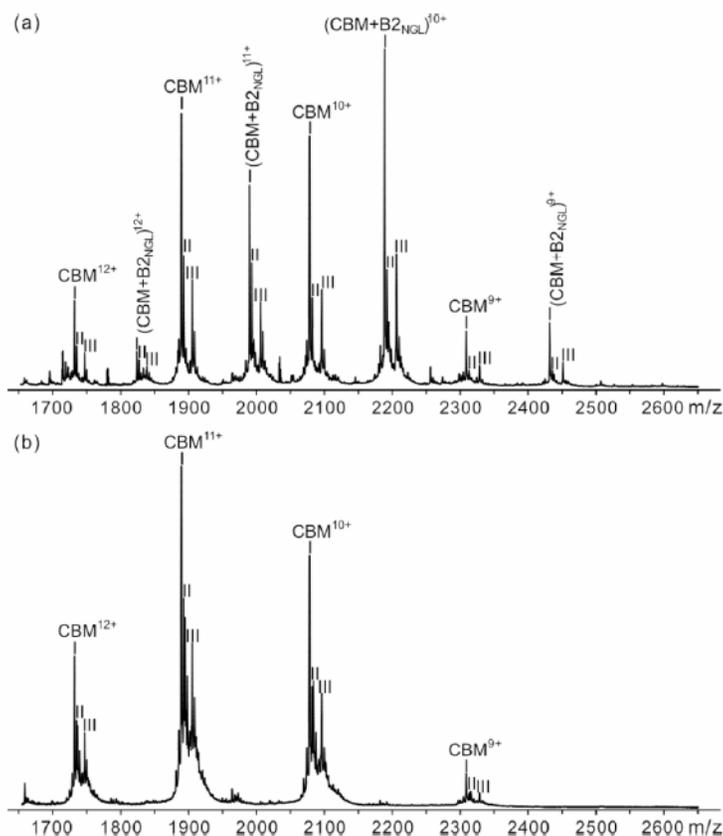
To demonstrate that ligand-bound CBM remains associated with the NDs in solution, ultracentrifugation analysis using a membrane filter with a MW cutoff of 100 kDa was carried out on an ammonium acetate solution (pH 6.8, 25 °C) of CBM (12 μM) with 10% B2<sub>NGL</sub> ND (21 μM). Because CBM cannot be reliably distinguished from the MSP used for the NDs by SDS-PAGE, ESI-MS measurements were carried out to analyze the supernatant and filtrate solutions. Shown in Figures 6.20a and 6.20b are mass spectra acquired for the supernatant and filtrate, respectively. It can be seen that free CBM is present in the filtrate, while (CBM + B2<sub>NGL</sub>) is only detected in the supernatant. This result, which is consistent with those obtained for solutions of CTB<sub>5</sub> and GM1 NDs, suggests that B2<sub>NGL</sub>-bound CBM is associated with the ND in solution and that the (CBM + B2<sub>NGL</sub>)<sup>n+</sup> ions detected by ESI-MS are the result of dissociation of the CBM-B2<sub>NGL</sub>-ND complexes in the gas phase.



**Figure 6.18.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12  $\mu\text{M}$ ) with (a) 8  $\mu\text{M}$  and (b) 30.8  $\mu\text{M}$  10% B2<sub>NGL</sub> ND (corresponding to 160  $\mu\text{M}$  and 616  $\mu\text{M}$  B2<sub>NGL</sub>, respectively). (c) Plots of fraction of ligand-bound CBM ( $f$ ) versus B2<sub>NGL</sub> concentration. The experimental conditions were the same as in (a) and (b), but with addition of 3.2 – 30.8  $\mu\text{M}$  ND containing 2.5% (●) or 10% (■) B2<sub>NGL</sub>. The dashed curve represents the theoretical plot calculated from the association constant reported in reference 38 for CBM binding to B2<sub>os</sub>. The error bars correspond to one standard deviation.



**Figure 6.19.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 μM) and (a) 2.5% B<sub>2NGL</sub> ND (14 μM) (corresponding to 70 μM B<sub>2NGL</sub>) or (b) 2.5% B<sub>2NGL</sub> ND (30.8 μM) (corresponding to 154 μM B<sub>2NGL</sub>).



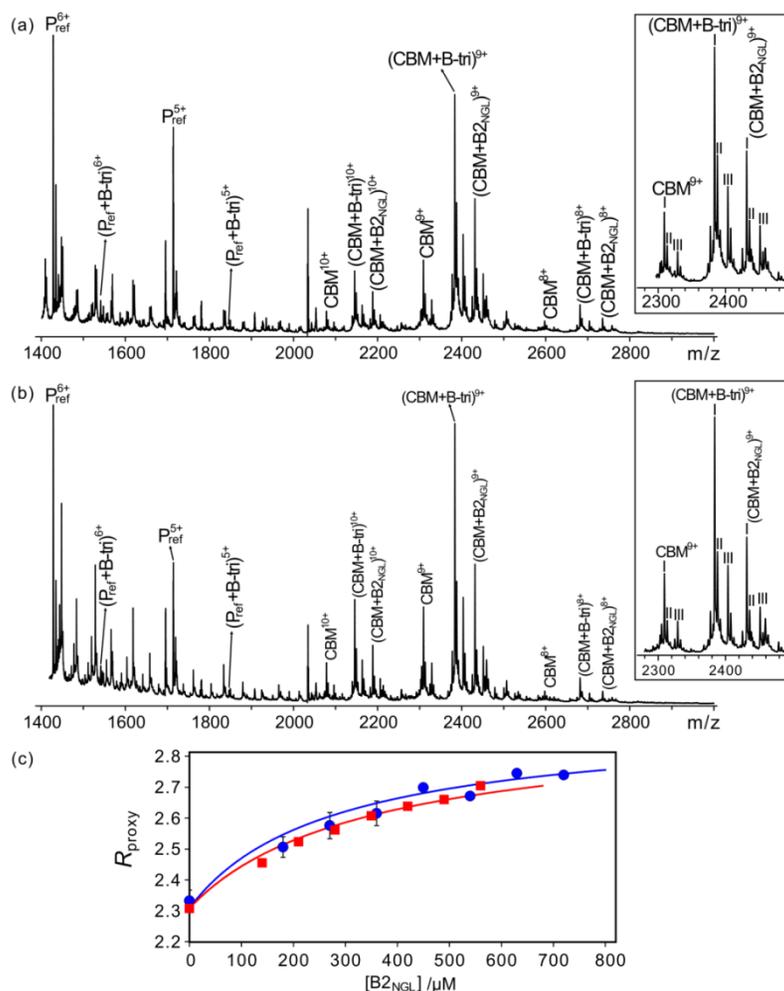
**Figure 6.20.** ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 μM) and 10% B2<sub>NGL</sub> ND (25 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa: (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa).

### 6.3.3 Protein affinities for glycolipids in nanodiscs – the *proxy ligand* ESI-MS assay

A weakness of the *direct* ESI-MS assay for quantifying protein-GL interactions involving NDs is that the detected protein-GL complexes result from dissociation of the protein-GL-ND complexes in the gas phase. Consequently, any differences in the ESI response factors for the free protein and the protein-GL complex ions will introduce errors into the affinity measurements. Given these limitations, a new ESI-MS binding assay, the *proxy ligand* method, was developed.

This assay, which combines *direct* ESI-MS measurements with competitive ligand-protein binding, was used to quantify the affinities of CBM for NDs containing 10% and 15% B<sub>2NGL</sub>. The B-tri ligand, which served as L<sub>proxy</sub> for these measurements, has an affinity for CBM of  $7.3 \times 10^4 \text{ M}^{-1}$ .<sup>38</sup> Shown in Figure 6.1b is a representative ESI mass spectrum acquired for the aqueous ammonium acetate solution (200 mM, pH 6.8, 25 °C) of 12  $\mu\text{M}$  CBM and 40  $\mu\text{M}$  B-tri. Ions corresponding to free CBM and CBM bound to B-tri were detected, i.e., CBM<sup>*n*+</sup> and (CBM + B-tri)<sup>*n*+</sup> at *n* = 8 to 10. The addition of 24  $\mu\text{M}$  of 15% B<sub>2NGL</sub> ND to the solution resulted in the appearance of ions corresponding to CBM bound to B<sub>2NGL</sub>, i.e., (CBM + B<sub>2NGL</sub>)<sup>*n*+</sup> at *n* = 8 to 10, also resulted in an increase in the abundance ratio of B-tri-bound to free CBM ions (i.e., *R*<sub>proxy</sub>) (Figure 6.21a). The increase in *R*<sub>proxy</sub> is consistent with a decrease in CBM available for binding to B-tri due to the competitive binding to B<sub>2NGL</sub>. Shown in Figure 6.21c is a plot of *R*<sub>proxy</sub> versus B<sub>2NGL</sub> concentration. The data were analyzed according to the procedure described in the Experimental section and an affinity of  $(1.4 \pm 0.1) \times 10^4 \text{ M}^{-1}$  was obtained by fitting eq 6.7 to the experimental data. Measurements carried out using 10% B<sub>2NGL</sub> ND yielded an affinity of  $(1.1 \pm 0.1) \times 10^4 \text{ M}^{-1}$  (Figures 6.21b and 6.21c). Notably, the B<sub>2NGL</sub> affinities measured using the *proxy ligand* ESI-MS assay are consistently higher (by a factor of ~5) than the values obtained by *direct* ESI-MS assays. The lower values measured directly by ESI-MS are attributed to non-uniform response factors for CBM and (CBM + B<sub>2NGL</sub>) species, *vide supra*. That the *K*<sub>a</sub> for the CBM-B<sub>2NGL</sub> interaction measured by *proxy ligand* method is lower (by a factor of ~5) than the value reported for B<sub>2os</sub> (*K*<sub>a</sub> =  $5.3 \times 10^4 \text{ M}^{-1}$ ) is also notable. This finding suggests that protein binding to GLs in NDs may be energetically less favorable than the interactions with the

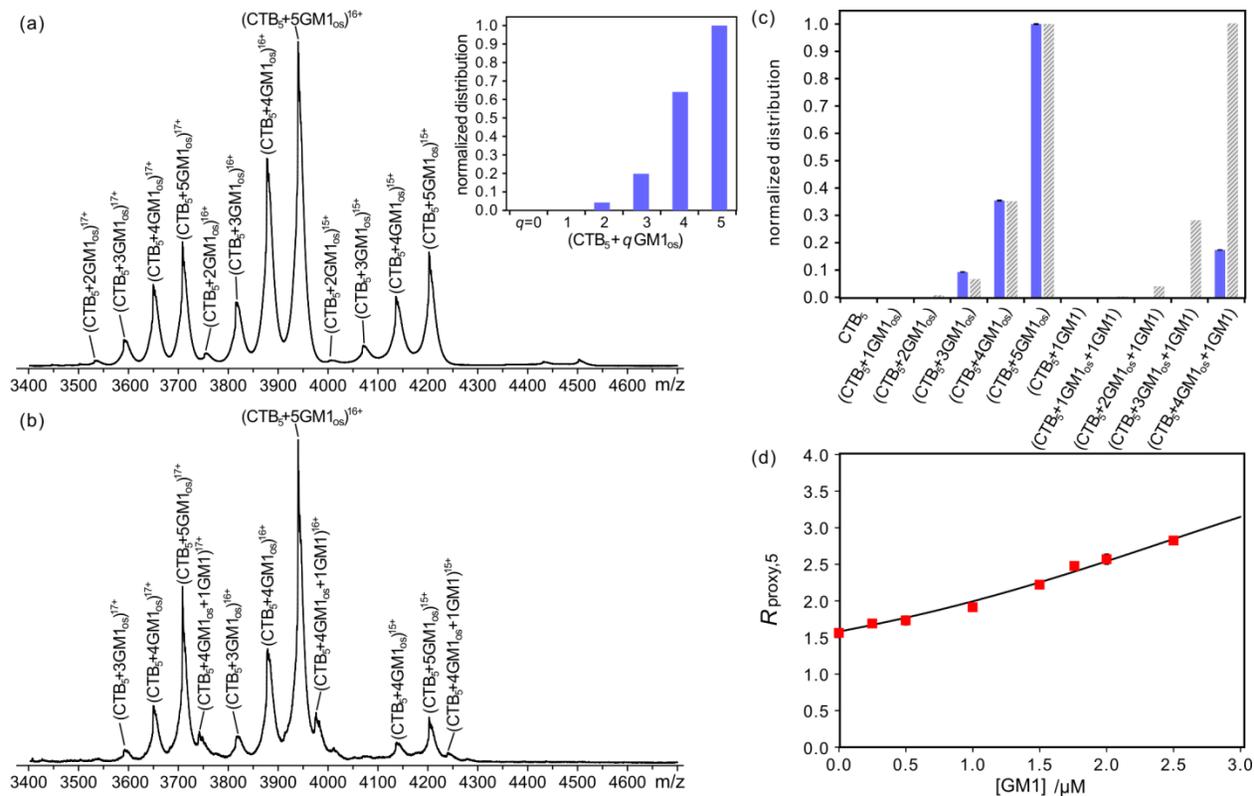
corresponding free oligosaccharides in solution.



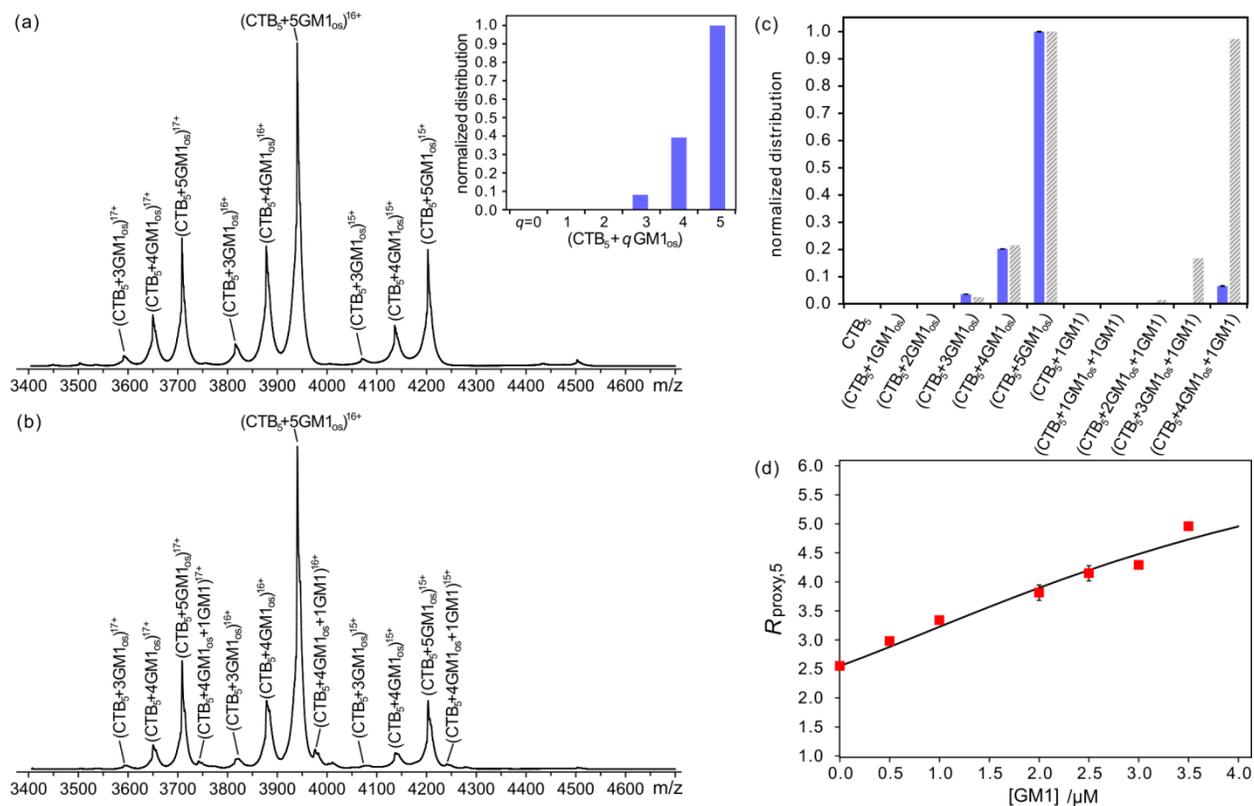
**Figure 6.21.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 12 μM CBM, 40 μM B-tri ( $L_{prox}$ ) with (a) 24 μM 15% B2<sub>NGL</sub> ND and (b) 21 μM 10% B2<sub>NGL</sub> ND (corresponding to 720 μM and 420 μM B2<sub>NGL</sub>, respectively); 5 μM  $P_{ref}$  (Ubq) was added to each solution to correct for the nonspecific ligand binding during ESI process. (c) Plots of  $R_{prox}$  ( $\equiv Ab(CBM + B-tri)/Ab(CBM)$ ) versus B2<sub>NGL</sub> concentration. The experimental conditions were the same as in (a), but with addition of 0 – 24 μM 15% B2<sub>NGL</sub> ND (●) or 0 – 28 μM 10% B2<sub>NGL</sub> ND (■). The error bars correspond to one standard deviation.

The *proxy ligand* ESI-MS method was also extended to evaluate the affinities of CTB<sub>5</sub> for GM1 NDs. However, because of the presence of multiple binding sites, the cooperative nature of GM1 binding and the possibility of multivalent binding effects, interpretation of binding data is generally more complicated than in the case of CBM. To minimize the occurrence of multivalent binding, measurements were carried out on solutions of CTB<sub>5</sub> with low concentrations of low percentage (0.5% and 1%) GM1 NDs and high concentrations of GM1<sub>os</sub>, which served as L<sub>proxy</sub>. Under these conditions, it is expected that CTB<sub>5</sub> will bind preferentially GM1<sub>os</sub> and will not interact with multiple GM1. Shown in Figure 6.22a is a representative ESI mass spectrum acquired for an aqueous ammonium acetate solution (200 mM, pH 6.8, 25 °C) of 4.4 μM CTB<sub>5</sub> and 20 μM GM1<sub>os</sub>. Ions corresponding to CTB<sub>5</sub> bound to between two and five GM1<sub>os</sub> were observed, with the (CTB<sub>5</sub> + 5GM1<sub>os</sub>) complex being the most abundant. The addition of 2.5 μM 0.5% GM1 ND to the solution resulted in the appearance of (CTB<sub>5</sub> + 4GM1<sub>os</sub> + GM1)<sup>n+</sup> ions, at  $n = 15$  to  $17$ , (Figure 6.22b) and a measurable increase of the abundance ratio of the (CTB<sub>5</sub> + 5GM1<sub>os</sub>) to (CTB<sub>5</sub> + 4GM1<sub>os</sub>) ions ( $\equiv R_{\text{proxy},5}$ ), which is consistent with CTB<sub>5</sub> binding to GM1 ND in solution. A plot of  $R_{\text{proxy},5}$  versus GM1 (in the ND) concentration is shown in Figure 6.22d. Using the binding model described in Experimental section, which is an extension of the Homans' model,<sup>33</sup> the association constants  $K_{a,1}$ ,  $K_{a,2}$  and  $K_{a,3}$ , corresponding to GM1 binding to CTB<sub>5</sub> sites with zero, one or two occupied nearest neighbour subunits, respectively, which gave the closest agreement to the experimentally determined  $R_{\text{proxy},5} - K_{a,1} = 2.8 \times 10^6 \text{ M}^{-1}$ ,  $K_{a,2} = 4.8 \times 10^6 \text{ M}^{-1}$  and  $K_{a,3} = 8.2 \times 10^6 \text{ M}^{-1}$ . Shown in Figure 6.22c is a comparison of the theoretical distribution of bound GM1<sub>os</sub> and GM1 (calculated using these  $K_{a,1}$ ,  $K_{a,2}$  and

$K_{a,3}$  values) with the experimentally-determined distribution determined from the mass spectrum in Figure 6.22b. Notably, there is excellent agreement in the distributions of bound GM1<sub>os</sub>. In contrast, the predicted distribution for bound GM1 does not resemble the experimental distribution. However, this disagreement can be explained in terms of non-uniform ESI-MS response factors for the (CTB<sub>5</sub> +  $q$ GM1<sub>os</sub>) and (CTB<sub>5</sub> +  $q$ GM1<sub>os</sub> + GM1) complexes, *vide supra*. Moreover, the concentration dependence of  $R_{\text{proxy},5}$  predicted theoretically agrees well with the experimental observations made over a range of concentrations (Figure 6.22d). Analogous measurements performed using 1% GM1 ND gave a similar affinities -  $K_{a,1} = 1.2 \times 10^6 \text{ M}^{-1}$ ,  $K_{a,2} = 2.0 \times 10^6 \text{ M}^{-1}$  and  $K_{a,3} = 3.5 \times 10^6 \text{ M}^{-1}$  (Figure 6.23). Notably, the measured affinities are slightly smaller than the values obtained for the corresponding CTB<sub>5</sub>-GM1<sub>os</sub> interactions<sup>34</sup> -  $K_{a,\text{proxy},1} = 3.2 \times 10^6 \text{ M}^{-1}$ ,  $K_{a,\text{proxy},2} = 5.5 \times 10^6 \text{ M}^{-1}$  and  $K_{a,\text{proxy},3} = 9.5 \times 10^6 \text{ M}^{-1}$ , a finding consistent with what was found for the CBM and B2<sub>NGL</sub> interaction.



**Figure 6.22.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 4.4 μM CTB<sub>5</sub> and 20 μM GM1<sub>os</sub> with (a) 0 μM and (b) 2.5 μM 0.5% GM1 ND. Inset shows the normalized distributions of free and GM1<sub>os</sub>-bound CTB<sub>5</sub>. (c) (■) Normalized distributions of free and ligand-bound CTB<sub>5</sub> measured from the mass spectrum in (b); (▨) theoretical distributions were calculated using association constants determined from the *proxy ligand* method and values reported in reference 34 for the stepwise binding of GM1<sub>os</sub> and GM1 to CTB<sub>5</sub>. (d) Plot of  $R_{\text{proxy},5}$  ( $\equiv Ab(\text{CTB}_5 + 5\text{GM1}_{\text{os}})/Ab(\text{CTB}_5 + 4\text{GM1}_{\text{os}})$ ) versus GM1 concentration. The experimental conditions were the same as in (a) and (b), but with addition of 0 – 2.5 μM 0.5% GM1 ND. The error bars correspond to one standard deviation.



**Figure 6.23.** ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB<sub>5</sub> (4.0 μM) and GM1<sub>os</sub> (20 μM) with (a) 0 μM and (b) 1.75 μM 1% GM1 ND (corresponding to 3.5 μM GM1). Insets show the normalized distributions of free and GM1<sub>os</sub>-bound CTB<sub>5</sub>. (c) (■) Normalized distributions of free and ligand-bound CTB<sub>5</sub> measured from mass spectrum in (b); (▨) theoretical distributions calculated from the association constants reported for the stepwise binding of GM1<sub>os</sub> (from reference 34) and 1% GM1 ND (determined in the present study) to CTB<sub>5</sub>. (d) Plot of  $R_{\text{proxy},5}$  ( $\equiv Ab(\text{CTB}_5+5 \text{GM1}_{\text{os}})/Ab(\text{CTB}_5+4 \text{GM1}_{\text{os}})$ ) versus GM1 concentration. The experimental conditions of are same as in (a) and (b), but with addition of 0 – 1.75 μM 1% GM1 ND. The error bars correspond to one standard deviation.

## 6.4 Conclusions

The present study represents the first detailed investigation into the mechanisms and energetics of protein interactions with GLs in NDs. The results of ESI-MS measurements performed on solutions of CTB<sub>5</sub> and GM1 NDs reveal that proteins bind reversibly to ND-bound GLs and, in the case of proteins with multiple ligand binding sites, are able to interact with GLs originating from different NDs. The results of ESI-MS measurements performed on solutions of NDs and picodiscs provide direct evidence for rapid GL diffusion between picodiscs and NDs. Based on this finding it is proposed that diffusion of GLs between NDs influences the nature of the protein-GL complexes detected. While ESI-MS serves as a convenient method for detecting protein interactions with GLs in NDs, the measured abundances of free and GL-bound protein ions do not necessarily reflect solution composition. There is overwhelming evidence that, in solution, the GL-bound proteins remain associated with NDs and are only released (as protein-GL complexes) in the gas phase. Consequently, different ESI-MS response factors are expected for the free proteins and GL-bound proteins. Finally, using the newly developed *proxy ligand* ESI-MS assay,  $K_a$  values for CBM-B2<sub>NGL</sub> and CTB<sub>5</sub>-GM1 interactions were quantified. A key finding of this study is that the affinities of the proteins for the GL ligands in the NDs are slightly lower (by a factor of  $\leq 5$ ) than those of the corresponding oligosaccharides in solution. Future efforts will exploit the *proxy ligand* ESI-MS method to study, in detail, the effects of ND composition on protein-GL binding.

## 6.5 References

- (1) Hakomori, S. *Curr. Opin. Hematol.* **2003**, *10*, 16.
- (2) Sharon, N.; Lis, H. *Sci. Am.* **1993**, *268*, 82.
- (3) Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Stanley, P.; Bertozzi, C.R.; Hart, G.W.; Etzler, M.E. *Essentials of Glycobiology* 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA 2009.
- (4) Holgersson, J.; Gustafsson, A.; Breimer, M. E. *Immunol. Cell Biol.* **2005**, *83*, 694.
- (5) Lopez P. H.; Schnaar, R. L. *Methods Enzymol.* **2006**, *417*, 205.
- (6) DeMarco, M. L. *Biochemistry* **2012**, *51*, 5725.
- (7) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.
- (8) Lingwood, C. A. *Glycoconjugate J.* **1996**, *13*, 495.
- (9) Lingwood, C. A.; Manis, A.; Mahfoud, R.; Khan, F.; Binnington, B.; Mylvaganam, M. *Chem. Phys. Lipids* **2010**, *163*, 27.
- (10) Lingwood, D.; Binnington, B.; Rog, T.; Vattulainen, I.; Grzybek, M.; Coskun, U.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.
- (11) MacKenzie, C. R.; Hiram, T.; Lee, K. K.; Altman, E.; Young, N. M. *J. Biol. Chem.* **1997**, *272*, 5533.
- (12) Lauer, S.; Goldstein, B.; Nolan, R. L.; Nolan, J. P. *Biochemistry* **2002**, *41*, 1742.
- (13) Gallegos, K. M.; Conrady, D. G.; Karve, S. S.; Gunasekera, T. S.; Herr, A. B.; Weiss, A. A. *PLoS One* **2012**, *7*, 10.
- (14) Vogel, J.; Bendas, G.; Bakowsky, U.; Hummel, G.; Schmidt, R. R.; Kettmann, U.;

- Rothe, U. *Biochim. Biophys. Acta.* **1998**, *1372*, 205.
- (15) Shi, J.; Yang, T.; Kataoka, S.; Zhang, Y.; Diaz, A. J.; Cremer, P. S. *J. Am. Chem. Soc.* **2007**, *129*, 5954.
- (16) Nath, A.; Atkins, W. M.; Sligar, S. G. *Biochemistry* **2007**, *46*, 2059.
- (17) Bayburt, T. H.; Sligar, S. G. *FEBS Lett.* **2010**, *584*, 1721.
- (18) Jayaraman, N.; Maiti, K.; Naresh, K. *Chem. Soc. Rev.* **2013**, *42*, 4640.
- (19) Borch, J.; Torta, F.; Sligar, S. G.; Roepstorff, P. *Anal. Chem.* **2008**, *80*, 6245.
- (20) Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618.
- (21) Leney, A. C.; Fan, X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271.
- (22) Sloan, C. D. K.; Marty, M. T.; Sligar, S. G.; Bailey, R. C. *Anal. Chem.* **2013**, *85*, 2970.
- (23) Vanheyne, S. *Science* **1974**, *183*, 656.
- (24) Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Ho, W. G. J. *J. Mol. Biol.* **1998**, *282*, 1043.
- (25) Kadioglu, A.; Andrew, P. W. *Trends Immunol.* **2004**, *25*, 143.
- (26) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.
- (27) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (28) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Nano Lett.* **2002**, *2*, 853.
- (29) Popovic, K.; Holyoake, J.; Pomès, R.; Privé, G. G. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 2908.

- (30) Leney, A. C.; Rezaei Darestani, R.; Li, J.; Nikjah, S.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Xiong, Z.; Prive, G. G.; Klassen, J. S. *Anal. Chem.* **2015**, *87*, 4402.
- (31) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.
- (32) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (33) Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1047.
- (34) Lin, H.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 104.
- (35) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477.
- (36) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Arch. Biochem. Biophys.* **2006**, *450*, 215.
- (37) Nakano, M.; Fukuda, M.; Kudo, T.; Miyazaki, M.; Wada, Y.; Matsuzaki, N.; Endo, H.; Handa, T. *J. Am. Chem. Soc.* **2009**, *131*, 8308.
- (38) Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinaige, B.; Boraston, A. B.; Klassen, J. S. *Glycobiology* **2015**, *25*, 170.
- (39) Moonens, K.; Bouckaert, J.; Coddens, A.; Tran, T.; Panjekar, S.; De Kerpel, M.; Cox, E.; Remaut, H.; De Greve, H. *Mol. Microbiol.* **2012**, *86*, 82.

## Chapter 7

### Conclusions and Future Work

#### 7.1 Conclusions

This work describes the development and application of ESI-MS methods to study noncovalent protein-carbohydrate interactions. The first four research projects (Chapter 2 – 5 in this thesis) focus on the characterization of binding between viral capsid proteins derived from human noroviruses and various carbohydrate ligands. The last research project (Chapter 6) not only elucidates a possible mechanism for protein-nanodisc containing glycolipids interactions, but also highlights the implementation of the ESI-MS based competitive binding assay, *proxy ligand* method, for quantifying protein-glycolipid binding affinities.

As a starting point, in Chapter 2, the *direct* ESI-MS assay was used to quantify the affinities of both full-length and C-terminal truncated P dimers (69 kDa) of a human norovirus (NoV) strain VA387 (GI.4 genogroup) with a panel of (47 oligosaccharides in total) soluble analogs of the histo-blood group antigens (HBGAs). A significant finding of this study is that all the HBGA oligosaccharides tested bind, albeit weakly (with intrinsic association constants,  $K_{a,int}$ ,  $\sim 2 \times 10^2$  to  $\sim 1.5 \times 10^3 \text{ M}^{-1}$ ), to the P dimers. Overall, the H and Lewis antigens exhibit weaker binding than the A and B antigens. The results are consistent with the binding model established based on X-ray crystallography, mutagenesis studies and molecular simulations, where VA387 P dimer can recognize HBGAs through either a  $\alpha$ -D-GalNAc /  $\alpha$ -D-Gal epitope, or a  $\alpha$ -L-Fuc epitope, using two nearby binding pockets. Consequently, the A and B oligosaccharides possess

both epitopes and thus are expected to exhibit stronger binding. The affinities are also affected by the precursor chain types of HBGAs, i.e. type 3 > 1  $\approx$  2  $\approx$  4  $\approx$  5  $\approx$  6 for H antigens; type 6 > 1  $\approx$  3  $\approx$  4  $\approx$  5 > 2 for A antigens; type 3 > 1  $\approx$  4  $\approx$  5  $\approx$  6 > 2 for B antigens. In contrast, the chain length of oligosaccharides, i.e., di- to hexa-saccharides, has little effect on the affinities. Among the HBGA oligosaccharides tested, B type 3 tetrasaccharide and A type 6 tetrasaccharide exhibit the highest affinities for the P dimer, with  $K_{a,int}$  values of  $(1.5 \pm 0.2) \times 10^3 \text{ M}^{-1}$  and  $(1.2 \pm 0.1) \times 10^3 \text{ M}^{-1}$ , respectively. Moreover, the truncated P dimer of VA387, which lacks a highly conserved Arg cluster at the C-terminus but is supposed to exist natively in vivo, shows similar affinities for the HBGA oligosaccharides to those for the full-length P dimer.

In Chapter 3, we reported on results of the catch-and-release (CaR)-ESI-MS assay for screening carbohydrate libraries against human NoV P particle (VA387, GII.4) to identify the specific interactions and to rank their relative affinities. Because of the high molecular weight of the P particle (~865 kDa), the identification of bound ligands directly from the m/z values measured by ESI-MS was not possible. Instead, all of the ligand-bound P particle ions at a given charge state (m/z) were isolated using a quadrupole mass filter and then activated using collision-induced dissociation (CID) to release these ligands (as ions) from the complex. The identity of the released ligands was then established from the measured m/z values, alone or in combination with ion mobility separation. The method was validated by successfully identifying HBGA ligands (which are known receptors for human NoVs) with  $K_{a,int} \geq 300 \text{ M}^{-1}$  from carbohydrate libraries for the P particle. More importantly, the relative abundances of the

released ligand ions are qualitatively consistent with the trends in their  $K_{a,int}$  values determined for the P dimer. Interestingly, the results of screening of a library containing 146 carbohydrates against the P particle reveal that, in addition to 16 oligosaccharides belonging to the family of HBGAs, another 12 oligosaccharides with the structures found in human milk as well as the cell walls of mycobacteria were identified as binders to the P particle. Notably, these newly found carbohydrate ligands show comparable affinities to those of HBGA receptors, as estimated from the relative abundance of released ligand ions. Taken together, the success of CaR-ESI-MS assay in this study represents an important extension of ESI-MS for detecting ligand interactions to large protein assemblies, where the mass spectrometer may not have sufficient resolution to distinguish the free and ligand-bound protein ions.

In Chapter 4, an ESI-MS based competitive binding assay, the *proxy protein* method was employed to quantify the intrinsic affinities of HBGA oligosaccharides for the virus-like particle (VLP, MW ~10.5 MDa) and P particle of NoV VA387. This method extends the applicability of ESI-MS to the cases where the protein-ligand complexes cannot be directly detected by ESI-MS either due to the protein's heterogeneity or due to its molecular weight being too large to obtain sufficient mass resolution. The *proxy protein* method relies on the uses of a proxy protein ( $P_{proxy}$ ) that binds to the ligand of interest with known affinity to monitor quantitatively the extent of ligand binding to the target protein. Using this method, the  $K_{a,int}$  values were measured for the P particle to thirteen HBGA oligosaccharides, which contain A, B and H epitopes, with variable sizes (disaccharide to tetrasaccharide) and different precursor chain types (type 1, 2, 3, 5 and 6),

as well as for the VLP to four HBGA oligosaccharides (including A and B trisaccharides and A and B type 6 tetrasaccharides). The intrinsic affinities of the HBGA oligosaccharides for the P particle range from 500 to 2300 M<sup>-1</sup>; while those for the VLP range from 1000 to 4000 M<sup>-1</sup>. To our knowledge, this is the first quantitative binding measurements for the carbohydrate interactions to the extremely large protein assemblies. Notably, the P particle exhibits affinities for the HBGAs ligands similar to those measured for the P dimer using the *direct* ESI-MS assay. However, the HBGA affinities for the VLP are consistently higher than those for the P dimer, but within a factor of three. Although the cause of the subtle binding affinity differences for these NoV capsid proteins is not clear, the results support the use of P dimers or P particles as surrogates to the VLP for human NoV-receptor binding studies.

In Chapter 5, we reported on the application of ESI-MS for screening of carbohydrate libraries against NoV capsid proteins (P dimer, P particle, VLP) originating from of two human NoV strains, VA387 (GII.4) and VA115 (GI.3), with the goal of discovering new carbohydrate ligands that may serve as alternative cellular receptors for human NoV infections. The results of CaR-ESI-MS screening against the P particle provided the first experimental evidence of interactions between human NoVs and gangliosides. In addition, the affinities measured for the ganglioside oligosaccharides to the P dimer, P particle and VLP of NoV VA387 as well as P dimer of NoV VA115 are comparable in magnitude ( $K_{a,int} \sim 10^2 \text{ M}^{-1} - 10^3 \text{ M}^{-1}$ ) to those reported for known HBGA receptors. Additional support of specific recognition of gangliosides by human NoVs was obtained by the binding measurements of the VLPs, P particles and glutathione

S-transferase (GST)-P domain fusion proteins to sialic acid-containing glycoconjugates, using enzyme-linked immunosorbent assays. The present finding suggests a new mechanism of human NoV-host interaction, one that involves HBGA and ganglioside receptors and co-receptors for attachment and penetration into host cells, and opens a new direction in human NoV research.

In Chapter 6, the results of a systematic ESI-MS study aimed at elucidating the processes that influence binding of water soluble proteins to glycolipid ligands incorporated into nanodiscs were described. ESI-MS measurements performed on solutions of cholera toxin B subunit homopentamer (CTB<sub>5</sub>) with nanodiscs containing GM1 demonstrated that proteins bind reversibly to glycolipid ligands associated nanodiscs and, in case of the protein possessing multiple ligand binding sites, are able to interact with glycolipids originating from different nanodiscs. In addition, the distribution of bound glycolipid is sensitive to both the nanodisc concentration and the number of glycolipid in each nanodisc. Moreover, the results of ESI-MS measurements performed on solutions of nanodiscs and picodiscs provided direct evidence for rapid glycolipid diffusion between nanodiscs and picodiscs. Based on this finding it is proposed that diffusion of glycolipids between nanodiscs can influence the nature of the protein-glycolipid complexes detected. While ESI-MS serves as a convenient method for detecting protein interactions with glycolipid in nanodiscs, the measured abundances of free and glycolipid-bound protein ions do not necessarily reflect solution composition. There is overwhelming evidence that, in solution, the glycolipid-bound proteins remain associated with nanodiscs and are only released (as protein-glycolipid complexes) in the gas phase. Consequently, different ESI-MS response

factors are expected for the free proteins and glycolipid-bound proteins.

Finally, a novel ESI-MS assay, *proxy ligand* method, which combines *direct* ESI-MS assay with competitive ligand-protein binding, was developed to quantify the protein affinities for the glycolipid ligand in nanodiscs. This method involves the introduction of a proxy ligand ( $L_{\text{proxy}}$ , which is usually an oligosaccharide) that binds to the protein (P) with known affinity and competes with the glycolipid ligand (L). The binding of P to L reduces the available amount of P in solution, leading to an increase in the relative concentration of  $PL_{\text{proxy}}$  complex. Consequently, the extent of PL binding can be deduced by monitoring the relative abundance of  $PL_{\text{proxy}}$  to P, from ESI-MS. Using the *proxy ligand* ESI-MS assay,  $K_a$  values for the interactions of CTB<sub>5</sub> with GM1, as well as a family 51 carbohydrate-binding module (CBM) with B type 2 neoglycolipid were quantified. A notable finding of this study is that the affinities of the proteins for the glycolipid ligands in nanodiscs are slightly lower (by a factor of  $\leq 5$ ) than those of the corresponding oligosaccharides in solution.

## 7.2 Future work

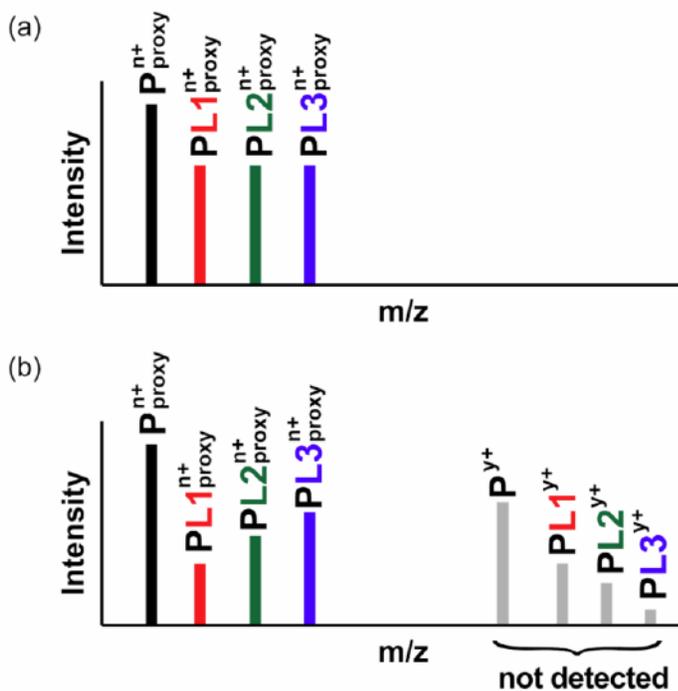
For over two decades' development, ESI-MS assays have been successfully implemented to studying many protein-ligand noncovalent interactions in vitro. However, there are still areas that require novel ESI-MS methods to be developed to overcome or compensate current limitations associated with the technique, and this is the driving force of our future research. Several possible extensions of the current studies are given in the following sections.

### 7.2.1 Rapid and quantitative screening of oligosaccharide libraries against carbohydrate-binding proteins using the *proxy protein* ESI-MS assay

The *proxy protein* method described in Chapter 4 extends the applicability of ESI-MS assay for quantifying carbohydrate binding to proteins that are not readily detected / resolved by ESI-MS, either due to the size or heterogeneity. Although this method was originally developed for studying interactions between the target protein (P) and a single ligand, the assay can, in principle, be extended to screening mixtures of carbohydrates, whereby the bindings of multiple carbohydrate ligands to P can be simultaneously detected and quantified. The key to this method is that the proxy protein ( $P_{\text{proxy}}$ ) used must show a broad specificity, as well as relatively high affinities for all the ligands in the library. Considering a simple situation where P is incubated with a carbohydrate library containing  $x$  ligands ( $L_1, L_2, L_3, \dots, L_x$ ) in the presence of a monovalent  $P_{\text{proxy}}$ , any ligand ( $L_x$ ) interacting with P results in a decrease in the relative concentration of  $P_{\text{proxy}}L_x$  complex while those non-binders of P have no such effect. Consequently, the relative ion abundance ratio for each ligand, i.e.,  $R_{\text{proxy},L_x} = Ab(P_{\text{proxy}}L_x) / Ab(P_{\text{proxy}})$ , which is directly determined from ESI-MS, enables the affinity of each  $PL_x$  interaction to be quantified. A schematic illustration of this method is given in Figure 7.1.

As a starting point, to demonstrate the application of the *proxy protein* ESI-MS assay for carbohydrate library screening, a control experiment was performed to screen a three-component library containing H type 6 trisaccharide, and B type 6 and A type 2 tetrasaccharides against the family 51 carbohydrate-binding module (CBM, MW 20,735 Da),<sup>3</sup> using the galectin-3C (Gal-3C, MW 16,330 Da) as  $P_{\text{proxy}}$ . Gal-3C binds to many human milk oligosaccharides (HMOs) as well

as A, B and H type 1, 2, 5, 6 HBGAs with relatively strong affinities ( $K_a \sim 10^4 - 10^5 \text{ M}^{-1}$ ),<sup>1,2</sup> The preliminary results indicate that the binding affinities obtained using this approach agree, within a factor of 2, with the values obtained from the individual ESI-MS measurements.<sup>2</sup> Current efforts are being directed toward using this assay to detect and quantify the bindings of the NoV P particle, VLP and other carbohydrate-binding proteins to carbohydrate libraries containing as many as 10 HMOs and HBGA oligosaccharides.



**Figure 7.1.** Schematic illustration of the *proxy protein* ESI-MS assay applied for quantifying bindings of a carbohydrate library containing 3 ligands (L1, L2 and L3) to the target protein (P). Simulation mass spectra of the proxy protein ( $P_{\text{proxy}}$ ) with the carbohydrates in the absence (a), and presence (b) of P.

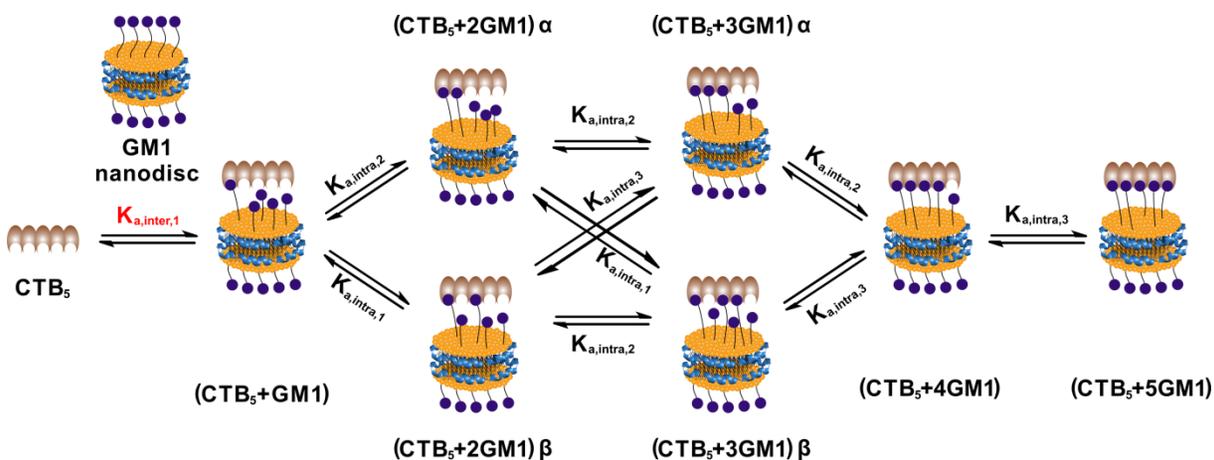
## 7.2.2 Quantifying protein-glycolipid interactions in different lipid environments

A relevant extension of the work in Chapter 6 is to investigate the energetics of protein-glycolipid interactions in different lipid environments using the *proxy ligand* ESI-MS method. In general, study of protein-glycolipid interaction in vitro requires a model membrane that can solubilize the glycolipid and mimic the environment in living cells. However, it is known that the protein-glycolipid interactions are sensitive to the lipid environment, i.e., the nature and composition of lipids, and membrane models used.<sup>4,5</sup> In this project, glycolipids would be incorporated into nanodiscs (consisting of the membrane scaffold protein (MSP1E1) and POPC or DMPC), picodiscs (consisting of the saposin A protein and POPC),<sup>6,7</sup> and POPC micelles/vesicles. The glycolipid affinities for the target protein would be measured using the *proxy ligand* ESI-MS assay as described in Chapter 6. The interaction of CBM with synthetic histo-blood group A type 2 tetrasaccharide neoglycolipid (A<sub>2</sub>NGL),  $\alpha$ -D-GalNAc-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-1,2-di-O-dodecyl-sn-glycero and B type 2 tetrasaccharide neoglycolipid (B<sub>2</sub>NGL),  $\alpha$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→2)]- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-1,2-di-O-dodecyl-sn-glycero would serve as the model protein-glycolipid complexes for this study. Preliminary results suggest that the affinities of protein-glycolipid binding are indeed affected by both the saccharide moiety of the glycolipid and the model membranes (i.e., nanodiscs, picodiscs and micelles/vesicles) that are used for presenting the glycolipid. However, the interactions are insensitive to the percentage of the glycolipid in a same model membrane. These results are consistent with the view that lipid environment can strongly influence protein binding. Future research is required to establish the effects of membrane

composition on the energetics of protein-glycolipid interactions.

### 7.2.3 Dissecting multivalent protein-glycolipid interactions

The application of the *proxy ligand* ESI-MS assay for quantifying the CTB<sub>5</sub>-GM1 nanodisc interactions was also described in Chapter 6, where the nanodisc concentrations and percentages of GM1 in the nanodisc are constrained to prevent the multivalent CTB<sub>5</sub>-GM1-nanodisc complexes formation. Actually, multiple glycolipids can be presented on both sides of the membrane bilayer of nanodiscs. Therefore, an extension of this work is to establish a more general binding model that allows study of the multivalent interactions<sup>8,9</sup> between CTB<sub>5</sub> and GM1 nanodiscs. An outstanding feature of multivalent protein-glycolipid interaction is that the binding is described by intermolecular ( $K_{a,inter}$ ) and intramolecular ( $K_{a,intra}$ ) association constants, which account for the first intermolecular attachment and subsequent intramolecular binding events, respectively. In addition, CTB<sub>5</sub>-GM1 binding exhibits slight positive allosteric cooperativity as demonstrated by the previous isothermal titration calorimetry (ITC)<sup>10</sup> and ESI-MS<sup>11</sup> analyses. Consequently, based on the binding model described in Chapter 6, which is an extension of the Homans' model,<sup>10</sup> an intrinsic intermolecular association constants  $K_{a,inter,1}$ , as well as three intrinsic intramolecular association constants  $K_{a,intra,1}$ ,  $K_{a,intra,2}$  and  $K_{a,intra,3}$  are required to characterize the multivalent CTB<sub>5</sub>-GM1-ND binding (Figure 7.2). The subscripts 1, 2 or 3 of the above association constants represent GM1 binding to CTB<sub>5</sub> sites with zero, one or two occupied nearest neighbour subunits, respectively.



**Figure 7.2.** Schematic representation of the multivalent CTB<sub>5</sub>-GM1-nanodisc interaction based on the cooperative binding model<sup>10</sup> proposed by Homan and coworkers. The multivalent binding is described by intermolecular ( $K_{a,inter,1}$ ) and intramolecular ( $K_{a,intra,1}$ ,  $K_{a,intra,2}$ ,  $K_{a,intra,3}$ ) association constants. The subscripts 1, 2 and 3 represent the case of GM1 binding to a subunit with zero, one or two ligand-bound nearest neighbour subunits, respectively. In this case, eight distinct species associated with CTB<sub>5</sub> are present in solution.

For this project, the *proxy ligand* ESI-MS assay and nanodisc technology would be combined to dissect the intermolecular and intramolecular intrinsic association constants for multivalent CTB<sub>5</sub>-GM1-ND interactions. Nanodiscs containing 5%, 8% and 10% GM1, which correspond to an average of 5, 8 and 10 molecules of GM1 on each side of a nanodisc, respectively, would be prepared. Following a similar procedure as described in Chapter 6, the GM1 nanodisc would be titrated into solutions containing fixed concentrations of CTB<sub>5</sub> and GM1 pentasaccharide (GM1<sub>os</sub>, L<sub>proxy</sub>). The change of ion abundance ratio of  $R_{proxy,5} = Ab(CTB_5+5GM1_{os}) / Ab(CTB_5+4GM1_{os})$ , measured by ESI-MS, would be used to monitor the

extent of multivalent CTB<sub>5</sub>-GM1-ND binding. Hence, the intrinsic association constants can be extracted by fitting the binding data to a suitable binding model. Notably, the *proxy ligand* ESI-MS assay combined with nanodisc technology provides an opportunity to look into the multivalent protein-glycolipid interactions and may also reveal new insights into the effects of the number of glycolipids in nanodiscs on protein binding.

### 7.3 References

- (1) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E. G.; Yagi, F.; Kasai, K. *Biochim. Biophys. Acta Gen. Subj.* **2002**, *1572*, 232.
- (2) Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. *Glycobiology* **2015**, *25*, 170.
- (3) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.
- (4) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.
- (5) Lingwood, D.; Binnington, B.; Róg, T.; Vattulainen, I.; Grzybek, M.; Coskun, Ü.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.
- (6) Popovic, K.; Holyoake, J.; Pomès, R.; Privé, G. G. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2908.
- (7) Leney, A. C.; Rezaei Darestani, R.; Li, J.; Nikjah, S.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Xiong, Z. J.; Privé, G. G.; Klassen, J. S. *Anal. Chem.* **2015**, *87*, 4402.

- (8) Kitov, P. I.; Bundle, D. R., *J. Am. Chem. Soc.* **2003**, *125*, 16271.
- (9) Sacchettini, J. C.; Baum, L. G.; Brewer, C. F. *Biochemistry* **2001**, *40*, 3009.
- (10) Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1047.
- (11) Lin, H.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 104.

## List of References

### Chapter 1

- (1) Ghazarian, H.; Idoni, B.; Oppenheimer, S. B. *Acta Histochem.* **2011**, *113*, 236.
- (2) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637.
- (3) Sharon, N.; Lis, H. *Sci. Am.* **1993**, *268*, 82.
- (4) Quioco, F. A. *Annu. Rev. Biochem.* **1986**, *55*, 287.
- (5) Holgersson, J.; Gustafsson, A.; Breimer, M. E. *Immunol. Cell Biol.* **2005**, *83*, 694.
- (6) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *65*, 441.
- (7) Sacchettini, J. C.; Baum, L. G.; Brewer, C. F. *Biochemistry* **2001**, *40*, 3009.
- (8) Palmer, R. A.; Niwa, H., *Biochem. Soc. Trans.* **2003**, *31*, 973.
- (9) Greco, A.; Ho, J. G. S.; Lin, S.-J.; Palcic, M. M.; Rupnik, M.; Ng, K. K.-S. *Nat. Struct. Mol. Biol.* **2006**, *13*, 460.
- (10) Ho, J. G. S.; Greco, A.; Rupnik, M.; Ng, K. K.-S. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18373.
- (11) Staunton, D.; Owen, J.; Campbell, I. D. *Acc. Chem. Res.* **2002**, *36*, 207.
- (12) Larsen, K.; Thygesen, M. B.; Guillaumie, F.; William, G. T.; Willats, W. G. T.; Jensen, K. J. *Carb. Res.* **2006**, *341*, 1209.
- (13) Wang, D. N.; Liu, S. Y.; Trummer, B. J.; Deng, C.; Wang, A. L. *Nat. Biotechnol.* **2002**, *20*, 275.
- (14) Liu, Y.; Palma, A. S.; Feizi, T. *Biol. Chem.* **2009**, *390*, 647.

- (15) Schuck, P. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 541.
- (16) Homola, J. *Anal. Bioanal. Chem.* **2003**, *377*, 528.
- (17) Daghestani, H. N.; Day, B. W. *Sensors* **2010**, *10*, 9630.
- (18) De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. *Cell. Mol. Bioeng.* **2008**, *1*, 204.
- (19) Karlsson, R. *J. Mol. Recognit.* **2004**, *17*, 151.
- (20) Disney, M. D.; Seeberger, P. H. *Drug Discov. Today: Targets* **2004**, *3*, 151.
- (21) Grant, O. C.; Smith, H. M.; Firsova, D.; Fadda, E.; Woods, R. J. *Glycobiology* **2014**, *24*, 17.
- (22) Fais, M.; Karamanska, R.; Russell, D. A.; Field, R. A. *J. Cereal Sci.* **2009**, *50*, 306–311.
- (23) Day, Y. S. N.; Baird, C. L.; Rich, R. L.; Myszka, D. G. *Protein Sci.* **2002**, *11*, 1017.
- (24) Moll, D.; Schweinsberg, S.; Hammann, C.; Friedrich, W.; Herberg, F. W. *Biol. Chem.* **2007**, *388*, 163.
- (25) Saboury, A. A. *J. Iran. Chem. Soc.* **2006**, *3*, 1.
- (26) Wilcox, D. E. *Inorg. Chim. Acta* **2008**, *361*, 857.
- (27) Velázquez Campoy, A.; Freire, E. *Biophys. Chem.* **2005**, *115*, 115.
- (28) Wishart, D. *Curr. Pharm. Biotechnol.* **2005**, *6*, 105.
- (29) Zuiderweg, E. R. P. *Biochemistry* **2001**, *41*, 1.
- (30) Goldflam, M.; Tarrago, T.; Gairi, M.; Giralt, E., *Methods Mol. Biol.* **2012**, *831*, 233.
- (31) Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.* **1999**, *38*, 1784.
- (32) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108.

- (33) Haselhorst, T.; Lamerz, A. C.; Itzstein, M., *Methods Mol. Biol.* **2009**, *534*, 375.
- (34) Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. *J. Chem. Educ.* **2011**, *88*, 990.
- (35) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294.
- (36) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818.
- (37) Van Dongen, W. D.; Heck, A. J. R. *Analyst* **2000**, *125*, 583.
- (38) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605.
- (39) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945.
- (40) Cederkvist, F. H.; Zamfir, A. D.; Bahrke, S.; Eijssink, V. G.; Sorlie, M.; Peter-Katalinic, J.; Peter, M. G., *Angew. Chem. Int. Ed.* **2006**, *45*, 2429.
- (41) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (42) Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587.
- (43) Jecklin, M. C.; Touboul, D.; Bovet, C.; Wortmann, A.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 332.
- (44) Jecklin, M. C.; Touboul, D.; Jain, R.; Toole, E. N.; Tallarico, J.; Drueckes, P.; Ramage, P.; Zenobi, R. *Anal. Chem.* **2009**, *81*, 408.
- (45) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1.
- (46) Kitova, E. N.; El-Hawiet, A.; Schnier, P.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.
- (47) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.

- (48) Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discov.* **2006**, *5*, 585.
- (49) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.
- (50) Feizi, T.; Chai, W. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 582.
- (51) Lopez, P. H.; Schnaar, R. L. *Methods Enzymol.* **2006**, *417*, 205.
- (52) Kleinschmidt, J. H. Ed. *Lipid-Protein Interactions: Methods and Protocols*, **2014**, *974*, Springer: New York.
- (53) Catimel, B.; Scott, A. M.; Lee, F. T.; Hanai, N.; Ritter, G.; Welt, S.; Old, L. J.; Burgess, A. W.; Nice, E. C. *Glycobiology* **1998**, *8*, 927.
- (54) Campanero-Rhodes, M. A.; Smith, A.; Chai, W.; Sonnino, S.; Mauri, L.; Childs, R. A.; Zhang, Y.; Ewers, H.; Helenius, A.; Imberty, A.; Feizi, T. *J. Virol.* **2007**, *81*, 12846.
- (55) Feizi, T. *Ann. N. Y. Acad. Sci.* **2013**, *1292*, 33.
- (56) Song, X.; Heimbürg-Molinario, J.; Cummings, R. D.; Smith, D. F. *Curr. Opin. Chem. Biol.* **2014**, *18*, 70.
- (57) Palma, A. S.; Feizi, T.; Childs, R. A.; Chai, W.; Liu, Y. *Curr. Opin. Chem. Biol.* **2014**, *18*, 87.
- (58) Lingwood, C. A.; Manis, A.; Mahfoud, R.; Khan, F.; Binnington, B.; Mylvaganam, M. *Chem. Phys. Lipids* **2010**, *163*, 27.
- (59) Hooper, N. M. *Curr. Biol.* **1998**, *8*, R114.
- (60) Lingwood, C. A. *Glycoconjugate J.* **1996**, *13*, 495.
- (61) Lingwood, D.; Binnington, B.; Rog, T.; Vattulainen, I.; Grzybek, M.; Coskun, U.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.

- (62) Mahfoud, R.; Manis, A.; Binnington, B.; Ackerley, C.; Lingwood, C. A. *J. Biol. Chem.* **2010**, *285*, 36049.
- (63) Yahi, N.; Aulas, A.; Fantini, J. *PLoS ONE* **2010**, *5*, e9079.
- (64) Gallegos, K. M.; Conrady, D. G.; Karve, S. S.; Gunasekera, T. S.; Herr, A. B.; Weiss, A. A. *PLoS ONE* **2012**, *7*, e30368.
- (65) Hammache, D.; Piéroni, G.; Yahi, N.; Delézay, O.; Koch, N.; Lafont, H.; Tamalet, C.; Fantini, J. *J. Biol. Chem.* **1998**, *273*, 7967.
- (66) Czogalla, A.; Grzybek, M.; Jones, W.; Coskun, U. *Biochim. Biophys. Acta.* **2014**, *1841*, 1049.
- (67) Nath, A.; Atkins, W. M.; Sligar, S. G. *Biochemistry* **2007**, *46*, 2059.
- (68) Bayburt, T. H.; Sligar, S. G. *FEBS Letters* **2010**, *584*, 1721.
- (69) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477.
- (70) Cho, H.; Wu, M.; Bilgin, B.; Walton, S. P.; Chan, C. *Proteomics* **2012**, *12*, 3273.
- (71) Sanghera, N.; Correia, B. E.; Correia, J. R.; Ludwig, C.; Agarwal, S.; Nakamura, H. K.; Kuwata, K.; Samain, E.; Gill, A. C.; Bonev, B. B.; Pinheiro, T. J. *Chem. Biol.* **2011**, *18*, 1422.
- (72) Shi, J.; Yang, T.; Kataoka, S.; Zhang, Y.; Diaz, A. J.; Cremer, P. S. *J. Am. Chem. Soc.* **2007**, *129*, 5954.
- (73) Rao, C. S.; Lin, X.; Pike, H. M.; Molotkovsky, J. G.; Brown, R. E. *Biochemistry* **2004**, *43*, 13805.

- (74) Chen, W. C.; Kawasaki, N.; Nycholat, C. M.; Han, S.; Pilotte, J.; Crocker, P. R.; Paulson, J. C. *PLoS One* **2012**, *7*, e39039.
- (75) Borch, J.; Torta, F.; Sligar, S. G.; Roepstorff, P. *Anal. Chem.* **2008**, *80*, 6245.
- (76) Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618.
- (77) Leney, A.; Fan, X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271.
- (78) Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 972A.
- (79) Fenn, J. B. *Angew. Chem. Int. Ed.* **2003**, *42*, 3871.
- (80) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804.
- (81) Kebarle, P.; Verkerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898.
- (82) Taylor, G. I. *J. Fluid Mech.* **1965**, *2*, 1.
- (83) Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. *J. Phys. Chem. A* **2002**, *106*, 9957.
- (84) Peschke, M.; Verkerk, U. H.; Kebarle, P. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1424.
- (85) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287;
- (86) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1979**, *71*, 4451.
- (87) Dole, M.; Mack, L. L.; Hines, R. L. *J. Chem. Phys.* **1968**, *49*, 2240.
- (88) de la Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93.
- (89) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2.
- (90) Ahadi, E.; Konermann, L. *J. Phys. Chem. B* **2012**, *116*, 104.
- (91) Konermann, L.; Rodriguez, A. D.; Liu, J. *Anal. Chem.* **2012**, *84*, 6798.
- (92) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1.

- (93) Karas, M.; Bahr, U.; Dulcks, T. *Fresenius J. Anal. Chem.* **2000**, *366*, 669.
- (94) Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300.
- (95) De Hoffmann, E.; Stroobant, V. *Mass Spectrometry Principles and Applications* 3th Ed. **2007**, John Wiley & Sons: New York.
- (96) March, R. E.; Hughes, R. J. *Quadrupole Storage Mass Spectrometry*, **1989**, John Wiley & Sons: New York.
- (97) Pedder, R. E. *Ardara Technologies Technical Note* **2009**, TN\_3005A.
- (98) Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrom. Rev.* **2000**, *19*, 65.
- (99) Mamyrin, B. A. *Int. J. Mass Spectrom.* **2001**, *206*, 251.
- (100) 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. *Int. J. Mass Spectrom.* **2007**, *261*, 1.
- (101) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401.
- (102) McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 599.
- (103) Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49.
- (104) Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339.
- (105) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265.
- (106) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857.
- (107) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528.

- (108) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859.
- (109) Polfer, N. C. *Chem. Soc. Rev.* **2011**, *40*, 2211.
- (110) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9*, 4205.
- (111) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135*, 12646.
- (112) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640.
- (113) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194.
- (114) Wytttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2007**, *58*, 511.
- (115) Versluis, C.; Heck, A. J. R. *Int. J. Mass Spectrom.* **2001**, *210*, 637.
- (116) McCammon, M. G.; Hernandez, H.; Sobott, F.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 5950.
- (117) Ilag, L. L.; Ubarretxena-Belandia, I.; Tate, C. G.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 14362.
- (118) Borch, J.; Jorgensen, T. J. D.; Roepstorff, P. *Curr. Opin. Chem. Biol.* **2005**, *9*, 509.
- (119) Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, *13*, 597.
- (120) Sharon, M.; Taverner, T.; Ambroggio, X. I.; Deshaies, R. J.; Robinson, C. V. *PLoS. Biol.* **2006**, *4*, 1314.
- (121) Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V. *Chem. Rev.* **2007**, *107*, 3544.

- (122) Lorenzen, K.; Vannini, A.; Cramer, P.; Heck, A. J. R. *Structure* **2007**, *15*, 1237.
- (123) Damoc, E.; Fraser, C. S.; Zhou, M.; Videler, H.; Mayeur, G. L.; Hershey, J. W. B.; Doudna, J. A.; Robinson, C. V.; Leary, J. A. *Mol. Cell. Proteomics* **2007**, *6*, 1135.
- (124) Tešić, M.; Wicki, J.; Poon, D. K. Y.; Withers, S. G.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 64.
- (125) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50.
- (126) Karasek, F. W. *Anal. Chem.* **1974**, *46*, 710A.
- (127) Creaser, C. S. ; Griffiths, J. R.; Bramwell, C. J.; Noreen, S.; Hill, C. A. ; Paul Thomas, C. L. *Analyst* **2004**, *129*, 984.
- (128) Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H. *J. Mass Spectrom.* **2008**, *43*, 1.
- (129) McCullough, B. J.; Kalapothakis, J.; Eastwood, H.; Kemper, P.; MacMillan, D.; Taylor, K.; Dorin, J.; Barran, P. E. *Anal. Chem.* **2008**, *80*, 6336.
- (130) Sergio, A.; Manel, A.; Marcelo, B. *Anal. Chim. Acta* **2011**, *703*, 114.
- (131) Cohen, M. J.; Karasek, F. W. *J. Chromatogr. Sci.* **1970**, *8*, 330.
- (132) Buryakov, I. A.; Krylov, E. V.; Nazarov, E. G.; Rasulev, U. K. *Int. J. Mass Spectrom. Ion Processes* **1993**, *128*, 143.
- (133) Shvartsburg, A. A.; Smith, R. D. *Anal. Chem.* **2008**, *80*, 9689.
- (134) Wildgoose, J.; McKenna, T.; Hughes, C.; Giles, K.; Pringle, S.; Campuzano, I.; Langridge, J.; Bateman, R. H. *Mol. Cell. Proteomics* **2006**, *5*, S14.
- (135) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat.*

*Protocols* **2008**, *3*, 1139.

- (136) Giles, K.; Williams, J. P.; Campuzano, I. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 1559.
- (137) Amster, I. J. *J. Am. Soc. Mass Spectrom.* **1996**, *31*, 1325.
- (138) Buchanan, M. V.; Hettich, R. L. *Anal. Chem.* **1993**, *65*, A245.
- (139) Marshall, A. G.; Grosshans, P. B. *Anal. Chem.* **1991**, *63*, A215.
- (140) Marshall, A. G.; Hendrickson, C. L. *Int. J. Mass Spectrom.* **2002**, *215*, 59.
- (141) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1.
- (142) El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S., *Anal. Chem.* **2012**, *84*, 3867.
- (143) Cheng, X. H.; Chen, R. D.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D.; Gao, J. M.; Sigal, G. B.; Mammen, M.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 8859.
- (144) Wortmann, A.; Rossi, F.; Lelais, G.; Zenobi, R. *J. Mass Spectrom.* **2005**, *40*, 777.
- (145) Wortmann, A.; Jecklin, M. C.; Touboul, D.; Badertscher, M.; Zenobi, R. *J. Mass Spectrom.* **2008**, *43*, 600.
- (146) Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 416.
- (147) Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801.
- (148) Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. *J. Am. Chem. Soc.* **1996**, *118*, 8646.

- (149) El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1893.
- (150) Hunter, E. P.; Lias, S. G. *Phys. Chem. Ref. Data* **1998**, *27*, 413.
- (151) Kitova, E. N.; Seo, M.; Roy, P. N.; Klassen, J. S. *J. Am. Chem. Soc.* **2008**, *130*, 1214.
- (152) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2005**, *77*, 3060.
- (153) Hossain, B. M.; Konermann, L. *Anal. Chem.* **2006**, *78*, 1613.
- (154) Daubenfeld, T.; Bouin, A. P.; van der Rest, G. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1239.
- (155) Lane, L. A.; Ruotolo, B. T.; Robinson, C. V.; Favrin, G.; Benesch, J. L. P. *J. Mass Spectrom.* **2009**, *283*, 169.
- (156) Shimon, L.; Sharon, M.; Horovitz, A. *Biophys. J.* **2010**, *99*, 1645.
- (157) Sun, N.; Sun, J.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1242.
- (158) Kitova, E. N.; Soya, N.; Klassen, J. S. *Anal. Chem.* **2011**, *83*, 5160.
- (159) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (160) Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 8301.
- (161) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472.
- (162) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2010**, *82*, 2170.
- (163) Chitta, R. K.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1031.
- (164) Wilcox, J. M.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **2008**, *80*, 2365.

- (165) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spectrom.* **2003**, *38*, 491.
- (166) Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708.
- (167) Mathur, S.; Badertscher, S.; Scott, M.; Zenobi, R. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6187.
- (168) Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708.
- (169) Van Berkel, G. J.; Asano, K. G.; Schnier, P. D. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 853.
- (170) Yang, P.; Cooks, R. G.; Ouyang, Z.; Hawkridge, A. M.; Muddiman, D. C. *Anal. Chem.* **2005**, *77*, 6174.
- (171) Yao, Y.; Shams-Ud-Doha, K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2015**, *26*, 98.

## Chapter 2

- (1) Lindesmith, L.; Moe, C.; Marionneau, S.; Ruvoen, N.; Jiang, X.; Lindbland, L.; Stewart, P.; LePendu, J.; Baric, R. *Nat. Med.* **2003**, *9*, 548.
- (2) Tan, M.; Jiang, X. *Expet. Rev. Mol. Med.* **2007**, *9*, 1.
- (3) Patel, M. M.; Widdowson, M.-A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (4) Tan, M.; Jiang, X. *Trends Microbiol.* **2005**, *13*, 285.
- (5) Yang, Y.; Xia, M.; Tan, M.; Huang, P.; Zhong, W.; Pang, X. L.; Lee, B. E.; Meller, J.;

- Wang, T.; Jiang, X. *J. Virol.* **2010**, *84*, 9595.
- (6) Baert, L.; Mattison, K.; Loisy-Hamon, F.; Harlow, J.; Martyres, A.; Lebeau, B.; Stals, A.; Van Coillie, E.; Herman, L.; Uyttendaele, M. *Int. J. Food Microbiol.* **2011**, *151*, 261.
- (7) Cannon, J. L.; Lindesmith, L. C.; Donaldson, E. F.; Saxe, L.; Baric, R. S.; Vinje, J. *J. Virol.* **2009**, *83*, 5363.
- (8) Johnston, C. P.; Qiu, H.; Ticehurst, J. R.; Dickson, C.; Rosenbaum, P.; Lawson, P.; Stokes, A. B.; Lowenstein, C. J.; Kaminsky, M.; Cosgrove, S. E.; Green, K. Y.; Perl, T. *M. Clin. Infect. Dis.* **2007**, *45*, 534.
- (9) Jiang, X.; Wang, M.; Graham, D. Y.; Estes, M. K. *J. Virol.* **1992**, *66*, 6527.
- (10) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (11) Cao, S.; Lou, Z.; Tan, M.; Chen, Y.; Liu, Y.; Zhang, Z.; Zhang, X. C.; Jiang, X.; Li, X.; Rao, Z. *J. Virol.* **2007**, *81*, 5949.
- (12) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (13) Tan, M.; Huang, P.; Xia, M.; Fang, P.; Zhong, W.; McNeal, M.; Wei, C.; Jiang, W.; Jiang, X. *J. Virol.* **2011**, *85*, 753.
- (14) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (15) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.
- (16) Bereszczak, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R.

*J. Struct. Biol.* **2012**, *177*, 273.

- (17) Hardy, M. E.; White, L. J.; Ball, J. M.; Estes, M. K. *J. Virol.* **1995**, *69*, 1693.
- (18) Greenberg, H. B.; Valdesuso, J. R.; Kalica, A. R.; Wyatt, R. G.; McAuliffe, V. J.; Kapikian, A. Z.; Chanock, R. M. *J. Virol.* **1981**, *37*, 994.
- (19) Tan, M.; Meller, J.; Jiang, X. *J. Virol.* **2006**, *80*, 7322.
- (20) Bu, W.; Mamedova, A.; Tan, M.; Xia, M.; Jiang, X.; Hegde, R. S. *J. Virol.* **2008**, *82*, 5340.
- (21) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendou, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.
- (22) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang, X. *J. Virol.* **2005**, *79*, 6714.
- (23) Oriol, R. *J. Immunogenet.* **1990**, *17*, 235.
- (24) Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1.
- (25) Tan, M.; Xia, M.; Cao, S.; Huang, P.; Farkas, T.; Meller, J.; Hegde, R. S.; Li, X.; Rao, Z.; Jiang, X. *Virology* **2008**, *379*, 324.
- (26) de Rougemont, A.; Ruvoen-Clouet, N.; Simon, B.; Estienney, M.; Elie-Caille, C.; Aho, S.; Pothier, P.; Le Pendu, J.; Boireau, W.; Belliot, G. *J. Virol.* **2011**, *85*, 4057.
- (27) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (28) Chen, Y.; Tan, M.; Xia, M.; Hao, N.; Zhang, X. C.; Huang, P.; Jiang, X.; Li, X.; Rao, Z.

- PLoS Pathog.* **2011**, *7*, e1002152.
- (29) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.
- (30) Tan, M.; Zhong, W.; Song, D.; Thornton, S.; Jiang, X. *J. Med. Virol.* **2004**, *74*, 641.
- (31) Tan, M.; Xia, M.; Chen, Y.; Bu, W.; Hegde, R. S.; Meller, J.; Li, X.; Jiang, X. *PLoS One* **2009**, *4*, e5058.
- (32) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.
- (33) Tan, M.; Jiang, X. *Nanomedicine* **2012**, *7*, 1.
- (34) Hansman, G. S.; Shahzad-ul-Hussan, S.; McLellan, J. S.; Chuang, G.-Y.; Georgiev, I.; Shimoike, T.; Katayama, K.; Bewley, C. A.; Kwong, P. D. *J. Virol.* **2012**, *86*, 284.
- (35) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945.
- (36) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423.
- (37) Meloncelli, P. J.; Lowary, T. L. *Aust. J. Chem.* **2009**, *62*, 558.
- (38) Meloncelli, P. J.; Lowary, T. L. *Carbohydr. Res.* **2010**, *345*, 2305.
- (39) Meloncelli, P. J.; West, L. J.; Lowary, T. L. *Carbohydr. Res.* **2011**, *346*, 1406.
- (40) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.
- (41) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (42) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.

- (43) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472.
- (44) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2010**, *82*, 2170.
- (45) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.
- (46) Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 416.
- (47) Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801.
- (48) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.

### Chapter 3

- (1) Patel, M. M.; Widdowson, M.-A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (2) Hutson, A. M.; Atmar, R. L.; Marcus, D. M.; Estes, M. K. *J. Virol.* **2003**, *77*, 405.
- (3) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendou, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.
- (4) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang, X. *J. Virol.* **2005**, *79*, 6714.
- (5) Tan, M.; Jiang, X. *Trends Microbiol.* **2005**, *13*, 285.
- (6) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem.*

*Int. Ed.* **2012**, *51*, 928.

- (7) Tan, M.; Jiang, X. *Expert. Rev. Mol. Med.* **2007**, *9*, 1.
- (8) Tan, M.; Jiang, X. *PLoS pathogens* **2010**, *6*, e1000983.
- (9) Tan, M.; Jiang, X. *Trends Microbiol.* **2011**, *19*, 382.
- (10) Oriol, R. *J. Immunogenet.* **1990**, *17*, 235.
- (11) Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1.
- (12) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. *Glycobiology* **2013**, *23*, 276.
- (13) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (14) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (15) Tan, M.; Huang, P.; Meller, J.; Zhong, W.; Farkas, T.; Jiang, X. *J. Virol.* **2003**, *77*, 12562.
- (16) Jiang, X.; Wang, M.; Graham, D. Y.; Estes, M. K. *J. Virol.* **1992**, *66*, 6527.
- (17) Shoemaker, G. K.; van Duijn, E.; Crawford, S. E.; Uetrecht, C.; Baclayon, M.; Roos, W. H.; Wuite, G. J. L.; Estes, M. K.; Prasad, B. V. V.; Heck, A. J. R. *Mol. Cell. Proteomics* **2010**, *9*, 1742.
- (18) Tan, M.; Meller, J.; Jiang, X. *J. Virol.* **2006**, *80*, 7322.
- (19) Tan, M.; Xia, M.; Cao, S.; Huang, P.; Farkas, T.; Meller, J.; Hegde, R. S.; Li, X.; Rao, Z.; Jiang, X. *Virology* **2008**, *379*, 324.
- (20) Bu, W.; Mamedova, A.; Tan, M.; Xia, M.; Jiang, X.; Hegde, R. S. *J. Virol.* **2008**, *82*,

5340.

- (21) Cao, S.; Lou, Z.; Tan, M.; Chen, Y.; Liu, Y.; Zhang, Z.; Zhang, X. C.; Jiang, X.; Li, X.; Rao, Z. *J. Virol.* **2007**, *81*, 5949.
- (22) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.
- (23) Chen, Y.; Tan, M.; Xia, M.; Hao, N.; Zhang, X. C.; Huang, P.; Jiang, X.; Li, X.; Rao, Z. *PLoS Pathog.* **2011**, *7*, e1002152.
- (24) Hansman, G. S.; Biertumpfel, C.; Georgiev, I.; McLellan, J. S.; Chen, L.; Zhou, T.; Katayama, K.; Kwong, P. D. *J. Virol.* **2011**, *85*, 6687.
- (25) Shanker, S.; Choi, J. M.; Sankaran, B.; Atmar, R. L.; Estes, M. K.; Prasad, B. V. V. *J. Virol.* **2011**, *85*, 8635.
- (26) Kubota, T.; Kumagai, A.; Ito, H.; Furukawa, S.; Someya, Y.; Takeda, N.; Ishii, K.; Wakita, T.; Narimatsu, H.; Shirato, H. *J. Virol.* **2012**, *86*, 11138.
- (27) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (28) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.
- (29) Tan, M.; Huang, P.; Xia, M.; Fang, P.; Zhong, W.; McNeal, M.; Wei, C.; Jiang, W.; Jiang, X. *J. Virol.* **2011**, *85*, 753.
- (30) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.
- (31) Tan, M.; Zhong, W.; Song, D.; Thornton, S.; Jiang, X. *J. Med. Virol.* **2004**, *74*, 641.
- (32) Tan, M.; Xia, M.; Chen, Y.; Bu, W.; Hegde, R. S.; Meller, J.; Li, X.; Jiang, X. *PLoS*

*One* **2009**, *4*, e5058.

- (33) Shang, J.; Cheng, F.; Dubey, M.; Kaplan, J. M.; Rawal, M.; Jiang, X.; Newburg, D. S.; Sullivan, P. A.; Andrade, R. B.; Ratner, D. M. *Langmuir* **2012**, *28*, 3338.
- (34) Zhang, X.; Dai, Y.; Zhong, W.; Tan, M.; Lv, Z.; Zhou, Y.; Jiang, X. *Bioorg. Med. Chem.* **2012**, *20*, 1616.
- (35) Tan, M.; Jiang, X. *Nanomedicine* **2012**, *7*, 1.
- (36) Xia, M.; Tan, M.; Wei, C.; Zhong, W.; Wang, L.; McNeal, M.; Jiang, X. *Vaccine* **2011**, *29*, 7670.
- (37) Fang, H.; Tan, M.; Xia, M.; Wang, L.; Jiang, X. *PLoS One* **2013**, *8*, e63269.
- (38) Dai, Y.; Wang, Y.; Zhang, X.; Tan, M.; Xia, M.; Wu, X.; Jiang, X.; Nie, J. *J. Virol. Methods* **2012**, *186*, 126.
- (39) Dai, Y.; Zhang, X.; Tan, M.; Huang, P.; Lei, W.; Fang, H.; Zhong, W.; Jiang, X. *Antivir. Res.* **2012**.
- (40) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.
- (41) Feng, X.; Jiang, X. *Antimicrob. Agents Chemother.* **2007**, *51*, 324.
- (42) Rademacher, C.; Guiard, J.; Kitov, P. I.; Fiege, B.; Dalton, K. P.; Parra, F.; Bundle, D. R.; Peters, T. *Chem. Eur. J.* **2011**, *17*, 7442.
- (43) Zhang, X.; Tan, M.; Chhabra, M.; Dai, Y.; Meller, J.; Jiang, X. *Plos One* **2013**, *8*, e69379.
- (44) Cederkvist, F.; Zamfir, A. D.; Bahrke, S.; Eijsink, V. G. H.; Sorlie, M.; Peter-Katalinic,

- J.; Peter, M. G. *Angew. Chem. Int. Ed.* **2006**, *45*, 2429.
- (45) Abzalimov, R. R.; Dubin, P. L.; Kaltashov, I. A. *Anal. Chem.* **2007**, *79*, 6055.
- (46) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50.
- (47) Bereszczak, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R. *J. Struct. Biol.* **2012**, *177*, 273.
- (48) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (49) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.
- (50) Hansman, G. S.; Shahzad-ul-Hussan, S.; McLellan, J. S.; Chuang, G.-Y.; Georgiev, I.; Shimoike, T.; Katayama, K.; Bewley, C. A.; Kwong, P. D. *J. Virol.* **2012**, *86*, 284.

## Chapter 4

- (1) Patel, M. M.; Widdowson, M.-A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (2) Jiang, X.; Wang, M.; Graham, D. Y.; Estes, M. K. *J. Virol.* **1992**, *66*, 6527.
- (3) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (4) White, L. J.; Hardy, M. E.; Estes, H. K. *J. Virol.* **1997**, *71*, 8066.
- (5) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (6) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.

- (7) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (8) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.
- (9) Bereszcak, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R. *J. Struct. Biol.* **2012**, *177*, 273.
- (10) Tamminen, K.; Huhti, L.; Koho, T.; Lappalainen, S.; Hytonen, V. P.; Vesikari, T.; Blazevic, V. *Immunology* **2012**, *135*, 89.
- (11) Oriol, R. *J. Immunogenet.* **1990**, *17*, 235.
- (12) Ravn, V.; Dabelsteen, E. *APMIS* **2000**, *108*, 1.
- (13) Hutson, A. M.; Atmar, R. L.; Graham, D. Y.; Estes, M. K. *J. Infect. Dis.* **2002**, *185*, 1335.
- (14) Hutson, A. M.; Atmar, R. L.; Marcus, D. M.; Estes, M. K. *J. Virol.* **2003**, *77*, 405.
- (15) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang, X. *J. Virol.* **2005**, *79*, 6714.
- (16) Tan, M.; Jiang, X. *Trends Microbiol.* **2005**, *13*, 285.
- (17) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (18) de Rougemont, A.; Ruvoen-Clouet, N.; Simon, B.; Estienney, M.; Elie-Caille, C.; Aho, S.; Pothier, P.; Le Pendu, J.; Boireau, W.; Belliot, G. *J. Virol.* **2011**, *85*, 4057.
- (19) Hansman, G. S.; Shahzad-ul-Hussan, S.; McLellan, J. S.; Chuang, G.-Y.; Georgiev, I.; Shimoike, T.; Katayama, K.; Bewley, C. A.; Kwong, P. D. *J. Virol.* **2012**, *86*, 284.

- (20) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. *Glycobiology* **2013**, *23*, 276.
- (21) El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 3867.
- (22) Jiang, X.; Zhong, W.; Farkas, T.; Huang, P.; Wilton, N.; Barrett, E.; Fulton, D.; Morrow, R.; Matson, D. O. *Arch. Virol* **2002**, *147*, 119.
- (23) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendu, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.
- (24) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.
- (25) Seto, N. O. L.; Palcic, M. M.; Compston, C. A.; Li, H.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1997**, *272*, 14133.
- (26) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423.
- (27) Meloncelli, P. J.; Lowary, T. L. *Aust. J. Chem.* **2009**, *62*, 558.
- (28) Meloncelli, P. J.; Lowary, T. L. *Carbohydr. Res.* **2010**, *345*, 2305.
- (29) Meloncelli, P. J.; West, L. J.; Lowary, T. L. *Carbohydr. Res.* **2011**, *346*, 1406.
- (30) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (31) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.

- (32) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.
- (33) Wang, L.; Xia, M.; Huang, P.; Fang, H.; Cao, D.; Meng, X.-J.; McNeal, M.; Jiang, X.; Tan, M. *Biomaterials* **2014**, *35*, 8427.
- (34) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1.
- (35) Shoemaker, G. K.; van Duijn, E.; Crawford, S. E.; Uetrecht, C.; Baclayon, M.; Roos, W. H.; Wuite, G. J. L.; Estes, M. K.; Prasad, B. V. V.; Heck, A. J. R. *Mol. Cell. Proteomics* **2010**, *9*, 1742.
- (36) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3*, 126.
- (37) Patenaude, S. I.; Seto, N. O. L.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol.* **2002**, *9*, 685.
- (38) Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587.
- (39) Soya, N.; Shoemaker, G. K.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2009**, *19*, 1224.
- (40) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E. G.; Yagi, F.; Kasai, K. *Biochim. Biophys. Acta Gen. Subj.* **2002**, *1572*, 232.
- (41) Collins, P. M.; Bum-Erdene, K.; Yu, X.; Blanchard, H. *J. Mol. Biol.* **2014**, *426*, 1439.
- (42) Shang, J.; Piskarev, V. E.; Xia, M.; Huang, P.; Jiang, X.; Likhoshesterov, L. M.; Novikova, O. S.; Newburg, D. S.; Ratner, D. M. *Glycobiology* **2013**, *23*, 1491.

- (43) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.
- (44) Bu, W.; Mamedova, A.; Tan, M.; Xia, M.; Jiang, X.; Hegde, R. S. *J. Virol.* **2008**, *82*, 5340.

## Chapter 5

- (1) Patel, M. M.; Widdowson, M. A.; Glass, R. I.; Akazawa, K.; Vinje, J.; Parashar, U. D. *Emerg. Infect. Dis.* **2008**, *14*, 1224.
- (2) Tan, M.; Jiang, X. *PLoS Pathog.* **2010**, *6*, e1000983.
- (3) Prasad, B. V. V.; Hardy, M. E.; Dokland, T.; Bella, J.; Rossmann, M. G.; Estes, M. K. *Science* **1999**, *286*, 287.
- (4) Tan, M.; Jiang, X. *Expert Rev. Mol. Med.* **2007**, *9*, 1.
- (5) Tan, M.; Jiang, X. *Trends Microbiol.* **2011**, *19*, 382.
- (6) Tan, M.; Jiang, X. *Expert Rev. Mol. Med.* **2014**, *16*, e5.
- (7) Bertolotti-Ciarlet, A.; White, L. J.; Chen, R.; Prasad, B. V. V.; Estes, M. K. *J. Virol.* **2002**, *76*, 4044.
- (8) Tan, M.; Hegde, R. S.; Jiang, X. *J. Virol.* **2004**, *78*, 6233.
- (9) Cao, S.; Lou, Z.; Tan, M.; Chen, Y.; Liu, Y.; Zhang, Z.; Zhang, X. C.; Jiang, X.; Li, X.; Rao, Z. *J. Virol.* **2007**, *81*, 5949.
- (10) Choi, J. M.; Hutson, A. M.; Estes, M. K.; Prasad, B. V. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9175.

- (11) Hansman, G. S.; Biertumpfel, C.; Georgiev, I.; McLellan, J. S.; Chen, L.; Zhou, T.; Katayama, K.; Kwong, P. D. *J. Virol.* **2011**, *85*, 6687.
- (12) Chen, Y.; Tan, M.; Xia, M.; Hao, N.; Zhang, X. C.; Huang, P.; Jiang, X.; Li, X.; Rao, Z. *PLoS Pathog.* **2011**, *7*, e1002152.
- (13) Tan, M.; Fang, P.; Xia, M.; Chachiyo, T.; Jiang, W.; Jiang, X. *Virology* **2011**, *410*, 345.
- (14) Tan, M.; Jiang, X. *J. Virol.* **2005**, *79*, 14017.
- (15) Tan, M.; Fang, P.; Chachiyo, T.; Xia, M.; Huang, P.; Fang, Z.; Jiang, W.; Jiang, X. *Virology* **2008**, *382*, 115.
- (16) Wang, L.; Huang, P.; Fang, H.; Xia, M.; Zhong, W.; McNeal, M. M.; Jiang, X.; Tan, M. *Biomaterials* **2013**, *34*, 4480.
- (17) Wang, L.; Xia, M.; Huang, P.; Fang, H.; Cao, D.; Meng, X.; McNeal, M.; Jiang, X.; Tan, M. *Biomaterials* **2014**, *35*, 8427.
- (18) Lindesmith, L.; Moe, C.; Marionneau, S.; Ruvoen, N.; Jiang, X.; Lindbland, L.; Stewart, P.; LePendou, J.; Baric, R. *Nat. Med.* **2003**, *9*, 548.
- (19) Hutson, A. M.; Atmar, R. L.; Graham, D. Y.; Estes, M. K. *J. Infect. Dis.* **2002**, *185*, 1335.
- (20) Tan, M.; Jin, M.; Xie, H.; Duan, Z.; Jiang, X.; Fang, Z. *J. Med. Virol.* **2008**, *80*, 1296.
- (21) Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1.
- (22) Huang, P.; Farkas, T.; Marionneau, S.; Zhong, W.; Ruvoen-Clouet, N.; Morrow, A. L.; Altaye, M.; Pickering, L. K.; Newburg, D. S.; LePendou, J.; Jiang, X. *J. Infect. Dis.* **2003**, *188*, 19.

- (23) Huang, P.; Farkas, T.; Zhong, W.; Thornton, S.; Morrow, A. L.; Jiang X. *J. Virol.* **2005**, *79*, 6714.
- (24) Shanker, S.; Choi, J.-M.; Sankaran, B.; Atmar, R. L.; Estes, M. K.; Prasad, B. V. V. *J. Virol.* **2011**, *85*, 8635.
- (25) Shirato, H.; Ogawa, S.; Ito, H.; Sato, T.; Kameyama, A.; Narimatsu, H.; Zheng, X.; Miyamura, T.; Wakita, T.; Ishii, K.; Takeda, N. *J. Virol.* **2008**, *82*, 10756.
- (26) Lindesmith, L.; Moe, C.; LePendu, J.; Frelinger, J. A.; Treanor, J.; Baric, R. S. *J. Virol.* **2005**, *79*, 2900.
- (27) Murakami, K.; Kurihara, C.; Oka, T.; Shimoike, T.; Fujii, Y.; Takai-Todaka, R.; Park, Y.; Wakita, T.; Matsuda, T.; Hokari, R.; Miura, S.; Katayama, K. *PLoS One* **2013**, *8*, e66534.
- (28) Rydell, G. E.; Dahlin, A. B.; Hook, F.; Larson, G. *Glycobiology* **2009**, *19*, 1176.
- (29) Bally, M.; Rydell, G. E.; Zahn, R.; Nasir, W.; Eggeling, C.; Breimer, M. E.; Svensson, L.; Hook, F.; Larson, G. *Angew. Chem. Int. Ed.* **2012**, *51*, 12020.
- (30) Tamura, M.; Natori, K.; Kobayashi, M.; Miyamura, T.; Takeda, N. *J. Virol.* **2004**, *78*, 3817.
- (31) de Rougemont, A.; Ruvoen-Clouet, N.; Simon, B.; Estienney, M.; Elie-Caille, C.; Aho, S.; Pothier, P.; Le Pendu, J.; Boireau, W.; Belliot, G. *J. Virol.* **2011**, *85*, 4057.
- (32) Fiege, B.; Rademacher, C.; Cartmell, J.; Kitov, P. I.; Parra, F.; Peters, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 928.
- (33) Taube, S.; Perry, J. W.; Yetming, K.; Patel, S. P.; Auble, H.; Shu, L. M.; Nawar, H. F.;

- Lee, C. H.; Connell, T. D.; Shayman, J. A.; Wobus, C. E. *J. Virol.* **2009**, *83*, 4092.
- (34) Stuart, A. D.; Brown, T. D. K. *J. Gen. Virol.* **2007**, *88*, 177.
- (35) Kim, D. S.; Hosmillo, M.; Alfajaro, M. M.; Kim, J. Y.; Park, J. G.; Son, K. Y.; Ryu, E. H.; Sorgeloos, F.; Kwon, H. J.; Park, S. J.; Lee, W. S.; Cho, D.; Kwon, J.; Choi, J. S.; Kang, M. I.; Goodfellow, I.; Cho, K. O. *PLoS Pathog.* **2014**, *10*, e1004172.
- (36) El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50.
- (37) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.
- (38) El-Hawiet, A.; Kitova, E. N.; Arutyunov, D.; Simpson, D. J.; Szymanski, C. M.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 3867.
- (39) Han, L.; Kitov, P. I.; Kitova, E. N.; Tan, M.; Wang, L.; Xia, M.; Jiang, X.; Klassen, J. S. *Glycobiology* **2013**, *23*, 276.
- (40) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423.
- (41) Meloncelli, P. J.; West, L. J.; Lowary, T. L. *Carbohydr. Res.* **2011**, *346*, 1406.
- (42) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (43) Jiang, Y.; Cole, R. B. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 60.
- (44) Zhang, Y.; Deng, L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1573.

- (45) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (46) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472.
- (47) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E. G.; Yagi, F.; Kasai, K. *Biochim. Biophys. Acta* **2002**, *1572*, 232.
- (48) Collins, P. M.; Bum-Erdene, K.; Yu, X.; Blanchard, H. *J. Mol. Biol.* **2014**, *426*, 1439.
- (49) Taube, S.; Jiang, M.; Wobus, C. E. *Viruses* **2010**, *2*, 1011.

## Chapter 6

- (1) Hakomori, S. *Curr. Opin. Hematol.* **2003**, *10*, 16.
- (2) Sharon, N.; Lis, H. *Sci. Am.* **1993**, *268*, 82.
- (3) Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Stanley, P.; Bertozzi, C.R.; Hart, G.W.; Etzler, M.E. *Essentials of Glycobiology* 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA 2009.
- (4) Holgersson, J.; Gustafsson, A.; Breimer, M. E. *Immunol. Cell Biol.* **2005**, *83*, 694.
- (5) Lopez P. H.; Schnaar, R. L. *Methods Enzymol.* **2006**, *417*, 205.
- (6) DeMarco, M. L. *Biochemistry* **2012**, *51*, 5725.
- (7) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.
- (8) Lingwood, C. A. *Glycoconjugate J.* **1996**, *13*, 495.
- (9) Lingwood, C. A.; Manis, A.; Mahfoud, R.; Khan, F.; Binnington, B.; Mylvaganam, M.

- Chem. Phys. Lipids* **2010**, *163*, 27.
- (10) Lingwood, D.; Binnington, B.; Rog, T.; Vattulainen, I.; Grzybek, M.; Coskun, U.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.
- (11) MacKenzie, C. R.; Hiramata, T.; Lee, K. K.; Altman, E.; Young, N. M. *J. Biol. Chem.* **1997**, *272*, 5533.
- (12) Lauer, S.; Goldstein, B.; Nolan, R. L.; Nolan, J. P. *Biochemistry* **2002**, *41*, 1742.
- (13) Gallegos, K. M.; Conrady, D. G.; Karve, S. S.; Gunasekera, T. S.; Herr, A. B.; Weiss, A. A. *PLoS One* **2012**, *7*, 10.
- (14) Vogel, J.; Bendas, G.; Bakowsky, U.; Hummel, G.; Schmidt, R. R.; Kettmann, U.; Rothe, U. *Biochim. Biophys. Acta.* **1998**, *1372*, 205.
- (15) Shi, J.; Yang, T.; Kataoka, S.; Zhang, Y.; Diaz, A. J.; Cremer, P. S. *J. Am. Chem. Soc.* **2007**, *129*, 5954.
- (16) Nath, A.; Atkins, W. M.; Sligar, S. G. *Biochemistry* **2007**, *46*, 2059.
- (17) Bayburt, T. H.; Sligar, S. G. *FEBS Lett.* **2010**, *584*, 1721.
- (18) Jayaraman, N.; Maiti, K.; Naresh, K. *Chem. Soc. Rev.* **2013**, *42*, 4640.
- (19) Borch, J.; Torta, F.; Sligar, S. G.; Roepstorff, P. *Anal. Chem.* **2008**, *80*, 6245.
- (20) Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618.
- (21) Leney, A. C.; Fan, X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271.
- (22) Sloan, C. D. K.; Marty, M. T.; Sligar, S. G.; Bailey, R. C. *Anal. Chem.* **2013**, *85*, 2970.
- (23) Vanheyne, S. *Science* **1974**, *183*, 656.

- (24) Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Ho, W. G. *J. Mol. Biol.* **1998**, *282*, 1043.
- (25) Kadioglu, A.; Andrew, P. W. *Trends Immunol.* **2004**, *25*, 143.
- (26) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.
- (27) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.
- (28) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Nano Lett.* **2002**, *2*, 853.
- (29) Popovic, K.; Holyoake, J.; Pomès, R.; Privé, G. G. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 2908.
- (30) Leney, A. C.; Rezaei Darestani, R.; Li, J.; Nikjah, S.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Xiong, Z.; Prive, G. G.; Klassen, J. S. *Anal. Chem.* **2015**, *87*, 4402.
- (31) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.
- (32) Kitova, E. N.; Kitov, P. I.; PaszkiewiCz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.
- (33) Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1047.
- (34) Lin, H.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 104.
- (35) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477.
- (36) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Arch. Biochem. Biophys.* **2006**, *450*, 215.
- (37) Nakano, M.; Fukuda, M.; Kudo, T.; Miyazaki, M.; Wada, Y.; Matsuzaki, N.; Endo, H.;

Handa, T. *J. Am. Chem. Soc.* **2009**, *131*, 8308.

- (38) Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. *Glycobiology* **2015**, *25*, 170.
- (39) Moonens, K.; Bouckaert, J.; Coddens, A.; Tran, T.; Panjikar, S.; De Kerpel, M.; Cox, E.; Remaut, H.; De Greve, H. *Mol. Microbiol.* **2012**, *86*, 82.

## Chapter 7

- (1) Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W. E. G.; Yagi, F.; Kasai, K. *Biochim. Biophys. Acta Gen. Subj.* **2002**, *1572*, 232.
- (2) Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. *Glycobiology* **2015**, *25*, 170.
- (3) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.
- (4) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.
- (5) Lingwood, D.; Binnington, B.; Róg, T.; Vattulainen, I.; Grzybek, M.; Coskun, Ü.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.
- (6) Popovic, K.; Holyoake, J.; Pomès, R.; Privé, G. G. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2908.
- (7) Leney, A. C.; Rezaei Darestani, R.; Li, J.; Nikjah, S.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Xiong, Z. J.; Privé, G. G.; Klassen, J. S. *Anal. Chem.* **2015**, *87*, 4402.

- (8) Kitov, P. I.; Bundle, D. R., *J. Am. Chem. Soc.* **2003**, *125*, 16271.
- (9) Sacchettini, J. C.; Baum, L. G.; Brewer, C. F. *Biochemistry* **2001**, *40*, 3009.
- (10) Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1047.
- (11) Lin, H.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 104.