

**Studying Non-Covalent Interactions Using Electrospray Ionization Mass  
Spectrometry**

by

Xuxin Fan

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

Master of Science

Department of Chemistry

University of Alberta

© Xuxin Fan, 2015

## Abstract

This thesis describes the application of electrospray ionization mass spectrometry (ESI-MS) to investigate non-covalent interactions. The first research chapter focuses on the protein-carbohydrate interaction. The binding affinities ( $K_a$ ) were measured for a library of thirty human oligosaccharides and recombinant sub-fragments of TcdA (TcdA-A2) and TcdB (TcdB-B3) using the direct ESI-MS assay. It is found that TcdA-A2 binds to seventeen of the oligosaccharides, although with uniformly low affinities,  $<2000 \text{ M}^{-1}$ . TcdB-B3 recognizes only four oligosaccharides, with low affinities. These binding data suggest a variety of glycan structures that could serve as the human cellular receptors for TcdA and TcdB. Affinity measurements were also carried on a series of Lewis X analogues that were previously identified as ligands for TcdA. Notably, the relative affinities measured by ESI-MS for TcdA-A2 are in reasonable agreement with the results determined from glycan microarray screening for the  $\text{Le}^x$  ligands. However, the glycan microarray data failed to identify binding for nine other structures (common to both libraries) that were found to exhibit low affinity interactions based on the ESI-MS data.

The second research chapter focuses on the peptide-membrane interaction. Using nanodisc as the model of biological membrane, ESI-MS was employed to investigate peptide-membrane interaction using antimicrobial peptide melittin and phospholipid nanodisc as model system. The ESI-MS study revealed melittin interact with phospholipid bilayer specifically. This result opens up further investigation on the development of the optimal instrument conditions to study peptide-membrane interaction.

## Preface

This thesis is an original work by Xuxin Fan. None of the text of this thesis has been previously published.

A version of Chapter 2 is being prepared for submission as an article for publication as X. Fan, E. N. Kitova, L. Eugenio, K. K. S. Ng and J. S. Klassen, “Carbohydrate Recognition by *Clostridium difficile* Toxins”. I was responsible for the data collection and analysis as well as the manuscript composition. E. N. Kitova contributed to manuscript edits. The proteins TcdA-A2 and TcdB-B3 used in Chapter 2 were prepared and provided by Luiz Eugenio and Dr. Kenneth K. S. Ng. Dr. John S. Klassen was the supervisory author and was involved with concept formation and manuscript composition.

## Acknowledgements

The work presented here could not have been accomplished without the support of the following individuals and funding agencies.

I would like to extend my gratitude to my supervisor, Dr. John S. Klassen for his leadership and guidance throughout my graduate studies. It was always a delightful experience to discuss science with him. I would like to thank my supervisor and defense committee members, Dr. Liang Li, Dr. Nils O. Petersen and Dr. Robert E. Campbell for their thoughtful and helpful comments into my research. I would like to thank Dr. Kenneth K. S. Ng for his collaboration, without which our work would not be possible.

During the past three years, I have the honor to encounter and work with many excellent colleagues in Klassen group. Among them, I would like to give my special thanks to Dr. Aneika C. Leney, Dr. Elena N. Kitova, and Dr. Rambod Daneshfar for giving me the training when I first arrived to this lab. I would also like to thank, in no particular order, Dr. Ling Han, Dr. Emma-Dune Leriche, Dr. Lu Deng, Dr. Amr El-Hawiet, Dr. Lan Liu, Yuyu Yao, Jingjing Zhang, Nobar Jalili, Reza Rezaei Darestani, Sanaz Nikjah, Km Shams-Ud-Doha, Jun Li and Yajie Chen for their helpful advice and continued encouragements.

I would also like to thank Mr. Gareth Lambkin in the Department of Chemistry's biological service for his assistance on protein purification.

I give sincere thanks to my friends, Tiffany Lai, Weixiao Zhao, Xiaobin Wang and many others for all their support.

I would like to thank my parents for providing me with education, emotional support and encouragement to pursue my degree.

I would like to acknowledge the financial support provided by the Department of Chemistry of University of Alberta.

## Table of Contents

<b>List of Tables</b> .....	viii
<b>List of Figures</b> .....	ix
<b>List of Abbreviations</b> .....	xv
<b>Chapter 1 Studying Non-covalent Interactions Using Electrospray Ionization Mass Spectrometry</b> ..	1
1.1 Introduction.....	1
1.2 Electrospray Ionization Mass Spectrometry .....	4
1.2.1 Electrospray Ionization .....	4
1.2.1.1 ESI mechanism .....	5
1.2.1.2 Nano ESI.....	11
1.2.2 MS Instrumentation.....	12
1.2.2.1 Quadrupole mass filter .....	13
1.2.2.2 Travelling wave ion guide (TWIG).....	14
1.2.2.3 Time of flight (TOF) analyzer .....	16
1.2.2.4 Tandem MS (MS/MS) .....	17
1.3 MS-based method for quantifying interactions of non-covalent complexes.....	18
1.3.1 Direct ESI-MS binding assay.....	18
1.3.2 Limitations of the direct ESI-MS binding assay .....	20
1.3.2.1 Non-uniform response factors.....	20
1.3.2.2 In-source dissociation .....	21
1.3.2.3 Nonspecific binding.....	23
1.4 Nanodisc technology.....	25
1.5 The present work.....	28
1.6 Literature cited .....	29
<b>Chapter 2 Identifying Binding Partners for <i>Clostridium difficile</i> Toxins using ESI-MS</b> .....	36
2.1 Introduction.....	36
2.2 Materials and Methods.....	43
2.2.1 Reagents.....	43
2.2.2 Expression and purification of MSP1E1 .....	46
2.2.3 Nanodisc Preparation .....	47

2.2.4	Mass spectrometry .....	48
2.3	Results and Discussion .....	51
2.3.1	Carbohydrate recognition by TcdA-A2 .....	51
2.3.2	Carbohydrate recognition by TcdB-B3 .....	62
2.3.3	Comparison of binding data from ESI-MS and glycan microarray screening .....	65
2.3.4	Glycosphingolipids recognition by <i>C. difficile</i> toxins.....	68
2.4	Conclusion .....	79
2.5	Literature cited .....	80
<b>Chapter 3 Interaction of Bee Venom Melittin with Zwitterionic Charged Phospholipid Nanodiscs: A Electrospray Ionization Mass Spectrometry Study .....</b>		<b>83</b>
3.1	Introduction.....	83
3.2	Material and Method.....	85
3.2.1	Materials .....	85
3.2.2	Preparation of Nanodisc.....	86
3.2.3	Mass spectrometry .....	87
3.3	Results and Discussion .....	88
3.3.1	Melitin interacts with phospholipid Nanodisc .....	88
3.3.2	CID of (Melittin+POPC ND) complex .....	92
3.4	Conclusion .....	94
3.5	Literature cited .....	94
<b>Chapter 4 Conclusions and Future Work.....</b>		<b>97</b>
4.1	Literature Cited .....	99
<b>Bibliography .....</b>		<b>100</b>

## List of Tables

- Table 1.1** Optimal ratios of phospholipids to MSPs, diameters and incubation temperature for nanodisc assembly. The diameter of nanodisc is determined by the length of MSP, while the incubation temperature is the phase transition temperature of the chosen lipid.
- Table 2.1** Apparent association constants ( $K_a$ , units of  $M^{-1}$ ) for the binding of TcdA-A2 and TcdB-B3 with oligosaccharides **L1 – L35** measured in an aqueous ammonium acetate (200 mM) solution at pH 7 and 25 °C using the direct ESI-MS assay.
- Table 2.2** Comparison between glycan microarray assay and direct ESI-MS assay binding results between TcdA-A2 and 14 oligosaccharides. Direct ESI assay's result is shown as Apparent association constants  $K_{a,app}$  ( $M^{-1}$ ) measured in aqueous ammonium acetate (200 mM) at pH 7 and 25 °C.<sup>a</sup> Glycan microarray's result is shown in relative fluorescence (RF).

## List of Figures

- Figure 1.1** Schematic representation of process in ESI-MS in positive ion mode. Figure was adapted from reference 34.
- Figure 1.2** Summary of different models for the formation of gas phase ions, (a) CRM: formation of a globular protein into the gas phase. (b) IEM: small ion ejection from a charged nanodroplet. (c) CEM: ejection of an unfolded protein.
- Figure 1.3** Schematic diagram of Synapt G2-S Q-IM-TOF mass spectrometer, adapted from the Waters user's manual.
- Figure 1.4** Diagram illustrating the quadrupole used in the Waters Synapt mass spectrometers.
- Figure 1.5** Diagram of a stacked ring ion guide used in Waters Synapt mass spectrometers.
- Figure 1.6** Cartoon of nanodisc. Blue belts represent two copies of MSP which surround the lipid bilayer without (left) or with (right) cellular receptors incorporated.
- Figure 2.1** Crystal structure of TcdA-A2, a sub-recombinant fragment of TcdA.

**Figure 2.2** Structures of the thirty-five oligosaccharides (**L1** – **L35**) investigated in the current study. The structures are shown here using the symbol nomenclature adopted by the Consortium for Functional Glycomics ([www.functionalglycomics.org](http://www.functionalglycomics.org)).

**Figure 2.3** Structures of the 7 gangliosides (GM1, GM2, GM3, GD1a, GT1b and GD2) that were incorporated in nanodiscs (NDs) for the present study.

**Figure 2.4** Cartoon representation of the CaR-ESI-MS assay for screening glycosphingolipids (GSLs) against target proteins. The protein and ND (lipid bilayer (red) with surrounding membrane scaffold protein (blue) containing different GSLs are incubated to form a complex. Using mild in-source dissociation conditions, the protein-GSL complexes are stripped out of the ND. The GSL ligands are then released (as ions) from the protein using collision-induced dissociation (CID) and mass analyzed.

**Figure 2.5** ESI-MS spectra acquired in positive-ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing TcdA-A2 (6  $\mu$ M),  $P_{ref}$  (2  $\mu$ M) and **L1** at (a) 10  $\mu$ M or (b) 80  $\mu$ M. Insets show the normalized distribution of free and ligand-bound TcdA-A2 after correction for nonspecific ligand binding using the reference protein

method.

**Figure 2.6** Plot of fraction of ligand-bound TcdA-A2 ( $R/(R+1)$ ) versus ligand concentration measured for (a) **L1**, (b) **L2**, (c) **L3**, (d) **L5**, (e) **L7**, (f) **L9**, (g) **L11**, (h) **L12**, (i) **L15**, (j) **L16**, (k) **L17**, (l) **L18**, (m) **L19**, (n) **L20**, (o) **L21**, (p) **L29**, (q) **L31**, (r) **L32**, (s) **L33**, (t) **L34** and (u) **L35**. The ESI-MS binding measurements were carried out on 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing TcdA-A2 (6  $\mu$ M), Lyz (2  $\mu$ M) and each ligand at a minimum of five different concentrations, ranging from 5 to 200  $\mu$ M. The solid curves correspond to the best fit of eq. 2.4 to the experimental data and the error bars correspond to one standard derivation.

**Figure 2.7** ESI-MS spectra acquired in positive-ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing TcdB-B3 (7  $\mu$ M),  $P_{ref}$  (5  $\mu$ M) and **L19** at (a) 40  $\mu$ M or (b) 80  $\mu$ M. Insets show the normalized distribution of free and ligand-bound TcdB-B3 after correction for nonspecific ligand binding using the reference protein method.

**Figure 2.8** Plot of fraction of ligand-bound TcdB-B3 ( $R/(R+1)$ ) versus ligand concentration measured for (a) **L19**, (b) **L20**, (c) **L22** and (d) **L25**. The ESI-MS binding measurements were carried out on 200 mM aqueous

ammonium acetate solutions (pH 7 and 25 °C) containing TcdB-B3 (7  $\mu$ M), Lyz (5  $\mu$ M) and each ligand at a minimum of five different concentrations, ranging from 10 to 200  $\mu$ M. The solid curves correspond to the best fit of eq 2.4 to the experimental data and the error bars correspond to one standard derivation.

**Figure 2.9** Relative affinities of fourteen oligosaccharides (**L1, L2, L3, L5, L6, L7, L9, L16, L17, L31, L32, L33, L34 and L35**) for TcdA-A2 measured using the direct ESI-MS assay. Also shown are the relative fluorescence intensities reported for the related glycans by the Consortium for Functional Glycomics ([www.functionalglycomics.org](http://www.functionalglycomics.org)).

**Figure 2.10** Characterization for the 7 ganglioside ND. (a) is the size exclusion chromatography for 7 ganglioside ND. The void volume ( $V_0$ ) and total column volume ( $V_c$ ) are shown as an example for the 7 ganglioside ND. (b) is the mass spectrum of 7 ganglioside ND under negative mode, a broad peak centered at mass-to-charge ratio ( $m/z$ )  $\sim$ 11000 was observed, which attributed to ND. (c) is the CID spectrum of 7 ganglioside ND under trap 100V after the ND is selected through quadrupole profile and subjected to trap region. All the 7 gangliosides are detected.

**Figure 2.11** ESI-MS spectra of TcdA-A2 6  $\mu\text{M}$  alone in aqueous ammonium acetate (200 mM) at pH 7 and 25  $^{\circ}\text{C}$  (a), upon incubation with 20  $\mu\text{M}$  2% 7 Ganglioside ND (b). Peaks labeled with \* corresponding to nonspecific interaction between target protein and DMPC, the major component of ND.

**Figure 2.12** ESI-CID spectra of TcdA-A2 6 $\mu\text{M}$  incubated with 20  $\mu\text{M}$  2% 7 Ganglioside ND selected at 3174 m/z using trap voltage 50V (a), under trap voltage 75V (b) and under trap voltage 100V (c).

**Figure 2.13** ESI-MS spectra of TcdB-B3 7  $\mu\text{M}$  alone in aqueous ammonium acetate (200 mM) at pH 7 and 25  $^{\circ}\text{C}$  (a), upon incubation with 20  $\mu\text{M}$  2% 7 Ganglioside ND (b). Peaks labeled with \* corresponding to nonspecific interaction between target protein and DMPC, the major component of ND.

**Figure 2.14** ESI-CID spectra of TcdB-B3 7 $\mu\text{M}$  incubated with 20  $\mu\text{M}$  2% 7 Ganglioside ND selected at 2874 m/z using trap voltage 5V (a), using trap voltage 50V (b), using trap voltage 100V (c).

**Figure 3.1** ESI mass spectra acquired in negative ion mode for aqueous ammonium acetate (200 mM) solutions of (18  $\mu\text{M}$ ) Melittin (a) or in the presence of (3.4  $\mu\text{M}$ ) POPC ND (b). Insulin chain A (8  $\mu\text{M}$ ) is

added to both solutions as reference peptide.

**Figure 3.2** CID mass spectra of (Melittin+ POPC ND) complex acquired in negative mode under different collision energy, trap 5V (a), trap 100V (b) and trap 200V (c). Complex ions were selected using quadrupole profile as described in experimental session.

**Figure 3.3** CID spectra of Melittin singly deprotonated ions acquired in negative mode at a collision energy of trap 5V (a), trap 40V (b), trap 50V (c), trap 75V(d). Quadrupole was set at  $m/z$  2845 to select [Mel-H]<sup>-</sup>.

**Figure 3.4** Energy profile plots of abundance of released melittin as the Trap collision energy increased from 5V to 200V at negative mode (a) and positive mode (b). The transfer energy is set as 2V for all the experiments.

## List of Abbreviations

<i>Ab</i>	Abundance of gas-phase ions
BIRD	Blackbody infrared radiative dissociation
CaR	Catch-and-release
CEM	Chain ejection model
CID	Collision induced dissociation
CRM	Charge residue model
Da	Dalton
DC	Direct current
DESI	Desorption electrospray ionization
DMPC	1,2-dimyristoyl- <i>sn</i> -glycero-3-phosphocholine
ECD	Electron capture dissociation
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
ESI-MS	Electrospray ionization mass spectrometry
FPLC	Fast performance liquid chromatography
Fuc	Fucose
Gal	Galactose
GalNAc	N-acetyl galactosamine
GSL	Glycosphingolipid
Glc	Glucose
GlcNAc	N-acetyl glucosamine
GM1 <sub>os</sub>	GM1 pentasaccharide

H	Hydrogen
HBGA	Histo-blood group antigen
H-bond	Hydrogen bond
HMOs	Human milk oligosaccharides
HPLC	High performance liquid chromatography
IEM	Ion evaporation model
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
$K_d$	Dissociation constant
$K_{eq}$	Equilibrium constant
L	Ligand
LC	Liquid chromatography
m/z	Mass-to-charge ratio
MS	Mass spectrometry
MSP	Membrane scaffold protein
MW	Molecular weight
MWCO	Molecular weight cut off
nanoESI	Nanoflow electrospray ionization
N	Nitrogen

ND	Nanodisc
Neu5Ac	N-acetyl neuraminic acid
NMR	Nuclear magnetic resonance
O	Oxygen
P	Protein
PL	Protein-ligand complex
POPC	1-palmitoyl-2-oleoyl- <i>sn</i> -glycero-3-phosphocholine
P <sub>ref</sub>	Reference protein
<i>R</i>	Abundance ratio
<i>RF</i>	Response factor
RF	Radio frequency
scFv	Single chain variable fragment
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SPR	Surface plasmon resonance spectroscopy
TcdA-A2	Sub-fragment of <i>Clostridium difficile</i> toxin A
TcdB-B3	Sub-fragment of <i>Clostridium difficile</i> toxin B
TLC	Thin layer chromatography
TOF	Time of flight
Tris	Tris(hydroxymethyl) aminomethane
TWIMS	Travelling wave ion mobility spectrometry

# Chapter 1

## Studying Non-covalent Interactions Using Electrospray Ionization Mass Spectrometry

### 1.1 Introduction

Protein-ligand interactions, such as protein-carbohydrate and protein-membrane interactions play essential roles in many important biochemical processes, including the immune response, cell-cell communications, bacterial toxin infections, signal transduction and ion transport.<sup>1-4</sup> The association of protein-ligand complexes is driven by intermolecular forces, such as hydrogen bond (H-bond) networking, ionic interactions, hydrophobic interactions and van der Waals force. In order to develop a complete understanding of many physiological and pathological cellular processes and to guide drug discovery and design efforts where therapeutic molecules work by blocking or modifying the biological function of proteins, detection and characterization (structure and thermodynamic and kinetic parameters) of protein-ligand interactions is of fundamental importance. There are a number of established analytical methods to identify and quantify protein-ligand interactions in vitro, each with their own particular strengths and weaknesses. A brief introduction of each technique is given below.

Spectroscopic approaches (e.g. fluorescence-based approaches)<sup>5-9</sup> are very sensitive and commonly used to characterize the interactions of macromolecular complexes. However, the spectroscopic methods require spectroscopic activity of the analyte, often in the fluorescence-based approaches. A chromophore or fluorophore is conjugated to the analyte

of interest which can lead to complications such as perturbation of the rate of reactions being investigated.<sup>10</sup>

X-ray crystallography represents the most widely used method to establish the high-order structure of biological complexes and to identify interactions between binding partners.<sup>11-12</sup>

However, not all biomolecules and their complexes are readily crystallizable. Furthermore, it is possible that the interactions present in the crystalline state are different from those in solution because proteins are constantly undergoing conformational fluctuations instead of being immobile.<sup>13</sup> X-ray crystallography cannot provide a direct measurement of the strength of non-covalent interactions.

Isothermal titration calorimetry (ITC) is the gold standard for the study of protein-ligand interactions. It directly measures the heat generated or absorbed when molecules interact. It collects data such as the number of binding sites and thermodynamics data.<sup>14-16</sup> However, conventional ITC method usually requires large amount (~mg) of pure protein and ligand for each analysis and suffers from low throughput (5-10 measurements/day). Recently, improvements in ITC technologies has introduced nanoITC which exhibits improved sensitivity and require lower sample amounts.<sup>17</sup>

Surface plasmon resonance (SPR) spectroscopy is another commonly used quantitative method for measuring both kinetic<sup>18-19</sup> and thermodynamic<sup>20-22</sup> parameters. This technique provides high sensitivity and requires a relatively small amount of sample (in the nanogram range). A limitation of SPR is the need to immobilize one of the binding partners (often the one with a lower molecular weight) on a sensor chip, which may affect the nature of the binding interaction.

Nuclear magnetic resonance (NMR) can also be used to characterize the structures of biological molecules and their complexes in solution and to quantify the strength of the interaction (usually reported as the dissociation constant,  $K_d$ ).<sup>13, 23-24</sup> However, NMR measurements are usually limited to relatively small proteins, with molecular weights lower than 40 kDa.<sup>24</sup> NMR measurements require a large amount of isotopically labeled proteins, typically in milligram quantities, which is unfavorable for biomolecule that are difficult to purify.

Small molecule microarrays have become a popular tool for the discovery of ligand interactions with proteins and proteins complexes.<sup>25-26</sup> Here, “small molecule” refers to oligosaccharides, peptide, or other ligands that are prepared using synthetic chemistry, rather than molecules that are obtained from cells or biochemical reactions, such as proteins. The arrays consist of a library of compounds that are attached to a solid support, usually through a covalent linker. Protein targets are incubated with the array and, following a washing step, specific interactions are identified, usually by a fluorescence-based detection.<sup>27</sup> However, this technique can only provide semi-quantitative binding data. Also, screening these microarrays can be problematic due to steric hindrance from the surface.<sup>28</sup> The length and the chemical properties of the linker may affect binding.<sup>29</sup>

Within the last two decades, electrospray ionization-mass spectrometry (ESI-MS) has emerged as a powerful tool for identifying and quantifying protein-ligand complexes *in vitro*.<sup>30-32</sup> ESI-MS assay offers several advantages, such as simplicity (no labeling or immobilization), speed (measurements normally can be completed in less than 1 min),

direct insight into binding stoichiometry and the ability to measure multiple binding equilibria simultaneously.

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

## **1.2 Electrospray Ionization Mass Spectrometry**

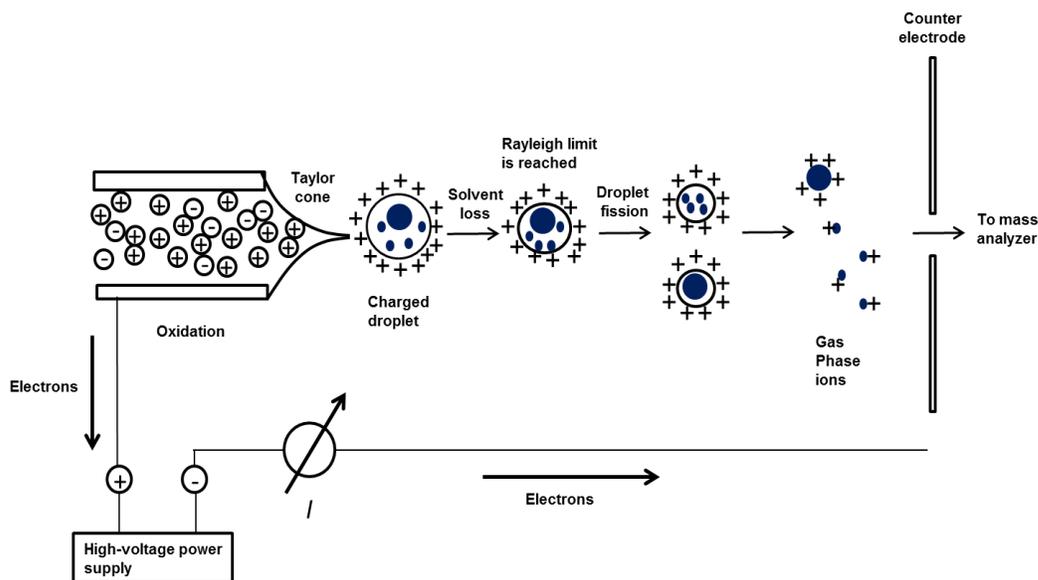
### **1.2.1 Electrospray Ionization**

Electrospray ionization (ESI) or its low flow variant nanoESI is considered as a soft ionization technique to convert solution-phase analytes into gas-phase ions, i.e., it induces little to no fragmentation which allows the production of intact gaseous ions from macromolecular analytes like proteins. This is in contrast to electron impact (EI) and other traditional approaches, where the rupture of covalent bonds is common place. ESI occurs at atmospheric pressure.<sup>33</sup>

Using ESI to form multiply charged ions has several advantages. Firstly, it allows the detection of large analytes on mass spectrometers with limited mass-to-charge ratio ( $m/z$ ) range. Also, high charge states facilitate ion dissociation in tandem MS experiments. The following includes an introduction about the mechanism about ESI and its low flow variant nanoESI.

### 1.2.1.1 ESI mechanism

There are three major processes in ESI-MS, shown in Figure 1.1 is a diagram describing the ESI process involved in the formation of gas phase ions:



**Figure 1.1** Schematic representation of process in ESI-MS in positive ion mode. Adapted from<sup>34</sup>

#### 1) The production of charged droplets

In ESI-MS, analyte solution is loaded into an ESI emitter, typically a stainless steel capillary. As shown in the schematic representation of the ESI event in the gas phase (Figure 1.1), a high voltage is applied to the emitter. An estimation of the voltage required for the onset of electrospray is given by eq. 1.1:<sup>35</sup>

$$V_0 = 2 \times 10^5 \sqrt{\sigma r_c} \ln\left(\frac{4d}{r_c}\right) \quad (1.1)$$

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

The imposed field will partially penetrate the liquid at the glass tip. To simplify the discussion, we assume all the experiments were carried in positive ion mode (as shown in Figure 1.1), which has a positive electrical potential on the glass tip. Thus, some positive ions in the liquid will move to the liquid surface while some negative ions will drift away from it until the imposed field inside the liquid surface is entirely removed by this charge distribution. However, enrichment of positive charge at the solution surface leads to surface destabilization. The positive ions are pulled downfield but cannot escape from the liquid. The surface is drawn out downfield into a cone called Taylor cone<sup>36</sup>. As the electrical field,  $E$ , increases, the cone become less stable and a liquid filament with a diameter of a few micrometers, whose surface are enriched on positive ions, is emitted from the Taylor cone tip. At some distance downstream, the liquid filament becomes unstable and forms separate droplets. The initial ESI droplets have radii in the micrometer range. Each droplet is positively charged due to the presence of excess ions. For example, if the major electrolyte present in the solution was ammonium acetate (used in this study), the excess positive ions at the surface will be mostly  $\text{NH}_4^+$ . The initial charged droplet can break up into small charged droplets which will be discussed in step 2.

## **2) Shrinkage of the charged droplets due to solvent evaporation and repeated charge induced droplet fissions**

Charged droplets in the micrometer size range emitted from the Taylor cone will undergo rapid solvent evaporation leading to the shrinkage of droplets while the charge of the

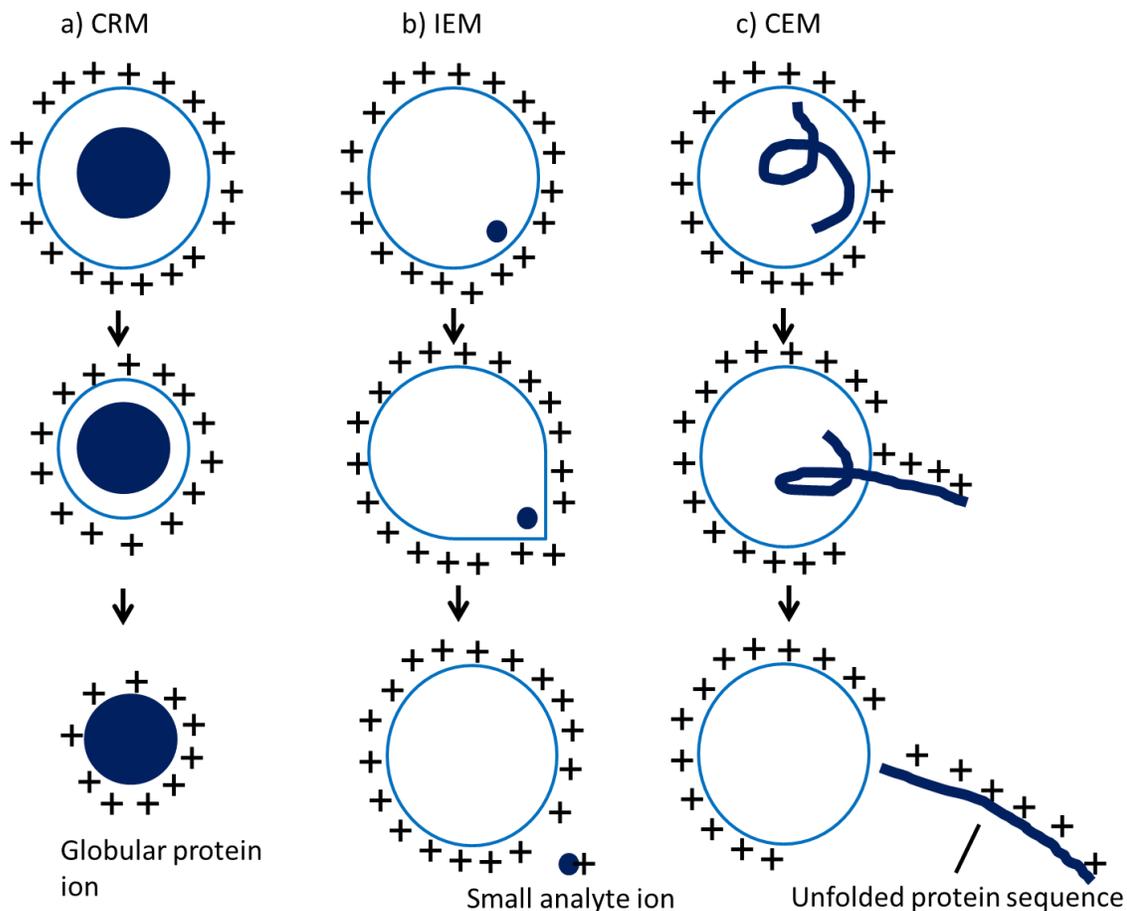
droplets' surface remains constant. The energy required for solvent evaporation is provided by the thermal energy of the ambient gas, air at atmospheric pressure in most cases. As the droplets size decrease, the surface ions would come into closer proximity to each other. When the Coulomb repulsion of the surface charges exceed the surface tension of the droplets over the same area, now referred to as the “Rayleigh limit”,<sup>34</sup> the droplet would increase the available surface area by undergoing Coulomb fission and thus resulting in the production of a jet of smaller and highly charged offspring droplets. The condition for the Coulomb fission is given by the Rayleigh equation:

$$Q_{Ry} = 8\pi\sqrt{\epsilon_0\gamma R^3} \quad (1.2)$$

where  $Q_{Ry}$  is the charge on the droplet,  $\epsilon_0$  is the electrical permittivity, and  $\gamma$  is the surface tension and  $R$  is the droplet radius. Repeated evaporation/fission events ultimately yield the final generation of ESI droplets with radii of a few nanometers.

### **3) The production of the gas-phase ions from these droplets**

The process of repeated droplet fissions of parent droplets, leading to smaller parent droplets and progeny droplets, will ultimately lead to very small charged droplets that are the precursors of the gas-phase ions. Three mechanisms, the charged residue model (CRM),<sup>37,38</sup> the ion evaporation model (IEM)<sup>39</sup> and the chain ejection model (CEM),<sup>40,41</sup> have been proposed to account for the formation of gas-phase ions from the very small and highly charged droplets. A detailed expression about each model is given below and shown in Figure 1.2.



**Figure 1.2** Summary of different models for the formation of gas phase ions, (a) CRM: formation of a globular protein into the gas phase. (b) IEM: small ion ejection from a charged nanodroplet. (c) CEM: ejection of an unfolded protein.

**a) Charged residue model (CRM):**<sup>37, 42</sup>

This model, which was originally proposed by Malcolm Dole et al. in 1968 and 1970, proposed an idea that a completion of subdivisions of the original droplets would eventually lead to the formation of “ultimate droplets” so small that each of them would contain only one solute molecule. As shown in Fig 1.2a, solvent evaporation from such a droplet will lead to a gas-phase analyte ion whose charge originates from the charges at the

surface of the vanishing droplet. It is widely accepted that large globular species such as natively folded proteins are released into the gas phase via this model.<sup>43,44-45</sup> This model has allowed quantitative predictions of the protein charge state in the gas phase using a simple correlation between charge state and protein mass and is well supported for proteins of widely varying mass.

**b) Ion evaporation model (IEM):**<sup>46,47</sup>

This model was proposed several years after the CRM model by Iribarne and Thomson<sup>46,47</sup> and offers a slightly different explanation for the possible production of gas-phase ions by evaporation of solvent from charged liquid droplets. They argued that before a charged droplet becomes small enough to contain only one solute molecule, the charge density on its surface would become so high that the resulting field would be sufficient to push one or more surface ions into the ambient gas, thereby forming gaseous ions of at least some of those solute molecules. Continued evaporation of solvent from successive generations of such charged droplets would ultimately result in driving many, if not most, of the surface cations on the original droplets into the gas phase. It predicts that direct ion emission from the droplets when the droplet size is small approximately <10 nm and the surface electric field is in the range of 1 V/nm. There are two forces contributing to the ion ejection. The leaving ion is repelled by the Coulomb repulsion from the remaining charges on the droplet. Also, the leaving ion is attracted to the droplet by the polarization of the droplet. The transition state is located where these two interactions become equal. IEM ejection requires ions to cross an activation barrier. This mechanism is used to account for the formation of

relatively small ions from ESI.<sup>46, 48</sup> However, this model does not apply for large ions such as proteins.

Interestingly, Michael Gross and co-workers<sup>49-50</sup> propose a theory of native protein ESI which combines both the CRM and the IEM, in which large proteins are charged residues, but their charge state is determined by the field emission of charged ions from highly charged droplets. This mechanism is expected to operate when salt additives (buffers) such as ammonium acetate or triethylacetate are present in millimolar concentrations in the solution that is electrosprayed.

**c) Chain Ejection model (CEM):**<sup>40, 41</sup>

This model was revealed by molecular dynamic simulations on unfolded proteins (Figure 1.2c). When protein unfolding is triggered, e.g., by the acidic pH change in surrounding solution, the conformation of the products are highly disordered, and nonpolar residues that were previously sequestered in the core become solvent accessible. Thus, unfolding switches the properties of the protein from compact/hydrophilic to extended/hydrophobic.<sup>51</sup> The hydrophobic character of unfolded protein makes it hard to stay within the droplet. Instead, the unfolded chain would migrate to the droplet surface when placed in a Rayleigh-charged nanodroplet. It is followed by one chain terminus's discharge to the gas phase. Then it undergoes stepwise ejection of the remaining protein and separation from the droplet. As reflected in the spectrum, unfolded proteins demonstrate a wider range of charge states centered at much lower  $m/z$  values. This CEM model also applies to polymer chains that are disordered, partially hydrophobic or capable of binding excess charge carrier.<sup>41, 52</sup>

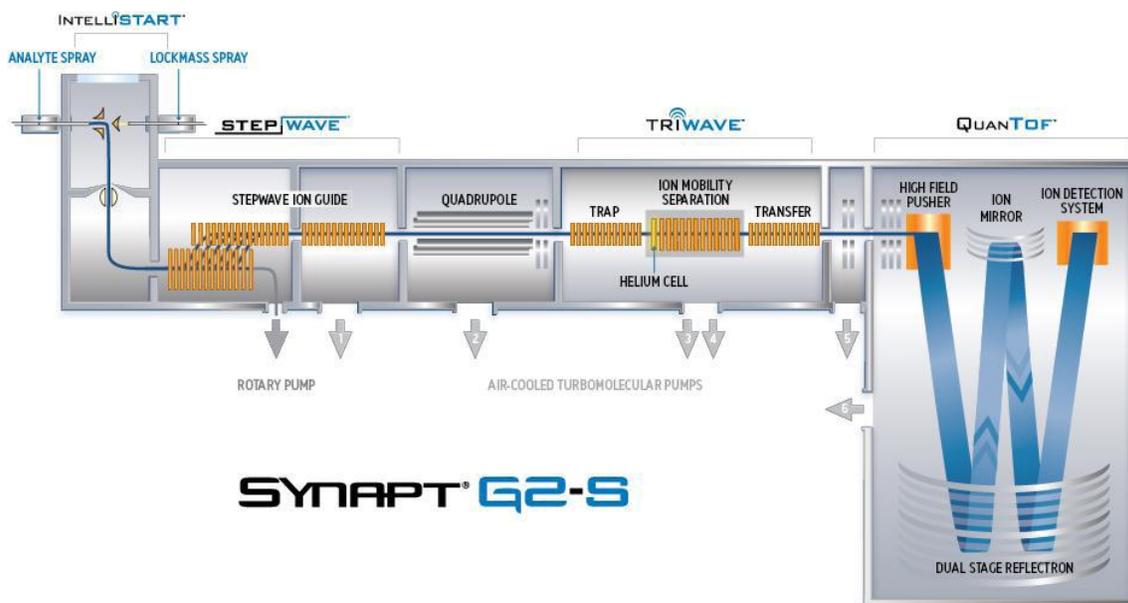
### ***1.2.1.2 Nano ESI***

In the present work, nanoESI, introduced by Matthias Wilm and Matthias Mann<sup>53-54</sup>, is applied in all the studies because it proves to be superior for the investigation of native protein complexes. The mechanism of nanoESI is the same as that of ESI, except for the use of a narrow glass tip which has an emitter tip opening of only a few  $\mu\text{m}$ , instead of approximately 100  $\mu\text{m}$  employed in conventional ESI. NanoESI has lower solution flow rates of 10-50 nL/min compared to 1-10  $\mu\text{L}/\text{min}$  of traditional ESI. What truly makes nanoESI the method of choice for investigating non-covalent protein-ligand complexes than conventional ESI are the following:

Importantly, only picomoles or less of the analyte are required per analysis, a very important feature in the analysis of biological molecules. Also, nonspecific aggregation which can occur during the ESI process is minimized as there are fewer analyte molecules per droplet.<sup>55-57</sup> Moreover, by using a smaller spray orifice, the size of the initially produced droplets is already reduced. As a result, a lower number of evaporation/ fission cycles are required before analyte ions are released in to the gas phase. By avoiding harsh desolvation conditions, the non-covalent protein-ligand complexes can be better maintained from buffered aqueous solutions to the gas phase. The spray tip being at a much smaller diameter results in a more efficient sampling of gas phase ions, hence, the tip can be placed closer to the sampling orifice resulting in more of the ions being sampled.

## 1.2.2 MS Instrumentation

In this study, the experiments were carried on Synapt G2-S hybrid Quadrupole Time of Flight mass spectrometer (HDMS) (Waters, UK) equipped with an ESI source shown in Figure 1.3.



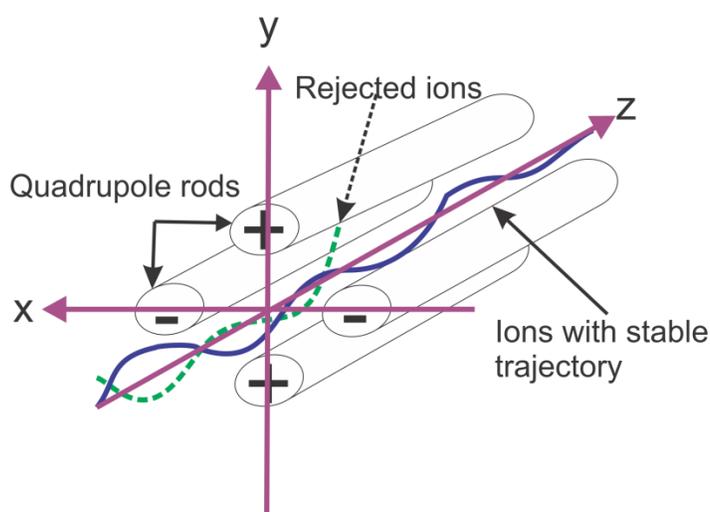
**Figure 1.3** Schematic diagram of Synapt G2-S Q-IM-TOF mass spectrometer, adapted from the Waters user's manual.

Briefly, gaseous ions are produced by nanoESI using borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.) pulled to  $\sim 5 \mu\text{m}$ . A platinum wire is inserted into the nanoESI tip and a capillary voltage of 1.0–1.3 kV is applied to carry out ESI. Formed gaseous ions are introduced into the mass spectrometer and entered into the StepWave transfer optic, which can minimize neutral contamination and enhance the signal-to-noise ratio. The resulting ions are then transmitted through a quadrupole mass filter to the Triwave section of the instrument which contains three traveling wave ion guides, Trap collision cell, Ion mobility

cell, Transfer collision cell, and before finally reaching the orthogonal acceleration reflectron TOF mass analyzer. To perform tandem MS, ions of interest can be isolated by the quadrupole mass filter and subjected to collision-induced dissociation (CID) in both Trap and Transfer regions. Detailed working mechanisms of quadrupole mass filter, traveling wave ion guides, TOF mass analyzer and tandem MS are given below.

### 1.2.2.1 *Quadrupole mass filter*

The quadrupole consists of four cylindrical metal rods which are accurately positioned in a radial array, with the diametrically opposed rods paired with each other. A Direct Current (DC) potential and a Radio Frequency (RF) potential, are applied to each pair of rods (the same absolute potential with different sign to each rod) to create a hyperbolic field.<sup>58</sup> By applying specific voltage and frequency, ions which possess a small range of mass-to-charge ratio ( $m/z$ ) values can be selected and transmitted through the quadrupole, whereas other ions with  $m/z$  values out of the small range will hit the rods and are discharged. The width of the bandpass region is controlled by the DC and RF potentials applied to the rods.

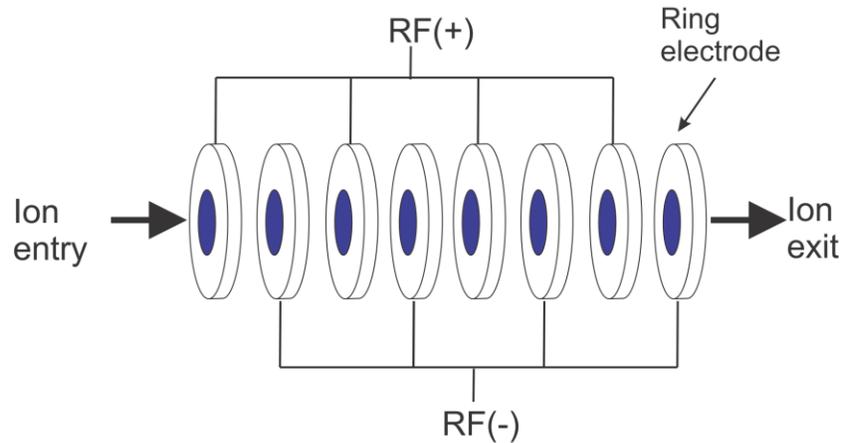


**Figure 1.4** Diagram illustrating the quadrupole used in the Waters Synapt mass spectrometers.

In the MS-TOF mode, the quadrupole operates only in the RF mode acting as a broad bandpass filter, which transmits and guide ions over a wide  $m/z$  range to other components of the apparatus. In the MS/MS-TOF mode, the quadrupole operates with both RF and DC. Depending on the specific voltage and frequency applied, ions of a particular  $m/z$  can be selected and move through the entire length of the rods; other ions outside the  $m/z$  range hit the rods and are expelled.

#### ***1.2.2.2 Travelling wave ion guide (TWIG)***

Travelling wave technology is applied on the StepWave ion guide and the TriWave section (Trap, IMS and Transfer) of Synapt G2S mass spectrometer. They function as low resolution mass analyzer to trap, focus, release, fragment and separate ions. These travelling-wave devices consist of a sequence of ring-shaped electrodes orthogonal to the ion transmission axis,<sup>59</sup> as shown in Figure 1.5. The ring electrodes are supported on printed circuit boards that deliver both RF and DC voltages. Adjacent rings have opposite phases of RF voltage applied to them, which radially confine the ions within the device while allowing them to pass unhindered along the axis. A DC voltage can be applied to a pair of adjacent rings, which produces a potential barrier within the device that the ions cannot cross. Subsequently, the DC voltage is 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 are consequently propelled through the device with the waves.



**Figure 1.5** Diagram of a stacked ring ion guide used in Waters Synapt mass spectrometers.

In the MS-TOF mode, the trap T-wave traps and accumulates ions, after which the ions are released in a packet into the IMS T-Wave, where IMS is performed. A high-pressure helium-filled cell at the front of the IMS T-wave cell is used to minimize scattering and /or fragmentation by collisionally cooling transferred ions. The transfer T-wave delivers the mobility separated ions to the QuanTOF for detection. In the MS/MS-TOF mode, collision-induced dissociation (CID) can be performed in both the trap and transfer regions ( $\sim 10^3$  mbar).

### 1.2.2.3 Time of flight (TOF) analyzer

TOF analyzers measure the flight time of the ions to move through a flight tube between the source and detector, and determine their  $m/z$  values using Equation 1.3:

$$\sqrt{\frac{m}{z}} = t \left( \frac{\sqrt{2eV_s}}{L} \right) \quad (1.3)$$

where  $m$  is the mass of the ion,  $z$  is the charge state of the ion,  $t$  is the flight time,  $e$  is the elementary charge,  $V_s$  is the acceleration potential, and  $L$  is the length of the flight tube. In general, time  $V_s$  and  $L$  are kept constant during analysis. According to this equation, the lower the  $m/z$  of the ion, the faster it will reach the detector.

There are two types of TOF analyzers: linear TOF analyzer and reflectron TOF analyzer. The linear TOF analyzer has the drawback that ions of the same  $m/z$  may reach the detector at different times, due to the initial energy distribution, resulting in peak broadening and poor resolution. Instead, Waters Synapt G2S equipped a reflectron TOF analyzer, as shown in Fig. 1.3. The energy distribution of ions is compensated by using an ion mirror which consists of successive sets of electric plates of increasing potential. The ion mirror can deflect the ions and reverse their flight direction. The fast ions penetrate deeper into the field and take a longer time to return than slow ions. Therefore, fast and slow ions with the same  $m/z$  arrive at the same time at the detector. The net effect is improved mass resolution with minimal losses in sensitivity.

#### ***1.2.2.4 Tandem MS (MS/MS)***

Tandem MS, also known as MS/MS, is an MS technique involving multiple sequential stages of mass spectrometry segments where ions are selected, energetically activated, dissociated and analyzed. It has emerged as a useful tool for probing the structure and stability of biomolecules and their complexes. For example, it has been used for the study of protein composition and topology,<sup>60-61</sup> and identification and localization the bound ligand.<sup>62-65</sup>

MS/MS experiments involve at least two stages of mass analysis on the gaseous ions. After isolation of precursor ions of interest in the first-stage of mass selection, the ions are then activated by various methods which will be elaborated on later and undergo fragmentation. In the second-stage of mass analysis, the product ions from the fragmentation are detected. The resulting product ions can provide structural insights regarding the parent ion. Various ion activation techniques are available including collision activated/induced dissociation (CAD and CID),<sup>66-68</sup> surface-induced dissociation (SID),<sup>69</sup> electron transfer/electron capture dissociation (ETD and ECD),<sup>70-71</sup> blackbody infrared radiative dissociation (BIRD),<sup>72</sup> and infrared multiphoton dissociation (IRMPD).<sup>73</sup> For the purpose of this study, CID techniques are discussed in a greater detail in the following sections.

CID can be performed in the Trap and Transfer cells of 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. The selected ions of interest are subjected to energetic collisions with a neutral background gas (often nitrogen or argon) within Tap and/or Transfer cells. During the collision, a

portion of the ion's kinetic energy is converted into internal energy. This results in the subsequent fragmentation if sufficient internal energy is accumulated, which ultimately leads to decomposition. The increase in internal energy is determined by many factors, including the number of collisions between ions and gas, the amount of time ions spend in the collision cell and the nature of the target ions.

A key feature of CID is that decomposition is slower than energy randomization. Thus the energy will be distributed across all internal modes of the ion with equal probability which leads to preferential decomposition at the weakest sites.<sup>74</sup> This feature has allowed CID to be used as a tool to probe the stability of non-covalent complexes in the gas phase. However, because ions do not reach thermal equilibrium during CID, the kinetic parameters for the dissociation process cannot be measured. In addition, non-covalent interactions are usually broken prior to covalent bond cleavage; the binding site cannot be localized by CID. Examples of CID applied for studying protein-ligand bindings can be found in Chapter 2 and 4 in this thesis.

After reviewing the basic principles and instrumentations of ESI-MS, the ESI-MS method used to quantify non-covalent protein-ligand interactions and its pitfalls are given in the following sections.

### **1.3 MS-based method for quantifying interactions of non-covalent complexes**

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

The direct ESI-MS assay is based on the detection and quantification of free and ligand-bound protein ions by ESI-MS. The association constant  $K_a$  for a given protein-ligand interaction is determined from the ratio ( $R$ ) of total abundance ( $Ab$ ) of all ligand-bound and

free protein ions, as measured by ESI-MS for solutions with known initial concentrations of protein ( $[P]_0$ ) and ligand ( $[L]_0$ ). For a 1:1 protein-ligand complex (eq. 1.4), Apparent association constant  $K_{a, app}$  can be expressed by eq. 1.5:



$$K_a = \frac{R}{[L]_0 - \frac{R}{1+R}[P]_0} \quad (1.5)$$

where  $[P]_0$  and  $[L]_0$  are the initial concentrations of the protein and ligand, respectively.  $R$  is given by eq 1.6:

$$\frac{[PL]_{eq}}{[P]_{eq}} = \frac{Ab(PL)}{Ab(P)} = R \quad (1.6)$$

Normally,  $K_a$  for a particular protein-ligand interaction is not determined at a single concentration of protein and the ligand but rather from measurements performed at different concentrations or from titration experiments, wherein the concentration of one analyte (usually P) is fixed and the concentration of another is varied. The value of  $K_a$  can be extracted using nonlinear regression analysis of the experimentally determined concentration-dependence of the fraction of ligand-bound protein, i.e.,  $R/(R+1)$ , which can be found from the following expression:

$$\frac{R}{R+1} = \frac{1 + K_a[L]_0 + K_a[P]_0 - \sqrt{(1 - K_a[L]_0 + K_a[P]_0)^2 + 4K_a[L]_0}}{2K_a[P]_0} \quad (1.7)$$

It follows that  $K_a$  values accessible with the direct ESI-MS binding assay range from approximately  $10^2$  to  $10^7 M^{-1}$ <sup>32</sup> which suits the study of protein-ligand interactions.

### 1.3.2 Limitations of the direct ESI-MS binding assay

As shown in eq.1.6, the successful implementation of the direct ESI-MS assay requires that the equilibrium abundance ratio of bound-to-free protein present in MS is the same as the equilibrium concentration ratio. Physical or chemical processes that alter this ratio will lead to incorrect  $K_a$  values and, potentially, obscure the true binding stoichiometry. Four common sources of error associated with ESI-MS measurements are: (1) non-uniform response factors, (2) in-source dissociation, (3) nonspecific ligand-protein binding, and (4) ESI-induced changes in solution, pH and temperature. For the purpose of this study, the first three sources of error are briefly described below, along with current strategies for minimizing their effects on the binding measurements.

#### 1.3.2.1 Non-uniform response factors

The abundance of each protein species measured by ESI-MS is related to the solution concentration by its response factor ( $RF$ ), which collectively accounts for the ionization and detection efficiencies:<sup>75</sup>

$$\frac{[PL]_{\text{eq}}}{[P]_{\text{eq}}} = \frac{Ab(PL) / RF_{\text{PL}}}{Ab(P) / RF_{\text{P}}} = RF_{\text{P/PL}} \frac{Ab(PL)}{Ab(P)} \quad (1.8)$$

An essential assumption in the above expression is that the  $RF$  values for P and PL ions are equal (i.e.,  $RF_{\text{P/PL}} \approx 1$ ). This assumption, which is applied in the present work, is valid in cases when the L is small compared to the P, such that the size and surface properties of P and PL are similar.<sup>75</sup> However, non-uniform ESI-MS response factors are generally expected in the case of protein-protein interactions.  $RF$ s depend on many factors, the size and structure of P and PL, the ESI conditions and the instrumental parameters used for the

measurements. It is often possible to “tune” the experimental conditions to achieve the correct  $R$  value, based on the known  $K_a$ . Such an approach can be used when investigating structurally-related interactions, but, must be used with caution.

Several strategies have been developed to minimize the effects of non-uniform  $RF$ s on the determination of  $K_a$  values.<sup>76-78</sup> One strategy involves fitting  $RF_{P/PL}$ , as an adjustable parameter, to the experimental titration data based on an appropriate binding model. This approach can account for in-source dissociation,<sup>79</sup> and can be used when complexes of variable stoichiometry are present in solution.<sup>76, 79</sup> However, this method requires high quality experimental data to obtain reliable  $K_a$  values due to the multiple adjustable parameters in the fitting model.<sup>80</sup> It is important to note that this approach assumes that  $RF_{P/PL}$  is independent of the range of concentration being investigated. While there have been several successful demonstrations of this strategy, the generality of the approach has not been established. Another method proposed involves the use of an internal standard (IS). An appropriate IS is one macromolecule that has similar MW and surface activity to the analyte of interest P, but which does not bind to L.<sup>76</sup> The advantage of this approach is that changes/fluctuations in  $RF_{P/PL}$  due to concentration, instability in the ESI, or other factors are reflected in the abundance of the IS.<sup>32</sup>

### ***1.3.2.2 In-source dissociation***

Collision-induced dissociation of the gaseous PL complexes in the ion source can alter the relative abundance of PL and P ions.<sup>81</sup> For a 1:1 P-L complex, in-source dissociation will necessarily decrease the magnitude of  $K_a$ . In extreme case, where no PL complexes survive for detection, this generates a false negative result. In-source dissociation depends on the

configuration of the ion source used, the choice of instrumental parameters and the gas-phase stability of the complex being investigated. Often, the occurrence of in-source dissociation can be identified from changes in  $R$  resulting from changes in ion source parameters, in particular voltage differences in regions of high pressure, that influence the internal energy of the ions.

In cases where the gaseous complexes are susceptible to in-source dissociation, low temperatures (drying gas, sampling capillary), low potentials across lens elements, and short accumulation times are essential for obtaining reliable binding constants. However, these conditions normally reduce signal intensities. Thus, a balance must be maintained between minimizing dissociation and obtaining mass spectra with sufficient signal-to-noise ratio. Addition of small organic molecules such as imidazole into the ESI solution can minimize the in-source dissociation.<sup>81-82</sup> The origin of the stabilizing effects of imidazole is hypothesized to be due to enhanced evaporative cooling from the dissociation of nonspecific imidazole adducts from the gaseous PL ions.<sup>81</sup> Additionally, imidazole, has a relatively high gas phase basicity and a relatively low gas phase acidity.<sup>83</sup> This may lead to a reduction in the charge states of the protein complex ions. Moreover, the introduction of imidazole vapor to the ESI ionization source was demonstrated to protect complexes against in-source dissociation.<sup>84</sup> These approaches have their limitations as labile gas phase complexes, rapidly dissociate at ambient temperature, which makes their detection by ESI-MS challenging.

### ***1.3.2.3 Nonspecific binding***

It is well established that free L can bind nonspecifically to P and PL due to concentration effects, which results in false positives during the ESI process. Consequently, the observation of gaseous ions corresponding to a particular PL complex does not, by itself, establish the presence of that interaction in solution. The observation of multiple ligands bound to the target protein with a Poisson-like distribution suggests the occurrence of nonspecific ligand binding. Changes in the magnitude of  $K_a$  with changes in ligand concentration may signify the occurrence of nonspecific ligand binding.

To understand the formation of nonspecific protein-ligand complexes, we need to examine the charge residue model (CRM). According to the CRM, initial ESI droplets undergo solvent evaporation until they approach the Rayleigh limit, at which point fission occurs, resulting in the release of several small multiply charged nanodroplets, which are often referred to as offspring droplets, containing no analyte or one or more molecules of analyte. Solvent evaporation from the nanodroplets ultimately yields gaseous ions. If a nanodroplet contains two or more analyte molecules, nonspecific intermolecular interactions can occur as the droplet evaporates to dryness, which leads to the formation of nonspecific complexes. Nonspecific binding of L to P and PL causes deviations of the binding stoichiometry from its true value in solution and introduces error into the  $K_a$  values measured by ESI-MS. Formation of nonspecific protein-ligand complexes can be minimized by decreasing initial concentration of ligand. For very weak ligand interactions,  $K_a < 10^4 \text{ M}^{-1}$ , however, high initial concentrations of ligand are required to produce detectable level of complexes. In such cases, nonspecific binding is often unavoidable.

A number of strategies have been developed to correct ESI mass spectra for the occurrence of nonspecific binding.<sup>81, 85-89</sup> The most straightforward approach is by the reference protein method which involves addition of a reference protein ( $P_{ref}$ ) that does not bind specifically to the protein and ligand of interest to the solution.<sup>86</sup>

The method assumes that nonspecific ligand binding is random, as suggested by the observation that the distribution of nonspecifically bound molecules often resembles that of a Poisson distribution, and affects equally all protein species present in the ESI droplets. The occurrence of nonspecific protein-ligand binding is monitored by the appearance of ions corresponding to nonspecific complexes of  $P_{ref}$  and L in the mass spectrum.<sup>86</sup> The fraction abundance of  $P_{ref}$  undergoing nonspecific ligand binding provides a quantitative measure of the contribution of nonspecific binding to the measured intensities of protein and specific protein-ligand complexes. For a given  $PL_i$  species, its apparent abundance ( $Ab_{app}(PL_i)$ ) measured by ESI-MS is composed of the true abundance of  $PL_i$  ( $Ab(PL_i)$ ) and the abundances of  $PL_{i-1}$ ,  $PL_{i-2}$ , ..., P that nonspecifically bind to 1, 2, ..., i molecules of L, respectively. Using this reference protein method, the “true” abundance of a given  $PL_i$  species ( $Ab(PL_i)$ ) can be calculated from  $Ab_{app}(PL_i)$  and the distribution of nonspecific  $P_{ref}L_i$  species using the following expression:<sup>86</sup>

$$Ab(PL_i) = [Ab_{app}(PL_i) - f_{1,P_{ref}}Ab(PL_{i-1}) - \dots - f_{i,P_{ref}}Ab(P)] / f_{0,P_{ref}} \quad (1.8)$$

Where  $f_{i,P_{ref}}$  is the fractional abundance of  $P_{ref}$  bound to i molecules of L as shown in eq.1.9:

$$f_{i,P_{ref}} = \frac{Ab(P_{ref}L_i)}{Ab(P_{ref}) + Ab(P_{ref}L_1) + Ab(P_{ref}L_2) + \dots + Ab(P_{ref}L_i) + Ab(P_{ref}L_{i+1}) + \dots} \quad (1.9)$$

In a more comprehensible way, eq 1.8 can be converted into eq 1.10

$$R_i = R_{i,\text{app}} - R_{1,\text{P}_{\text{ref}}} R_{i-1} - R_{2,\text{P}_{\text{ref}}} R_{i-2} - \dots - R_{i-1,\text{P}_{\text{ref}}} R_1 - R_{i,\text{P}_{\text{ref}}} \quad (1.10)$$

Where  $R_i$  is the “true” abundance ratio of  $i$ -ligand-bound to free protein ions; and  $R_{i,\text{app}}$  and  $R_{i,\text{P}_{\text{ref}}}$  are defined in eqs 1.11a and 1.11b, respectively, are values that can be measured from mass spectra.

$$R_{i,\text{app}} = \frac{Ab_{\text{app}}(\text{PL}_i)}{Ab_{\text{app}}(\text{P})} \quad (1.11a)$$

$$R_{i,\text{P}_{\text{ref}}} = \frac{Ab(\text{P}_{\text{ref}}\text{L}_i)}{Ab(\text{P}_{\text{ref}})} \quad (1.11b)$$

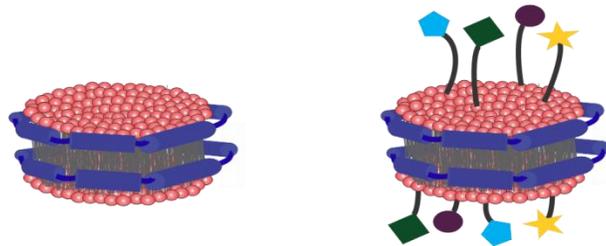
The underlying assumption of the reference protein method is that the nonspecific ligand binding is a random process and affects equally to all protein species present in the ESI droplets. This assumption has been rigorously tested and proven to valid for the nonspecific binding of neutral, acidic and basic biomolecules to proteins during ESI-MS analysis.<sup>86-88</sup> Notably, the corrected binding constants are in good agreement with what have been reported using other methods.

#### 1.4 Nanodisc technology

To investigate protein-glycolipid interaction in vitro, model membrane systems are essential to solubilize glycolipids and maintain natural environment for them since glycolipids are naturally embedded in a dynamic lipid bilayer. Currently, there are a variety of model membranes available, including detergent micelles, of which the hydrophilic head groups are exposed to solvent which hydrophobic tails are toward the center, liposomes

which is a lipid bilayer rolled up into a hollow spherical shell and enclosing a small region of water, bicelles, of which the bilayer centre is shielded by another micelle-like assembly formed by detergent and nanodisc which is a lipid bilayer surrounded by two copies of membrane scaffold proteins (MSPs).

Compared to detergent micelles, liposomes and bicelles, nanodisc (ND) possesses the advantages of self-assembly, solubility, size-control and stability. For the purpose of this study, a brief presentation of the ND technology will be given. As shown in Figure 1.6, each nanodisc is composed by a certain number of phospholipids in the form of a phospholipid bilayer, which is enclosed by two MSPs. The two MSPs situate themselves at the rim of the bilayer, shielding the hydrophobic alkyl chains of the phospholipids of each leaflet from the aqueous environment, thereby creating a soluble nanoscale lipid bilayer.



**Figure 1.6** Cartoon of nanodisc. Blue belts represent two copies of MSP which surround the lipid bilayer without (left) or with (right) cellular receptors incorporated.

The membrane scaffold protein has been developed by Sligar et al., and is based upon a modified sequence of the human serum apolipoprotein AI (APO-A1).<sup>90-91</sup> The genetic modifications have been performed on APO-A1 to enhance monodispersity of the formed disc, and to ensure optimal expression in *E. coli*. The MSP constructs have been engineered

in different lengths, to allow the formation of nanodiscs of different diameters. Additionally, the constructs have been outfitted with different affinity tags at the end of the sequence, such as 6\*His-tag, FLAG-tag and Cysteine. Specific protease cleavage sites have also been added to this part of the sequence in order to remove affinity tag after purification. The first paper introducing the nanodisc concept was published by Bayburt and Sligar in 2002, and presented two MSP constructs called MSP1 & MSP2.<sup>90</sup> After that, several other MSPs were developed to serve different needs.<sup>92</sup>

NDs are formed from a mixture containing an optimal ratio of detergent-solubilized MSP molecules and phospholipids. By conducting empirical experiments, the optimal ratios were determined for assembling nanodiscs composed of different MSPs with the lipids POPC, DPPC, and DMPC. Different ratios have been analyzed by size exclusion chromatography (SEC) in order to determine the optimal ratio between MSP and phospholipids. The results shown in the following Table 1.1 are reproduced from the assembly guidelines provided by Sligar et. al.<sup>90-91</sup>

**Table 1.1** Optimal molar ratios of phospholipids to MSPs, diameters and incubation temperature for nanodisc assembly. The diameter of nanodisc is determined by the length of MSP, while the incubation temperature is the phase transition temperature of the chosen lipid.<sup>91-92</sup>

Optimal Ratios	MSP1D1	MSP1E1	MSP1E1D1	MSP1E2D1	MSP1E3D1	Incubation Temperature
POPC	61	79	79	103	125	4°C
DPPC	82	106	106	134	167	37°C
DMPC	77	100	102	122	148	23°C

Diameter of nanodisc (Å)	98	105	106	119	129
-----------------------------------	----	-----	-----	-----	-----

---

Data adapted from references 91 and 92.

The distinct characteristics of ND offer diverse biochemical applications. To date, it has been applied to receptor studies, enzyme studies, ligand binding, channels and transporters, and structural investigations.<sup>93-95</sup> Our group is the first to report the combination of ND and ESI to study protein-glycolipid interactions.<sup>96</sup> With the purpose to apply this technique, two different phospholipids (DMPC, POPC) and MSP1E1 were used to prepare ND separately in the present study. NDs with varying degrees of glycosphingolipids (GSLs), the receptors studied in Chapter 2, were prepared by mixing the desired ratios of DMPC and GSL, keeping the total GSL/DMPC POPC molar amount constant. Also, in Chapter 3, empty ND is served as membrane mimic to study protein-membrane interactions.

### 1.5 The present work

This thesis focuses on the development and application of ESI-MS to study non-covalent interactions in the gas phase. Chapter 2 focuses on identifying binding partners for *Clostridium difficile* toxins. First, the application of the direct ESI-MS assay for screening and quantifying the stoichiometry and absolute affinity between *C. difficile* toxins and its carbohydrate ligands is demonstrated. Then it demonstrates the application of nanodisc-based catch and release assay to screen glycolipids for *C. difficile* toxins' binding receptors. Chapter 3 extends ESI-MS assay to study peptide-membrane interactions using melittin and POPC phospholipid bilayer as model system.

## 1.6 Literature cited

1. Cleveland, D. W.; Mao, Y. H.; Sullivan, K. F. *Cell* **2003**, *112*, 407-421.
2. Igakura, T.; Stinchcombe, J. C.; Goon, P. K. C.; Taylor, G. P.; Weber, J. N.; Griffiths, G. M.; Tanaka, Y.; Osame, M.; Bangham, C. R. M. *Science* **2003**, *299*, 1713-1716.
3. Sanderson, J. M. *Org. Biomol. Chem.* **2005**, *3*, 201-212.
4. Matsuzaki, K. *Biochem. Soc. Trans.* **2001**, *29*, 598-601.
5. Sugawara, T.; Kuwajima, K.; Sugai, S. *Biochemistry* **1991**, *30*, 2698-2706.
6. Srisa-Art, M.; Dyson, E. C.; deMello, A. J.; Edell, J. B. *Anal. Chem.* **2008**, *80*, 7063-7067.
7. Malnasi-Csizmadia, A.; Pearson, D. S.; Kovacs, M.; Woolley, R. J.; Geeves, M. A.; Bagshaw, C. R. *Biochemistry* **2001**, *40*, 12727-12737.
8. Bizzarri, A. R.; Cannistraro, S. *J. Phys. Chem. B* **2009**, *113*, 16449-16464.
9. Hill, J. J.; Royer, C. A. In *Fluorescence Spectroscopy*, Brand, L.; Johnson, M. L. **1997**, *278*, 390-416.
10. Bao, J.; Krylova, S. M.; Wilson, D. J.; Reinstein, O.; Johnson, P. E.; Krylov, S. N. *Chembiochem.* **2011**, *12*, 2551-2554.
11. Ho, J. G. S.; Greco, A.; Rupnik, M.; Ng, K. K. S. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 18373-18378.
12. Greco, A.; Ho, J. G.; Lin, S. J.; Palcic, M. M.; Rupnik, M.; Ng, K. K. *Nat. Struct. Mol. Biol.* **2006**, *13*, 460-1.
13. Sachchidanand; Lequin, O.; Staunton, D.; Mulloy, B.; Forster, M. J.; Yoshida, K.; Campbell, I. D. *J. Biol. Chem.* **2002**, *277*, 50629-50635.

14. Saboury, A. A. *J. Iran. Chem. Soc.* **2006**, *3*, 1-21.
15. Wilcox, D. E. *Inorg. Chim. Acta* **2008**, *361*, 857-867.
16. Velazquez-Campoy, A.; Freire, E. *Biophys. Chem.* **2005**, *115*, 115-124.
17. Utsuno, K.; Uludag, H. *Biophys. J.* **2010**, *99*, 201-207.
18. Daghestani, H. N.; Day, B. W. *Sensors* **2010**, *10*, 9630-9646.
19. De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. *Cell. Mol. Bioeng.* **2008**, *1*, 204-215.
20. Piliarik, M.; Vaisocherova, H.; Homola, J. *Methods Mol. Bio.* **2009**, *503*, 65-88.
21. Karlsson, R. *J. Mol. Recog.* **2004**, *17*, 151-161.
22. Schuck, P. *Annu. Rev. Bioph. Biom.* **1997**, *26*, 541-566.
23. Gilbert, M.; Albala, J. S. *Curr. Opin. Chem. Biol.* **2002**, *6*, 102-105.
24. Wishart, D. *Curr. Pharm. Biotechnol.* **2005**, *6*, 105-120.
25. Horlacher, T.; Seeberger, P. H. *Chem. Soc. Rev.* **2008**, *37*, 1414-1422.
26. Laurent, N.; Voglmeir, J.; Flitsch, S. L. *Chem. Commun.* **2008**, *37*, 4400-4412.
27. Liu, Y.; Palma, A. S.; Feizi, T. *Biol. Chem.* **2009**, *390*, 647-656.
28. He, X. G.; Gerona-Navarro, G.; Jaffrey, S. R. *J. Pharmacol. Exp. Ther.* **2005**, *313*, 1-7.
29. Barnes-Seeman, D.; Park, S. B.; Koehler, A. N.; Schreiber, S. L. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 2376-2379.
30. Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.
31. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611.

32. Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431-441.
33. Covey, T. R.; Thomson, B. A.; Schneider, B. B. *Mass Spectrom. Rev.* **2009**, *28*, 870-897.
34. Kebarle, P.; Verkerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
35. Smith, D. P. H. *IEEE Trans. Ind. Appl.* **1986**, *22*, 527-535.
36. Taylor, G. *P. Roy. Soc. A.* **1964**, *280*, 383-397.
37. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.
38. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Chem. Phys.* **1970**, *52*, 4977-4986.
39. Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.
40. Konermann, L.; Rodriguez, A. D.; Liu, J. *Anal. Chem.* **2012**, *84*, 6798.
41. Ahadi, E.; Konermann, L. *J. Phys. Chem. B* **2012**, *116*, 104.
42. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Chem. Phys.* **1970**, *52*, 4977-4986.
43. De La Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93-104.
44. Fenn, J. B.; Rosell, J.; Meng, C. K. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 1147-1157.
45. Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804-817.
46. Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287.
47. Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, *71*, 4451-4463.
48. Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 3654-3668.
49. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 6926-6927.

50. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *Anal. Chem.* **2009**, *81*, 369-377.
51. Creighton, T. E., *Proteins: Structures and Molecular Properties*. W. H. Freeman: **1993**.
52. Chung, J. K.; Consta, S. *J. Phys. Chem. B* **2012**, *116*, 5777-5785.
53. Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390-4399.
54. Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.
55. Lorenzen, K.; Duijn, E. V. *Current Protocols in Protein Science*, John Wiley & Sons, Inc.: **2001**.
56. Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300-308.
57. Karas, M.; Bahr, U.; Dulcks, T. *Fresenius J. Anal. Chem.* **2000**, *366*, 669-676.
58. Hoffmann, E. d.; Stroobant, V., *Mass Spectrometry: Principles and Applications, 3rd Edition*. Wiley: **2007**, *2*, 502-503.
59. 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*, 1-12.
60. Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, *13*, 597-605.
61. Baird, N. J.; Ferre-D'Amare, A. R. *RNA-Publ. RNA Soc.* **2013**, *19*, 167-176.
62. El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50-58.
63. Borch, J.; Jorgensen, T. J. D.; Roepstorff, P. *Curr. Opin. Chem. Biol.* **2005**, *9*, 509-516.

64. Ilag, L. L.; Ubarretxena-Belandia, I.; Tate, C. G.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 14362-14363.
65. Heath, B. L.; Jockusch, R. A. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1911-1920.
66. McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614.
67. Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851-1858.
68. March, R. E.; Todd, J. F. J., *Practical Aspects of Trapped Ion Mass Spectrometry Volume IV 1<sup>st</sup> Edition*. CRC Press: **2010**, *4*, 739-767.
69. Jones, C. M.; Beardsley, R. L.; Galhena, A. S.; Dagan, S.; Cheng, G.; Wysocki, V. *H. J. Am. Chem. Soc.* **2006**, *128*, 15044-15045.
70. Tang, X. J.; Brewer, C. F.; Saha, S.; Chernushevich, I.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 750-754.
71. Myers, S. A.; Daou, S.; Affar, E. B.; Burlingame, A. *Proteomics* **2013**, *13*, 982-991.
72. Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.
73. McLafferty, F. W.; Kelleher, N. L.; Begley, T. P.; Fridriksson, E. K.; Zubarev, R. A.; Horn, D. M. *Curr. Opin. Chem. Biol.* **1998**, *2*, 571-578.
74. Wyttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.*, **2007**, *58*, 511-533.
75. Tjernberg, A.; Carno, S.; Oliv, F.; Benkestock, K.; Edlund, P. O.; Griffiths, W. J.; Hallen, D. *Anal. Chem.* **2004**, *76*, 4325-4331.
76. Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708-6715.
77. Chitta, R. K.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1031-1038.
78. Wilcox, J. M.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **2008**, *80*, 2365-2371.

79. Chitta, R.; Rempel, D.; Gross, M. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1031-1038.
80. Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spec.* **2003**, *38*, 491-501.
81. Sun, N.; Sun, J. X.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1242-1250.
82. Liu, L.; Bagal, D.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2009**, *131*, 15980-15981.
83. Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413-656.
84. Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801-7806.
85. Hossain, B. M.; Konermann, L. *Anal. Chem.* **2006**, *78*, 1613-1619.
86. Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010-3018.
87. Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 8301-8311.
88. Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472-481.
89. Kitova, E. N.; Soya, N.; Klassen, J. S. *Anal. Chem.* **2011**, *83*, 5160-5167.
90. Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Nano Lett.* **2002**, *2*, 853-856.
91. Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477-3487.
92. Bayburt, T. H.; Sligar, S. G. *FEBS letters* **2010**, *584*, 1721-1727.
93. Kobashigawa, Y.; Harada, K.; Yoshida, N.; Ogura, K.; Inagaki, F. *Anal. Biochem.* **2011**, *410*, 77-83.

94. Leitz, A. J.; Bayburt, T. H.; Barnakov, A. N.; Springer, B. A.; Sligar, S. G. *Biotechniques* **2006**, *40*, 601-602.
95. Alami, M.; Dalal, K.; Lelj-Garolla, B.; Sligar, S. G.; Duong, F. *Embo Journal* **2007**, *26*, 1995-2004.
96. Zhang, Y. X.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C. S.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618-7621.

## Chapter 2

### Identifying Binding Partners for *Clostridium difficile* Toxins using ESI-MS

#### 2.1 Introduction

*Clostridium difficile* is a Gram-positive, spore forming, strict anaerobic bacterium responsible for a variety of toxin-mediated gastrointestinal diseases that range in severity from antibiotic-associated diarrhea to pseudomembranous colitis.<sup>97,98</sup> Each year in the United States, *C. difficile* causes around 250,000 clinically diagnosed cases of disease, contributing to the cause of thousands of deaths and costing the health-care system over \$8 billion.<sup>99</sup> Healthy individuals rarely develop *C. difficile* associated diseases, but patients whose healthy levels of bacteria in the gut are at low numbers, *i.e.*, during common broad-spectrum antibiotic treatment, are highly susceptible. Although *C. difficile* can produce up to six different toxins,<sup>100-101</sup> the main virulence factors are the two exotoxins, TcdA and TcdB, 250 kDa to 308 kDa in size, respectively. They are the largest bacterial toxins, with structures that can be organized into four regions, each with its unique function: a N-terminal domain contains glucosyltransferase activity, a cysteine protease domain, a delivery/pore forming domain, and a C-terminal domain consisting of combined repetitive oligopeptides (CROPs).<sup>100,102</sup> During infection, the CROPs region of Tcd binds to carbohydrate receptors present on the epithelium of the large intestine initiates ligand-mediated endocytosis followed by auto-catalytic cleavage of the glucosyltransferase domain, which is translocated to the cytosol and glucosylates target proteins.<sup>103</sup> The toxins covalently modify large number of proteins including Rho GTPases causing their functional inactivation and inducing disassembly of the actin

cytoskeleton, which in turn leads to morphological changes and death of susceptible cells, and ultimately leads to downstream effects such as inflammation and diarrhea.<sup>104, 105</sup>

The standard treatment for *C. difficile*-associated diseases is to terminate the original antibiotic treatment and administer either metronidazole or vancomycin.<sup>106</sup> However, *C. difficile* strains resistant to these antibiotics are beginning to emerge.<sup>107</sup> Additionally, both metronidazole and vancomycin are also broad-spectrum antibiotics and their use will disrupt normal colonic bacterial populations. Consequently, reemergence of *C. difficile* infection after termination of antibiotic is quite common (up to 20% of patients) and can be very difficult to treat.<sup>98</sup> Several new approaches to treat *C. difficile*-associated diseases are currently being developed, one is using narrow scope antibiotic Fidaxomicin,<sup>108</sup> however its use has not been studied in patients with multiple recurrences and is limited by its cost.<sup>109, 110</sup> Another emerging treatment is fecal microbiota transplantation, but practical barrier and safety concerns have prevented its widespread use.<sup>111</sup> Therefore, there is a clear need for more effective therapeutics treating *C. difficile* infection. It has been proposed that such a therapy may tend to provide host cell receptor analogs in various form to competitively binds to TcdA and TcdB, divert the toxin from their native receptors on the host cell surface, thus preventing *C. difficile* from colonizing the intestinal tract, suppressing the cytotoxic effects of *C. difficile* toxins and facilitating their elimination from the body.

At present, the human cellular receptors for TcdA and TcdB in humans have not been conclusively identified.<sup>112</sup> The only known native receptors for TcdA contains terminal  $\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)-D-GlcNAc sequence.<sup>113</sup> The related trisaccharide  $\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)-D-Glc and its analogs have also been

found bind weakly to fragments of TcdA and TcdB.<sup>114</sup> While such structures are found in rodent such as rabbits and hamsters,<sup>113</sup>  $\alpha$ -anomeric galactose bonds are not generally found in human tissue.<sup>115</sup> Instead, it has been suggested that the glycan receptors in humans contain the sequence  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-GlcNAc.<sup>116</sup> This sequence is found in many human glycans, including such as glycosphingolipids and human histo-blood group antigens. Indeed it has been reported that TcdA binds to the glycosphingolipid  $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-cer<sup>117-118</sup> and to several Lewis X, Y and I glycan sequences, which are known to be present on the surface of human large intestine epithelial cells, although with low affinities.<sup>116-119</sup> Screening of a glycan microarray (glycan array version3.0, Consortium for Functional Glycomics) against fragments of TcdA and TcdB revealed that all the highest binding ligands share the  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-GlcNAc (LacNAc) motif. The strongest binding ligand is identified as LacNAc-Le<sup>x</sup>, which contains a 1,3 linked  $\alpha$ -Fuc (Lewis X epitope), along with the LacNAc motif.<sup>120-121</sup> Some of other stronger binding ligands contained 2,3 linked sialic acid. It should be noted that glycan array screening provided relative binding affinity, while absolute binding affinity values cannot be determined; no ligands were identified for TcdB. Some neutral and acidic human milk oligosaccharides have also been reported to bind *C. difficile* toxin A and toxin B, albeit weakly.<sup>114, 122</sup>

With the goal of establishing a more complete view of the carbohydrate binding specificity of TcdA and TcdB, a library of thirty oligosaccharides was screened against recombinant sub-fragments of TcdA and TcdB (as shown in Figure 2.1) using electrospray ionization mass spectrometry (ESI-MS) and the affinities of identified ligands were quantified. The oligosaccharides investigated have structures found in

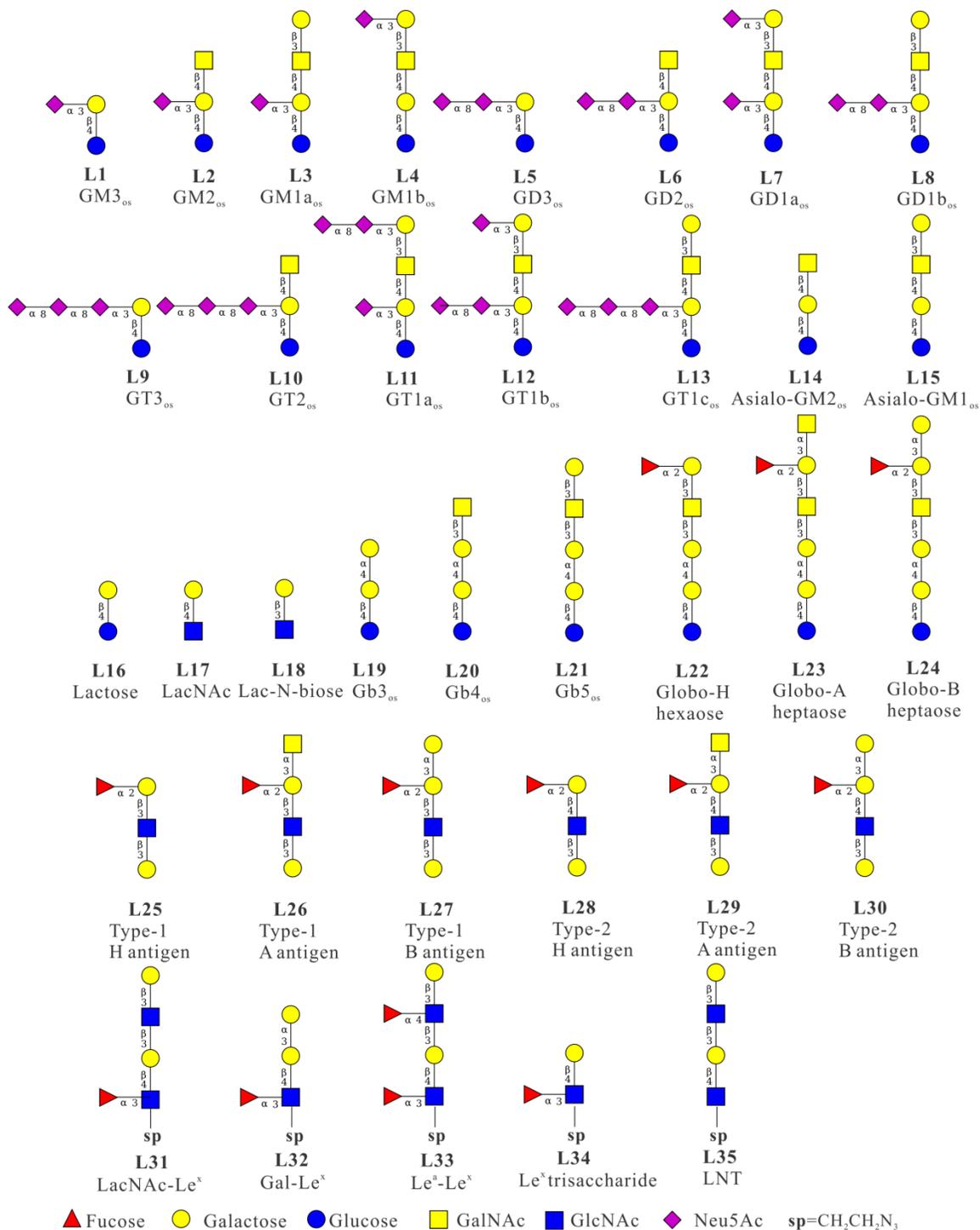
human glycans, namely gangliosides, globosides and human histo-blood group antigens (HBGA) (Structures were shown in Figure 2.2). Gangliosides and globosides are two subfamilies of glycosphingolipids (GSL) and are found in the cell membranes. Gangliosides are acidic (sialyted) GSL ( $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-ceramide), while globosides are neutral GSL and contain  $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc-ceramide. Some bacterial toxins, including cholera toxin and Shiga toxins, are known to exploit binding to glycosphingolipids to gain entry to host cells.<sup>123,124,125</sup> The HBGAs 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>126</sup> HBGA phenotype is determined by the terminal part of the oligosaccharide chain linked to protein or lipid, they all share  $\alpha$ -L-Fucose structure. Depending on how many fucose residues they carry and by what linkage is fucose attached to the precursor chain, they are divided into H, A, B and Lewis antigens. The antigen determinants can be associated with six different carbohydrate structures, i.e., precursor chain types. The present study focused on type 1 and type 2 HBGA oligosaccharides since these are widely found in red blood cells, mucosal epithelia and different organs.<sup>127</sup> Affinity measurements were also carried on TcdA and TcdB sub-fragments and a series of Lewis X analogues, which had been screened previously using a glycan microarray (Structures were shown in Figure 2.2).<sup>121</sup> These binding data allowed for a direct comparison of the ESI-MS and glycan microarray methods for screening lectins.

Additionally, with the purpose to examine the effect of glycosphingolipids' ceramide chain on protein-carbohydrate binding, 7 commercially available gangliosides,

whose oligosaccharide parts were found to interact with the *C. difficile* toxin A are chosen to screen with the same toxins TcdA-A2 and TcdB-B3 (Structure shown in Figure 2.3).

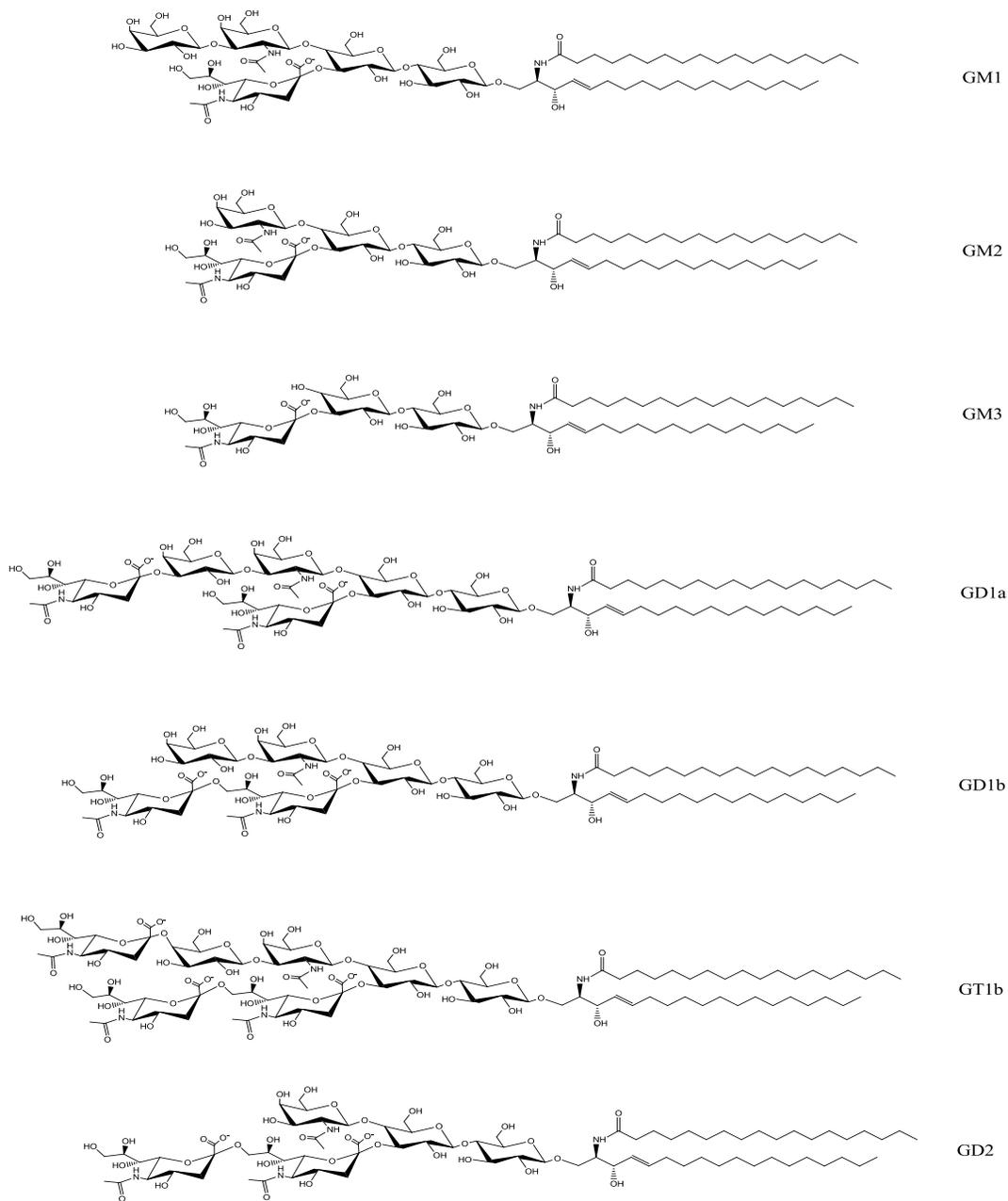


**Figure 2.1** Crystal structure of TcdA-A2, a sub-recombinant fragment of TcdA. No crystal structure is available for TcdB-B3.



**Figure 2.2** Structures of the thirty-five oligosaccharides (L1 – L35) investigated in the current study. The structures are shown here using the symbol nomenclature

adopted by the Consortium for Functional Glycomics  
([www.functionalglycomics.org](http://www.functionalglycomics.org)).



**Figure 2.3** Structures of the 7 gangliosides (GM1, GM2, GM3, GD1a, GT1b and GD2) that were incorporated in nanodiscs (NDs) for the present study.

## 2.2 Materials and Methods

### 2.2.1 Reagents

#### *Proteins*

Sub-fragments of TcdA (TcdA-A2, molecular weight (MW) 29,575 Da) and TcdB (TcdB-B3, MW 30,241 Da) were expressed in *E. coli* and purified as described previously<sup>12</sup>. The purified proteins were concentrated and exchanged into 200 mM ammonium acetate (pH 6.8) using Amicon Ultra centrifugal filters with a MW cutoff of 10 kDa and 30 kDa (Millipore, MA) for TcdA-A2 and TcdB-B3, respectively, and stored at -80°C until needed. Lysozyme (Lyz, MW 14,310 Da), which served as reference protein ( $P_{ref}$ ) for the ESI-MS binding measurements, was purchased from Sigma-Aldrich (Oakville, ON, Canada) and used after purification. The concentrations of the protein stock solutions were determined by measuring absorbance at 280 nm using a NanoDrop 1000 Spectrophotometer (Thermo Sci Inc. CA) and extinction coefficients (50,770 M<sup>-1</sup> cm<sup>-1</sup>, TcdA-A2; 49,740 M<sup>-1</sup> cm<sup>-1</sup>, TcdB-B3; 36000 M<sup>-1</sup> cm<sup>-1</sup>, Lyz) calculated from amino acid composition using the ExPASy webserver (<http://web.expasy.org/protparam/>).

#### *Carbohydrate ligands and gangliosides*

##### *Carbohydrate ligands*

The thirty compound carbohydrate library consisted of fifteen ganglioside oligosaccharides ( $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc, **L1** ( $\equiv$  GM3<sub>os</sub>), MW 633.21 Da;  $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc, **L2** ( $\equiv$  GM2<sub>os</sub>), MW 836.29 Da;  $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc, **L3** ( $\equiv$  GM1a<sub>os</sub>), 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, **L4** ( $\equiv$  GM1b<sub>os</sub>), MW 998.34 Da;  $\alpha$ -

$\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L5** ( $\equiv$  GD3<sub>os</sub>), MW 924.31 Da);  $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L6** ( $\equiv$  GD2<sub>os</sub>), MW 1127.39 Da;  $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -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, **L7** ( $\equiv$  GD1a<sub>os</sub>), MW 1289.44 Da;  $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L8** ( $\equiv$  GD1b<sub>os</sub>), MW 1289.44 Da;  $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L9** ( $\equiv$  GT3<sub>os</sub>), MW 1215.40 Da;  $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L10** ( $\equiv$  GT2<sub>os</sub>), MW 1418.48Da;  $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -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, **L11** ( $\equiv$  GT1a<sub>os</sub>), MW 1580.53 Da;  $\alpha$ -Neu5Ac-(2-3)- $\beta$ -D-Galp-(1-3)- $\beta$ -D-GalNAc-(1-4)-[ $\alpha$ -Neu5Ac-(2-8)- $\alpha$ -Neu5Ac-(2-3)]- $\beta$ -D-Galp-(1-4)-D-Glc, **L12** ( $\equiv$  GT1b<sub>os</sub>), MW 1581.39 Da;  $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)-[ $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 8)- $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L13** ( $\equiv$  GT1c<sub>os</sub>), MW 1580.53 Da;  $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L14** ( $\equiv$  asialo-GM2<sub>os</sub>), MW 545.20 Da;  $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L15** ( $\equiv$  asialo-GM1<sub>os</sub>), MW 707.25 Da), nine globoside oligosaccharides ( $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L16** ( $\equiv$  Lac), MW 342.30 Da;  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-GlcNAc, **L17** ( $\equiv$  LacNAc), MW 383.35 Da;  $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc, **L18** ( $\equiv$  Lacto-N-biose), MW 383.35 Da;  $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L19** (Gb3<sub>os</sub>), MW 504.17 Da;  $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc, **L20** ( $\equiv$  Gb4<sub>os</sub>), MW 707.25 Da;  $\beta$ -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, **L21** ( $\equiv$  Gb5<sub>os</sub>), MW 869.76 Da;  $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-

(1→4)-β-D-Gal-(1→4)-D-Glc, **L22** (≡ Globo-H hexaose), MW 1015.90 Da; α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc, **L23** (≡ Globo-A heptaose), MW 1219.09 Da; α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc, **L24** (≡ Globo-B heptaose), MW 1178.04 Da; and six HBGA oligosaccharides (α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-D-Gal, **L25** (≡ H type 1 tetrasaccharide), MW 691.62 Da; α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→4)-D-Gal, **L26** (≡ H type 2 tetrasaccharide), MW 691.62 Da; α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-D-Gal, **L27** (≡ A type 1 pentasaccharide), MW 894.82 Da; α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-D-Gal, **L28** (≡ A type 2 pentasaccharide), MW 894.82 Da; α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-D-Gal, **L29** (≡ B type 1 pentasaccharide), MW 853.76 Da; α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-D-Gal, **L30** (≡ B type 2 pentasaccharide), MW 853.76 Da). Five other synthetic oligosaccharide analogs were investigated (β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-CH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub>, **L31** (≡ LacNAc-Le<sup>x</sup>), MW 963.89 Da; α-D-Gal-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-CH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub>, **L32** (≡ Galα3Le<sup>x</sup>), MW 760.70 Da; β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[(α-L-Fuc-(1→3)]-β-D-GlcNAc-CH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub>, **L33** (≡ Le<sup>a</sup>-Le<sup>x</sup>), MW 1110.03 Da; β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-CH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub>, **L34** (≡ Le<sup>x</sup> trisaccharide), MW 817.75 Da; β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-CH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub>, **L35** (≡ LNT), MW 598.56 Da). The structures of **L1** - **L35** are given in Figure 2.2.

The oligosaccharides **L1 - L11**, **L13 - L15** and **L19 - L30** were purchased from Elicityl SA (Crolles, France); **L12** was purchased from Carbosynth Ltd (Berkshire, UK), **L16** and **L17** were purchased from Dextra Laboratories (Reading, UK) and **L18** was purchased from Sigma-Aldrich Canada (Oakville, Canada). **L31 - L35** were obtained from the Consortium for Functional Glycomics (CFG). All compounds were stored at -20 °C in Milli-Q water (Millipore, MA) at a concentration of 1 mM prior to use.

The structures of 7 gangliosides used to be incorporated into nanodisc are shown in Figure 2.3. GM1 (1545.8 Da (d18:1-18:0), 1573.9 Da (d20:1-18:0)), GM2 (1383.7 Da, 1411.7 Da) and GM3 (1180.5 Da) were purchased from Axxora LLC (San Diego, CA). GD1a (1836.1 Da), GD1b (1836.1 Da) and GT1b (2126.4 Da) were purchased from Sigma-Aldrich Canada (Oakville, Canada). GD2 (1674.0 Da) was purchased from MyBioSource, Inc (San Diego, CA). 1, 2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC, MW 677.9 Da) dissolved in chloroform was purchased from Avanti Polar Lipids (Alabaster, AL).

### **2.2.2 Expression and purification of MSP1E1**

Recombinant membrane scaffold protein MSP1E1 (MW 27 494 Da) was prepared using plasmid pMSP1E1 acquired from Addgene (Cambridge, MA). Protein expression and purification was adapted from the procedure described at <http://sligarlab.life.uiuc.edu/nanodisc.html>. Briefly, protein was expressed using the pET28a system (Novagen, Madison, WI) with the *E. coli* BL-21(DE3) strain as a host and expressed on lysogeny broth (LB) agar plate with 30 µg/mL kanamycin. After an overnight incubation of the plate at 37 °C, a colony was picked and placed into 20 mL LB medium subculture. The subculture was placed into a shaker at 200 rpm at 37 °C and the

OD<sub>600</sub> was monitored until it reached ~0.6-0.8. All the subculture was then transferred to 1000 ml terrific broth (TB) scale-up culture which was again shook at 250 rpm at 37 °C and induced when the OD<sub>600</sub> reached ~0.6-0.8 with 1 mM Isopropyl β-D-1 thiogalactopyranoside (IPTG). The temperature was then lowered to 28 °C and the 250 rpm speed was maintained overnight. The next day, the cultures were centrifuged at 8000 rpm at 4 °C for 20 minutes and the resulting cell pellets were collected. The cell pellets can be stored at -80 °C up to several months.

For purification, about 15 grams of cell pellets were resuspended in 100 mL 20 mM sodium phosphate (pH 7.4) and disrupted in the presence 1% Triton X-100, inhibitor and RNase by sonication. The lysate was cleared by centrifugation at 36,000 rpm for 60 min at 4 °C. The supernatant was passed through a 2.5 x 6 cm column containing HisPur Ni-NTA resin (Thermo Sci Inc. CA). MSP1E1 fractions were collected and evaluated for purity by both sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and mass spectrometry, then dialyzed against nanodisc buffer (20 mM Tris, 0.1 M NaCl, 0.5 mM EDTA, pH 7.4) for 8 hrs and buffer was changed 3 times. The dialyzed protein was then aliquoted and stored at -80°C. The MSP1E1 concentration was determined by absorbance at 280 nm using calculated extinction coefficient.

### **2.2.3 Nanodisc Preparation**

Nanodiscs were prepared as previously described.<sup>128, 129</sup> Briefly, DMPC lipid (dissolved in chloroform) and 7 gangliosides GM1, GM2, GM3, GD1a, GD1b, GT1b, GD2 (dissolved in chloroform: methanol 2:1) were mixed in the desired molar ratios, 2% for each ganglioside and 86% for DMPC, dried under nitrogen, and placed in a vacuum desiccator overnight to form a lipid film. Lipids were then re-suspended in 20 mM

TrisHCl, 0.5 mM EDTA, 100 mM NaCl, 25 mM sodium cholate pH 7.4 by sonication for 15 min and MSP1E1 added to give a final molar ratio of 100:1 lipid: MSP1E1. After incubation of the cholate:lipid:MSP1E1 mixture for 15 min at room temperature, ND self-assembly was initiated by the addition of Bio-Beads (Bio-Rad Laboratories Ltd., Canada) (0.5 g per mL of reconstitution mixture) and the solution incubated for a further 3 h to ensure all detergent had been removed. Finally, NDs were purified using a Superdex 200 10/300 size exclusion column (GE Healthcare Bio-Sciences, Uppsala, Sweden) equilibrated in 200 mM ammonium acetate pH 6.8. Both size exclusion chromatography and mass spectrometry were used to confirm ND's composition. The total number of lipid molecules (DMPC and GSL) per ND was taken to be 205.<sup>130</sup> Nanodiscs were then concentrated to approximately 60  $\mu$ M and stored at -80 °C until needed.

#### **2.2.4 Mass spectrometry**

Direct ESI-MS assay and Catch-and-Release (CaR) ESI-MS assay were carried out using a Synapt G2S and Synapt G2 quadrupole-ion mobility separation-time of flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK), respectively. The direct ESI-MS assay was implemented in positive ion mode, whereas the Catch-and-Release (CaR)-ESI-MS assay was operated in negative ion mode. NanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.) pulled to  $\sim$ 5 $\mu$ 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 0.80-1.20 kV,

source temperature 60 °C, cone voltage 30 V, trap voltage 5 V, and transfer voltage 2V. Data acquisition and processing were performed using MassLynx software (version 4.1)

*Direct ESI-MS assay*

The direct ESI-MS binding measurements were carried out on two virulence factor of *C. difficile* toxins TcdA and TcdB with a library of thirty reducing oligosaccharides (**L1 - L30**) in aqueous ammonium acetate solutions (200 mM, pH 7) at 25 °C. In all cases, the protein concentration was maintained at one concentration and the oligosaccharide concentration was varied from 5 μM to 200 μM. The concentration of P<sub>ref</sub> (Lyz) was fixed at 2 μM for TcdA-A2 measurements and 5 μM for TcdB-B3 measurements.

A complete description of the implementation of the assay and data analysis procedures for determining protein-ligand association constants ( $K_a$ ) can be found elsewhere.<sup>32, 122, 131</sup> Briefly, for a 1:1 protein-ligand complex, the abundance ( $Ab$ ) ratio ( $R$ ) of the ligand-bound protein (PL)-to-free protein (P) ions measured by ESI-MS (after correction of the mass spectrum for nonspecific ligand-protein binding) is taken to be equal to the equilibrium concentration ratio in the solution, eq 2.2:



$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]_{eq}}{[P]_{eq}} \quad (2.2)$$

and  $K_a$  can be calculated by eq 2.3:

$$K_a = \frac{R}{[L]_0 - \frac{R}{1+R}[P]_0} \quad (2.3)$$

where  $[P]_0$  and  $[L]_0$  are the initial concentrations of the P and L, respectively. Due to the low affinities of the interactions identified in the present study, the titration method,

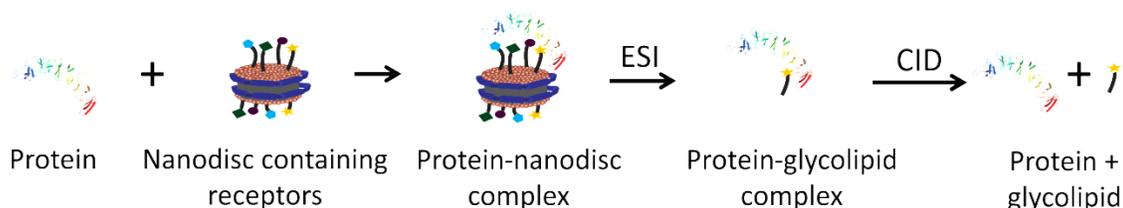
wherein the concentration of P was fixed and the concentration of the L varied, was employed to measure  $K_a$ . The value of  $K_a$  was determined by fitting eq 2.4 to a plot of  $(R/R+1)$  – which corresponds to the fraction of occupied binding sites – versus the concentration of L. At least five different concentrations of each oligosaccharide were used and measurements were carried out in triplicates. The errors were reported as the pooled standard deviation. The reference protein method, which involves the addition of a  $P_{ref}$  to the solution, was used to correct the mass spectra for the occurrence of non-specific carbohydrate-protein binding during the ESI process.<sup>86</sup>

$$\frac{R}{R+1} = \frac{1+K_a[L]_0+K_a[P]_0-\sqrt{(1-K_a[L]_0+K_a[P]_0)^2+4K_a[L]_0}}{2K_a[P]_0} \quad (2.4)$$

#### *Catch and Release (CaR) ESI-MS assay*

The CaR-ESI-MS assay (shown in Figure 2.5) was performed to identify binding partners of *C. difficile* toxins. Since the poor solubility of gangliosides, nanodisc is applied to solubilize different gangliosides receptors and maintain their native environment. Then *C. difficile* toxins were screened separately with nanodiscs containing different gangliosides. The toxin is able to strip its binding partners out of nanodiscs forming protein-ganglioside complexes. To release the protein-GSL complex from the NDs in the electrospray ion source, a source temperature of 60 °C and a cone voltage 50V were used; the backing pressure was 3.4 mbar. The pressure in the Trap and Transfer region was maintained at  $2.77 \times 10^{-2}$  mbar and  $2.84 \times 10^{-2}$  mbar, respectively, and the Trap and Transfer voltages were 5 and 2 V, respectively. To identify which ligand bind to the target toxin, ions corresponding to protein-ganglioside complexes were isolated using the

quadrupole mass filter (Low Mass resolution set to 4). The complexes were then subjected to CID in the Trap region of the Synapt G2S by increasing the trap voltage from 5 V to 150 V. Argon (trap gas flow rate here 8 mLmin<sup>-1</sup>) was used for CID in the Trap region. In most instances, the deprotonated ligands released from the complexes could be identified from their MWs.



**Figure 2.4** Cartoon representation of the CaR-ESI-MS assay for screening glycosphingolipids (GSLs) against target proteins. The protein and ND (lipid bilayer (red) with surrounding membrane scaffold protein (blue) containing different GSLs are incubated to form a complex. Using mild in-source dissociation conditions, the protein-GSL complexes are stripped out of the ND. The GSL ligands are then released (as ions) from the protein using collision-induced dissociation (CID) and mass analyzed.

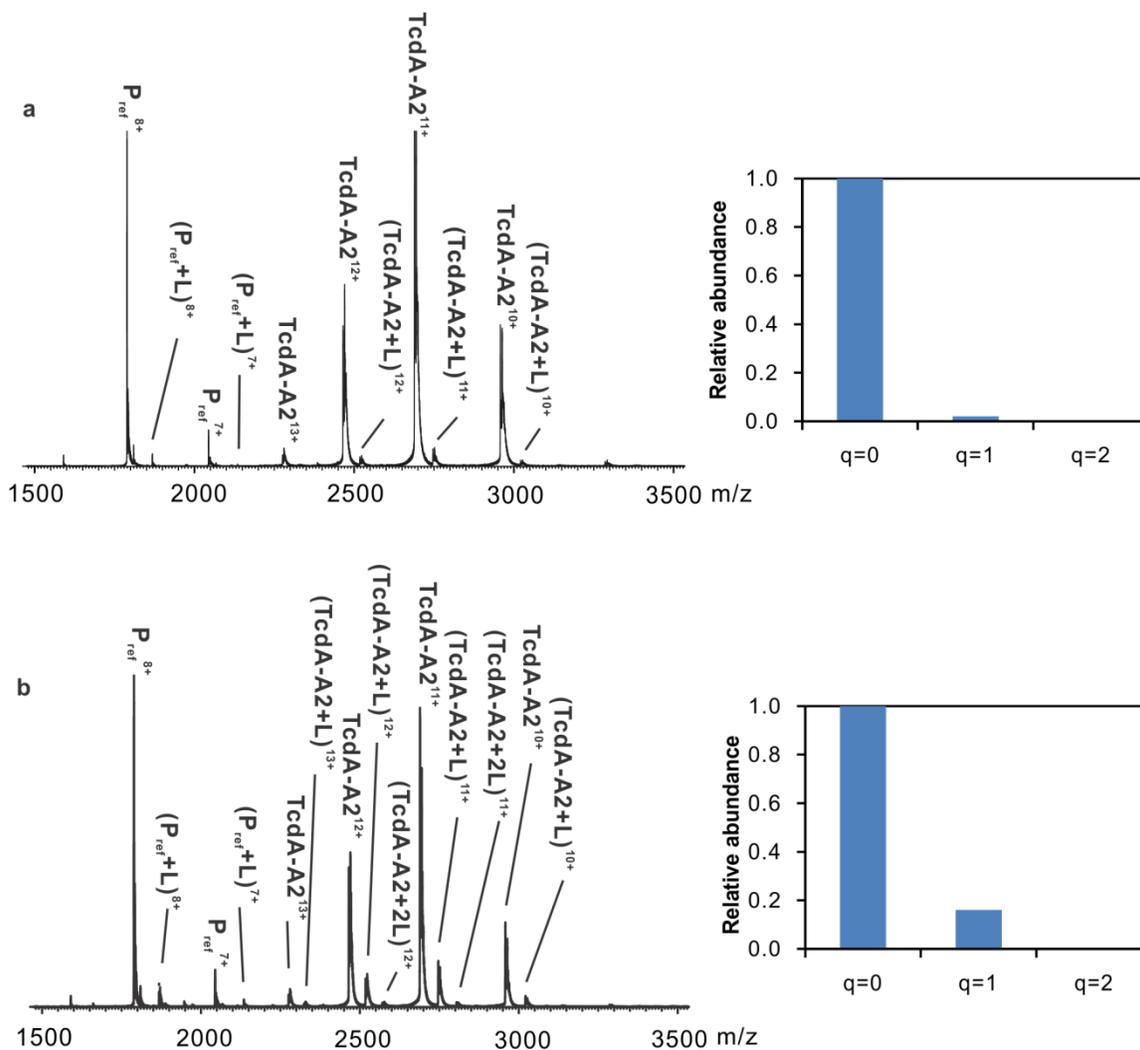
## 2.3 Results and Discussion

The main objective of this study is to find out if *C. difficile* toxins demonstrate binding to a selected library.

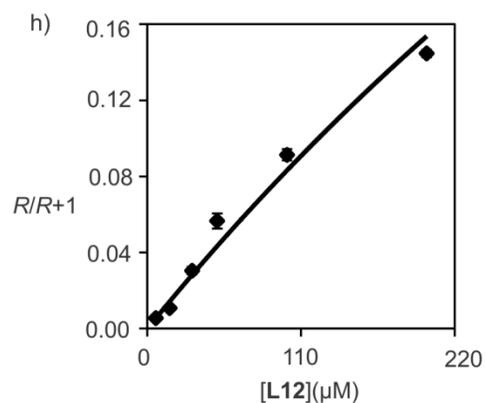
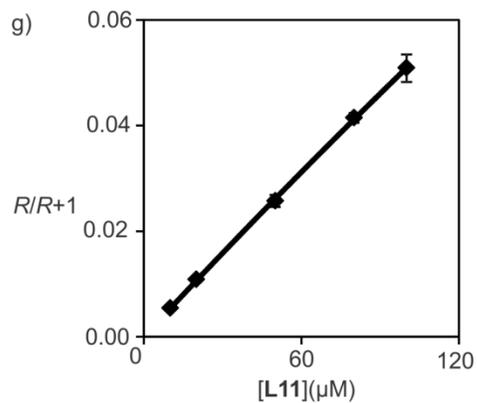
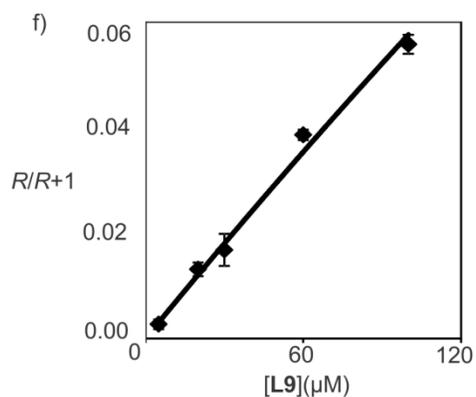
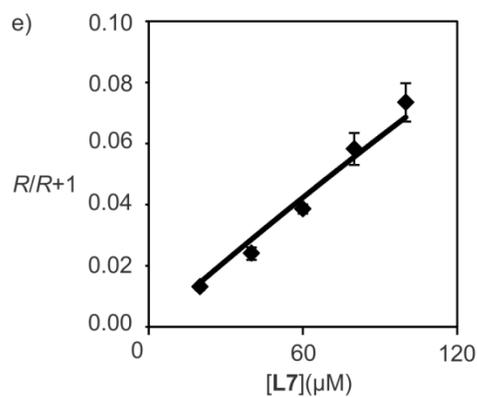
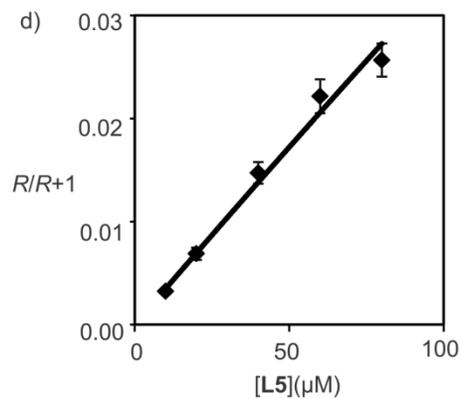
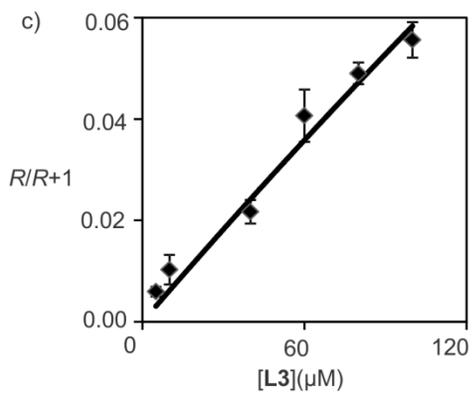
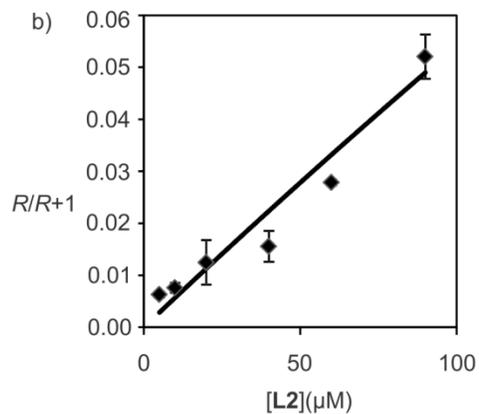
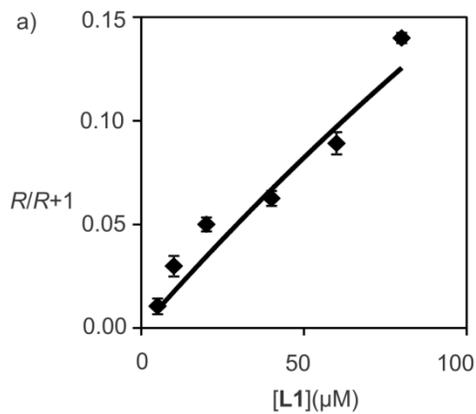
### 2.3.1 Carbohydrate recognition by TcdA-A2

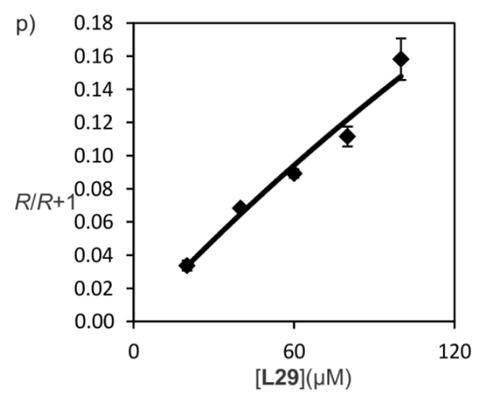
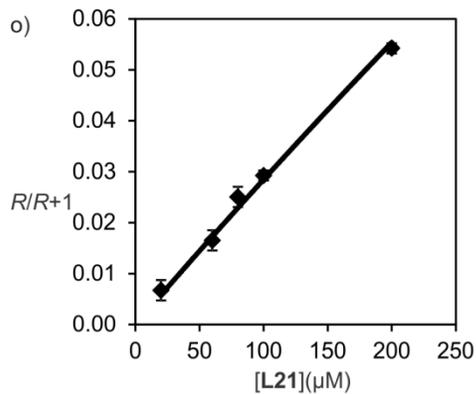
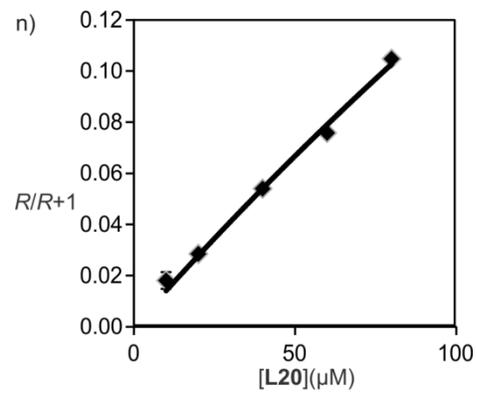
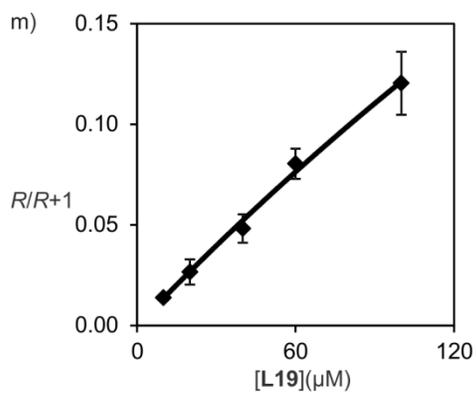
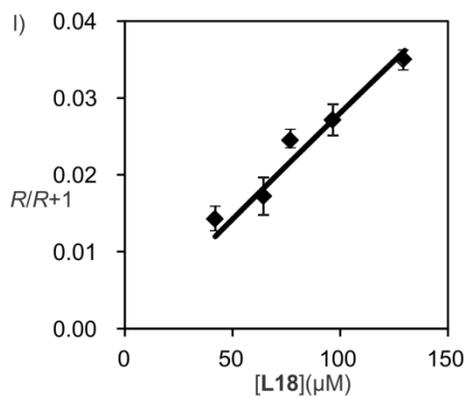
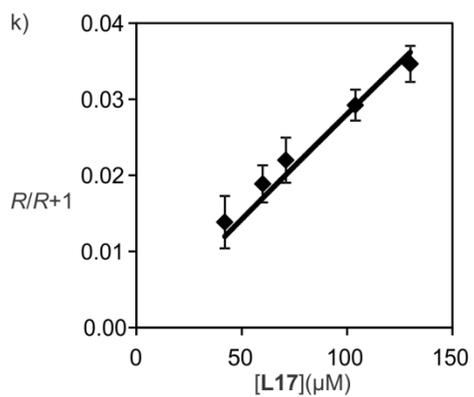
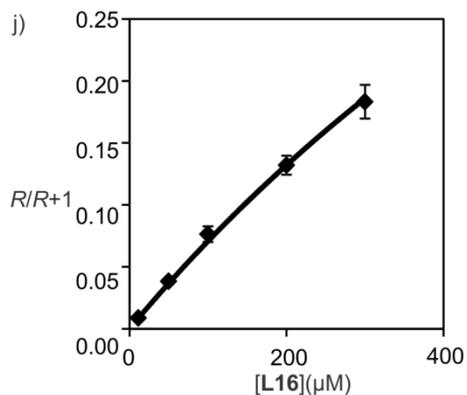
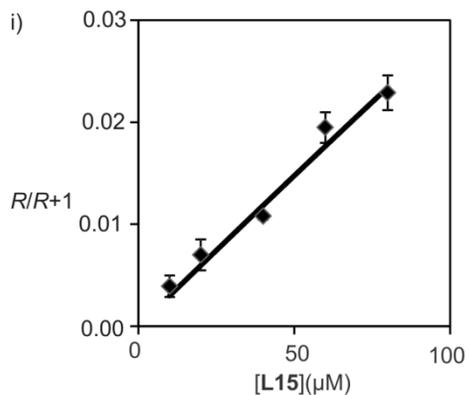
The binding measurements were carried between the sub-fragments of *C. difficile* toxin A, TcdA-A2 and a library of 30 human oligosaccharides (**L1 - L30**). Shown in Figure 2.6 are representative ESI-MS spectra acquired for solution of TcdA-A2 and **L1** ( $\alpha$ -D-

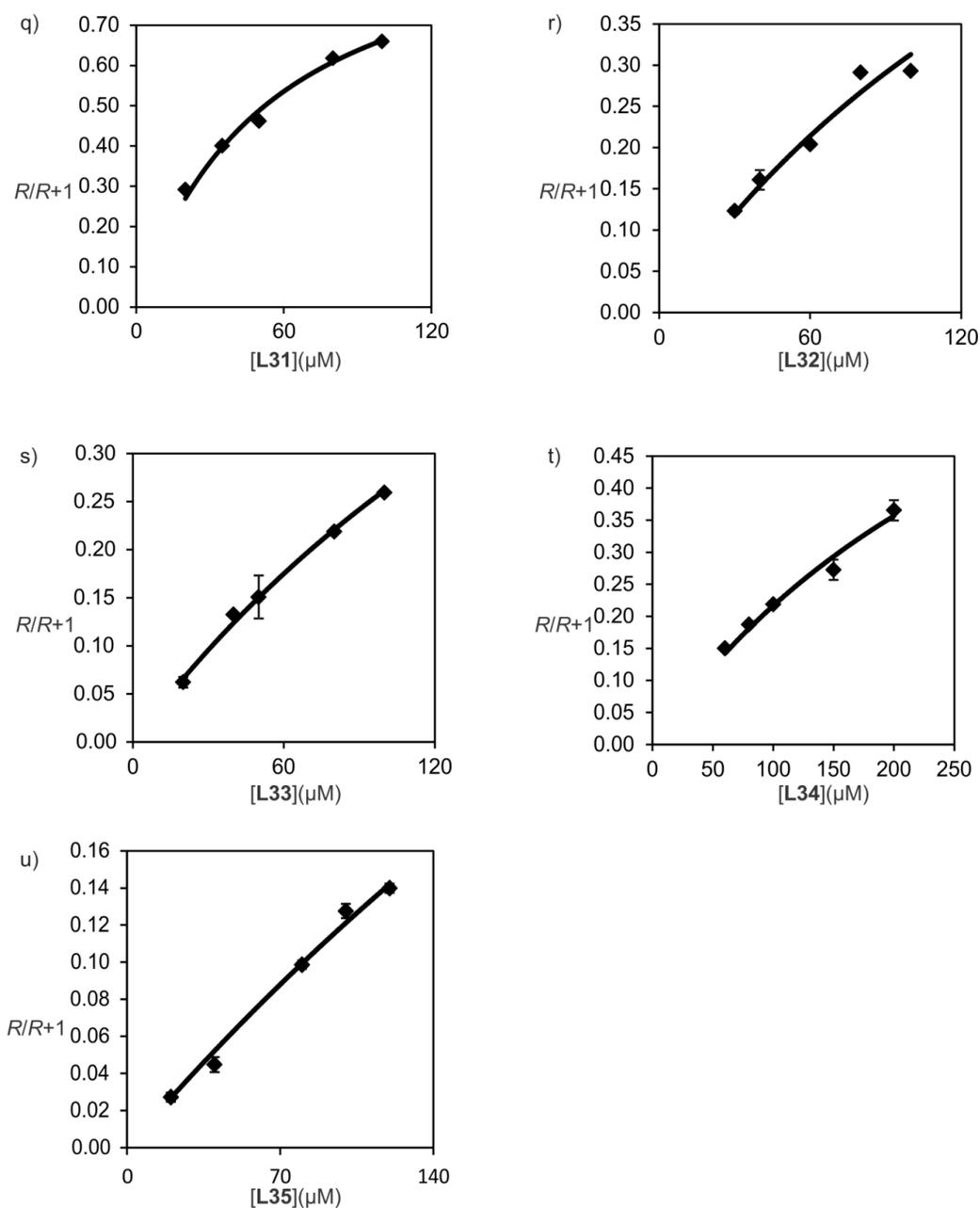
Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc) at 10 μM and 80 μM, respectively. Inspection of the ESI-MS spectra reveals signals corresponding to protonated ions of free (unbound) TcdA-A2, as well as TcdA-A2 bound to one or two molecules of **L1**, i.e.,  $(\text{TcdA-A2} + q\mathbf{L1})^{n+}$ , where  $q = 0 - 2$  and  $n = 10 - 13$ . At 10 μM only one molecule of **L1** is bound, while at 80 μM, two molecules are bound. In both cases, free and **L1**-bound  $P_{\text{ref}}$  ions were also detected, i.e.,  $(P_{\text{ref}} + q\mathbf{L1})^{n+}$ , where  $q = 0 - 2$  and  $n = 7$  and  $8$ . The presence of **L1**-bound  $P_{\text{ref}}$  ions indicates that nonspecific binding of **L1** to TcdA-A2 occurred during the ESI process and influenced the relative abundances of the  $(\text{TcdA-A2} + q\mathbf{L1})^{n+}$  ions. Using the distributions of measured distribution of  $(P_{\text{ref}} + q\mathbf{L1})^{n+}$  ions, the contribution of nonspecific was subtracted from the distributions of  $(\text{TcdA-A2} + q\mathbf{L1})^{n+}$  ions. After correction for nonspecific binding, TcdA-A2 is found to bind to, at most, one molecule of **L1** under these solution conditions (Figures 1a and 1b). Fitting eq 3 to the binding data (plotted as  $(R/R+1)$  vs **L1** concentration) yields a  $K_a$  of  $1800 \pm 170 \text{ M}^{-1}$ . Because TcdA-A2 is known to possess two carbohydrate binding sites, this value corresponds to the apparent affinity. Analogous measurements were performed for the other twenty nine oligosaccharides (**L1**, **L2** and **L4 - L30**). The results of the titration experiments are given in Figures 2.7 and the corresponding values of  $K_a$  are listed in Table 2.1



**Figure 2.5** ESI-MS spectra acquired in positive-ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing TcdA-A2 (6 μM),  $P_{ref}$  (2 μM) and **L1** at (a) 10 μM or (b) 80 μM. Insets show the normalized distribution of free and ligand-bound TcdA-A2 after correction for nonspecific ligand binding using the reference protein method.







**Figure 2.6** Plot of fraction of ligand-bound TcdA-A2 ( $R/(R+1)$ ) versus ligand concentration measured for (a) L1, (b) L2, (c) L3, (d) L5, (e) L7, (f) L9, (g) L11, (h) L12, (i) L15, (j) L16, (k) L17, (l) L18, (m) L19, (n) L20, (o) L21, (p) L29, (q) L31, (r) L32, (s) L33, (t) L34 and (u) L35. The ESI-MS binding measurements were carried out on 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing TcdA-A2

(6  $\mu\text{M}$ ), Lyz (2  $\mu\text{M}$ ) and each ligand at a minimum of five different concentrations, ranging from 5 to 200  $\mu\text{M}$ . The solid curves correspond to the best fit of eq. 2.4 to the experimental data and the error bars correspond to one standard derivation.

**Table 2.1** Apparent association constants ( $K_a$ , units of  $M^{-1}$ ) for the binding of TcdA-A2 and TcdB-B3 with oligosaccharides **L1 – L35** measured in an aqueous ammonium acetate (200 mM) solution at pH 7 and 25 °C using the direct ESI-MS assay.<sup>a</sup>

	Oligosaccharides	Structure	$K_a$ (A2)	$K_a$ (B3)
<b>L1</b>	GM3os	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc	1800±170	NB
<b>L2</b>	GM2os	$\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	570±30	NB
<b>L3</b>	GM1aos	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	620±70	NB
<b>L4</b>	GM1bos	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	NB <sup>b</sup>	NB
<b>L5</b>	GD3os	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc	310±50	NB
<b>L6</b>	GD2os	$\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	NB	NB
<b>L7</b>	GD1aos	$\alpha$ -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	740±40	NB
<b>L8</b>	GD1bos	$\beta$ -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	NB	NB
<b>L9</b>	GT3os	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc	610±20	
<b>L10</b>	GT2os	$\beta$ -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	NB	NB
<b>L11</b>	GT1aos	$\alpha$ -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	530±60	
<b>L12</b>	GT1bos	$\alpha$ -Neu5Ac-(2→3)- $\beta$ -D-Galp-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -Neu5Ac-(2→8)- $\alpha$ -Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc	910±50	NB
<b>L13</b>	GT1cos	$\beta$ -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	NB	NB
<b>L14</b>	Asialo-GM2os	$\beta$ -D-GalNAc-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	NB	NB
<b>L15</b>	Asialo-GM1os	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)- $\beta$ -D-Gal-(1→4)-D-Glc	300±20	NB
<b>L16</b>	Lactose	$\beta$ -D-Gal-(1→4)-D-Glc	760±190	NB

<b>L17</b>	LacNAc	$\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-GlcNAc	290 $\pm$ 60	NB
<b>L18</b>	Lacto-N-biose	$\beta$ -D-Gal-(1 $\rightarrow$ 3)-D-GlcNAc	240 $\pm$ 110	NB
<b>L19</b>	Gb3	$\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	1380 $\pm$ 20	600 $\pm$ 280
<b>L20</b>	Gb4	$\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc	1400 $\pm$ 30	230 $\pm$ 150
<b>L21</b>	Gb5	$\beta$ -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	270 $\pm$ 50	NB
<b>L22</b>	Globo-H	$\alpha$ -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	110 $\pm$ 20	570 $\pm$ 40
<b>L23</b>	Globo-A	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -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	NB	NB
<b>L24</b>	Globo-B	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -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	NB	NB
<b>L25</b>	H type 1	$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)-D-Gal	NB	560 $\pm$ 70
<b>L26</b>	H type 2	$\alpha$ -L-Fuc-(1 $\rightarrow$ 2)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-D-Gal	NB	NB
<b>L27</b>	A type 1	$\alpha$ -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	NB	NB
<b>L28</b>	A type 2	$\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-D-Gal	1700 $\pm$ 30	NB
<b>L29</b>	B type 1	$\alpha$ -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	NB	NB
<b>L30</b>	B type 2	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)-D-Gal	NB	NB
<b>L31</b>	LacNAc-Le <sup>x</sup>	$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 3)]- $\beta$ -D-GlcNAc-Sp <sup>c</sup>	20000 $\pm$ 76	NB
<b>L32</b>	Gal $\alpha$ 3Le <sup>x</sup>	GlcNAc-Sp <sup>c</sup>	0	NB
<b>L33</b>	Le <sup>a</sup> -Le <sup>x</sup>	$\alpha$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 3)]- $\beta$ -D-GlcNAc $\beta$ -Sp <sup>c</sup>	4700 $\pm$ 400	NB
<b>L34</b>	Le <sup>x</sup> trisaccharide	$\beta$ -D-Gal-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 4)]- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-[( $\alpha$ -L-Fuc-(1 $\rightarrow$ 3)]- $\beta$ -D-GlcNAc-Sp <sup>c</sup>	3600 $\pm$ 200	NB
<b>L35</b>	LNT	$\beta$ -D-Gal-(1 $\rightarrow$ 4)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 3)]- $\beta$ -D-GlcNAc-Sp <sup>c</sup>	2800 $\pm$ 200	NB
		$\beta$ -D-Gal-(1 $\rightarrow$ 3)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-Sp <sup>c</sup>	1400 $\pm$ 100	

a. The reported errors are one standard deviation.

b. NB $\equiv$  no binding detected.

c. Sp $\equiv$ CH<sub>2</sub>CH<sub>2</sub>N<sub>3</sub>

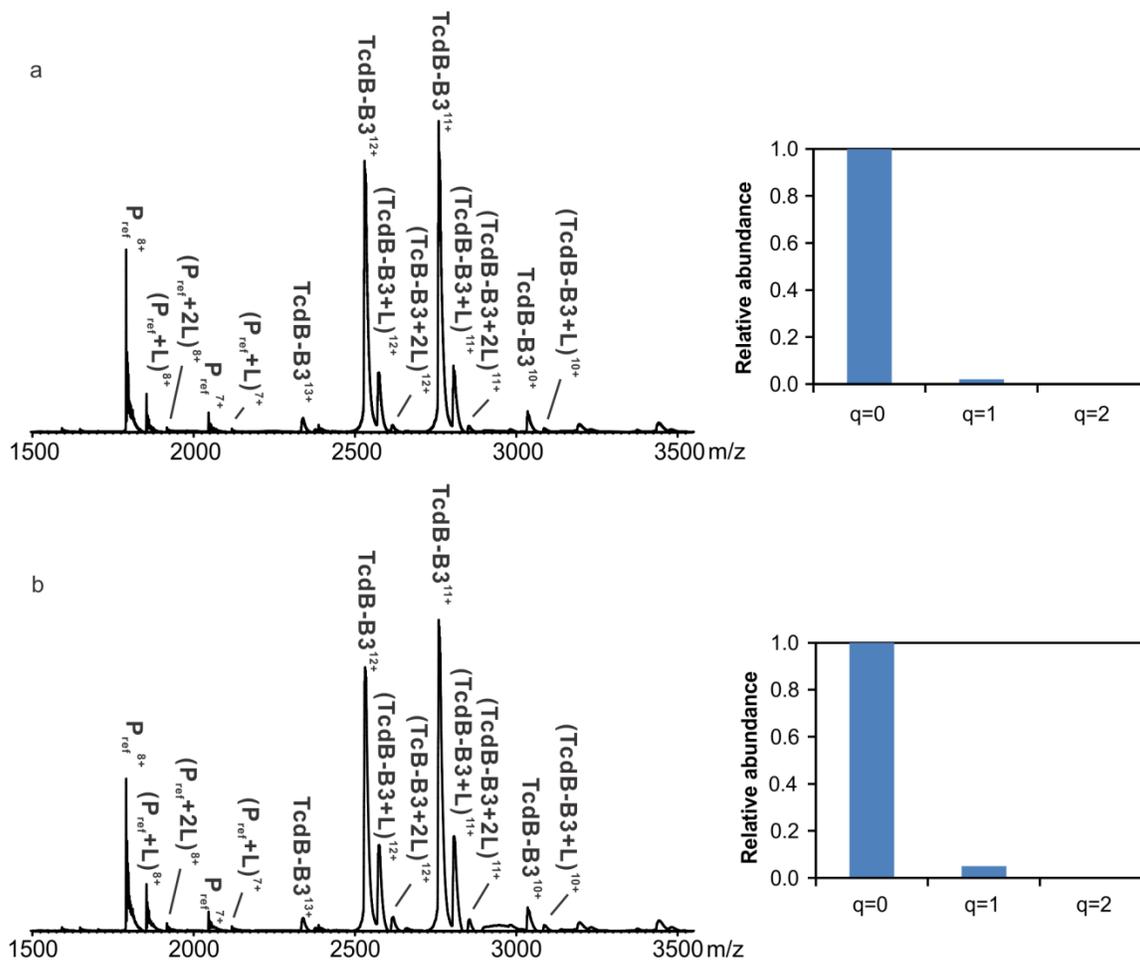
Inspection of Table 2.1 reveals that, among the 30 oligosaccharides (**L1- L30**) tested, TcdA-A2 bind to ~57% of them. However, and as is the case for most protein-carbohydrate interactions, the measured affinities are relatively low, with values ranging from 110 to 1800 M<sup>-1</sup>. Of the ganglioside oligosaccharides investigated, TcdA-A2 binds to nine of them (**L1 -L3, L5, L7, L9, L11, L12** and **L15**). To our knowledge, these are the first data implicating gangliosides as possible cellular receptors for TcdA. Among them, **L1**, which is the GM3 trisaccharide, exhibits the highest affinity (1800±170 M<sup>-1</sup>); **L12**, the GT1b heptasaccharide exhibits the second highest affinity (910±50 M<sup>-1</sup>). Of the neutral oligosaccharides investigated, eight were found to bind. Lactose (**L16**) and its analogs (**L17, L18**) were found to bind, as were four of the globosides (**L19 - L22**). Notably, only one of the HBGA oligosaccharides (**L28**) was recognized by TcdA-A2; this pentasaccharide exhibited an affinity of 1700 ±30 M<sup>-1</sup>. It should be noted that, in a previous ESI-MS study of TcdA-A2, no binding was detected with **L1** or **L16**.<sup>122</sup> This discrepancy may be related to differences in the sensitivity of the mass spectrometers used in the two studies and the fact that the titration method, which used in the present study and generally leads to more reliable affinities for weak interactions,<sup>132</sup> was not used in the earlier study. It must also be emphasized that the existence of the (TcdA-A2 + **L1**) complex is supported by results obtained by x-ray crystallography, *vide infra*.

Analysis of the binding data provides new insights into the structural motifs recognized by TcdA. All of the ganglioside oligosaccharides possess β-D-Gal-(1→4)-D-Glc (lactose) disaccharide at the reducing end. Consequently, since TcdA-A2 binds to **L16** with a measurable affinity (760±190 M<sup>-1</sup>), this disaccharide moiety serves as a recognition element for the gangliosides. The higher affinity exhibited by **L1** indicates a

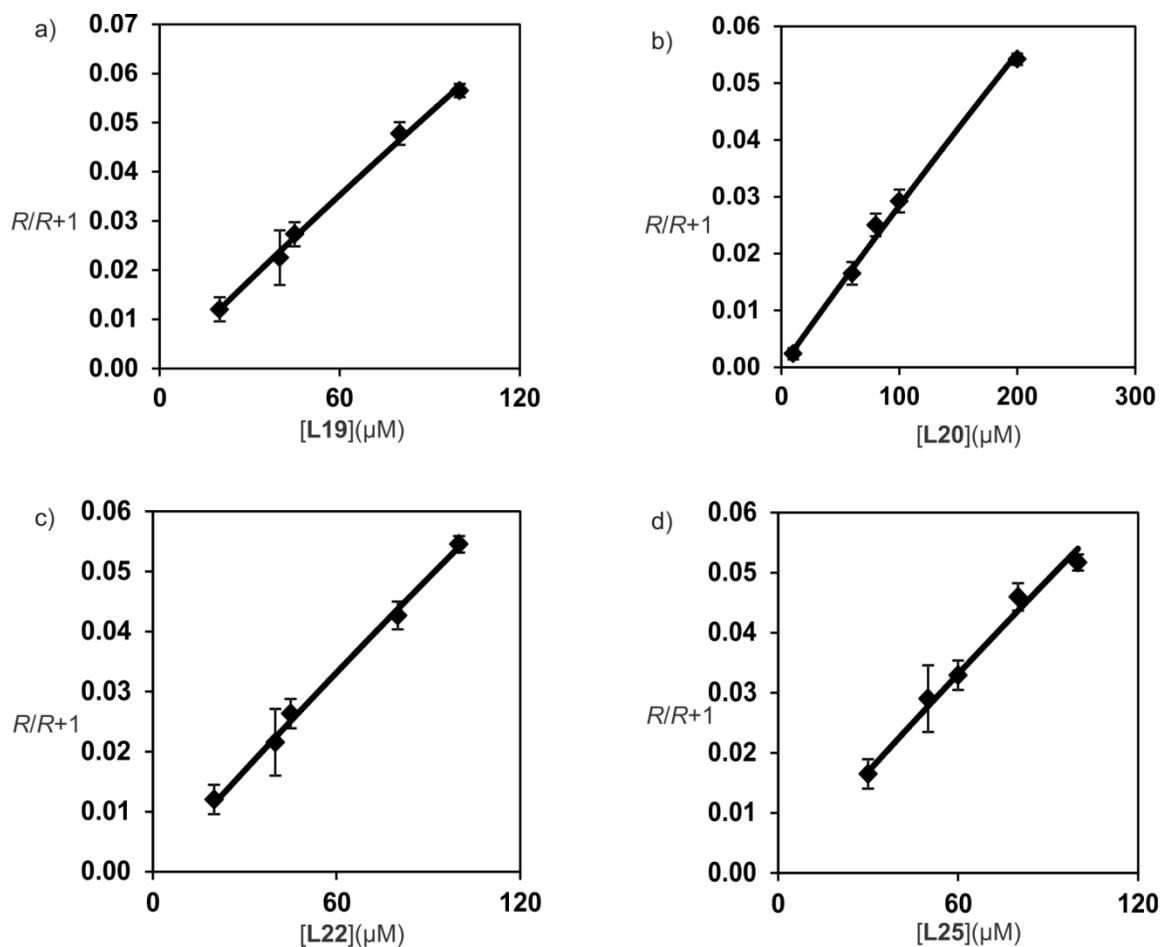
favourable role of the ( $\alpha$ 2-3)-linked sialic acid; notably the  $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc motif is present in all of the ganglioside oligosaccharide ligands (**L1** - **L3**, **L7** and **L11**). The presence of additional residues linked to Gal influences binding and results in either a loss of binding or a reduction in affinities to values similar to **L16**. Moreover, the presence of additional sialic acids, ( $\alpha$ 2-3)- or ( $\alpha$ 2-8)-linked, is tolerated by TcdA-A2, provided there is no chain extension beyond GalNAc. However, in all cases, the affinities were lower than that measured for **L1**. Although all of the globoside oligosaccharides tested (**L19** - **L24**) possess the a  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc disaccharide at the reducing end, only **L19** – **L22** are found to bind. In contrast to the ganglioside oligosaccharides, extension of the terminal  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc disaccharide with  $\alpha$ -D-Gal or  $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal leads to enhanced affinities ( $\sim 1400 \text{ M}^{-1}$ ), compare to **L18**. However, further extension results in a reduction (**L21**, **L22**) or loss of affinity (**L23**, **L24**), suggesting the possibility of unfavourable contacts arising from the additional residues. Unlike the other oligosaccharides tested, the type 1 HBGA oligosaccharides all possess  $\beta$ -D-Gal-(1 $\rightarrow$ 3)-D-GlcNAc disaccharide and the type 2 HBGA all possess  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-GlcNAc disaccharide in the middle of the sequence instead of at the reducing end. Of the structures tested only **L29** type-2 A pentasaccharide ( $\alpha$ -D-GalNAc-(1 $\rightarrow$ 3)-[ $\alpha$ -L-Fuc-(1 $\rightarrow$ 2)]- $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-D-Gal) binds; the affinity is relatively high ( $1700 \pm 30 \text{ M}^{-1}$ ) and similar to that of **L1**. The absence of binding in all but **L29** indicates that the LacNAc moiety is not sufficient for binding. Based on these data alone it is not possible to rationalize the observation that only **L29** binds.

### 2.3.2 Carbohydrate recognition by TcdB-B3

Unlike TcdA-A2, which binds to a majority of the oligosaccharides tested, the ESI-MS binding measurements revealed that TcdB-B3 is much less promiscuous in terms of the carbohydrate structures it recognizes. Of the thirty oligosaccharides investigated, binding was detected for only four of them (**L19**, **L20**, **L22** and **L25**). Shown in Figures 2.8a and 2.8b are representative ESI-MS spectra acquired for solutions of TcdB-B3 and **L19** at 40  $\mu\text{M}$  and 80  $\mu\text{M}$ , respectively. Inspection of the ESI-MS spectra reveals signals corresponding to protonated ions of free (unbound) TcdB-B3, as well as TcdB-B3 bound to one or two molecule of **L19**, i.e.,  $(\text{TcdB-B3} + q\text{L19})^{n+}$ , where  $q = 0 - 2$  and  $n = 10 - 13$ . Free and **L19**-bound  $\text{P}_{\text{ref}}$  ions were also detected, i.e.,  $(\text{P}_{\text{ref}} + q\text{L19})^{n+}$ , where  $q = 0 - 2$  and  $n = 7$  and 8, indicating the occurrence of nonspecific binding. After correction for nonspecific binding (insets, Figures 3a and 3b), TcdB-B3 is found to bind to one molecule of **L19** under these solution conditions. Fitting eq 3 to the binding data ( $(R/R+1)$  vs **L19** concentration) yields a  $K_a$  of  $600 \pm 280 \text{ M}^{-1}$  (Figure 2.9). The results of the titration experiments obtained measured for **L20**, **L22** and **L25** are given in Figures 2.9 and the corresponding values of  $K_a$  are listed in Table 2.1. It can be seen that **L19**, **L22** and **L25** exhibit similar affinities  $\sim 600 \text{ M}^{-1}$  while **L20** forms a weaker interaction,  $\sim 200 \text{ M}^{-1}$ . It is notable that, although both **L19** and **L20** have the  $\beta\text{-D-Gal-(1}\rightarrow\text{4)-D-GlcNAc}$  disaccharide at the reducing end, **L19** exhibits a higher affinity than **L20**, suggesting the possibility of unfavourable contacts arising from the additional residues. Also, TcdB-B3 binds to **L25** (H tetrasaccharide type 1) in a similar affinity as **L22** (Globo-H hexaose) ( $570 \pm 40 \text{ M}^{-1}$  and  $560 \pm 70 \text{ M}^{-1}$ , respectively), suggesting that H-epitope (1,2 linked  $\alpha\text{-L-Fuc}$ ) plays an important role in TcdB's binding.



**Figure 2.7** ESI-MS spectra acquired in positive-ion mode for aqueous ammonium acetate solutions (10 mM) at pH 7 and 25 °C containing TcdB-B3 (7 μM), P<sub>ref</sub> (5 μM) and L19 at (a) 40 μM or (b) 80 μM. Insets show the normalized distribution of free and ligand-bound TcdB-B3 after correction for nonspecific ligand binding using the reference protein method.



**Figure 2.8** Plot of fraction of ligand-bound TcdB-B3 ( $R/(R+1)$ ) versus ligand concentration measured for (a) L19, (b) L20, (c) L22 and (d) L25. The ESI-MS binding measurements were carried out on 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing TcdB-B3 (7  $\mu\text{M}$ ), Lyz (5  $\mu\text{M}$ ) and each ligand at a minimum of five different concentrations, ranging from 10 to 200  $\mu\text{M}$ . The solid curves correspond to the best fit of eq 2.4 to the experimental data and the error bars correspond to one standard derivation.

### 2.3.3 Comparison of binding data from ESI-MS and glycan microarray screening

Glycan microarrays represent the dominant technology for screening carbohydrate libraries against lectins.<sup>25-26</sup> The TcdA and TcdB fragments (TcdA-A2, TcdB-B1 and TcdB-B2) were previously submitted to the Consortium for Functional Glycomics and screened against a carbohydrates library using glycan microarrays (Glycan microarray version 3). The arrays at the Consortium for Functional Glycomics utilize amine-reactive *N*-hydroxysuccinimide (NHS)-activated glass slides to couple amine-functionalized oligosaccharides to the surface. The Glycan microarray version 3 contains three hundred and twenty different oligosaccharide structures, with either a -CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub> (Sp0) or -CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub> (Sp8) group at the reducing end. Following incubation of the target protein with the array and a subsequent washing step, interactions were identified, through detection of bound protein using a fluorescently labelled antibody. Of the three hundred and twenty oligosaccharides tested, nine specific carbohydrate structures are also found in the library of thirty human oligosaccharides used in the current study (**L1 - L3, L5 - L7, L9, L16 and L17**).

Notably, in the case of TcdA-A2, glycan microarray screening failed to identify binding to any of these (**L1 - L3, L5 - L7, L9, L16 and L17**) oligosaccharides. In contrast, low but detectable binding was identified by ESI-MS. These differences could be due to the fast dissociation (off) kinetics generally associated with low affinity interactions could result in the loss of binding during the washing step and result in false negatives. Additionally, the presence of the aglycone linkers used in the glycan microarray assay, could interfere with binding through steric clashes for binding of the oligosaccharides containing (lactose) at the reducing end. Additional, albeit qualitative,

support for the latter explanation can be found in a comparison of binding data acquired for a series of Lewis X analogs (**L31** – **L35**), which were identified as ligands from glycan microarray screening. Using the same approach as described above, affinities for these oligosaccharides were measured by ESI-MS (Table 2.1). **L31** is the highest affinity ligand, with a  $K_a$  of  $20\,000\text{ M}^{-1}$ . Significantly lower affinities were measured for **L32** ( $4700 \pm 400\text{ M}^{-1}$ ), **L33** ( $3600 \pm 200\text{ M}^{-1}$ ), **L34** ( $2800 \pm 200\text{ M}^{-1}$ ) and **L35** ( $1400 \pm 100\text{ M}^{-1}$ ). The relative affinities measured for these structures, as well as for the other ligands identified by ESI-MS are shown in Table 2.2 and Figure 2.10, along with the relative fluorescence intensities (which are taken to reflect relative affinities) measured in the glycan array screening. It can be seen that, overall, there is reasonable agreement between the ESI-MS and Consortium for Functional Glycomics results in the case of **L31** – **L35**. The exception being **L32**, which exhibits a higher relative affinity in the Consortium for Functional Glycomics data than expected based on the  $K_a$  value measured by ESI-MS. The origin of this discrepancy is not known but could potentially be the result of multivalent binding of the protein to the immobilized **L32**.

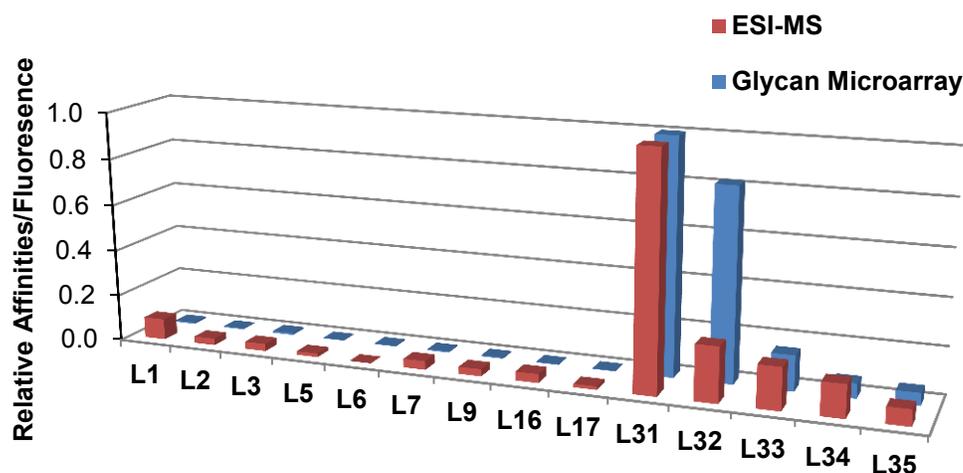
**Table 2.2** Comparison between glycan microarray assay and direct ESI-MS assay binding results between TcdA-A2 and 14 oligosaccharides<sup>a</sup>. Direct ESI assay's result is shown as apparent association constants  $K_{a,app}$  ( $M^{-1}$ ) measured in aqueous ammonium acetate (200 mM) at pH 7 and 25 °C. Glycan microarray's result is shown in relative fluorescence (RF).

	Oligosaccharides	Structure used in Glycan Microarray	$K_{a,app}(A2)$	RF <sup>a</sup>
<b>L1</b>	GM3os	$\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc-Sp0 <sup>c</sup>	1800±170	326±31
<b>L2</b>	GM2os	$\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc-Sp0 <sup>c</sup>	570±30	217±53
<b>L3</b>	GM1aos	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc-Sp0 <sup>c</sup>	620±70	439±98
<b>L5</b>	GD3os	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc-Sp0 <sup>c</sup>	310±50	155±74
<b>L6</b>	GD2os	$\beta$ -D-GalNAc-(1→4)-[ $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)]- $\beta$ -D-Gal-(1→4)-D-Glc-Sp0 <sup>c</sup>	NB <sup>b</sup>	224±36
<b>L7</b>	GD1aos	$\alpha$ -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-Sp0 <sup>c</sup>	740±40	256±64
<b>L9</b>	GT3os	$\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→8)- $\alpha$ -D-Neu5Ac-(2→3)- $\beta$ -D-Gal-(1→4)-D-Glc-Sp0 <sup>c</sup>	610±20	244±85
<b>L16</b>	Lactose	$\alpha$ -D-Gal-(1→4)- $\beta$ -D-Glc-Sp0 <sup>c</sup>	760±190	184±24
<b>L17</b>	LacNAc	$\alpha$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-Sp0 <sup>c</sup>	290±60	180±76
<b>L31</b>	LacNAc-Le <sup>x</sup>	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-Sp0 <sup>c</sup>	20000±760	53030±9104
<b>L32</b>	Gal $\alpha$ 3Le <sup>x</sup>	$\alpha$ -D-Gal-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc $\beta$ -Sp8 <sup>c</sup>	4700±400	43372±9697
<b>L33</b>	Le <sup>a</sup> -Le <sup>x</sup>	$\beta$ -D-Gal-(1→3)-[ $\alpha$ -L-Fuc-(1→4)]- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-Sp0 <sup>c</sup>	3600±200	8034±1477
<b>L34</b>	Le <sup>x</sup> trisaccharide	$\beta$ -D-Gal-(1→4)-[ $\alpha$ -L-Fuc-(1→3)]- $\beta$ -D-GlcNAc-Sp8 <sup>c</sup>	2800±200	3349±389
<b>L35</b>	LNT	$\beta$ -D-Gal-(1→3)- $\beta$ -D-GlcNAc-(1→3)- $\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc-Sp0 <sup>c</sup>	1400±100	2844±610

a. RF≡ Relative Fluorescence, value is obtained from CFG website.

b. NB≡ no binding detected.

c. Sp0≡CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>; Sp8≡CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>.



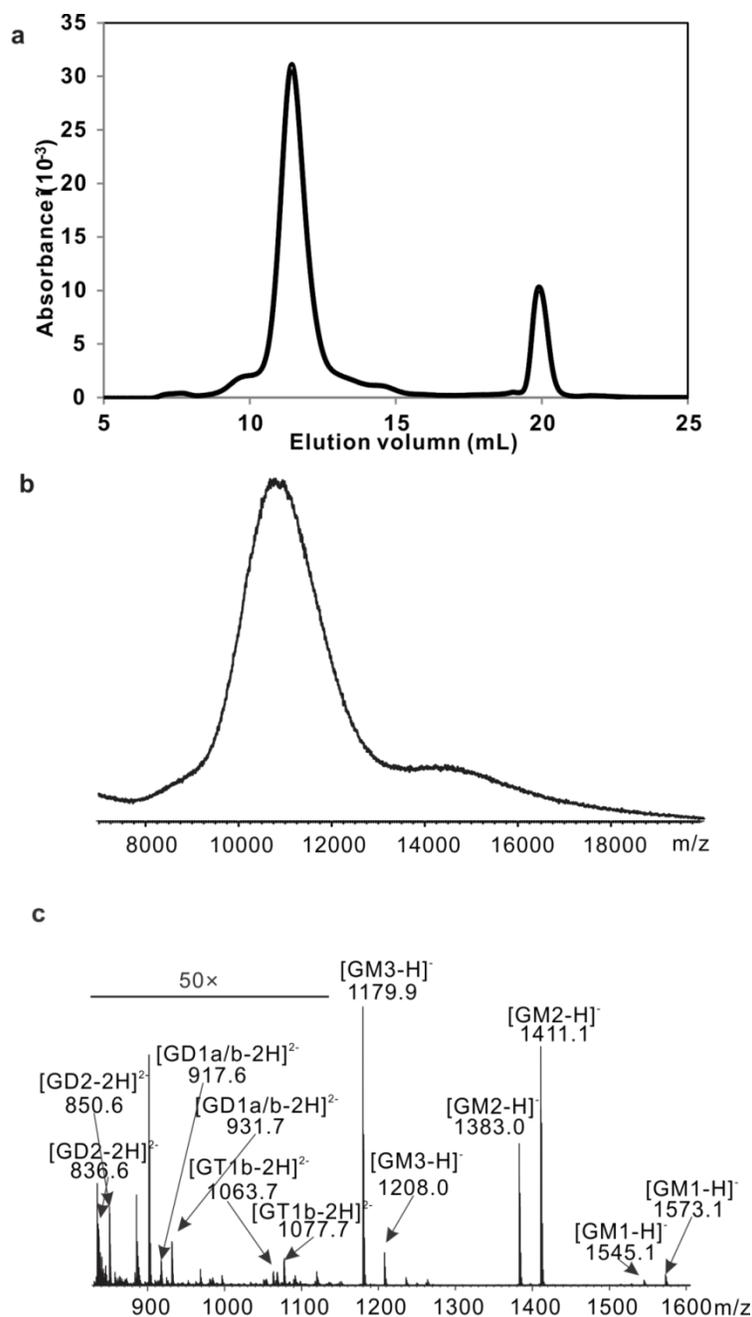
**Figure 2.9** Relative affinities of fourteen oligosaccharides (L1, L2, L3, L5, L6, L7, L9, L16, L17, L31, L32, L33, L34 and L35) for TcdA-A2 measured using the direct ESI-MS assay. Also shown are the relative fluorescence intensities reported for the related glycans by the Consortium for Functional Glycomics ([www.functionalglycomics.org](http://www.functionalglycomics.org)).

Two shorter (compared to TcdB-B3) fragments of TcdB, referred to as TcdB-B1 and TcdB-B2, were also submitted to Consortium for Functional Glycomics for screening. However, no evidence of strong binding to any of the structures was obtained. For TcdB-B1, only three structures exhibited relative fluorescence between 1000 and 2000 RFU, the remainder had signals <500 RFU. For TcdB-B2, fluorescence was <400 RFU for all glycans. The absence of any strong binders is qualitatively consistent with present ESI-MS data.

### 2.3.4 Glycosphingolipids recognition by *C. difficile* toxins

Gangliosides belong to a type named glyosphingolipid (GSLs). They are found in the cell membrane of organisms from bacteria to human, and are the major glycans of the vertebrate

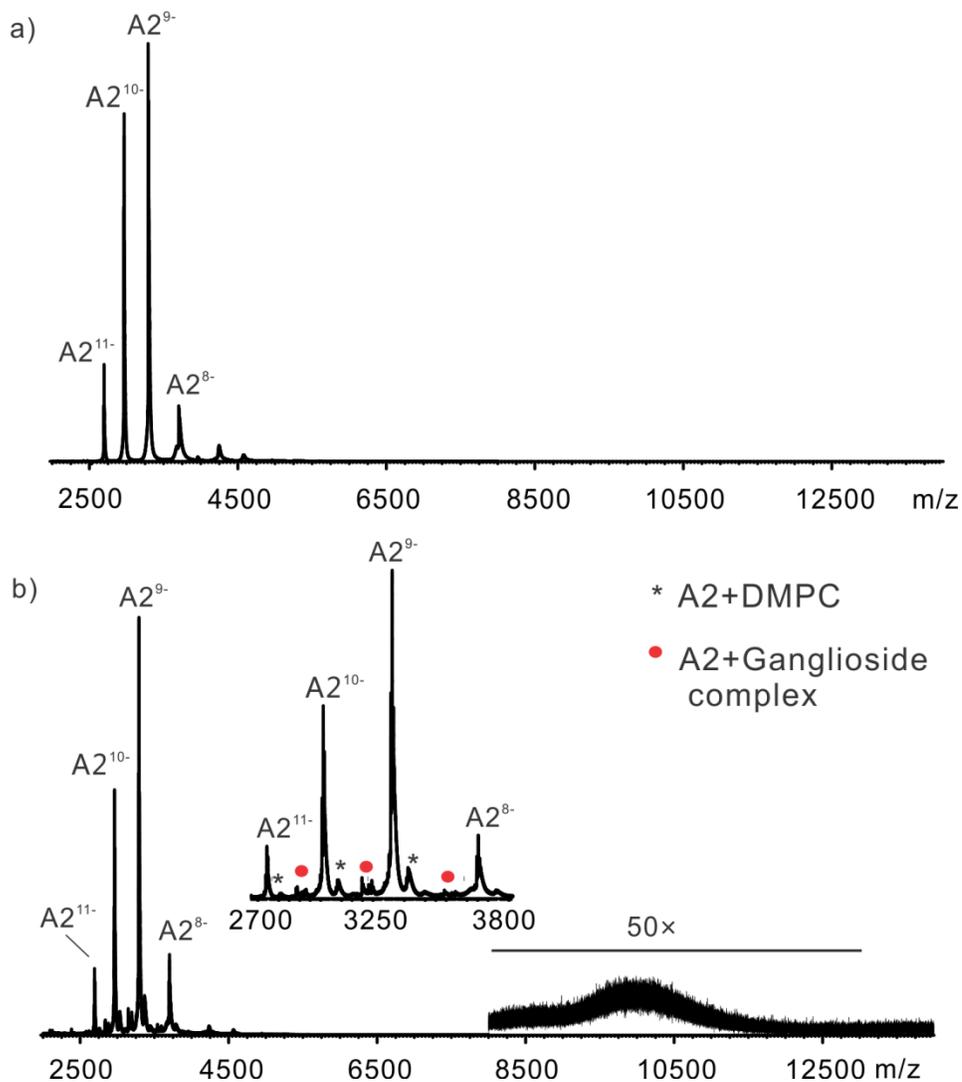
brain. All GSLs share a ceramide lipid moiety that consist a long-chain amino alcohol (sphingosine) in amide linkage to a fatty acid and a diversity of glycans attached to ceramide. Ganglioside is also named sialyted-glycosphingolipids meaning it contains negatively charged sialic acid. They are primarily expressed in the outer leaflet of the plasma membrane of cells, with their glycans facing the external milieu. The oligosaccharide parts of ganglioside (**L1- L15**) have been studied in the previous chapter. Among them, **L1**(≡ GM3<sub>os</sub>), **L2**(≡ GM2<sub>os</sub>), **L3**(≡ GM1<sub>os</sub>), **L5**(≡ GD3<sub>os</sub>), **L7**(≡ GD1a<sub>os</sub>), **L11**(≡ GT1a<sub>os</sub>), **L12**(≡ GT1b<sub>os</sub>) and **L15**(≡ Asialo-GM1<sub>os</sub>) are found to bind to TcdA-A2 while no binders are found for TcdB-B3. To examine the effect of the ceramide lipid moiety's effect on binding, 7 different ganglioside (GM1, GM2, GM3, GD1a, GD1b, GD2, GT1b) which contains 5 ganglioside whose oligosaccharide parts have been identified to bind to TcdA-A2 were chosen to be incorporated into nanodiscs. These NDs are referred to as the 7 G ND and each GSL was present at 2% in these NDs. The ND's composition has been confirmed with size exclusion chromatography and mass spectrometry. A single peak with an elution volume at approximately 12 mL confirmed ND formation, as shown in Figure 2.4a. The ND is further injected into the mass spectrometer. A broad peak centered at mass-to-charge ratio (m/z) ~10000 was observed (Figure 2.4b). This spectral feature is attributed to the gaseous ions of intact ND. This intact ND is further selected by changing quadrupole profile and subjected to undergo CID fragmentation. As shown in Figure 2.4c, all the 7 ganglioside can be detected, meaning 7 ganglioside are successfully incorporated into the ND.



**Figure 2.10** Characterization for 7 ganglioside ND which is prepared using GM3, GM2, GM1, GD2, GD1a, GD1b, GT1b. (a) is the size exclusion chromatography for 7 ganglioside ND. (b) is the mass spectrum of 7 ganglioside ND under negative mode, a broad peak centered at mass-to-charge ratio (m/z)  $\sim$ 11000 was observed, which attributed to ND. (c) is the CID spectrum of 7

ganglioside ND under trap 100V after the ND is selected through quadrupole profile and subjected to trap region. All the 7 gangliosides are detected.

The fragments of *C. difficile* toxins' receptor binding regions (TcdA-A2 and TcdB-B3) was chosen to screen with 7 G ND through CaR-ESI-MS assay. Shown in Figure 2.11a are a ESI mass spectra acquired in negative ion mode for a 200 mM aqueous ammonium acetate solution solutions of TcdA-A2 alone (6 $\mu$ M). The major TcdA-A2 ions detected by ESI-MS are deprotonated ions  $A2^{n-}$ , where n range from 11- to 9-. Upon addition of 7 G ND to the TcdA-A2 solution, as shown in Figure 2.11b, a broad peak centered at mass-to-charge ratio ( $m/z$ )  $\sim$ 10000 was observed which attributes to intact ND. Also, the introduction of 7 G ND resulted in the appearance of  $(A2+iGSL)^{n-}$  complex which is detected after the protein peak. Presumably, TcdA-A2 was able to strip its binding partners out of the 7 G ND forming protein-glycolipid complex and the complex will be stripped out of the ND during electrospray process. The exact identity and amount of GSL attached to the target protein is hard to tell due to poor resolution. It is important to note that labelling the complex ions as  $(A2 +iGSL)^{n-}$  is more appropriated. Interestingly, the distribution of the complex's abundance is not the same as the protein peak's abundance, e.g.  $(A2+iGSL)^{10-}$  is the highest abundance peak among complex peaks while  $A2^{9-}$  is the highest abundance peak. Also, some non-specific binding between A2 and DMPC, the major component of ND, can be observed. Those peaks can be distinguished from the complex peak due to the  $m/z$  deference and were labelled with \*.

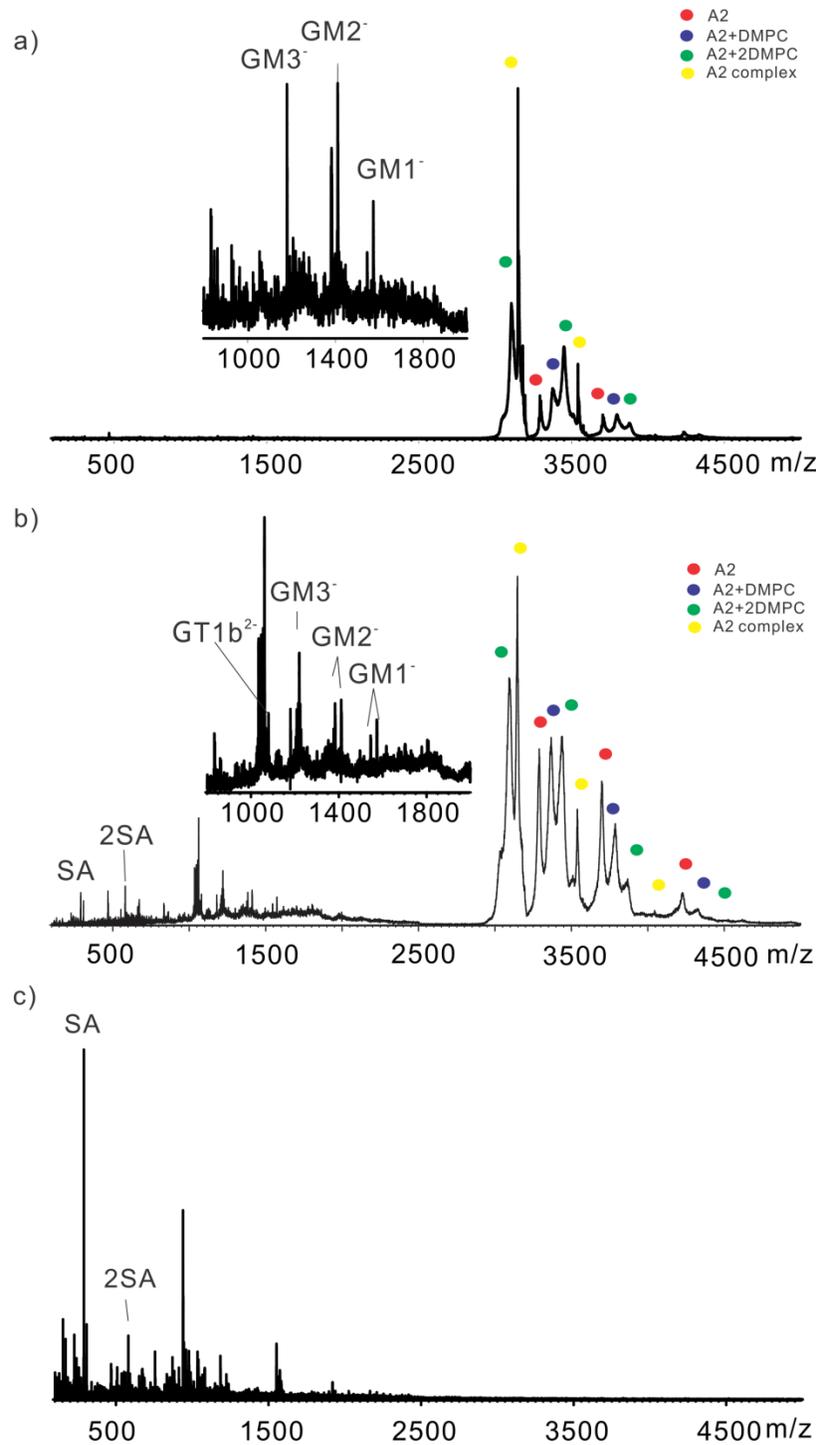


**Figure 2.11** ESI-MS spectra of TcdA-A2 6  $\mu\text{M}$  alone in aqueous ammonium acetate (200 mM) at pH 7 and 25  $^{\circ}\text{C}$  (a), upon incubation with 20  $\mu\text{M}$  2% 7 Ganglioside ND (b). Peaks labeled with \* corresponding to nonspecific interaction between target protein and DMPC, the major component of ND.

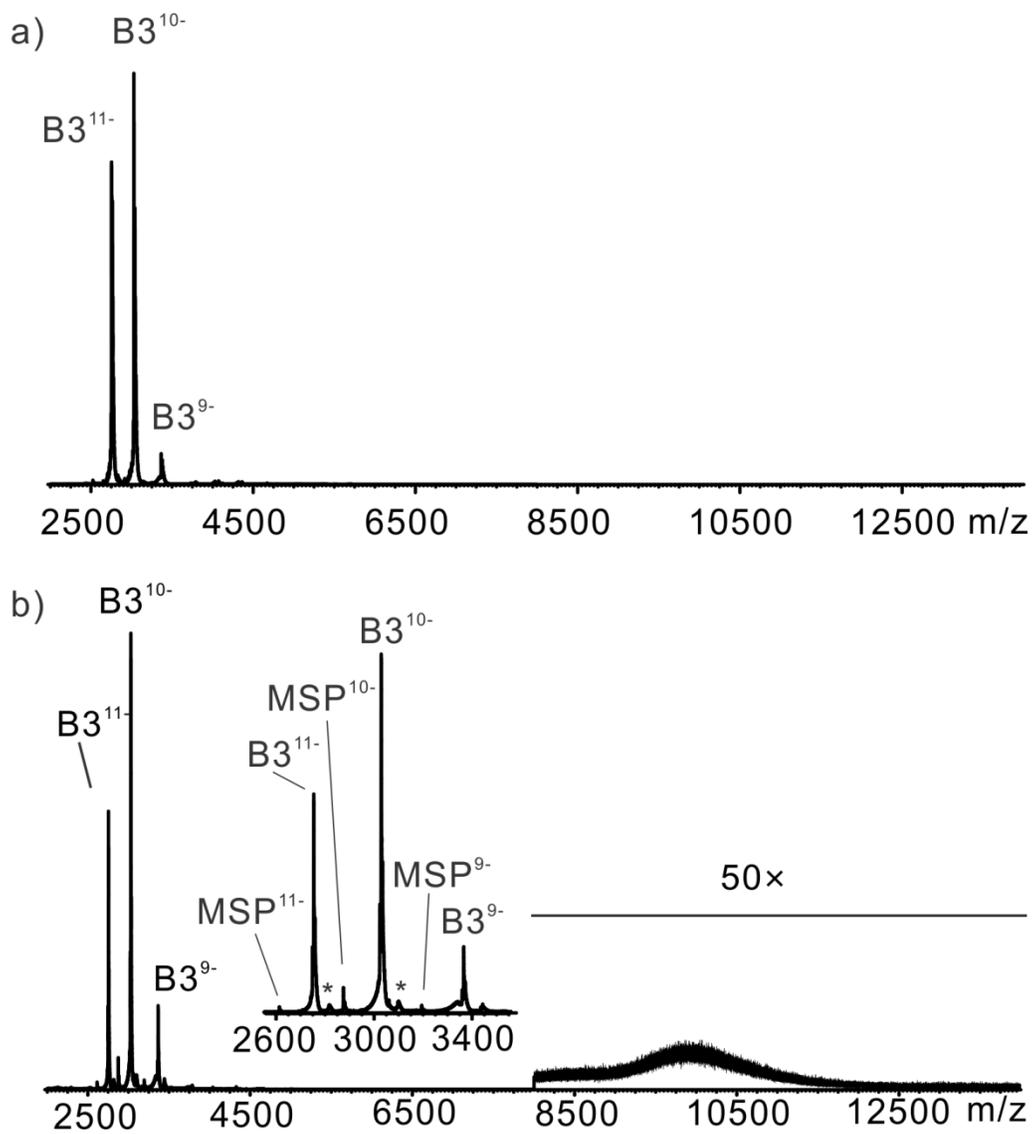
To identify what ligands TcdA-A2 binds to, CaR-ESI-MS assay was applied to isolate the A2-GSL complex. The quadrupole mass filter was centered at 2881 m/z, the predicted m/z value for  $(\text{A2}+\text{GT1b})^{11-}$  (using a MW of 2126 Da for GT1b) and the LM window was set as 4 to ensure the 11- charge state of all the A2-GSL complex were isolated and subjected to the trap region.

CID in the voltage range from 5V to 150V was performed to release the potential binding partners to TcdA-A2. Upon increasing the trap voltage to 50V, shown in Figure 2.12a, deprotonated ligands appeared at  $m/z$  1179, 1383, 1411, 1545, 1573. According to the measured  $m/z$  values, those ligands correspond to the deprotonated ions of GM1 (two major forms d18:1-18:0 and d20:1-18:0), GM2 (two major forms d18:1-18:0 and d20:1-18:0) and GM3 (d18:1-18:0). This result is in good agreement with the direct ESI-MS study done on oligosaccharide parts of ganglioside. While increasing the energy from 50V to 75V, single sialic acid ( $m/z$  290) was observed in the spectra, which indicated further dissociation of ganglioside after them dissociating from the complex. Interestingly,  $m/z$  580 peak appears which correspond to double sialic acid, this suggests higher order gangliosides could be released from the complex cause based on the structure of ganglioside (Figure 2.12b). Also, higher order ganglioside doubly charged GT1b peak ( $m/z$  1162) was detected; reason being a larger ligand requires more energy to dissociate from the complex. Notably, none of GD2, GD1a, GD1b deprotonated peaks can be detected. However, the possibility of those ligands being the binding partners of A2 cannot be ruled out, because they may undergo secondary fragmentation during the CID process under higher voltage. Upon increasing the trap energy to 100 V, the protein peaks disappear and all the ligands related to ganglioside undergo a strong fragmentation. Only single sialic acid peak and double sialic acid peak can be observed (Figure 2.12c). This experiment revealed among 7 gangliosides, 4 of them (GM3, GM2, GM1, GT1b) can be recognized by TcdA-A2, which is in good agreement of the direct ESI-MS study done on the oligosaccharide part of ganglioside, in which GM3<sub>os</sub> and GT1b<sub>os</sub> are the highest-binding ligands. It is also worth noting, dissociation of gangliosides added difficulties to arrange the binding affinities between gangliosides based on their abundances in mass spectra in the CaR-ESI-MS study. For negative-charged monosialyted

ganglioside, it became neutral after losing one sialic acid and wasn't present in mass spectra. For higher order gangliosides, like di-sialyted or tri-sialyted gangliosides, it became mono-sialic acid once it loose sialic acids. Therefore, the abundance of ganglioside shown in CID mass spectra only indicates the presence of negative-charged ganglioside, not including the ganglioside which loose sialic acid.



**Figure 2.12** ESI-CID spectra of TcdA-A2 6 $\mu$ M incubated with 20  $\mu$ M 2% 7 Ganglioside ND selected at 3174 m/z using trap voltage 50V (a), under trap voltage 75V (b) and under trap voltage 100V (c).

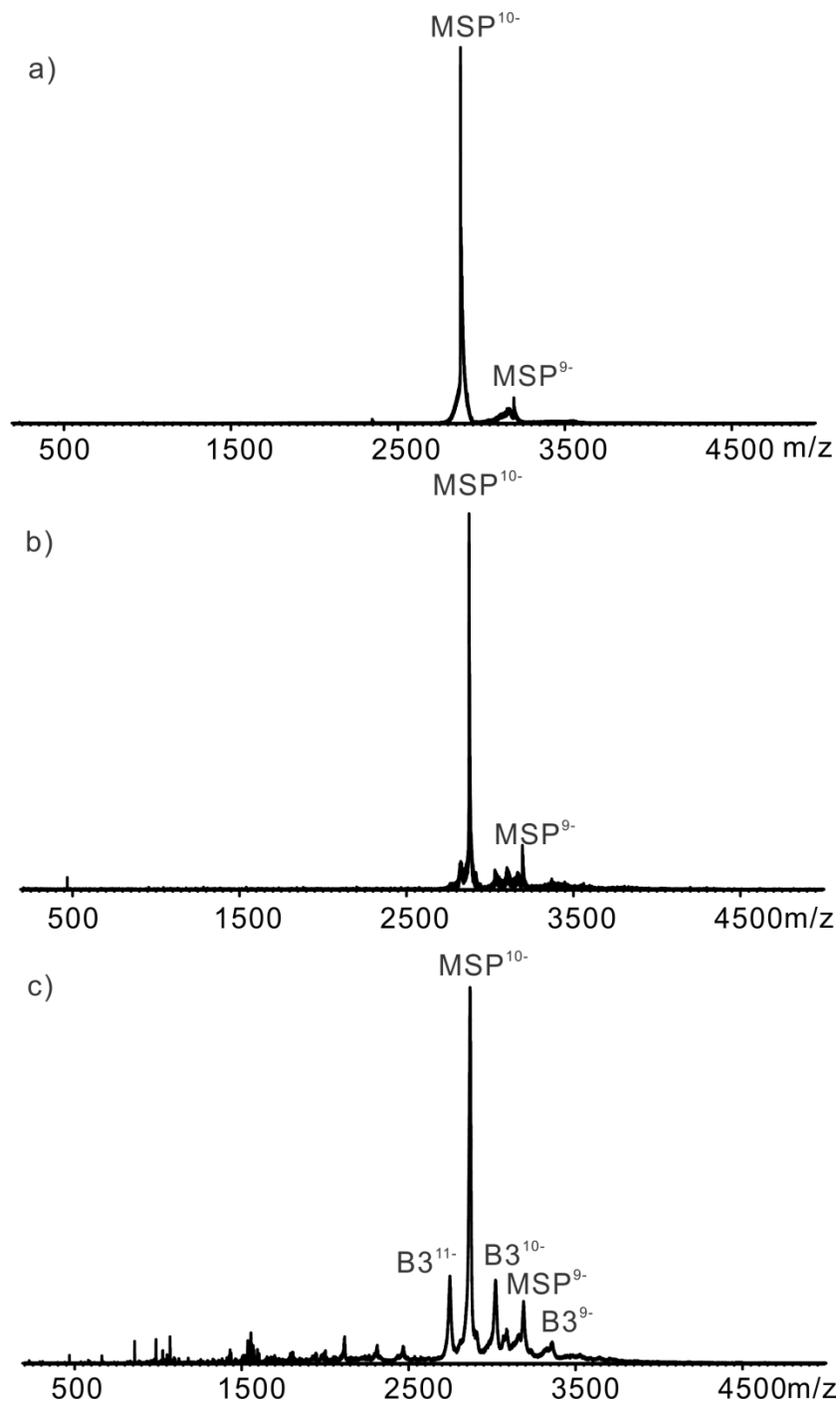


**Figure 2.13** ESI-MS spectra of TcdB-B3 7  $\mu\text{M}$  alone in aqueous ammonium acetate (200 mM) at pH 7 and 25  $^{\circ}\text{C}$  (a), upon incubation with 20  $\mu\text{M}$  2% 7 Ganglioside ND (b). Peaks labeled with \* corresponding to nonspecific interaction between target protein and DMPC, the major component of ND.

Same experiments have been carried on to TcdB-B3. Shown in Figure 2.13a is the ESI-MS spectra of TcdB-B3 alone. The major TcdB-B3 ions detected by ESI-MS in negative mode are deprotonated ions carrying 11-, 10- and 9- charges. After mixing with the 7 G ND, a broad

peak centered at  $\sim 10000$  m/z appeared which confirms the existence of nanodisc. Also, membrane scaffold protein peak is observed presumably due to nanodisc is slightly dissociated under the cone voltage 100V or after storage for a period of time (Figure 2.13b). The difference between Figure 2.11b (A2 mix with 7 G ND) and Figure 2.13b (B3 mix with 7 G ND) is that in Figure 2.11b a series of complex peaks could be observed and none of them can be attributed to membrane scaffold protein peaks, while in Figure 2.13b no complex peaks could be observed relates to  $(B3+iGSL)^{n-}$  and only membrane protein peak could be observed.

To further identify whether any GSL ligands bind to TcdB-B3, quadruple centered at m/z 2874, the predicted m/z value for  $(B3+GM2)^{11-}$  (using a MW of 1383 Da for GM2), while LM window was set as 4 to make sure a wide range ions around m/z 2874 were selected to trap region which may include  $MSP^{10-}$  (m/z 2748) and  $B3^{11-}$  (m/z 2748). Trap energy ranging from 5V to 100V was applied to release the potential binding partners. Figure 2.14a shows the MS-CID spectra under trap 5V,  $MSP^{10-}$  and  $MSP^{9-}$  peaks were the major peaks detected. No deprotonated peaks according to GSL is detected. Upon increasing the energy to 50V, no GSL ligands were observed releasing from the region selected (Figure 2.14b). Upon increasing the trap energy to 100V, a small amount of B3 released from the complex (Figure 2.14c). Some fragment ions are also observed, but none of those ions matched to the known GSL ligands incorporated into nanodiscs. Presumably, those fragment ions come from the fragmentation of DMPC or MSP. Besides, either sialic acid peak (m/z 290) or double sialic acid peak(m/z 580) was not observed from the fragment ions, which confirmed that none of the GSL ligands specifically bind to TcdB-B3.



**Figure 2.14** ESI-CID spectra of TcdB-B3 7 $\mu$ M incubated with 20  $\mu$ M 2% 7 Ganglioside ND selected at 2874 m/z using trap voltage 5V (a), using trap voltage 50V (b), using trap voltage 100V (c). Peaks labeled with \* corresponding to nonspecific interaction between target protein and DMPC, the major component of ND.

## 2.4 Conclusion

The results of quantitative binding measurements conducted on two fragments of *C. difficile* toxins, TcdA and TcdB, and a library of thirty human oligosaccharides, including structures found in gangliosides, globosides and the histo-blood antigens are reported. TcdA-A2 binds to seventeen of the oligosaccharides tested, while TcdB-B3 binds to only four. All of the ganglioside oligosaccharide ligands identified for TcdA-A2 share a  $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc motif. The highest-affinity ligands for TcdA-A2 are **L1** ( $\alpha$ -D-Neu5Ac-(2 $\rightarrow$ 3)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc) with  $K_a$  values of  $1800 \pm 170 \text{ M}^{-1}$ , the presence of additional sialic acids, ( $\alpha$ 2-3)- or ( $\alpha$ 2-8)-linked, reduces or eliminates binding. Additional experiments have been done to confirm that **L1** with a ceramide lipid moiety (GM3) can also be recognized by TcdA-A2. All the globoside oligosaccharide ligands share either a  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc or  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-GlcNAc motif. The highest-affinity ligands for TcdA-A2 are **L19** ( $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc) and **L20** ( $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc), with  $K_a$  values of  $1380 \pm 30 \text{ M}^{-1}$  and  $1400 \pm 30 \text{ M}^{-1}$ , respectively; further extension with  $\alpha$ -D-Gal or  $\beta$ -D-GalNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-Gal resulted in reduction or loss of binding. TcdB-B3 is found bind to **L18-L20**, **L22**, **L25**. They either contain a  $\beta$ -D-Gal-(1 $\rightarrow$ 4)-D-Glc or 1,2-linked  $\alpha$ -Fuc (H epitope). Among all the human cellular oligosaccharides, binding is weak, with apparent affinities  $< 2000 \text{ M}^{-1}$ . However, the affinities are similar in magnitude to that measured for the native carbohydrate receptor of TcdA found in rabbits and other rodents. Taken together, these findings suggest a variety of glycan structures that could serve as the human cellular receptors for TcdA and, possibly, TcdB.

Affinity measurements were also carried on a series of Lewis X analogues **L31-L35**, containing a synthetic aglycone, that were previously identified as ligands for TcdA. Notably, the

relative affinities measured by ESI-MS for TcdA-A2, which range from 1400 to 20000 M<sup>-1</sup>, are in reasonable qualitative agreement with results determined from glycan microarray screening. However, the glycan microarray data failed to identify binding for nine other structures (common to both libraries) that were found to exhibit low affinity interactions based on the ESI-MS data. Based on this finding it is recommended that glycan array screening be complemented, when possible, with binding assays which has lower limit of detection, such as ESI-MS.

## 2.5 Literature cited

97. Bartlett, J. G. *Clin. Infect. Dis.* **1994**, *18*, 265-272.
98. Kachrimanidou, M.; Malisiovas, N. *CRC Cr. Rev. Microbiol.* **2011**, *37*, 178-187.
99. Dubberke, E. R.; Olsen, M. A. *Clin. Infect. Dis.* **2012**, *55*, 88-92.
100. Voth, D. E.; Ballard, J. D. *Clin. Microbiol. Rev.* **2005**, *18*, 247-263.
101. Curry, S. R.; Marsh, J. W.; Muto, C. A.; O'Leary, M. M.; Pasculle, A. W.; Harrison, L. H. *J. Clin. Microbiol.* **2007**, *45*, 215-221.
102. Ballard, J. D. *Nature* **2010**, *467*, 665-666.
103. Jank, T.; Giesemann, T.; Aktories, K. *Glycobiology* **2007**, *17*, 15-22.
104. Just, I.; Hofmann, F.; Genth, H.; Gerhard, R. *Int. J. Med. Microbiol.* **2001**, *291*, 243-250.
105. Just, I.; Gerhard, R. *Reviews of physiology, biochemistry and pharmacology*, Springer Berlin Heidelberg: **2005**, *152*, 23-47.
106. Kuijper, E. J.; van Dissel, J. T.; Wilcox, M. H. *Curr. Opin. Infect. Dis.* **2007**, *20*, 376-83.
107. Petrella, L. A.; Sambol, S. P.; Cheknis, A.; Nagaro, K.; Kean, Y.; Sears, P. S.; Babakhani, F.; Johnson, S.; Gerding, D. N. *Clin. Infect. Dis.* **2012**, *55*, 351-357.
108. Mullane, K. *Ther. Adv. Chro. Dis.* **2014**, *5*, 69-84.

109. Crook, D. W.; Walker, A. S.; Kean, Y.; Weiss, K.; Cornely, O. A.; Miller, M. A.; Esposito, R.; Louie, T. J.; Stoesser, N. E.; Young, B. C.; Angus, B. J.; Gorbach, S. L.; Peto, T. E. *Clin. Infect. Dis.* **2012**, *55*, 93-103.
110. Bartsch, S. M.; Umscheid, C. A.; Fishman, N.; Lee, B. Y. *Clin. Infect. Dis.* **2013**, *57*, 555-61.2.
111. Youngster, I.; Russell, G. H.; Pindar, C.; Ziv-Baran, T.; Sauk, J.; Hohmann, E. L. *J. A.M.A.* **2014**, *312*, 1772-1778.
112. Dallas, S. D.; Rolfe, R. D. *J. Med. Microbiol.* **1998**, *47*, 879-888.
113. Krivan, H. C.; Clark, G. F.; Smith, D. F.; Wilkins, T. D. *Infect. Immun.* **1986**, *53*, 573-581.
114. Dingle, T.; Wee, S.; Mulvey, G. L.; Greco, A.; Kitova, E. N.; Sun, J. X.; Lin, S. J.; Klassen, J. S.; Palcic, M. M.; Ng, K. K. S.; Armstrong, G. D. *Glycobiology* **2008**, *18*, 698-706.
115. Larsen, R. D.; Riveramarrero, C. A.; Ernst, L. K.; Cummings, R. D.; Lowe, J. B. *J. Biol. Chem.* **1990**, *265*, 7055-7061.
116. Tucker, K. D.; Wilkins, T. D. *Infect. Immun.* **1991**, *59*, 73-78.
117. Karlsson, K. A. *Curr. Opin. Struc. Biol.* **1995**, *5*, 622-635.
118. Teneberg, S.; Lonroth, I.; Lopez, J. F. T.; Galili, U.; Halvarsson, M. O.; Angstrom, J.; Karlsson, K. A. *Glycobiology* **1996**, *6*, 599-609.
119. Smith, J. A.; Cooke, D. L.; Hyde, S.; Borriello, S. P.; Long, R. G. *J. Med. Microbiol.* **1997**, *46*, 953-958.
120. Zhang, P.; Razi, N.; Eugenio, L.; Fentabil, M.; Kitova, E. N.; Klassen, J. S.; Bundle, D. R.; Ng, K. K. S.; Ling, C.-C. *Chem. Commun.* **2011**, *47*, 12397-12399.
121. CFG <http://www.functionalglycomics.org/>.

122. El-Hawiet, A.; Kitova, E. N.; Kitov, P. I.; Eugenio, L.; Ng, K. K. S.; Mulvey, G. L.; Dingle, T. C.; Szpacenko, A.; Armstrong, G. D.; Klassen, J. S. *Glycobiology* **2011**, *21*, 1217-1227.
123. Angstrom, J.; Teneberg, S.; Karlsson, K. A. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11859-11863.
124. Merritt, E. A.; Sixma, T. K.; Kalk, K. H.; Vanzanten, B. A. M.; Hol, W. G. J. *Mol. Microbiol.* **1994**, *13*, 745-753.
125. Ling, H.; Boodhoo, A.; Hazes, B.; Cummings, M. D.; Armstrong, G. D.; Brunton, J. L.; Read, R. J. *Biochemistry* **1998**, *37*, 1777-1788.
126. Oriol, R. *J. Immunogene.* **1990**, *17*, 235-245.
127. Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1-28.
128. Ritchie, T. K.; Grinkova, Y. V.; Bayburt, T. H.; Denisov, I. G.; Zolnerciks, J. K.; Atkins, W. M.; Sligar, S. G. *Method. Enzymol.* **2009**, *464*, 211-231.
129. Bayburt, T. H.; Sligar, S. G. *FEBS letters* **2010**, *584*, 1721-1727.
130. Bayburt, T. H.; Leitz, A. J.; Xie, G.; Oprian, D. D.; Sligar, S. G. *J. Biol. Chem.* **2007**, *282*, 14875-14881.
131. Wang, W. J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955.
132. Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discov.* **2006**, *5*, 585-595.

## Chapter 3

### Interaction of Bee Venom Melittin with Zwitterionic Charged Phospholipid Nanodiscs:

#### A Electrospray Ionization Mass Spectrometry Study

##### 3.1 Introduction

Peptide-membrane interactions are an active research topic due to the heavy role it plays in numerous crucial biological processes, such as antimicrobial peptide action, hormone-receptor interactions, drug bioavailability across the blood-brain barrier and viral fusion processes. Many techniques have been used to elucidate peptide-membrane interaction, including chromatographic retention,<sup>133</sup> surface-plasmon resonance,<sup>133-135</sup> calorimetry,<sup>136-141</sup> X-ray diffraction,<sup>140, 142</sup> fluorescence spectroscopy,<sup>133, 138</sup> circular dichroism,<sup>133, 137-139, 142-144</sup> NMR,<sup>137, 139, 143, 145</sup> AFM,<sup>136, 144</sup> vibrational spectroscopy,<sup>133, 144, 146</sup> Langmuir-Blodgett pressure-area isotherm<sup>141, 144</sup> and single-molecule fluorescence imaging.<sup>147-148</sup>

While valuable information on peptide-membrane interactions has been gained using the above methods, it is noted that these aforementioned studies are all studied using a model membrane, most often detergent micelles or liposomes, but these in vitro systems have a number of limitations. As model membranes, micelles have the inherent disadvantage of being unilamellar, which can produce a poorly defined membrane environment. And liposome's preparation often leads to a population of vesicles with non-uniform size.

Recently, synthetic phospholipid bilayers nanodiscs (ND)<sup>92</sup> emerged as model membrane systems. NDs are more stable and monodisperse than the aforementioned model membranes. In addition, the scaffold proteins on the bilayer allow the lipids to display a broad phase transition, which simulates more of the natural biological membranes.<sup>149</sup> Thus, ND have been used to study membrane associated peptide by different techniques for structural and binding information.<sup>150</sup>

However, to the best of our knowledge, they have not been previously used in conjunction with direct ESI-MS binding measurements for peptide-membrane interaction study. Here, we report on the first Electrospray Ionization Mass Spectrometry (ESI-MS) study on the interaction between antimicrobial peptide melittin and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC), a zwitterionic charged phospholipid nanodisc.

Melittin is the major toxin of the venom of the honey bee *Apis mellifera*<sup>151</sup> and has hemolytic<sup>152</sup> and antimicrobial properties.<sup>153</sup> The molecular interaction between a melittin monomer and a lipid bilayer was first described by Terwilliger et al.<sup>154</sup> The binding of cationic melittin to cell membrane has several steps. Firstly, the binding is initiated by an electrostatic attraction to the membrane surface.<sup>155</sup> The concentration of melittin near the plane of binding is considerably higher than in bulk solution. Melittin begins to occupy space in the head-group region of the phospholipid molecules in the bilayer. Electrostatic attraction is then followed by chemical adsorption where melittin inserts into the lipid bilayer and partially compensates for the electric charge of the membrane surface. Melittin exists mainly as random coil in buffer and forms 50-65%  $\alpha$ - helix when bound to membrane. It is believed that the distribution of melittin between the model membrane bilayers depends on the molar peptide-to-lipid ratio (P/L).<sup>154</sup> At low P/L's, the helical orientation of melittin is parallel to the plane of the bilayers and above a defined threshold concentration, the orientation begins to become perpendicular to the bilayer. Then aggregation of melittin occurs inside the lipid bilayer leading to spontaneous formation of lipid pores.<sup>156</sup> It has been reported, that for melittin and POPC bilayer, at P/L=1/128, melittin starts to reorient inside of bilayer. Upon reorientation, at P/L=1/32, melittin starts to induce pore formation.<sup>157</sup> The inner diameter of melittin pore was found to be ~4.4nm at a P/L=1/15.<sup>158</sup>

In the present study, we were interested in observing the strong binding of melittin to POPC nanodisc, even the formation of pores. In order to do so, the concentration of the melittin was kept high enough for melittin to insert into the nanodisc. To describe the current study on melittin-lipid interaction the following topics are discussed: (1) would melittin bind to a lipid bilayer mimic-the ND? (2) If so, would this interaction be preserved in gas phase? (3) If the interaction can be preserved, can ESI-MS be used to probe this interaction?

## **3.2 Material and Method**

### **3.2.1 Materials**

Melittin from honey bee venom (Mel, MW 2846 Da), Insulin chain A from bovine pancreas (ISA, MW 2531 Da), Insulin chain B from bovine pancreas (ISB, MW 3496 Da) were purchased from Sigma-Aldrich (Oakville, ON, Canada) and used without further purification. Melittin contains 26 amino acids. Its sequence is GIGAV LKVLT TGLPA LISWI KRKRQ Q; Insulin chain A contains 21 amino acids. Its sequence is GIVEQ CCASV CSLYQ LENYC N; Insulin chain B contains 30 amino acids with sequence as FVNQH LCGSH LVEAL YLVCG ERGFF YTPKA. The protein concentration in the aqueous solution was determined by UV spectroscopy at 280 nm using an absorption coefficient of 5500  $M^{-1} cm^{-1}$  for Melittin, 2980  $M^{-1} cm^{-1}$  for insulin chain A and 3230  $M^{-1} cm^{-1}$  for insulin chain B. 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC, 760.1 Da) dissolved in chloroform was purchased from Avanti Polar Lipids (Alabaster, AL).

Dextran polysaccharide, with average MWs of 100 kDa, was purchased from Sigma-Aldrich Canada (Oakville, Canada). Recombinant membrane protein MSP1E1 (MW 27 494 Da) was prepared using plasmid pMSP1E1 acquired from Addgene (Cambridge, MA). Protein

expression and purification was then carried out using the procedure described at <http://sligarlab.life.uiuc.edu/nanodisc.html>.

Stock solutions of dextran polysaccharides were prepared by dissolving a known amount of the solid sample in ultrafiltered water (Milli-Q, Millipore) to yield a final concentration of 1 mM. The solutions were stored at -20 °C until needed. A ND stock solution was concentrated and dialyzed against 200 mM ammonium acetate (pH 7.0) using an amicon microconcentrator with a MW cut-off of 30 kDa and stored at -80 °C until needed. The concentration of the ND solutions was determined by absorbance at 280 nm using the extinction coefficient of MSP1E1. Stock solutions of each surfactant was prepared by dissolving a known amount of the sample in ultra-filtered water (Milli-Q, Millipore) to yield a final concentration which is smaller than its CMC value. The solutions were stored at -20 °C until needed. The ND used in the present study, which was composed of the phospholipid POPC and recombinant membrane scaffold protein MSP1E1, has an estimated MW of 190 kDa and Stokes radius of 10.4 nm.<sup>92, 96, 159</sup> The total number of lipid molecules (POPC) per ND was taken to be 205.<sup>160</sup>

### **3.2.2 Preparation of Nanodisc**

Empty nanodisc was prepared as previously described.<sup>96, 159</sup> Briefly, a defined amount of POPC lipid was dried from a stock solution in chloroform under a gentle stream of nitrogen followed by a high vacuum overnight, and then placed in a vacuum desiccator overnight to form a lipid film. Lipids were then re-suspended in 20 mM TrisHCl, 0.5 mM EDTA, 100 mM NaCl, 25 mM sodium cholate pH 7.4 by sonication for 15 min and MSP1E1 added to give a final molar ratio of 80:1 lipid: MSP1E1. After incubation of the cholate: lipid: MSP1E1 mixture for 15 min at room temperature, ND self-assembly was initiated by the addition of Bio-Beads (Bio-Rad laboratories Ltd, Canada) ( 0.5g per mL of reconstitution mixture) and the solution incubated for a further 3 h

to ensure all detergent had been removed. Finally, NDs were purified using a Superdex 200 10/300 size exclusion column (GE Healthcare Bio-Sciences, Uppsala, Sweden) equilibrated in 200 mM ammonium acetate pH 6.8. A single peak with an elution volume at approximately 12 mL confirmed ND formation, the ND composition is further confirmed by mass spectrometry. The formed ND has diameter around 10 nm and Nanodiscs were then concentrated to approximately 60  $\mu$ M and stored at -80 °C until needed.

### **3.2.3 Mass spectrometry**

All the mass spectrometry experiments were carried out using a Synapt G2S quadrupole-ion mobility separation-time of flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK) equipped with a nanoESI source.

Capillary voltage was applied through a platinum wire inserted into a glass capillary (1.0 mm o.d., 0.68 mm i.d.) pulled in-house to  $\sim$ 5  $\mu$ M using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). Negative and positive ionization was applied throughout with a capillary voltage ranging from 0.8-1.0 kV. The source parameters for both negative and positive ion modes were: capillary voltage 0.80-1.20 kV, source temperature 60 °C, cone voltage 30 V, trap voltage 5V, and transfer voltage 2V. Data acquisition and processing were performed using MassLynx software (version 4.1).

All the MS experiments on isolation of Melittin-Nanodisc complex was carried out under TOF-MS mode, a.k.a. the Radio Frequency only Mode, to allow a wide range of ions through to the TOF section of the mass spectrometer. The  $m/z$  position of the transmission range is set via the Radio Frequency Setup dialog window under Quadrupole MS profile. The quadrupole mass profile was set at  $m/z$  steps of 7000, 13000 and 20000 with dwell times at 25%, 25%, 50%, respectively. Ramp time for  $m/z$  7000 and 13000 were set at 20% and 40%. According to the

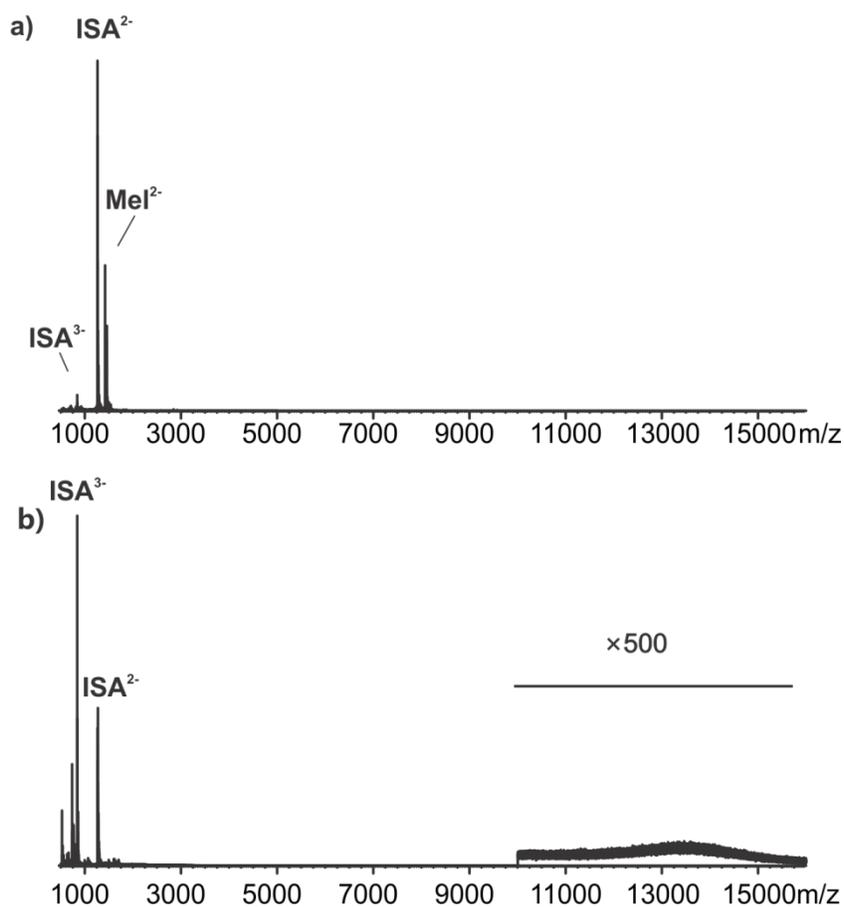
transmission efficiency curve provided by Waters, any ions present <5600 Da (80% of 7000) will not be transmitted by the quad. Any ions >50000 Da (2.5 times the set mass 20000) would be transmitted poorly. This parameter excludes the possibility of free melittin peptide ions passing through quadrupole, only allowing ions including melittin-nanodisc complex to pass through quadrupole. After quadrupole isolation, CID energy was added in the trap region of the Synapt G2s filled with Argon and the resulting mass spectra of dissociated melittin-ND complex extracted using MassLynx software v4.1.

### **3.3 Results and Discussion**

#### **3.3.1 Melittin interacts with phospholipid Nanodisc**

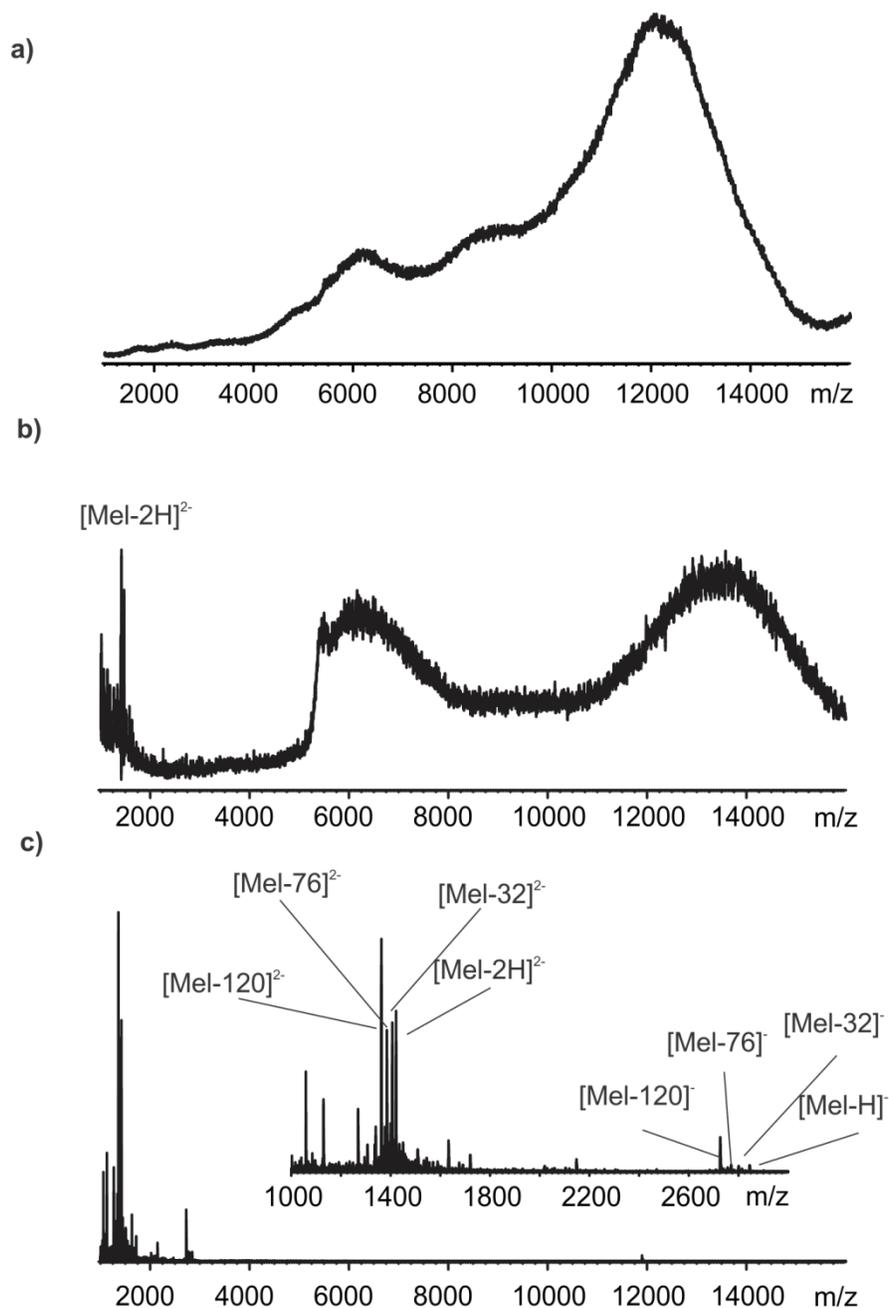
This mass spectrometry assay involves direct ESI-MS analysis of antimicrobial peptides and phospholipid nanodisc in this assay, in aqueous solution. The ND consists of a discoidal phospholipid bilayer surrounded by two copies of an amphipathic helical scaffold protein named MSP1E1 in the present study. First, we want to explore whether antimicrobial peptide melittin, is going to interact with phospholipid nanodisc. Shown in Figure 3.1 are ESI mass spectra acquired in negative ion mode for aqueous neutral solution of Melittin alone and in the presence of POPC ND. To exclude non-specific interaction, insulin chain A is added in both solution to serve as reference peptide, to the best our knowledge, insulin chain A doesn't interact with phospholipids. In Figure 3.1a, the major melittin ions detected by ESI- MS correspond to the doubly deprotonated ions, i.e.,  $\text{Mel}^{2-}$ . The reference peptide insulin chain A is detected as losing two or three protons,  $\text{ISA}^{2-}$  and  $\text{ISA}^{3-}$ . Upon addition of POPC ND to the solution, a broad peak centered at mass-to-charge ratio ( $m/z$ )  $\sim 13000$  was observed (Figure 4.1b). This spectral feature is attributed to the gaseous ions of intact ND. Interestingly, the ratio between two different

charge states of insulin chain A ions varies, meaning  $ISA^{3-}$  is more abundant than  $ISA^{2-}$ . Also, the introduction of POPC ND leads to the total disappearance of the  $Mel^{2-}$  ions. This may suggest all the melittin attach to the POPC ND in the solution, thus no free melittin ions can be observed on the mass spectra. However, this conclusion is not solid because research has shown that the response factor of a protein will change after the introduction of a large molecular weight protein.<sup>161</sup> In this case, POPC ND would affect the response factor of melittin.



**Figure 3.1** ESI mass spectra acquired in negative ion mode for aqueous ammonium acetate (200 mM) solutions of (18  $\mu$ M) Melittin (a) or in the presence of (3.4  $\mu$ M) POPC ND (b). Insulin chain A (8  $\mu$ M) is added to both solutions as reference peptide.

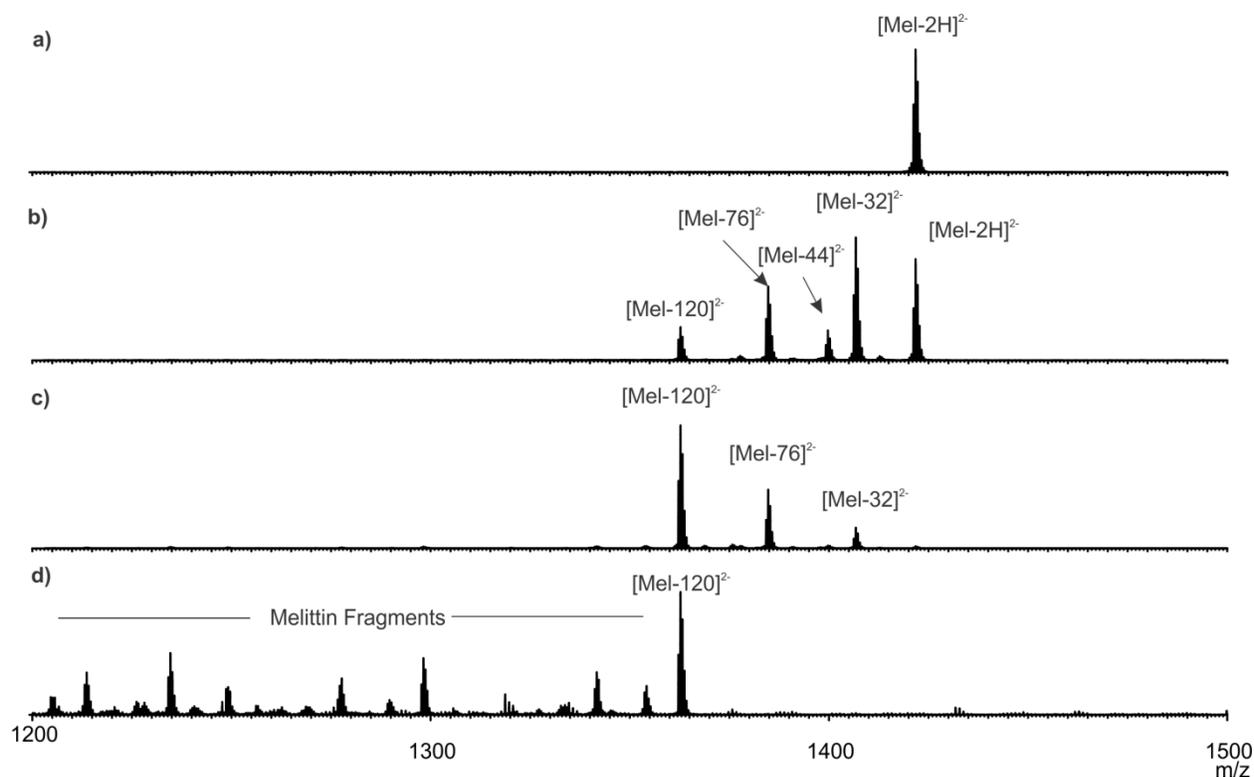
To make sure melittin is attached to ND, the ND ion was then isolated through quadruple mass filter and CID was performed on the isolated complexes as shown in Figure 3.2.



**Figure 3.2** CID mass spectra of (Melittin+ POPC ND) complex acquired in negative mode under different collision energy, trap 5V (a), trap 100V (b), trap 200V (c). Complex ions were selected using quadrupole profile as described in experimental session.

At low energy just sufficient enough for ions to pass the trap cell, as shown in Figure 3.2a, a broad peak centered at  $m/z$  12000 was observed, which has the same characteristic as the nanodisc peak. It indicates melittin binds to the nanodisc under this energy. After increasing the trap voltage to 100V, as shown in Figure 3.2b, nanodisc starts to dissociate from one broad peak to two separate peaks. Also, a doubly deprotonated melittin peak can be observed, indicating melittin is dissociated from the complex. Upon increasing the voltage to trap 200V, the broad peak disappeared indicating the melittin-ND complex is all completely dissociated. The remaining peaks observed include a singly deprotonated melittin peak and a doubly deprotonated melittin peak. Interestingly, there are some fragments of melittin peaks observed and same pattern has been observed on singly deprotonated melittin peaks and doubly deprotonated peaks, which indicated melittin ions undergo second fragmentation. And this fragmentation is different than usual peptide fragmentation observed in MALDI since the lost molecular weight is smaller than an amino acid.

To confirm these fragments actually emerged from melittin, efforts have been made to study the fragmentation pattern of melittin in CID mode. Shown in Figure 3.3 is CID mass spectra acquired for the  $[\text{Mel-H}]^-$  ions. When investigating at low collision energy (trap=5V), the major peak observed are  $[\text{Mel-2H}]^-$  ions, which indicates Melittin easily loses one more proton under low energy. Upon increasing the collision energy to Trap 40V (Fig. 3.3b), Melittin shows the same fragmentation as observed in Figure 3.2. Keeping increasing the energy (Fig. 3.3c and 3.3d), doubly deprotonated ions disappeared, melittin become more fragmented. Even though the exact identity of melittin fragments remains unknown, this experiment confirms melittin ions can undergo fragmentation in CID mode and this fragmentation pattern is the same as melittin which were released from the POPC ND.



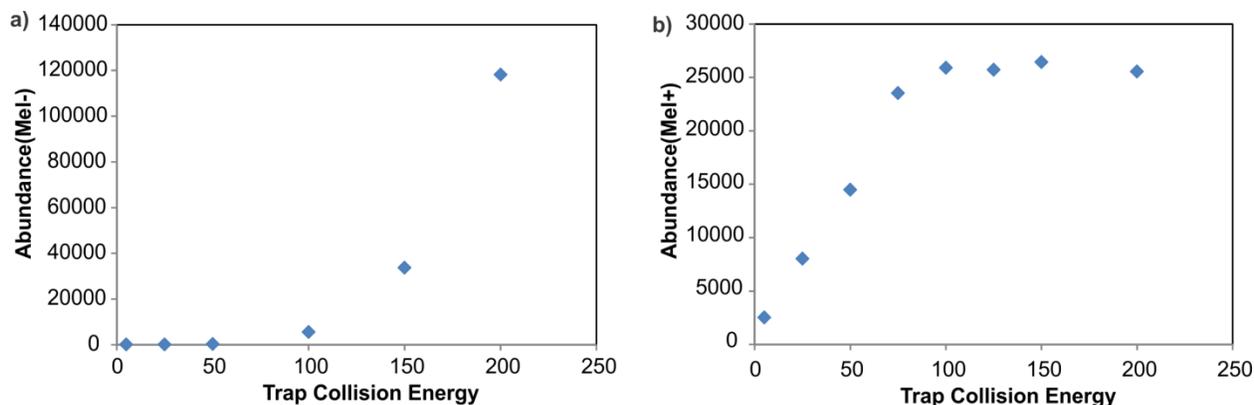
**Figure 3.3** CID spectra of Melittin singly deprotonated ions acquired in negative mode at a collision energy of trap 5V (a), trap 40V (b), trap 50V (c), trap 75V(d). Quadrupole was set at  $m/z$  2845 to select  $[\text{Mel-H}]^-$ .

### 3.3.2 CID of (Melittin+POPC ND) complex

To further study how melittin release from (Melittin+POPC ND) complex, a mixture of 44  $\mu\text{M}$  melittin and 8.4  $\mu\text{M}$  POPC ND is incubated in 200 mM AmAc for 5 min to reach equilibrium state and then sprayed in the mass spectrometry in positive mode. Considering each nanodisc has 205 lipids, the molar ratio between melittin and POPC lipids (P/L) is 1/39, which ensures melittin is inserted into the POPD lipid bilayer. The abundance of melittin ions ( $Ab_{\text{Mel}^-}$ ) includes the abundance of two deprotonated melittin peaks ( $\text{Mel}^-$ ,  $\text{Mel}^{2-}$ ) and fragment ions we have discussed above. As shown in Figure 3.4a, melittin ions start to release from the complex.

However, considering melittin is a positive charged peptide. It would be easier to detect melittin in positive mode. Similar experiments have been carried in positive mode. A mixture of 70  $\mu\text{M}$  melittin and 7.7  $\mu\text{M}$  POPC ND is incubated in 200 mM AmAc to maintain P/L as 1/23. This ratio could also ensure melittin is inserted into the POPD lipid bilayer. The (Melittin+POPC ND) complex is selected through quadrupole profile as described in experimental section. A range of energy is applied on the (Melittin+POPC ND) complex from collision energy of 5V to 200V. However, in positive mode, the released melittin ions we observed are different protonated melittin ions ( $\text{Mel}^+$ ,  $\text{Mel}^{2+}$ ,  $\text{Mel}^{3+}$  and  $\text{Mel}^{4+}$ ). Therefore, the abundance of melittin ions ( $\text{Ab}_{\text{Mel}^+}$ ) includes the abundance of different protonated melittin peaks.

Plots of the abundance ( $\text{Ab}_{\text{Mel}^+}$ ) of the release melittin ions versus energy are shown in Figure 3.4b. Interestingly, we can see as the energy applied to (Melittin+POPC ND), melittin starts to release from the nanodisc, and reach to plateau at Trap 100V. Compared the energy profile plots acquired in negative mode and positive mode, it is curious to know why melittin can't reach plateau under negative mode even apply same energy on the complex. It may be due to melittin is a cationic peptide, meaning melittin contains 5 positively charged residues (Arginine+Lysine) and 0 negative charged residues (Aspartic acid and Glutamic Acid). Under negative mode, there may be some released melittin that couldn't carry any charges, which can be proven by observing trap 5V to trap 100V region. In this region, limited amount of melittin could be observed under negative mode while under positive mode melittin is released rapidly. This comparison indicates positive mode suits the need for this experiment. And performing this experiment at a range of different P/L molar ratios will allow further study of the complex binding observed in Melittin-membrane interactions.



**Figure 3.4** Energy profile plots of abundance of released melittin as the Trap collision energy increased from 5V to 200V at negative mode (a) and positive mode (b). The transfer energy is set as 2V for all the experiments.

### 3.4 Conclusion

The present study represents the first ESI-MS investigation into melittin and POPC lipid bilayer mimic, POPC NDs. This study indicates melittin interact with POPC NDs and this non-covalent interaction can be preserved in gas phase. Also, upon increasing the energy applied to (Melittin+ POPC ND) complex, certain amounts of melittin can be observed to release from the complex and positive mode is the suitable condition to perform experiment on cationic peptide. Taken together, combination of ESI-MS and nanodisc can serve as an effective tool to study the binding between peptide-membrane interactions. However, to study the complex binding patterns between nanodisc and peptides, further experiments are needed.

### 3.5 Literature cited

133. Blondelle, S. E.; Lohner, K.; Aguilar, M. *Biochim. Biophys. Acta* **1999**, *1462*, 89-108.
134. Mozsolits, H.; Wirth, H. J.; Werkmeister, J.; Aguilar, M. I. *Biochim. Biophys. Acta* **2001**, *1512*, 64-76.
135. Papo, N.; Shai, Y. *Biochemistry* **2003**, *42*, 458-466.

136. Pedersen, T. B.; Kaasgaard, T.; Jensen, M. O.; Frokjaer, S.; Mouritsen, O. G.; Jorgensen, K. *Biophys. J.* **2005**, *89*, 2494-2503.
137. Dave, P. C.; Billington, E.; Pan, Y. L.; Straus, S. K. *Biophys. J.* **2005**, *89*, 2434-2442.
138. Poklar, N.; Fritz, J.; Macek, P.; Vesnaver, G.; Chalikian, T. V. *Biochemistry* **1999**, *38*, 14999-15008.
139. Hunter, H. N.; Jing, W.; Schibli, D. J.; Trinh, T.; Park, I. Y.; Kim, S. C.; Vogel, H. J. *Biochim. Biophys. Acta* **2005**, *1668*, 175-189.
140. Lohner, K.; Prenner, E. J. *Biochim. Biophys. Acta* **1999**, *1462*, 141-156.
141. Castillo, J. A.; Pinazo, A.; Carilla, J.; Infante, M. R.; Alsina, M. A.; Haro, I.; Clapes, P. *Langmuir* **2004**, *20*, 3379-3387.
142. Hristova, K.; Dempsey, C. E.; White, S. H. *Biophys. J.* **2001**, *80*, 801-811.
143. Beschiaschvili, G.; Seelig, J. *Biochemistry* **1990**, *29*, 52-58.
144. Deshayes, S.; Plenat, T.; Aldrian-Herrada, G.; Divita, G.; Le Grimellec, C.; Heitz, F. *Biochemistry* **2004**, *43*, 7698-7706.
145. Bechinger, B. *Biochim. Biophys. Acta* **1999**, *1462*, 157-183.
146. Williams, R. W.; Starman, R.; Taylor, K. M.; Gable, K.; Beeler, T.; Zasloff, M.; Covell, D. *Biochemistry* **1990**, *29*, 4490-4496.
147. Fox, C. B.; Wayment, J. R.; Myers, G. A.; Endicott, S. K.; Harris, J. M. *Anal. Chem.* **2009**, *81*, 5130-5138.
148. Myers, G. A.; Gacek, D. A.; Peterson, E. M.; Fox, C. B.; Harris, J. M. *J. Am. Chem. Soc.* **2012**, *134*, 19652-19660.
149. Denisov, I. G.; McLean, M. A.; Shaw, A. W.; Grinkova, Y. V.; Sligar, S. G. *J. Phys. Chem. B* **2005**, *109*, 15580-15588.

150. Denisov, I. G.; Baas, B. J.; Grinkova, Y. V.; Sligar, S. G. *J. Biol. Chem.* **2007**, *282*, 7066-7076.
151. Haberman, E. *Science* **1972**, *177*, 314-322.
152. Haberman, E.; Kowallek, H. *H-S Z. Fur. Physiol. Chem.* **1970**, *351*, 884-890.
153. Brogden, K. A. *Nat. Rev. Microbio.* **2005**, *3*, 238-250.
154. Terwilliger, T. C.; Weissman, L.; Eisenberg, D. *Biophys. J.* **1982**, *37*, 353-361.
155. Boman, H. G. *J. Intern. Med.* **2003**, *254*, 197-215.
156. Lee, M. T.; Hung, W. C.; Chen, F. Y.; Huang, H. W. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5087-5092.
157. Irudayam, S. J.; Berkowitz, M. L. *BBA-Biomembranes* **2012**, *1818*, 2975-2981.
158. Yang, L.; Harroun, T. A.; Ding, L.; Huang, H. W. *Biophys. J.* **2001**, *81*, 1475-1485.
159. Leney, A. C.; Fan, X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271-5277.
160. Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Arch. Biochem. Biophys.* **2006**, *450*, 215-222.
161. Lin, H.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2013**, *85*, 8919-8922.

## Chapter 4

### Conclusions and Future Work

This work describes the development and application of ESI-MS methods to study non-covalent protein-carbohydrate and peptide-membrane interactions. The first research project focuses on identification of the binding partners for *C. difficile* toxins and quantification of the binding affinities between the found ligands and *C. difficile* toxins. The second research project explores ESI-MS combined with nanodisc technology to study peptide-membrane interactions and investigates the optimal experimental conditions.

In the first part of Chapter 2, direct ESI-MS assay was used to identify and quantify the affinities between a library of thirty human oligosaccharides and two recombinant sub-fragments of *C. difficile* toxins TcdA (TcdA-A2) and TcdB (TcdB-B3). A significant finding of this part is that TcdA-A2 binds to seventeen of the oligosaccharides, although with uniformly low affinities,  $<2000 \text{ M}^{-1}$ . The highest affinity ligands are the oligosaccharides of the GM3 ganglioside ( $1800 \pm 170 \text{ M}^{-1}$ ), the globoside Gb3 ( $1380 \pm 30 \text{ M}^{-1}$ ), Gb4 ( $1400 \pm 30 \text{ M}^{-1}$ ) and the histo-blood group A type 2 antigen ( $1700 \pm 30 \text{ M}^{-1}$ ). On the other hand, TcdB-B3 recognized only four oligosaccharides, with low affinities-the oligosaccharides of the globosides Gb3 ( $600 \pm 280 \text{ M}^{-1}$ ) and Gb4 ( $230 \pm 150 \text{ M}^{-1}$ ) and Globo-H hexaose ( $570 \pm 40 \text{ M}^{-1}$ ) and H type 1 tetrasaccharide ( $560 \pm 70 \text{ M}^{-1}$ ). Taken together, these binding data suggest a variety of glycan structures that could serve as the human cellular receptors for TcdA and TcdB. In the second part, affinity measurements were also carried on a series of Lewis X analogues that were previously identified as ligands for TcdA. Notably, the relative affinities measured by ESI-MS for TcdA-A2 are in reasonable agreement with the result determined from the glycan microarray screening for the  $\text{Le}^x$  ligands. However, the glycan microarray data failed to identify binding for nine other

structures (common to both libraries) that were found to exhibit low affinity interactions based on the ESI-MS data. Oligosaccharide portion of several gangliosides were found as the ligands for TcdA-A2, so in the last part of this chapter, nanodisc were used as the tool to solubilize 7 gangliosides and ESI-MS was used to screen both TcdA-A2 and TcdB-B3 to examine the acyl chain's effect on binding. It is found the ganglioside recognized by TcdA-A2 is in strong agreement with the ganglioside oligosaccharide data. Also, no ganglioside is recognized by TcdB-B3, which is also consistent with the previous finding.

The possible extension of Chapter 2's study is exploring *C. difficile* toxins' interaction with other group of oligosaccharides which we haven't investigated in this current study. It is shown in the literature TcdA binds to several Lewis X, Y and I glycan sequences.<sup>116-119</sup> In our study, we have confirmed Lewis X analogs bind to TcdA-A2 with a significant binding affinity ( $K_a \sim 20000 \text{ M}^{-1}$ ). It would be worth measuring the TcdA-A2's binding affinities with Lewis Y and Lewis I glycan sequence and compare those values with current results.

In Chapter 3, ESI-MS was combined with nanodisc technology to study the penetration and interaction of bee venom melittin with 1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC) bilayer nanodisc. The experiment showed that melittin adsorbed onto the phospholipid bilayer of POPC nanodisc but did not disrupt the phospholipid bilayer or produce "cracking" of the POPC nanodisc. Also, it is found the positive mode is the optimal condition to study melittin's interaction with phospholipids. However, more experiments need to be done to establish the method to study peptide-membrane interaction using ESI and nanodisc technology.

The interaction of melittin and zwitterionic POPC phospholipid has been studied using isothermal titration calorimetry<sup>162</sup> and computer simulation.<sup>157</sup> The possible extension of Chapter

3's work is to measure melittin's adsorption coefficient on POPC nanodisc and compare the result to the reported literature result. The possible way to do it is to quantify unbound melittin's concentration based on melittin's abundance in ESI-MS. Internal standard method could be investigated to establish the relationship between melittin's concentration and its abundance in ESI.<sup>163</sup> Also future work needs to be carried out to fully characterize the parameters that control the peptide-membrane interactions, such as the peptide inherent properties, composition and size of the nanodiscs, composition and charge of the nanodisc rim, ionic strength of the bulk solution, etc.

#### **4.1 Literature Cited**

116. Tucker, K. D.; Wilkins, T. D. *Infect. Immun.* **1991**, *59*, 73-78.
117. Karlsson, K. A. *Curr. Opin. Struc. Biol.* **1995**, *5*, 622-635.
118. Teneberg, S.; Lonroth, I.; Lopez, J. F. T.; Galili, U.; Halvarsson, M. O.; Angstrom, J.; Karlsson, K. A. *Glycobiology* **1996**, *6*, 599-609.
119. Smith, J. A.; Cooke, D. L.; Hyde, S.; Borriello, S. P.; Long, R. G. *J. Med. Microbiol.* **1997**, *46*, 953-958.
157. Irudayam, S. J.; Berkowitz, M. L. *BBA-Biomembranes* **2012**, *1818*, 2975-2981.
162. Klocek, G.; Schulthess, T.; Shai, Y.; Seelig, J. *Biochemistry* **2009**, *48*, 2586-2596.
163. Wortmann, A.; Rossi, F.; Lelais, G.; Zenobi, R. *J. Mass Spec.* **2005**, *40*, 777-784.

## Bibliography

1. Cleveland, D. W.; Mao, Y. H.; Sullivan, K. F. *Cell* **2003**, *112*, 407-421.
2. Igakura, T.; Stinchcombe, J. C.; Goon, P. K. C.; Taylor, G. P.; Weber, J. N.; Griffiths, G. M.; Tanaka, Y.; Osame, M.; Bangham, C. R. M. *Science* **2003**, *299*, 1713-1716.
3. Sanderson, J. M. *Org. Biomol. Chem.* **2005**, *3*, 201-212.
4. Matsuzaki, K. *Biochem. Soc. Trans.* **2001**, *29*, 598-601.
5. Sugawara, T.; Kuwajima, K.; Sugai, S. *Biochemistry* **1991**, *30*, 2698-2706.
6. Srisa-Art, M.; Dyson, E. C.; deMello, A. J.; Edel, J. B. *Anal. Chem.* **2008**, *80*, 7063-7067.
7. Malnasi-Csizmadia, A.; Pearson, D. S.; Kovacs, M.; Woolley, R. J.; Geeves, M. A.; Bagshaw, C. R. *Biochemistry* **2001**, *40*, 12727-12737.
8. Bizzarri, A. R.; Cannistraro, S. *J. Phys. Chem. B* **2009**, *113*, 16449-16464.
9. Hill, J. J.; Royer, C. A. In *Fluorescence Spectroscopy*, Brand, L.; Johnson, M. L. **1997**, 278, 390-416.
10. Bao, J.; Krylova, S. M.; Wilson, D. J.; Reinstein, O.; Johnson, P. E.; Krylov, S. N. *Chembiochem.* **2011**, *12*, 2551-2554.
11. Ho, J. G. S.; Greco, A.; Rupnik, M.; Ng, K. K. S. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 18373-18378.
12. Greco, A.; Ho, J. G.; Lin, S. J.; Palcic, M. M.; Rupnik, M.; Ng, K. K. *Nat. Struct. Mol. Biol.* **2006**, *13*, 460-1.
13. Sachchidanand; Lequin, O.; Staunton, D.; Mulloy, B.; Forster, M. J.; Yoshida, K.; Campbell, I. D. *J. Biol. Chem.* **2002**, *277*, 50629-50635.
14. Saboury, A. A. *J. Iran. Chem. Soc.* **2006**, *3*, 1-21.
15. Wilcox, D. E. *Inorg. Chim. Acta* **2008**, *361*, 857-867.

16. Velazquez-Campoy, A.; Freire, E. *Biophys. Chem.* **2005**, *115*, 115-124.
17. Utsuno, K.; Uludag, H. *Biophys. J.* **2010**, *99*, 201-207.
18. Daghestani, H. N.; Day, B. W. *Sensors* **2010**, *10*, 9630-9646.
19. De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. *Cell. Mol. Bioeng.* **2008**, *1*, 204-215.
20. Piliarik, M.; Vaisocherova, H.; Homola, J. *Methods Mol. Bio.* **2009**, *503*, 65-88.
21. Karlsson, R. *J. Mol. Recog.* **2004**, *17*, 151-161.
22. Schuck, P. *Annu. Rev. Bioph. Biom.* **1997**, *26*, 541-566.
23. Gilbert, M.; Albala, J. S. *Curr. Opin. Chem. Biol.* **2002**, *6*, 102-105.
24. Wishart, D. *Curr. Pharm. Biotechnol.* **2005**, *6*, 105-120.
25. Horlacher, T.; Seeberger, P. H. *Chem. Soc. Rev.* **2008**, *37*, 1414-1422.
26. Laurent, N.; Voglmeir, J.; Flitsch, S. L. *Chem. Commun.* **2008**, *37*, 4400-4412.
27. Liu, Y.; Palma, A. S.; Feizi, T. *Biol. Chem.* **2009**, *390*, 647-656.
28. He, X. G.; Gerona-Navarro, G.; Jaffrey, S. R. *J. Pharmacol. Exp. Ther.* **2005**, *313*, 1-7.
29. Barnes-Seeman, D.; Park, S. B.; Koehler, A. N.; Schreiber, S. L. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 2376-2379.
30. Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.
31. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611.
32. Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431-441.
33. Covey, T. R.; Thomson, B. A.; Schneider, B. B. *Mass Spectrom. Rev.* **2009**, *28*, 870-897.
34. Kebarle, P.; Verkerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898-917.
35. Smith, D. P. H. *IEEE Trans. Ind. Appl.* **1986**, *22*, 527-535.

36. Taylor, G. P. *Roy. Soc. A.* **1964**, *280*, 383-397.
37. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240-2249.
38. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Chem. Phys.* **1970**, *52*, 4977-4986.
39. Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.
40. Konermann, L.; Rodriguez, A. D.; Liu, J. *Anal. Chem.* **2012**, *84*, 6798.
41. Ahadi, E.; Konermann, L. *J. Phys. Chem. B* **2012**, *116*, 104.
42. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Chem. Phys.* **1970**, *52*, 4977-4986.
43. De La Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93-104.
44. Fenn, J. B.; Rosell, J.; Meng, C. K. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 1147-1157.
45. Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804-817.
46. Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287.
47. Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, *71*, 4451-4463.
48. Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 3654-3668.
49. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 6926-6927.
50. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *Anal. Chem.* **2009**, *81*, 369-377.
51. Creighton, T. E., *Proteins: Structures and Molecular Properties*. W. H. Freeman: **1993**.
52. Chung, J. K.; Consta, S. *J. Phys. Chem. B* **2012**, *116*, 5777-5785.
53. Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390-4399.
54. Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.

55. Lorenzen, K.; Duijn, E. V. *Current Protocols in Protein Science*, John Wiley & Sons, Inc.: **2001**.
56. Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300-308.
57. Karas, M.; Bahr, U.; Dulcks, T. *Fresenius J. Anal. Chem.* **2000**, *366*, 669-676.
58. Hoffmann, E. d.; Stroobant, V., *Mass Spectrometry: Principles and Applications, 3rd Edition*. Wiley: **2007**, *2*, 502-503.
59. 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*, 1-12.
60. Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, *13*, 597-605.
61. Baird, N. J.; Ferre-D'Amare, A. R. *RNA-Publ. RNA Soc.* **2013**, *19*, 167-176.
62. El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 50-58.
63. Borch, J.; Jorgensen, T. J. D.; Roepstorff, P. *Curr. Opin. Chem. Biol.* **2005**, *9*, 509-516.
64. Ilag, L. L.; Ubarretxena-Belandia, I.; Tate, C. G.; Robinson, C. V. *J. Am. Chem. Soc.* **2004**, *126*, 14362-14363.
65. Heath, B. L.; Jockusch, R. A. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1911-1920.
66. McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614.
67. Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851-1858.
68. March, R. E.; Todd, J. F. J., *Practical Aspects of Trapped Ion Mass Spectrometry Volume IV 1<sup>st</sup> Edition*. CRC Press: **2010**, *4*, 739-767.
69. Jones, C. M.; Beardsley, R. L.; Galhena, A. S.; Dagan, S.; Cheng, G.; Wysocki, V. H. *J. Am. Chem. Soc.* **2006**, *128*, 15044-15045.

70. Tang, X. J.; Brewer, C. F.; Saha, S.; Chernushevich, I.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 750-754.
71. Myers, S. A.; Daou, S.; Affar, E. B.; Burlingame, A. *Proteomics* **2013**, *13*, 982-991.
72. Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.
73. McLafferty, F. W.; Kelleher, N. L.; Begley, T. P.; Fridriksson, E. K.; Zubarev, R. A.; Horn, D. M. *Curr. Opin. Chem. Biol.* **1998**, *2*, 571-578.
74. Wytttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.*, **2007**, *58*, 511-533.
75. Tjernberg, A.; Carno, S.; Oliv, F.; Benkestock, K.; Edlund, P. O.; Griffiths, W. J.; Hallen, D. *Anal. Chem.* **2004**, *76*, 4325-4331.
76. Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708-6715.
77. Chitta, R. K.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1031-1038.
78. Wilcox, J. M.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **2008**, *80*, 2365-2371.
79. Chitta, R.; Rempel, D.; Gross, M. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1031-1038.
80. Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; De Pauw, E. *J. Mass Spec.* **2003**, *38*, 491-501.
81. Sun, N.; Sun, J. X.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1242-1250.
82. Liu, L.; Bagal, D.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2009**, *131*, 15980-15981.
83. Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413-656.

84. Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801-7806.
85. Hossain, B. M.; Konermann, L. *Anal. Chem.* **2006**, *78*, 1613-1619.
86. Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010-3018.
87. Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 8301-8311.
88. Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472-481.
89. Kitova, E. N.; Soya, N.; Klassen, J. S. *Anal. Chem.* **2011**, *83*, 5160-5167.
90. Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Nano Lett.* **2002**, *2*, 853-856.
91. Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477-3487.
92. Bayburt, T. H.; Sligar, S. G. *FEBS letters* **2010**, *584*, 1721-1727.
93. Kobashigawa, Y.; Harada, K.; Yoshida, N.; Ogura, K.; Inagaki, F. *Anal. Biochem.* **2011**, *410*, 77-83.
94. Leitz, A. J.; Bayburt, T. H.; Barnakov, A. N.; Springer, B. A.; Sligar, S. G. *Biotechniques* **2006**, *40*, 601-602.
95. Alami, M.; Dalal, K.; Lelj-Garolla, B.; Sligar, S. G.; Duong, F. *Embo Journal* **2007**, *26*, 1995-2004.
96. Zhang, Y. X.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C. S.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618-7621.
97. Bartlett, J. G. *Clin. Infect. Dis.* **1994**, *18*, 265-272.
98. Kachrimanidou, M.; Malisiovas, N. *CRC Cr. Rev. Microbiol.* **2011**, *37*, 178-187.
99. Dubberke, E. R.; Olsen, M. A. *Clin. Infect. Dis.* **2012**, *55*, 88-92.

100. Voth, D. E.; Ballard, J. D. *Clin. Microbiol. Rev.* **2005**, *18*, 247-263.
101. Curry, S. R.; Marsh, J. W.; Muto, C. A.; O'Leary, M. M.; Pasculle, A. W.; Harrison, L. H. *J. Clin. Microbiol.* **2007**, *45*, 215-221.
102. Ballard, J. D. *Nature* **2010**, *467*, 665-666.
103. Jank, T.; Giesemann, T.; Aktories, K. *Glycobiology* **2007**, *17*, 15-22.
104. Just, I.; Hofmann, F.; Genth, H.; Gerhard, R. *Int. J. Med. Microbiol.* **2001**, *291*, 243-250.
105. Just, I.; Gerhard, R. *Reviews of physiology, biochemistry and pharmacology*, Springer Berlin Heidelberg: **2005**, *152*, 23-47.
106. Kuijper, E. J.; van Dissel, J. T.; Wilcox, M. H. *Curr. Opin. Infect. Dis.* **2007**, *20*, 376-83.
107. Petrella, L. A.; Sambol, S. P.; Cheknis, A.; Nagaro, K.; Kean, Y.; Sears, P. S.; Babakhani, F.; Johnson, S.; Gerding, D. N. *Clin. Infect. Dis.* **2012**, *55*, 351-357.
108. Mullane, K. *Ther. Adv. Chro. Dis.* **2014**, *5*, 69-84.
109. Crook, D. W.; Walker, A. S.; Kean, Y.; Weiss, K.; Cornely, O. A.; Miller, M. A.; Esposito, R.; Louie, T. J.; Stoesser, N. E.; Young, B. C.; Angus, B. J.; Gorbach, S. L.; Peto, T. E. *Clin. Infect. Dis.* **2012**, *55*, 93-103.
110. Bartsch, S. M.; Umscheid, C. A.; Fishman, N.; Lee, B. Y. *Clin. Infect. Dis.* **2013**, *57*, 555-61.2.
111. Youngster, I.; Russell, G. H.; Pindar, C.; Ziv-Baran, T.; Sauk, J.; Hohmann, E. L. *J. A.M.A.* **2014**, *312*, 1772-1778.
112. Dallas, S. D.; Rolfe, R. D. *J. Med. Microbiol.* **1998**, *47*, 879-888.
113. Krivan, H. C.; Clark, G. F.; Smith, D. F.; Wilkins, T. D. *Infect. Immun.* **1986**, *53*, 573-581.

114. Dingle, T.; Wee, S.; Mulvey, G. L.; Greco, A.; Kitova, E. N.; Sun, J. X.; Lin, S. J.; Klassen, J. S.; Palcic, M. M.; Ng, K. K. S.; Armstrong, G. D. *Glycobiology* **2008**, *18*, 698-706.
115. Larsen, R. D.; Riveramarrero, C. A.; Ernst, L. K.; Cummings, R. D.; Lowe, J. B. *J. Biol. Chem.* **1990**, *265*, 7055-7061.
116. Tucker, K. D.; Wilkins, T. D. *Infect. Immun.* **1991**, *59*, 73-78.
117. Karlsson, K. A. *Curr. Opin. Struc. Biol.* **1995**, *5*, 622-635.
118. Teneberg, S.; Lonroth, I.; Lopez, J. F. T.; Galili, U.; Halvarsson, M. O.; Angstrom, J.; Karlsson, K. A. *Glycobiology* **1996**, *6*, 599-609.
119. Smith, J. A.; Cooke, D. L.; Hyde, S.; Borriello, S. P.; Long, R. G. *J. Med. Microbiol.* **1997**, *46*, 953-958.
120. Zhang, P.; Razi, N.; Eugenio, L.; Fentabil, M.; Kitova, E. N.; Klassen, J. S.; Bundle, D. R.; Ng, K. K. S.; Ling, C.-C. *Chem. Commun.* **2011**, *47*, 12397-12399.
121. CFG <http://www.functionalglycomics.org/>.
122. El-Hawiet, A.; Kitova, E. N.; Kitov, P. I.; Eugenio, L.; Ng, K. K. S.; Mulvey, G. L.; Dingle, T. C.; Szpacenko, A.; Armstrong, G. D.; Klassen, J. S. *Glycobiology* **2011**, *21*, 1217-1227.
123. Angstrom, J.; Teneberg, S.; Karlsson, K. A. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11859-11863.
124. Merritt, E. A.; Sixma, T. K.; Kalk, K. H.; Vanzanten, B. A. M.; Hol, W. G. J. *Mol. Microbiol.* **1994**, *13*, 745-753.
125. Ling, H.; Boodhoo, A.; Hazes, B.; Cummings, M. D.; Armstrong, G. D.; Brunton, J. L.; Read, R. J. *Biochemistry* **1998**, *37*, 1777-1788.
126. Oriol, R. *J. Immunogene.* **1990**, *17*, 235-245.

127. Ravn, V.; Dabelsteen, E. *Apmis* **2000**, *108*, 1-28.
128. Ritchie, T. K.; Grinkova, Y. V.; Bayburt, T. H.; Denisov, I. G.; Zolnerciks, J. K.; Atkins, W. M.; Sligar, S. G. *Method. Enzymol.* **2009**, *464*, 211-231.
129. Bayburt, T. H.; Sligar, S. G. *FEBS letters* **2010**, *584*, 1721-1727.
130. Bayburt, T. H.; Leitz, A. J.; Xie, G.; Oprian, D. D.; Sligar, S. G. *J. Biol. Chem.* **2007**, *282*, 14875-14881.
131. Wang, W. J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2003**, *75*, 4945-4955.
132. Hofstadler, S. A.; Sannes-Lowery, K. A. *Nat. Rev. Drug Discov.* **2006**, *5*, 585-595.
133. Blondelle, S. E.; Lohner, K.; Aguilar, M. *Biochim. Biophys. Acta* **1999**, *1462*, 89-108.
134. Mozsolits, H.; Wirth, H. J.; Werkmeister, J.; Aguilar, M. I. *Biochim. Biophys. Acta* **2001**, *1512*, 64-76.
135. Papo, N.; Shai, Y. *Biochemistry* **2003**, *42*, 458-466.
136. Pedersen, T. B.; Kaasgaard, T.; Jensen, M. O.; Frokjaer, S.; Mouritsen, O. G.; Jorgensen, K. *Biophys. J.* **2005**, *89*, 2494-2503.
137. Dave, P. C.; Billington, E.; Pan, Y. L.; Straus, S. K. *Biophys. J.* **2005**, *89*, 2434-2442.
138. Poklar, N.; Fritz, J.; Macek, P.; Vesnaver, G.; Chalikian, T. V. *Biochemistry* **1999**, *38*, 14999-15008.
139. Hunter, H. N.; Jing, W.; Schibli, D. J.; Trinh, T.; Park, I. Y.; Kim, S. C.; Vogel, H. J. *Biochim. Biophys. Acta* **2005**, *1668*, 175-189.
140. Lohner, K.; Prenner, E. J. *Biochim. Biophys. Acta* **1999**, *1462*, 141-156.
141. Castillo, J. A.; Pinazo, A.; Carilla, J.; Infante, M. R.; Alsina, M. A.; Haro, I.; Clapes, P. *Langmuir* **2004**, *20*, 3379-3387.
142. Hristova, K.; Dempsey, C. E.; White, S. H. *Biophys. J.* **2001**, *80*, 801-811.

143. Beschiaschvili, G.; Seelig, J. *Biochemistry* **1990**, *29*, 52-58.
144. Deshayes, S.; Plenat, T.; Aldrian-Herrada, G.; Divita, G.; Le Grimellec, C.; Heitz, F. *Biochemistry* **2004**, *43*, 7698-7706.
145. Bechinger, B. *Biochim. Biophys. Acta* **1999**, *1462*, 157-183.
146. Williams, R. W.; Starman, R.; Taylor, K. M.; Gable, K.; Beeler, T.; Zasloff, M.; Covell, D. *Biochemistry* **1990**, *29*, 4490-4496.
147. Fox, C. B.; Wayment, J. R.; Myers, G. A.; Endicott, S. K.; Harris, J. M. *Anal. Chem.* **2009**, *81*, 5130-5138.
148. Myers, G. A.; Gacek, D. A.; Peterson, E. M.; Fox, C. B.; Harris, J. M. *J. Am. Chem. Soc.* **2012**, *134*, 19652-19660.
149. Denisov, I. G.; McLean, M. A.; Shaw, A. W.; Grinkova, Y. V.; Sligar, S. G. *J. Phys. Chem. B* **2005**, *109*, 15580-15588.
150. Denisov, I. G.; Baas, B. J.; Grinkova, Y. V.; Sligar, S. G. *J. Biol. Chem.* **2007**, *282*, 7066-7076.
151. Haberman, E. *Science* **1972**, *177*, 314-322.
152. Haberman, E.; Kowallek, H. *H-S Z. Fur. Physiol. Chem.* **1970**, *351*, 884-890.
153. Brogden, K. A. *Nat. Rev. Microbio.* **2005**, *3*, 238-250.
154. Terwilliger, T. C.; Weissman, L.; Eisenberg, D. *Biophys. J.* **1982**, *37*, 353-361.
155. Boman, H. G. *J. Intern. Med.* **2003**, *254*, 197-215.
156. Lee, M. T.; Hung, W. C.; Chen, F. Y.; Huang, H. W. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5087-5092.
157. Irudayam, S. J.; Berkowitz, M. L. *BBA-Biomembranes* **2012**, *1818*, 2975-2981.
158. Yang, L.; Harroun, T. A.; Ding, L.; Huang, H. W. *Biophys. J.* **2001**, *81*, 1475-1485.

159. Leney, A. C.; Fan, X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271-5277.
160. Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Arch. Biochem. Biophys.* **2006**, *450*, 215-222.
161. Lin, H.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2013**, *85*, 8919-8922.
162. Klocek, G.; Schulthess, T.; Shai, Y.; Seelig, J. *Biochemistry* **2009**, *48*, 2586-2596.
163. Wortmann, A.; Rossi, F.; Lelais, G.; Zenobi, R. *J. Mass Spec.* **2005**, *40*, 777-784.