Characterization of Noncovalent Interactions between

Human Milk Oligosaccharides and Proteins

by

Yajie Chen

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

Master of Science

Department of Chemistry
University of Alberta

Abstract

This thesis focuses on the characterization of noncovalent interactions between human milk oligosaccharides (HMOs) and proteins using electrospray ionization mass spectrometry (ESI-MS) based assays. HMOs play critical roles in protecting infants from viral and bacterial infections through noncovalent interactions with lectins. An method based on catch-and-release (CaR) ESI-MS assay was developed using three human galectin proteins as model systems. The method used three dimensional information that was extracted from purified HMOs: molecular weight (MW), ion-mobility separation arrival time (IMS-AT), and collision induced dissociation (CID) fingerprints to identify the binders. The result showed this method is able to simultaneously screen mixtures of free HMOs for binding to lectins in vitro.

The method then was extended to screening HMO binders from natural libraries that extracted from pooled human milk. The HMO binders to a C-terminal fragment of human galectin 3 (hGal-3C) from natural libraries (human milk fractions) were identified. The assay was able to identify at least 35 HMO structures presented in the fractions. After screening against hGal-3C, the assay discovered a total of 17 molecular weights as the binder, and 11 of them correspond to 21 HMO structures that were previously shown to be ligands for this lectin.

Discovery of the HMO binders to viral lectins is often a key for understanding the defensive mechanisms of HMO. Here, we identified and analyzed the interactions between the rotavirus spike protein VP8* domain and HMOs using a direct ESI-MS assay. The results showed that the binding patterns between different human rotavirus strains varies. Overall, the interactions of VP8* to HMOs are weak ($K_a < 10^4$ M^{-1}). G10P[11] showed preference in binding to type 2 lactosamine (LacNAc) and α2-6 sialic acid; G1P[8] showed almost no binding to HMOs; G2P[4] has various binding affinities to different structures in a range of 10^3 M^{-1} $< K_a < 10^4$ M^{-1} and seems

to prefer structures with one or more fucoses. The results suggested that the defensive mechanisms

of HMOs could be both structure-depended and strain-depended.

Preface

This thesis contains research work under the supervision of Dr. John S. Klassen in the Department of Chemistry, University of Alberta, between January 2015 and August 2018. As the supervisory author, he was involved throughout all the projects in the concept formation and manuscript composition.

Chapter 2 has been published as: El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E.N.; St-Pierre, Y.; Klassen, J.S. "High Throughput, Label- and Immobilization-Free Screening of Human Milk Oligosaccharides Against Lectins" *Anal. Chem.* **2017**, *89*, 8713. I was responsible for performing concept-proving experiments. A. El-Hawiet performed most experiments and data analysis as well as prepared the manuscript. K. Shams-Ud-Doha and I were responsible for preparing the human milk oligosaccharide (HMO) library. E.N. Kitova assisted with the manuscript preparation.

Part of Chapter 3 has been published as: El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E.N.; Kitov, P.I.; Bode, L.; Hage, N.; Falcone, F.H.; Klassen, J.S. "Screening natural libraries of human milk oligosaccharides against lectins using CaR-ESI-MS" *Analyst* **2018**, *143*, 536. A. El-Hawiet and I together performed the experiments and data analysis as well as prepared the manuscript. E.N. Kitova provided support in fragmentation fingerprints of HMOs and helped in preparing the human milk fractions. We acknowledge L. Bode for providing the human milk fractions and F.H. Falcone for providing the bacterial proteins, and both of them provided input in the manuscript.

Chapter 4 showed preliminary data of an on-going project. I was responsible for data collection and analysis and the writing.

Yajie Chen

July 2018

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Professor John S. Klassen, for his patient guidance and support and his inspiration on all the projects throughout my graduate study. I would like to thank my supervisors and defense committee members, Professor Robert E. Campbell, Professor Liang Li, and Professor Florence Williams for their help and support during my research and the comments to the thesis. I am particularly grateful for the encouragement and help given by Dr. Amr El-hawiet.

I also wish to acknowledge the help provided by Dr. Elena N. Kitova, Dr. Pavel I. Kitov, and Dr. Lin Han, for their expert advice, technical supports, and guidance during my research.

I would like to extend my thanks to both previous and present colleagues in the Klassen group for their support, encouragement, and friendship during my program: Dr. Heajin Park, Dr. Mickey Richards, Jun Li, Jianing Li, Hong Lin, Yilin Wang, Zhixiong Li, and Erick Baez.

I would like to express my special thanks to Dr. Anna D. Jordan for her great help in preparing the thesis and the support during the time of preparation. I wish to acknowledge the financial support provided by the Department of Chemistry, University of Alberta.

Lastly, I am grateful to my family for their continuous and unparalleled love.

Table of Contents

Chapter 1 Characterization of Noncovalent Interactions between Human Milk Oligosaccharides and Proteins

Chapter 4 Identifying Interactions between Human Milk Oligosaccharides and VP8* Domain of Rotavirus

Chapter 5 Conclusions and Future Work

Table of Contents (*cont'd*)

Bibbiography

List of Tables

List of Figures

ammonium acetate solution (20 mM, pH 6.8) of *Fr2* (0.05 µg µL^{-1}).

Figure 3.5. Representative IMS-ATDs acquired in negative ion mode for aqueous ammonium acetate solution (40 mM, pH 6.8) of *Fr2* (0.05 µg µL^{-1}) for deprotonated HMO ions at (a) m/z 608.12; (b) m/z 690.62; (c) m/z 706.25; (d) m/z 836.25; (e) m/z 852.31; and (f) m/z 998.34. (g) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 852.31.

Figure 3.6. (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of *Fr3* (0.05 µg µL^{-1}). IMS-ATDs measured for deprotonated HMO ions with (b) m/z 341.31, (c) m/z 487.17, and (d) m/z 706.25.

Figure 3.7. (a) Representative ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (20 mM, pH 6.8) of *Fr4* (0.05 µg µL^{-1}). IMS-ATDs measured for deprotonated HMO ions with (b) m/z 632.21; (c) m/z 643.44; (d) m/z 753.67; (e) m/z 826.17; (f) m/z 835.20; and (g) m/z 899.18. (i) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 997.34.

Figure 3.8. CID mass spectrum acquired for deprotonated HMO ions at m/z 835.20. Fragmentation scheme shown for (b) β- GlcNAc-(1→3)-[α-Neu5Ac-(2→6)]-β-Gal-(1→4)-β-Glc and (c) α-Neu5Ac-(2→6)-β-GlcNAc-(1→3/6)-β-Gal-(1→4)-β-Glc.

Figure 3.9. CID mass spectrum acquired for deprotonated HMO ions at m/z 836.25. (b) Fragmentation scheme shown for α-L-Fuc-(1→3)-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc.

Figure 3.10. CID mass spectrum acquired for deprotonated HMO ions at m/z 1201.30. (b) Fragmentation scheme shown for β-D-GlcNAc-(1→3)-[L-Fuc-(1→3)]-β-D-Gal-(1→4)-[L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-β-D-Gal-(1→4)-β-D-Glc.

β-D-GlcNAc-(1→4)-[L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc.

Figure 3.18. (a) Representative ESI mass spectrum acquired in the negative ion mode for 40 mM aqueous ammonium acetate solutions (pH 6.8) of P_{ref} (5 µM), *Fr2* (0.05 µg µL^{-1}), and hGal-3C (15 µM), (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at a Trap voltage of 40 V showing the released HMOs ligands; IMS-ATDs of (c) m/z 690.62; (d) m/z 706.25; (e) m/z 836.25; (f) m/z 852.31; and (g) m/z 998.34. (h) CID mass spectrum acquired for released HMO anions with IMS-AT of 12.43 ms using a Transfer voltage of 30 V.

Figure 3.19. (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of P_{ref} (5 µM), *Fr1* (0.05 µg µL^{-1}), and hGal-3C (15 µM). (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at Trap voltage 40 V. IMS-ATDs of released HMO ions at (c) m/z 632.21, (d) m/z 836.25, (e) m/z 852.31, (f) m/z 998.34, (g) m/z 1055.32, (h) m/z 1071.38, (i) m/z 1217.46, and (j) m/z 1363.25. (k) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 852.31.

Figure 3.20. (a) Representative ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (20 mM, pH 6.8) of P_{ref} (5 µM), *Fr3* (0.05 µg µL^{-1}), and hGal-3C (15 µM), (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at Trap voltage 40 V. IMS-ATDs of released HMO ions at (c) m/z 487.17 and (d) m/z 706.25.

Figure 3.21. (a) Representative ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (20 mM, pH 6.8) of P_{ref} (5 µM), *Fr4* (0.05 µg µL^{-1}), and hGal-3C (15 µM). (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at Trap voltage 40 V. IMS-ATDs of released HMO ions at (c) m/z 632.21, (d) m/z 643.44, (e) m/z 753.67, (f) m/z 826.17, (g) m/z 899.18, and (h) m/z 997.34. (i) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 997.34.

List of Abbreviations

Ab	Abundance of gas-phase ions
AEAB	2-(N-aminoethyl)amino benzamide
AT	Arrival time
ATD	Arrival time distribution
BabA	Blood group antigen-binding adhesion
CA	Carbonic anhydroase
CaR	Catch and release
CEM	Chain ejection model
CID	Collision induced dissociation
CRM	Charged residue model
Da	Dalton
DC	Direct current
EI	Electron impact ionization
ELISA	Enzyme-linked immunosorbent assay
Eq	Equation
ESI	Electrospray ionization
Fr	Fraction
Fuc	Fucose
FUT2	Fucosyltransferase 2
FUT3	Fucosyltransferase 3
FWHM	Full width at half maximum
Gal	Galactose

Glc	Glucose
GlcNAc	N-acetylglucosamine
hGal	Human galectin
hGal-1	Human galectin 1
hGal-3C	The C terminal of human galectin 3
hGal-7	Human galectin 7
HGBA	Histo-Blood Group Antigens
HMG	Human milk glycan
HMO	Human milk oligosaccharide
IEM	Ion evaporation model
IMS	Ion mobility separation
ITC	Isothermal titration calorimetry
K_a	Association constant
L	Ligand
LabA	LacDiNAc-specific binding adhesin
Lac	Lactose
LacNAc	Lactosamine
LNT	Lacto-N-tetraose
m/z	Mass-to-charge ratio
MALDI	Matrix assisted laser desorption ionization
MS	Mass spectrometry
MW	Molecular weight
nanoESI	Nanoflow electrospray ionization

P	Protein
pHMO	Pooled human milk oligosaccharide
PL	Protein-ligand complex
P_{ref}	Reference protein
$P_{ref}L$	Reference protein-ligand complex
Q-	Quadrupole-
R	Abundance ratio
R_{app}	Apparent binding affinity
RF	Radio Frequency
R_{ref}	Reference protein abundance ratio
SGM	Shotgun Glycan microarray
Sia	Sialic acid
ToF	Time-of-flight
T-wave	Traveling wave
TWIMS	Traveling wave ion mobility separation
WHO	World health organization

<center>**Chapter 1**</center>

<center>**Characterization of Noncovalent Interactions between Human Milk**</center>

<center>**Oligosaccharides and Proteins**</center>

1.1 Introduction

1.1.1 Human milk oligosaccharides

In the early 19th century, people preferred cow milk based formula for newborns instead of breastfeeding. With the development of research on the benefits of human milk, people started to understand and embrace the fact that human milk is a gift for babies. Nowadays, a number of health organizations, including the World Health Organization (WHO), recommend breastfeeding as the best source of nutrients for babies for the first six months of their life.[1] A number of research showed that breastfeeding brings health benefits, more balanced nutrition, and better developmental outcomes for infants. The breastfed babies have a lower mortality rate and a lower diarrhea and respiratory infection rate than the formula-fed babies.[2,3] Breastfeeding even results in long-term effects regarding cardiovascular disease and obesity in adulthood.[4] Figure 1.1 shows the composition of human and cow milk. The percentages of lactose and fats are similar. However, cow milk has more protein than human milk, and it has only a trace amount of oligosaccharides compared to human milk.
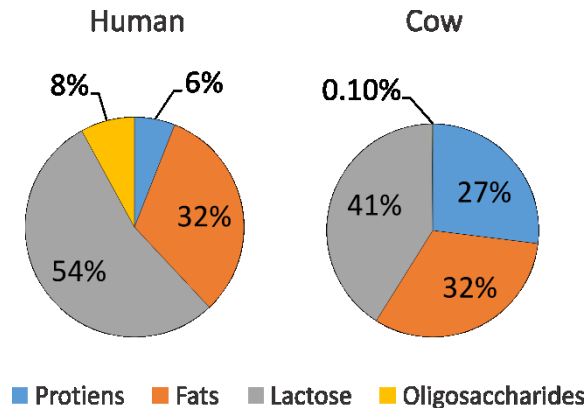
Figure 1.1. Composition of human and cow milk, adapted from reference 5.

Human milk oligosaccharides (HMOs) are the third largest component in human milk. They present in milk in a high concentrations (approximate 7–23 g/L) and are composed of hundreds of different structures of oligosaccharides.[5] HMOs commonly start with a lactose core and includes five building blocks: glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), fucose (Fuc), and sialic acid (Sia). The basic structure composition of HMO and some common HMOs are shown in Figure 1.2.
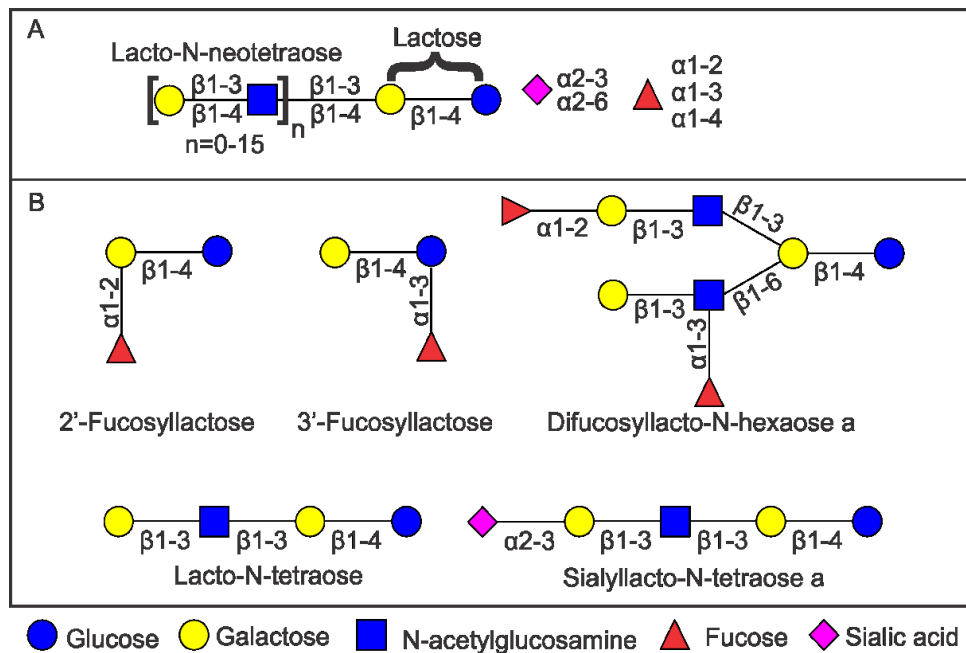
Figure 1.2. A) Basic structure of HMO and B) some common HMOs, adapted from reference 5.

The structure can be elongated by multiple LacNAc motifs with the addition of Fuc and/or Sia. The linkages between monosaccharides are generated enzymatically, therefore, a large number of structural isomers exist among HMOs.[6] Some of the linkages between monosaccharides depend on individual gene; for example, the linkage of Fuc largely depends on the Lewis group type. The 'Secretors', meaning people who have the fucosyltransferase 2 (FUT2) enzyme, have

α1-2 linkage of Fuc onto the HMO backbone.[7] On the other hand, people without the FUT2 enzyme will not have this particular link. Similarly, Fuc with α1-3 or 1-4 linkages only exist when one has the fucosylatranferase 3 (FUT3) enzyme.[8] Sia brings negative charge to the HMOs. There are two possible linkages for Sia, including α2-3 and α2-6. It is possible that one HMO has more than one Fuc and/or Sia. With combinations of building blocks and linkages, more than 200 structures of HMO have been discovered. More than 80% of HMOs are neutral, and about 70% of HMOs are fucoylated.[5]

Since 1954 when HMOs were found the first time,[9] several lines of evidence showed HMOs have contributions in infant health. They act as prebiotics by promoting the growth of beneficial microorganisms in the intestine of breast-fed neonates.[10] The presence of HMOs can change the cell surface glycan profiles in vitro, leading to a reduction in binding of bacteria.[11] Although most HMOs were found mostly in the intestinal environment, some systemic HMOs were found, too (100–200 µL per mL of milk in estimation). Many immune proteins are considered potential binders of HMOs, such as selectins, siglecs, and galectins.[12] Through the interaction with these human lectins, HMOs act as immune modulators that promote the infant's native immune system. In addition, HMOs have been shown to be related to brain development. Researches indicated breastfed babies have higher concentration of sialic acid in their brain, which is connected to learning ability.[13] Also, meta-analysis showed breastfed children have higher IQ on average.[14]

Another important role of HMOs is the behavior of acting as antimicrobials. Instead of changing the glycan profile, this defends the infants in a more direct manner. Many bacteria and viruses start their infection through interaction with the glycans on the epithelial cells of the host (infants). Due to the high concentration of HMOs, and their structural similarity to host glycans, researchers believe that HMOs block the pathogen interaction by acting as soluble decoys.[15,16] For

example, with addition of 2'-FL and 3'-FL, the infectivity of norovirus was reduced, compared to the addition of lactose.[17]

Central to the varied biological roles played by HMOs are the specific noncovalent interactions they form with endogenous and exogenous protein receptors. While the importance of these interactions is appreciated, the molecular details of these interactions are understood poorly. The large number of HMO structures found in human milk, including the presence of many structural isomers and a wide range of concentrations, represent significant challenges to the comprehensive analysis of HMO interactions with proteins.[18]

1.1.2 Current techniques to Characterize HMO–protein Interactions

To understand the underlying mechanism of how HMOs bring health benefits to infants, characterizing the interactions between HMOs and its receptors is important. The current techniques used for carbohydrates–protein interaction studies includes isothermal titration calorimetry (ITC), enzyme-linked immunosorbent assay (ELISA), glycan microarray, and electrospray ionization mass spectrometry (ESI-MS). Some of these tools detect free ligands in solution whereas others rely on immobilization of one of the binding partners. Each of these techniques has its own advantages and limitations.

ITC is one of the most frequently used methods for measuring the binding affinity of protein–small molecule interactions. Often, it is considered the 'gold standard' approach. ITC determines the binding affinity and binding stoichiometry through enthalpy changes.[19] Perret and colleagues demonstrated using ITC for measuring the interactions between two HMOs and a bacterial lectin.[20] The result showed an association constant is in a range of $\sim 10^7$ M^{-1}. One of the limitations of ITC is the inability to obtain reliable measurements for low binding affinities (K_a $< 10^4$ M^{-1}). In addition, the experiments are usually time-consuming (hours); and it requires a large amount of protein and ligand (in miligrams).[21]

ELISA is used widely to study the inhibition of ligands as a competitive assay in HMO–protein interaction studies. A common setup for a competitive ELISA assay is to immobilize the target protein on a plate, as it allows different concentrations or types of ligands to compete with the biotinylated native ligand.[22] The color intensity developed from the bound native ligands reversely corresponds to the binding of the tested ligands. Although ELISA cannot measure association or disassociation directly, it provides information on the inhibition ability between different ligands to the target protein. ELISA is limited to measure high affinity interactions, due to multiple washing steps. In addition, the immobilization of the protein impacts the ligand interaction potentially. Other ELISA based methods were developed by attaching oligosaccharides covalently to a plate, yet they require chemical modification of the glycans.[23]

A Glycan microarray is another technique that requires anchoring ligands onto a solid surface. Human milk glycan (HMG) microarrays developed by Cummings and co-workers is a dominate technique for screening large amount of HMO binders to lectins.[24] The steps for preparing a microarray include conjugating purified glycans with 2-(N-aminoethyl)amino benzamide (AEAB) at their reducing ends and 'print' them onto a solid surface.[25] The current shotgun HMG microarray contains 247 oligosaccharides fractions from pooled human milk.[26] With a large number of ligands, a glycan microarray is a great tool for discovering the binding partners of a protein in a high throughput manner. However, it bears the same limitations as ELISA. Because of the washing steps, it detects high affinity interactions only. Additionally, the modification of glycans for immobilization and the modification on protein for signal detecting could alter the nature of the interactions. A recent study showed a poor agreement on results between glycan microarray and an ESI-MS based method.[27]

ESI-MS has been used for studying ligand–protein interactions since the 1990s.[28] Several advantages of ESI-MS suggests that this method is suitable for HMO–protein interaction studies.

5

First, it requires no labeling or immobilization of the ligands or the proteins; second, the amount of sample required is sufficiently low (in picomole), that the necessary amount of purified protein and HMO often are unavailable; lastly, it is able to detect low affinities ($K_a < 10^4$ M^{-1}) which is a shared feature for most HMO–protein interactions. Through a direct ESI-MS assay, the interaction affinity can be measured.[29] With an Car-ESI-MS ESI-MS assay, identification of binders in a single screening can be achieved in little time (<1hr). In this thesis, both assays are used to characterize HMO–protein interactions. The detailed working principle of the assays and their limitations are described in the following sections.

1.2 Electrospray Ionization Ion Mobility Time of Flight Mass Spectrometry

A Synapt G2 quadrupole-ion mobility separation-time-of-flight (Q-IMS-ToF) mass spectrometer (Waters UK Ltd., Manchester, UK), equipped with a nanoflow electrospray ionization (nanoESI) source was used in this thesis. Figure 1.3 shows a schematic diagram of a Synapt G2 mass spectrometer. Its major components include a nanoESI, a quadrupole, a Triwave, and a ToF. Information on each component will be described.
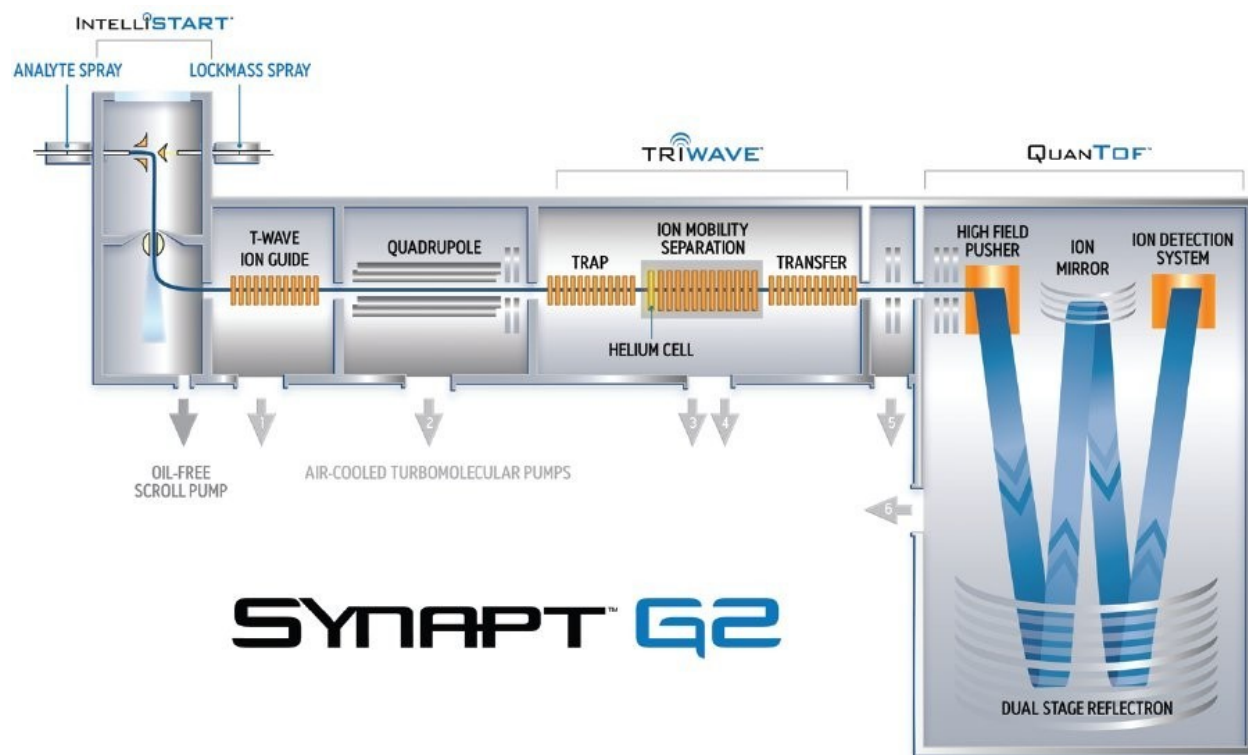
Figure 1.3. A schematic diagram of the Synapt G2 HDMS system (http://www.waters.com).

1.2.1 Electrospray ionization

Since a mass spectrometer measures charged residuals, analyte ionization is critical. There are several methods of ionizing a molecule, including electron impact ionization (EI), matrix assisted laser desorption ionization (MALDI), and electrospray ionization (ESI). To study the noncovalent HMO–protein interactions, an ionization method that is gentle and does not destroy protein–ligand complex is required. Currently ESI is the most commonly used ionization method, and it is one of the most 'soft' methods.[30] NanoESI is a minimized-flow ESI with several benefits in studying protein–carbohydrate interactions. It provides a more efficient method of introducing a solution sample, especially for the ones with a high salt contamination level because of the smaller droplet size.[31,32] It limits the sample consumption and reduces the formation of nonspecific binding. Despite the advantages of nanoESI, it shares the same fundamental process as the conventional

7

ESI. The sample is sprayed from a glass tube with a tip diameter of ~5 µM by applying a voltage to the solution.

Three major steps of ionization are shown in Figure 1.4.[33–36] Using the positive ion mode as an example, first, at high voltage, the positive ions start to accumulate at the tip of the capillary and form a Taylor cone, whereas the negative ions are left behind.[37] Once the electrostatic forces at the surface of the Taylor cone overcome the solvent adhesion force, droplets with positive analyte ions are formed. Next, the diameter of analyte droplets shrinks until the repulsion between the charges at the surface overcomes the cohesive force of the surface tension. Then, the droplets go through fission repeatedly until they attain nanometer diameters. In nanoESI, because the initial size of the droplets are smaller (starting with micrometers), the droplets go through fewer cycles of shrinkage, and this leads to a shorter time between solution analytes and gas phase ions. This helps to maintain the solution composition and to reduce the chance of having nonspecific binding.[38,39] Once all the solvent is evaporated completely, gas phase ions are generated and proceed to the downstream analyzers.
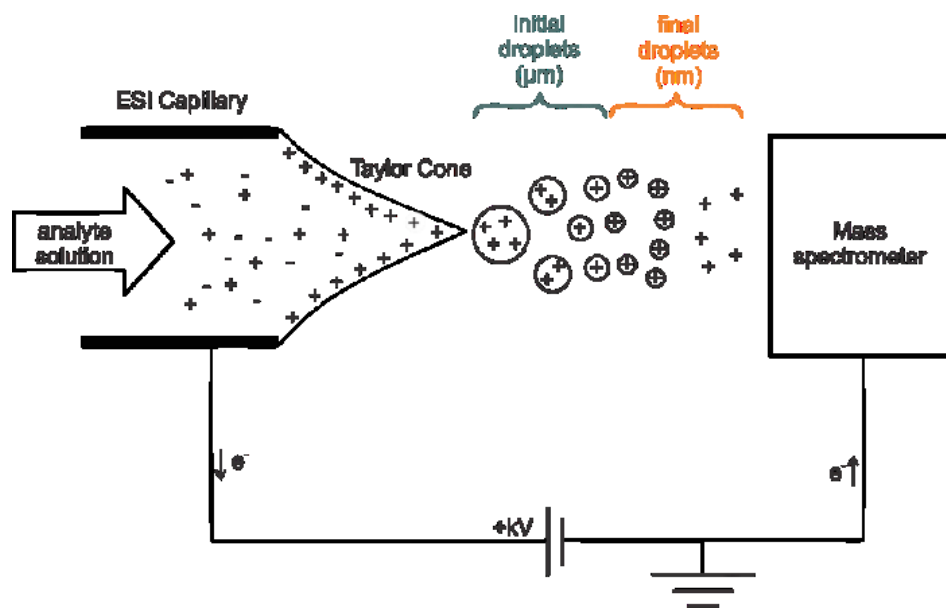
Figure 1.4. Illustration of an ESI process in the positive ion mode, adapted from reference 40.

To date, three proposed mechanisms of gas phase ion generation are involved and shown in Figure 1.5.[40] They are the ion evaporation model (IEM), the charged residue model (CRM), and the chain ejection model (CEM). CRM was suggested to explain the formation of gas phase ions of proteins and HMO–protein complexes.[41,42] In this model, after leaving the Taylor cone, the droplets go through shrinking-fission cycles and end in highly charged nano-droplets containing one macromolecule. The charges on the droplet surface are transferred onto the surface of the molecules when all solution are evaporated completely.
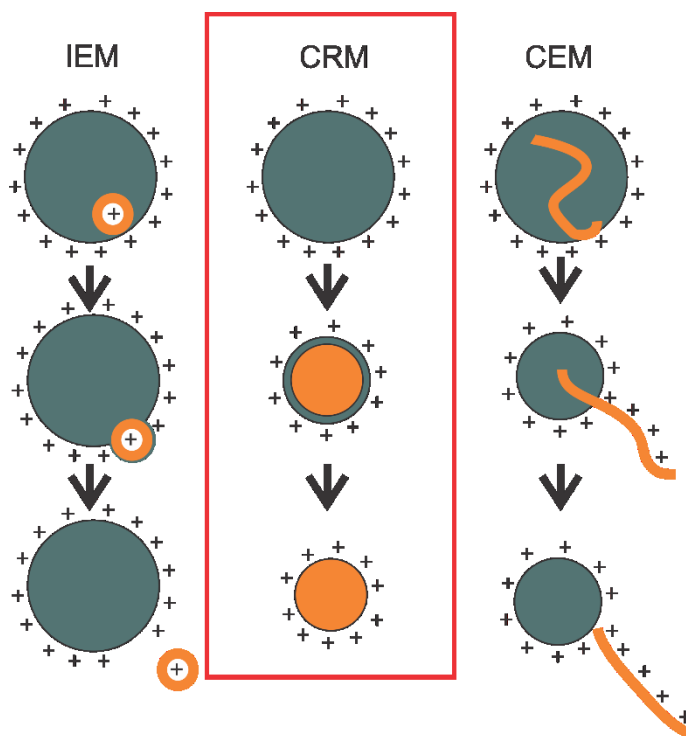
Figure 1.5. Three principle theories of gas phase ion generation mechanism, adapted from reference 40.

Due to the ion formation mechanism, nonspecific binding occurs during the ionization step. It could be a potential problem for both direct and CaR-ESI-MS assays, which will be discussed

in Section 1.4.1. Once the analytes gaseous ions form, they are injected into mass filters and analyzers.

1.2.2 Quadrupole

A quadrupole is used commonly as an ion filter. As the name suggests, a quadrupole is composed of four robs, shown in Figure 1.6.
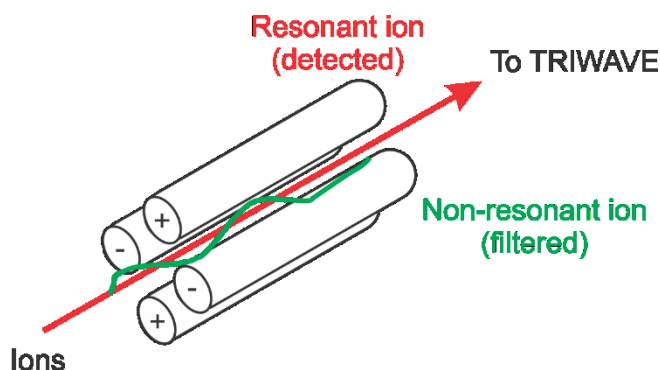
Figure 1.6. Diagram of rod arrangement of a quadrupole, adapted from reference 43.

The four rods are positioned accurately in a radical array, with two rods connected with positive voltages and two rodes with negative voltages.[44] A direct current (DC) potential and a radio frequency (RF) potential are applied to each pair and generate an electrical field. A quadrupole only allows ions with a particular *m/z* ratio to pass through, based on the voltages applied. Figure 1.7 explains the selectivity of a quadrupole. With a constant DC/RF ratio, a straight mass scan line can be obtained. By increasing the DC and RF voltages, the ions with *m/z* from low to high values will pass the quadrupole in turns. The ratio of DC/RF also determines the resolution. If the slope of the mass scan line becomes greater (dashed line in Figure 1.7), a smaller *m/z* range within the stable area is selected, meaning a higher resolution. When the slope is zero, the quadrupole operates in a RF only mode and allows ions above a certain *m/z* threshold to pass.
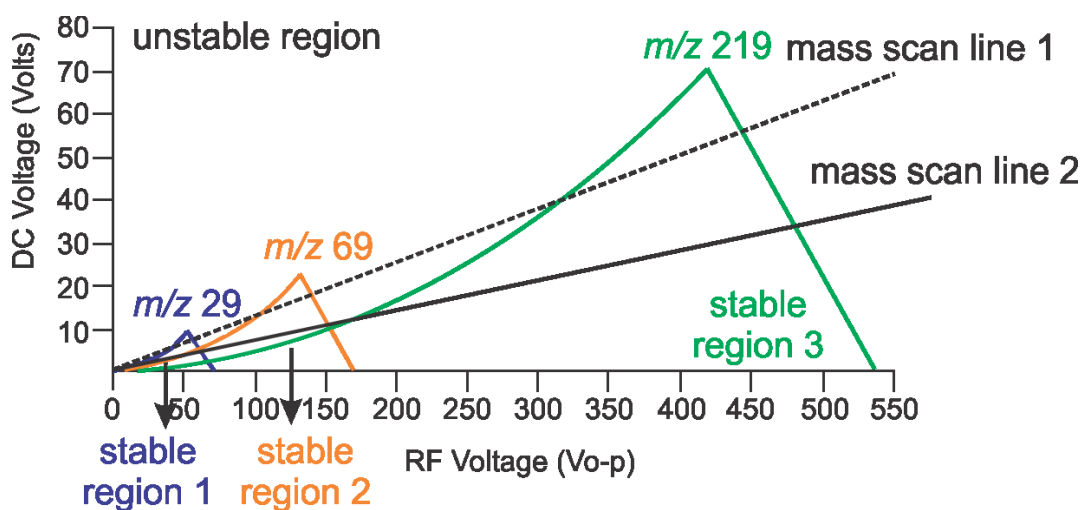
Figure 1.7. Principles of selectivity of a quadrupole for different *m/z* ions, adapted from reference 45.

1.2.3 TriWave–Ion mobility separation and collision induced dissociation

Once ions are transmitted through the quadrupole mass filter, they enter the TriWave section. It consists of three traveling (T-wave) ion guides (Trap, ion mobility separation, and Transfer), shown in Figure 1.2. The ion guides are composed of non-uniform, moving electric field /voltage pulses that push ions through the chambers. Each T-wave has a stack of ring electrodes where positive and negative RF voltages are applied on the adjacent rings, as shown in Figure 1.8.[46,47]
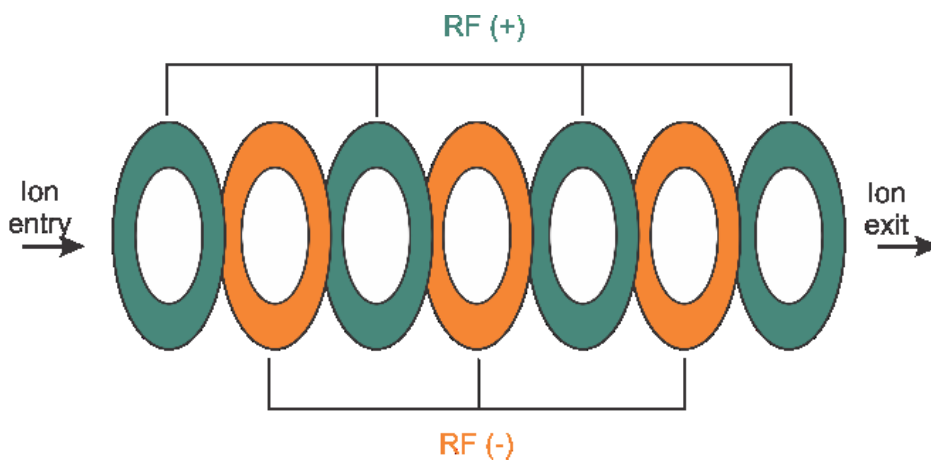
Figure 1.8. Illustration of T-Wave ring electrodes, adapted from reference 48.

A continuous series DC voltage is applied to provide the pulses to push ions moving forward. Meanwhile, the IMS cell is filled with nitrogen gas as the buffer gas that flow against the ions' motion. The mobility of each ion depends on the cross-section (i.e., the size and the shape) of the ion. With the presence of the waves from DC pulses, larger ions with lower mobility will move slower than the high-mobility ions. Therefore, the gas phase ions are separated based on their cross section area instead of their m/z ratio.[49] It adds a dimension to protein–ligand interaction analysis. When needed, Trap and Transfer T-wave ion guides are available for pre- or post-IMS collision induced dissociation (CID) fragmentation.[41] This setting provides a unique possibility when designing an experiment.

CID is also known as collision activated dissociation. It is critical for releasing ligands from the protein–ligand complexes, and the fragmentation in the precursor ions. In Synapt G2, both Trap and Transfer ion-guide are able to perform CID by applying a constant DC voltage as the collision energy to each ring electrode. In the same time, a transient DC voltages is applied too for 'pushing' ions through the next stage of the instrument. Once protein–ligand complex ions enter the Trap/Transfer region, they will collide with the background gases (Argon in Synapt G2) that accompanied by a high internal energy, results in an increased internal energy in the target ions.[50] At first, the noncovalent interactions between protein and ligands will be destroyed; once the process provides enough internal energy, fragmentation of the covalent bond will occur. For the released ligand ions, further CID can be applied in the Transfer region, resulting the covalent bond fragmentation. The applications of using CID to study the noncovalent interaction and to generate fragmentation fingerprints are described in Chapter 2 and 3.

1.2.4 Time of flight

A Time of Flight (ToF) analyzer has a high scan speed and a large mass range.[51] It usually is

composed of one or more flight tubes between 0.5 and 2 m. The ions are accelerated by an electric

field with a certain voltage (U), and the potential energy, which is the charge of the particle (q)

multiplied by the voltage (U), is converted to kinetic energy (E_k). For an ion with a certain mass,

the kinetic energy is determined by the velocity (v) and voltage (U) (Eq 1.1). Since velocity is

determined by the flight path length (D) and flight time (t), the flight time (t) can be related to the

mass (m) (Eqs 1.2, 1.3):

$$E_k = qU = \frac{1}{2} mv^2 \tag{1.1}$$

$$v = \frac{D}{t} \tag{1.2}$$

$$t = \sqrt{\frac{mD^2}{2qU}} \tag{1.3}$$

According to the above equations, ions with different *m/z* will arrive at the detector at a

different time, with larger ions having a longer flight time.[52] In Synapt G2, the ToF analyzer is

composed of four flight tubes plus two stage reflectrons and one ion mirror. The mechanisms of

ion mirror and reflectron are the same. Both are composed of multiple electric plates with a

negative voltage applied. Thus, faster ions with more positive charges penetrate deeper into the

field, i.e., have a longer distance, allowing the same *m/z* ions with different speeds to reach the

detector at the same time, meaning a better resolution. This setting allows for a maximum flight

path without occupying a big space, and it compensates for the different initial energy between

ions with the same *m/z* ratio.[53]

1.3 Direct ESI-MS Assay

The binding stoichiometry and affinity of HMO–protein complexes can be determined from the

direct ESI-MS assay by quantifying the ratio (*R*) of the total abundance (*Ab*) between the free (P)

and ligand-bound protein (PL) ions (Eqs 1.4,1.5). The binding association constant (K_a) can be

determined from R, the initial concentration of ligand ($[L]_0$), and the initial concentration of protein ($[P]_0$) in the solution, using Eq 1.6:[54]

$$P + L \rightleftharpoons PL \tag{1.4}$$
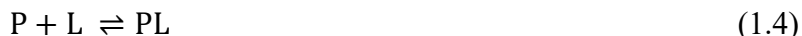
$$\frac{[PL]_{eq}}{[P]_{eq}} = \frac{Ab(PL)}{Ab(P)} = R \tag{1.5}$$

$$K_a = \frac{R}{[L]_0 - \frac{R}{1+R}[P]_0} \tag{1.6}$$

To increase the reliability of the affinity measurements, a K_a value can be obtained from a nonlinear regression analysis of a series of R values using Eq 1.7:

$$\frac{R}{R+1} = \left\{ 1 + K_a[L]_0 + K_a[P]_0 - [(1 - K_a[L]_0 + K_a[P]_0)^2 + 4K_a[L]_0]^{\frac{1}{2}} \right\} / (2K_a[P]_0) \tag{1.7}$$

In the case of multiple ligands ($L_1, L_2, L_3 \ldots L_x$), which is common in screening libraries, the K_a of an individual molecular weight ($K_{a,Lx}$) is calculated from Eq 1.8:

$$K_{a,L_x} = \frac{R_{PL_x}}{[L_x]_0 - \frac{R_{PL_x}}{1 + R_{PL_1} + \cdots + R_{PL_x}}[P]_0} \tag{1.8}$$

It should be stressed that these equations are based on the single binding stoichiometry shown in Eq 1.4.

1.4 Catch and Release Electrospray Ionization Ion Mobility Mass Spectrometry Assay

Direct ESI-MS assay is a powerful tool for providing quantitative information on interactions, but it is limited when the target protein has heterogeneity or gives a low quality spectrum due to a high molecular weight. In this work, catch and release (CaR) ESI-MS assay coupled with IMS-MS was used for richer interaction characterization information.[55] In brief, the target protein 'catches' its ligands in solution, and the HMO–protein complex ions are isolated in the quadrupole. Followed by collisional induced dissociation (CID) carried by the Trap region, the binders are released and identified by their molecular weights (MWs). In the cases where MWs are not sufficient to

recognize the released ligands, IMS and post-IMS CID fragmentation were used for more dimensional information on the ligand characterization, which makes it possible to distinguish between isomer binders. The utility of the CaR-ESI-MS assay for defined carbohydrate library screening has been demonstrated in Chapter 2 and 3.

1.5 Limitations of ESI-MS Assays

1.5.1 Nonspecific Binding

Using ESI-MS method to obtain accurate quantitative affinity measurements requires careful controlling in the level of nonspecific binding, as it is generated artificially in the process of ESI. Due to the mechanism of protein–ligand complex formation (described in Section 1.2.1), one or more ligands that do not interact with the target protein naturally may be trapped within the same nano-droplet, and it leads to stable gaseous complexes, as described in Figure 1.9. This causes higher associate constant and potentially wrong binding stoichiometry in the solution.
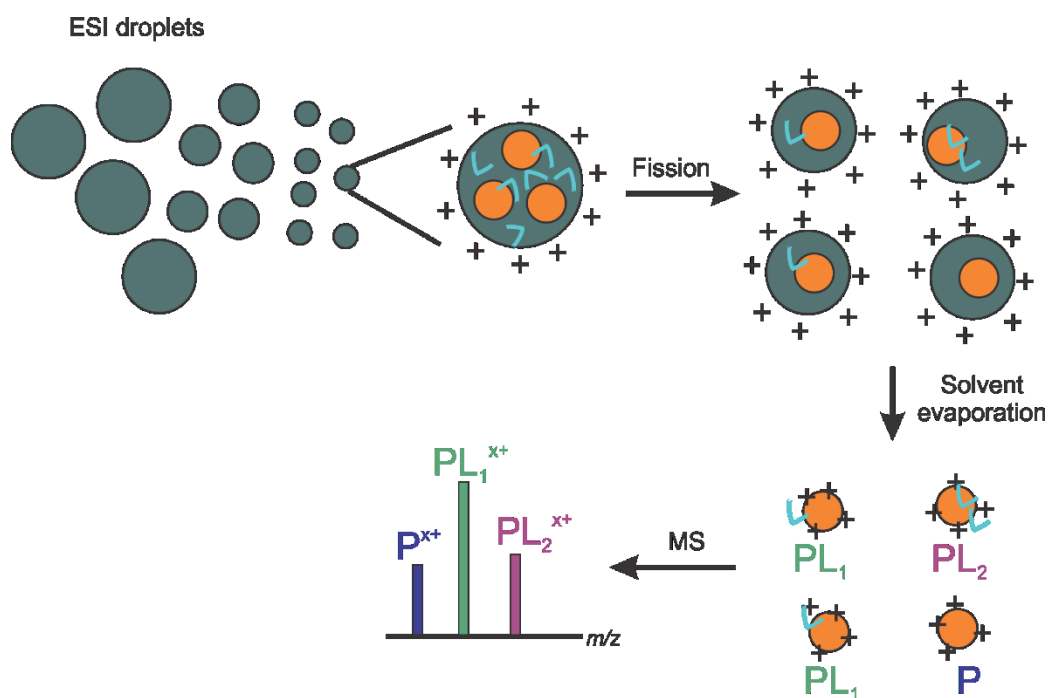
Figure 1.9. Illustration of the formation of nonspecific binding complexes during an ESI process under the positive ion mode, adapted from reference 54.

15

The factors affecting the level of nonspecific binding were discussed by Wang et al.[56] The size, structure and the charge state of the protein has a very minimal effect on the extent of nonspecific binding. In contrast, it overall depends on the concentration, structure (size) and hydrophobicity of the ligands. With a high ligand concentration, there is a higher chance of the protein encounters more ligand molecules in one nano-droplet, therefore, forming nonspecific complexes. Although by lowering the ligand concentration could reduce the artificial binding, it is almost unavoidable in weak binding measurements, because a large quantity of ligands is required to obtain a reasonable signal of protein-ligand complex.

For both direct and CaR-ESI-MS assays, a method of monitoring and measuring the level of nonspecific binding is essential. To correct the ESI mass spectra for the occurrence of nonspecific protein–ligand interactions in the ESI droplets, the reference protein method was used in this work. A reference protein (P_{ref}), which does not bind specifically to the target protein or any of the ligands, was introduced in the analyte solution.[47] The abundance ratio (R_{ref}) between P_{ref}-ligand to free P_{ref} can be used to monitor the level of nonspecific binding of ligand to a target protein in CaR-ESI-MS assay. In direct ESI-MS, the *Ab ratio* of apparent binding affinity (R_{app}) of protein–ligand can be corrected by Eqs 1.9–1.11:

$$R_{ref} = \frac{Ab(\mathrm{P_{ref}L})}{Ab(\mathrm{P_{ref}})} \tag{1.9}$$

$$R_{app} = \frac{Ab(\mathrm{PL})}{Ab(\mathrm{P})} \tag{1.10}$$

$$R = R_{app} - R_{ref} \tag{1.11}$$

An example of using the reference protein method to minimize false positives caused by nonspecific binding was showed by Sun et al.[57] The complex investigated was carbohydrate-binding antibody sigle-chain fragment (ScFv) and its ligands, a trisaccharides, with bovine

carbonic anhydrase II (CA) served as the reference protein. At 15 μM ligand concentration, clear nonspecific binding was observed. Without correction, the measured K_a were almost as twice high as the previous ITC reported affinity. With the application the reference protein correction, the K_a was in an excellent agreement with the ITC value.

1.5.2 Non-uniform response factors

The response factors play a critical role in the relationship between the abundances of P or PL measured by ESI-MS and the solution concentration (Eq 1.12): [54]

$$\frac{[PL]}{[P]} = \frac{RF_P.Ab(PL)}{RF_{PL}.Ab(P)} = RF_{PL/P}\frac{Ab(PL)}{Ab(P)} \tag{1.12}$$

where RF_P and RF_{PL} represent the response factor of P and PL, respectively. Only if the relative response factor ($RF_{PL/P}$) is approximately equal to one, do the abundances (Ab) measured represent the concentrations in solution correctly, which is essential for the direct ESI-MS assay. RF values depend on various factors, including the structure, size, surface properties of the PL and P, as well as the solution composition and instrumental parameters.[57–61] Due to the small size of the HMO ligands (MW <2000 Da), we assume that PL and P share similar size and surface properties.[54] Therefore, the interference caused by non-uniform response factors are limited.

1.5.3 In-source dissociation

For a gas phase protein–ligand complex ion, one of the major reasons of a false negative result comes from collision-induced dissociation in the ion source (i.e. in-source dissociation). If the level of in-source dissociation is high enough, no protein–ligand complex can be detected, or the abundance of the complex is reduced artificially, leading to unreliable K_a values.[60] A number of reasons contribute to this problem, including instrument parameters, the choice of ion source, and the size and gas phase stability of the complexes. Normally, the presence of in-source dissociation can be detected by the change of R (the ratio between PL complexes to P) with different ion source parameters.

Although it is challenging to remove this effect completely, there are several strategies to reduce it. A low temperature spray, a short accumulation time, or applying a low voltage give 'gentle' conditions, which lower the degree of dissociation. However, these approaches compromise the intensity of the ion signal. Therefore, a balance between minimizing in-source dissociation and obtaining an adequate ion signal must be achieved.

1.6 The Present Work

Chapter 2 and 3 focused on the interactions between HMOs and human galectins, which are potential targets for HMOs when they act as immune modulators. Chapter 2 describes work on developing a CaR-ESI-MS assay based method for identifying ligands of three human galectins (h-Gal1, h-Gal7, and h-Gal3C) from a purified HMO library. First, IMS arrival times (IMS-ATs) and CID fingerprints of a library containing 31 HMOs were obtained by ESI-IMS-MS. Next, the library was screening against three human galectins. Using molecular weight, IMS-ATs, and CID fingerprints, the HMO binders of target proteins were identified. In Chapter 3, the method was extended to screen natural HMO libraries that extracted from pooled human milk fractions. With the method established in Chapter 2, the compositions of the fractions were identified. Next, h-Gal3C was served as a model system and was screened against the fractions, and the HMO binders to Gal3C were identified. Overall, the method established a HMO library with multi-dimensional identification information, including molecular weight, IMS-ATs, and CID fragmentation patterns. Then, we used this information directly to identify the binders for target proteins, or to identify the HMO binders from human milk fractions.

Rotavirus is one of the major reason causes gastrointestinal disease in children, and studies showed that HMOs have inhibition effects in the infectivity of rotavirus. The goal of Chapter 4 is to reveal the binding pattern of the spike proteins of rotavirus from different strains to HMOs, using a direct ESI-MS assay. Seven HMO libraries, composed of 35 purified HMOs, were screened

against a rotavirus outer layer spike proteins from three strains. Using the direct ESI-MS assay, the affinity of each HMO ligand to these proteins can be determined. The binding affinities are generally weak ($K_a < 10^4 \, M^{-1}$), yet each protein has a different preference to the binding motif.

Chapter 5 is a summary this thesis, and it describes the future work that should be carried out and some preliminary data.

1.7 Literature cited

1. World Health Organization, Global strategy for infant and young child feeding. Geneva, Switzerland: World Health Organization and UNICEF, **2003**.

2. WHO Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality. Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. Lancet 2000, *355*, 451.

3. Horta, B.L.; Victora, C.G. Short-term effects of breastfeeding: a systematic review of the benefits of breastfeeding on diarrhea and pneumonia mortality. Geneva: World Health Organization, **2013**.

4. Horta, B.L.; de Mola, C.L.; Victora, C.G. *Acta Paediatr Suppl* **2015**, *104*, 30.

5. Bode L., *Glycobiology* **2012**, *22*, 1147.

6. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. *Annu. Rev. Nutr.* **2000**, *20*, 699.

7. Kumazaki, T.; Yoshiba, A. *Proc. Natl. Acad. Sci.* **1984**, *81*, 4193.

8. Johnson, P.H.; Watkins, W.M.; *Glycoconj. J.* **1992**, *9*, 241.

9. György, P.N.R.; Rose, C.S. *Arch Biochem. Biophys.* **1954**, *48*, 193.

10. Morelli, L.; *J Nutr.* **2008**, *138*, 1791S.

11. Angeloni, S.; Ridet, J.L.; Kusy, N.; Gao, H.; Crevoisier, F.; Guinchard, S. *Glycobiology* **2005**, *15*, 31.

12. Bode, L. *J. Nutr.* **2006**, *136*, 2127.

13. Wang, B.; McVeagh, P.; Petocz, P.; Brand-Miller, J. *Am. J. Clin. Nutr.* **2003**, *78*, 1024.

14. Kramer, M.S.; Aboud, F.; Mironova, E.; the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group *Arch. Gen. Psychiatry.* **2008**, *65*, 578.

15. Bode, L. *Nutr. Rev.* **2009**, *67*, S183.

16. Coppa, G.V.; Zampini, L.; Galeazzi, T.; Facinelli, B.; Ferrante, L.; Capretti, R. *Pediatr. Res.* **2006**, *59*, 377.

17. Weichert, S. *J. Virol.* **2016**, *90*, 4843.

18. Wu, S.; Tao, N.; German, J.B.; Grimm, R.; Lebrilla, C.B. *J. Proteome Res.* **2010**, *9*, 4138.

19. Pierce, M.M.; Raman, C.S.; Nall, B.T. *Methods* **1999**, *19*, 213.

20. Perret, S.; Sabin, C.; Dumon, C. *Biochem. J.* **2005**, *389*, 325.

21. Utsuno, K.; Uludag, H. *Biophys. J.* **2010**, *99*, 201.

22. Koromyslova, A.; Tripathi, S.; Morozov, V.; Schroten, H.; Hansman, G.S. *Virology* **2017**, *508*, 81.

23. Satoh, A.; Kojima, K.; Koyama, T.; Ogawa, H.; Matsumoto, I. *Anal. Biochem.* **1998**, *260*, 96.

24. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Mickum, M.L.; Ashline, D.J.; Prasad, B.V.; Estes, M.K.; Reinhold, V.N.; Cummings, R.D.; Smith, D.F. *Mol. Cell Proteomics* **2014**, *13*, 2944.

25. Song, X.; Lasanajak, Y.; Xia, B.; Smith, D.F.; Cummings, R.D. *ACS Chem. Biol.* **2009**, *4*, 741.

26. Ashline, D.J.; Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Prasad, V.; Estes, M.K.; Cummings, R.D.; Smith, D.F.; Reinhold, V.N. *Mol. Cell Proteomics* **2014**, *13*, 2961.

27. Shams-Ud-Doha, K.; Kitova, E.N.; Kitov, P.I.; St-Pierre, Y.; Klassen, J.S. *Anal. Chem.* **2017**, *89*, 4914.

28. Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818.

29. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605.

30. Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M. *Science* **1989**, *246*, 64.

31. Karas, M.; Bahr, U.; Dulcks, T. *J. Anal. Chem.* **2000**, *366*, 669.

32. Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.***1999***, 10*, 300.

33. Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A.

34. Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804.

35. Kebarle, P.; Verkerk, U.H. *Mass Spectrom. Rev.* **2009**, *28*, 898.

36. Fenn, J.B. *Angew Chem Int Ed Engl.* **2003**, *42*, 3871.

37. Wu, X.; Oleschuk, R.D.; Cann, N.M. *The Analyst* **2012**, *137*, 4150.

38. Jecklin, M. C.; Touboul, D.; Jain, R.; Toole, E. N.; Tallarico, J.; Drueckes, P.; Ramage, P.; Zenobi, R. *Anal. Chem.* **2009**, *81*, 408.

39. Karas, M.; Bahr, U.; Dulcks, T.; Fresenius J. *Anal. Chem.* **2000**, *366*, 669

40. Konermann, L.; Ahadi, E.; Rodriguez, A.D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2.

41. Iavarone, A.T.; Williams, E.R. *J. Am. Chem. Soc.* **2003**, *125*, 2319.

42. Dole, M.; Mark, L.L.; Hines, R.L.; Mobley, R.C.; Ferguson, L.D.; Alice, M.B. *J. Chem. Phys.* **1968**, *49*, 2240.

43. Batey, J.H. *Vacuum* **2014**, *101*, 410.

44. De Hoffmann, E.; Stroobant, V. Mass Spectrometry Principles and Applications 3th Ed. John Wiley & Sons.: New York, **2007**.

45. Quadrupole mass analyzers: theoretical and practical considerations. In Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics.

46. Scarff, C.A.; Patel, V.J.; Thalassinos, K.; Scrivens, J.H. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 625.

47. Giles, K.; Pringle, S.D.; Worthington, K.R.; Little, D.; Wildgoose, J.L.; Bateman, R.H. *Rapid Comm. Mass Spectrom.* **2004**, *18*, 2401.

48. Wu, L.; Vogt, F.G. *J. Pharm. Biomed. Anal.* **2012**, *69*, 133.

49. Zandkarimi, F.; Wickramasekara, S.; Morré, J.; Stevens, J.F.; Maier, C.S. Electrospray Ionization Traveling Wave Ion Mobility Spectrometry Mass Spectrometry for the Analysis

of Plant Phenolics: An Approach for Separation of Regioisomers. D.R. Gang Ed.; Springer Cham: Switzerland **2013**. 21.

50. McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 599.

51. Wiley, W.C.; McLaren, L.H. *Rev. Sci. Instrum.* **1955**, *26*, 1150.

52. Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrom. Rev.* **2000**, *19*, 65.

53. Mamyrin, B.A.; Karataev, V.I.; Shmikk, D.V.; Zagulin, V.A. *Sov. Phys. JETP.* **1973**, *37*, 45.

54. Kitova, E.N.; El-Hawiet, A.; Schnier, P.D.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.

55. El-Hawiet, A.; Shoemaker, G.K.; Daneshfar, R.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2012**, *84*, 50.

56. Wang, W.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2005**, *77*, 3060.

57. Sun, J.; Kitova, E.N.; Wang, W.; Klassen, J.S. *Anal. Chem.* **2006**, *78*, 3010.

58. Jecklin, M.C.; Touboul, D.; Bovet, C.; Wortmann, A.; Zenobi, R. *J. Am Soc. Mass Spectrom.* **2008**, *19*, 332.

59. Yu, Y.; Kirkup, C.E.; Pi, N.; Leary, J.A. *J. Am Soc. Mass Spectrom.* **2004**, *15*, 1400.

60. Sun, J.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2007**, *79*, 416.

61. Bagal, D.; Kitova, E.N.; Liu, L.; El-Hawiet, A.; Schnier, P.D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801.

Chapter 2

High-Throughput, Label- and Immobilization-Free Screening of Human Milk Oligosaccharides Against Lectins*

2.1 Introduction

Oligosaccharides represent the third largest solid component in human milk.[1] Human milk oligosaccharides (HMOs) are unique among oligosaccharides found in mammalian milk, not only because of their complex structures and high concentrations (7 g L^{-1} to 23 g L^{-1}), but also because of their diverse beneficial health effects.[2,3] For example, they act as prebiotics,[4] intestinal epithelial cell and immune system modulators,[5] essential nutrients for brain development,[6] antiadhesive antimicrobial agents,[7-9] and protective agents against necrotizing enterocolitis.[10] HMOs are composed of five monosaccharide building blocks: D-glucose (Glc), D-galactose (Gal), N-acetylglucosamine (GlcNAc), L-fucose (Fuc), and N-acetylneuraminic acid (NeuAc), and possess lactose (Lac, β-Gal-(1→4)-Glc) at the reducing end.[11] Enzymatic elongation of the HMO structure involves the addition of GlcNAc to Gal through β1→3 or β1→6 linkages, followed by Gal addition through β1→3 or β1→4 linkages. Further decoration with lactosamine (LacNAc, β-Gal-(1→4)-GlcNAc), Fuc in α-(1→2), α-(1→3), or α-(1→4) linkages, and NeuAc in α-(2→3) or α-(2→6) linkages, results in a large number of structural isomers.[12] To date, more than 200 different HMOs have been detected,[13,14] with fucosylated, sialylated, and non-fucosylated neutral oligosaccharides representing (by mass) 35–50%, 12–14%, and 42–55%, respectively, of HMOs.[14-16] Approximately 50 structures account for ~99% (by mass) of the HMOs present in human milk.[14]

Elucidating the molecular interactions between HMOs and their protein receptors is essential to understanding their diverse regulatory and protective roles.[3,14,17] Given that HMOs exist naturally as free oligosaccharides, that the most abundant HMOs found in milk are relatively small (ranging from tri- to pentasaccharide), and that the reducing end (terminal lactose moiety) is often implicated in binding, the lectin binding properties of HMOs should be carried under equilibrium conditions using the free oligosaccharides, without any labeling (for detection) or derivatization (for immobilization).[18,19] The use of unlabeled lectins is also preferable as the introduction of affinity tags or chromophores may also influence HMO binding.

A number of binding assays, including frontal affinity chromatography and surface plasmon resonance spectroscopy, have been used to detect and quantify individual lectin interactions with modified HMOs.[18,20] Applications of these methods to free, unmodified HMOs also have been reported, although these measurements were done using a competitive binding strategy.[20,21] Currently, human milk glycan (HMG) microarrays, including the shotgun and defined HMG arrays, wherein HMOs are immobilized on a glass slide through their terminal Glc residue, represent the dominant technology for high-throughput screening of HMOs against lectins.[22,23] The most recent version of the shotgun HMG (HM-SGM-v2), which consists of 247 oligosaccharide fractions prepared from pooled human milk,[21] has been used to evaluate the HMO-binding specificities of a number of human, bacterial and viral lectins.[21-25] Although the HMG microarrays represent a convenient and high-thoroughput approach to evaluating HMO specificities, they have several well-known limitations, such as the influence of the chemical modification of the reducing residue Glc required to immobilize the oligosaccharide on the surface and the length and nature of the linker used on lectin binding.[18,19,25,26] Additionally, because of the washing steps involved, many low affinity interactions are missed.[28-30] Notably, a recent comparison of the trends in HMO affinities measured using a ESI-MS method for a series of human

galectins (hGal) and results obtained with the HM-SGM-v2 microarray showed poor agreement.[31] It was found also that derivatization of the Glc residue (required for immobilization) of HMOs altered both their absolute and relative affinities. This finding suggests that the discrepancies between HMO specificities, as determined by HM-SGM-v2 microarray data, and relative affinities are due, at least in part, to the chemical modification of the HMOs.[31]

At present, isothermal titration calorimetry (ITC) and electrospray ionization mass spectrometry (ESI-MS) represent the most widely used assays for quantifying lectin interactions with free (unmodified) HMOs.[21,32,33] The ITC assay has the unique advantage that it measures the enthalpy of association directly and the affinity indirectly. However, it requires relatively large quantities of protein and HMO (~mg of each), is not well-suited to low affinity ($<10^3$ M^{-1}) interactions or interactions that produce small changes in heat, and cannot be applied to mixtures of HMOs.[34] In the ESI-MS assay, affinities are determined from the relative abundances of free and ligand-bound protein ions measured in the mass spectrum.[35,36] Given the speed, the absence of any labeling or immobilization requirements, and the ability to study multiple equilibria simultaneously, the ESI-MS assay is well-suited for HMO library screening.[37,38] However, the prevalence of structural isomers, which cannot be distinguished based on molecular weight (MW) alone, represents a significant challenge to screening (simultaneously) HMO mixtures by ESI-MS and one that requires additional dimensions of analysis (e.g. ion mobility separation (IMS) and collision-induced dissociation (CID)) to overcome. [39-41]

Here, we describe a high-throughput and label- and immobilization-free ESI-MS assay for screening libraries of free HMOs (reducing sugars) against lectins *in vitro*. The method is based on catch-and-release (CaR)-ESI-MS, whereby ligands are identified following their release, as ions, from lectin-HMO complexes upon collisional activation in the gas phase, and allows for the simultaneous screening of a mixture of HMOs against one or more target lectins.[35-37] A key feature

of the assay is the use of IMS arrival times (ATs) and CID fingerprinting to distinguish between all structural isomers present in the library. To demonstrate the implementation of the assay, a library of 31 HMOs, comprised of the most abundant oligosaccharides found in human milk, was screened against three hGal proteins, which are expressed and secreted from intestinal epithelial cells.[42] Comparison of the CaR-ESI-MS results with recently reported affinity data served to establish the reliability of the assay.[31]

2.2 Experimental

2.2.1 Proteins

The C-terminal fragment (residues 107–250) of human galectin-3 (hGal-3C, MW 16 330 Da) was a gift from Prof. C. Cairo (University of Alberta), S-carboxyamidomethylated oxidation resistant (C2S substituted to improve stability) recombinant hGal-1 (dimer MW 29 235 Da) was a gift from S. Sato (Laval University), and recombinant hGal-7 (dimer MW 29 888 Da) was produced and purified as described previously.[43] Ubiquitin (MW 8 565 Da), which served as the reference protein (P_{ref}), was purchased from Sigma-Aldrich Canada (Oakville, Canada). To produce protein stock solutions, a solution containing ~0.5 mg of protein was dialyzed against 200 mM aqueous ammonium acetate (pH 6.8) using a 10 kDa (for the hGal proteins) or 5 kDa (for ubiquitin) MW cut-off Amicon Ultra-4 centrifugal filter (Millipore Corp, Bedford, MA); the final concentration was adjusted to 200 μM. All stock solutions were stored at -20°C until used.

2.2.2 Human milk oligosaccharides

The structures of the HMOs (**L1** - **L31**) are shown in Figure 2.1 and listed in Table 2.1.
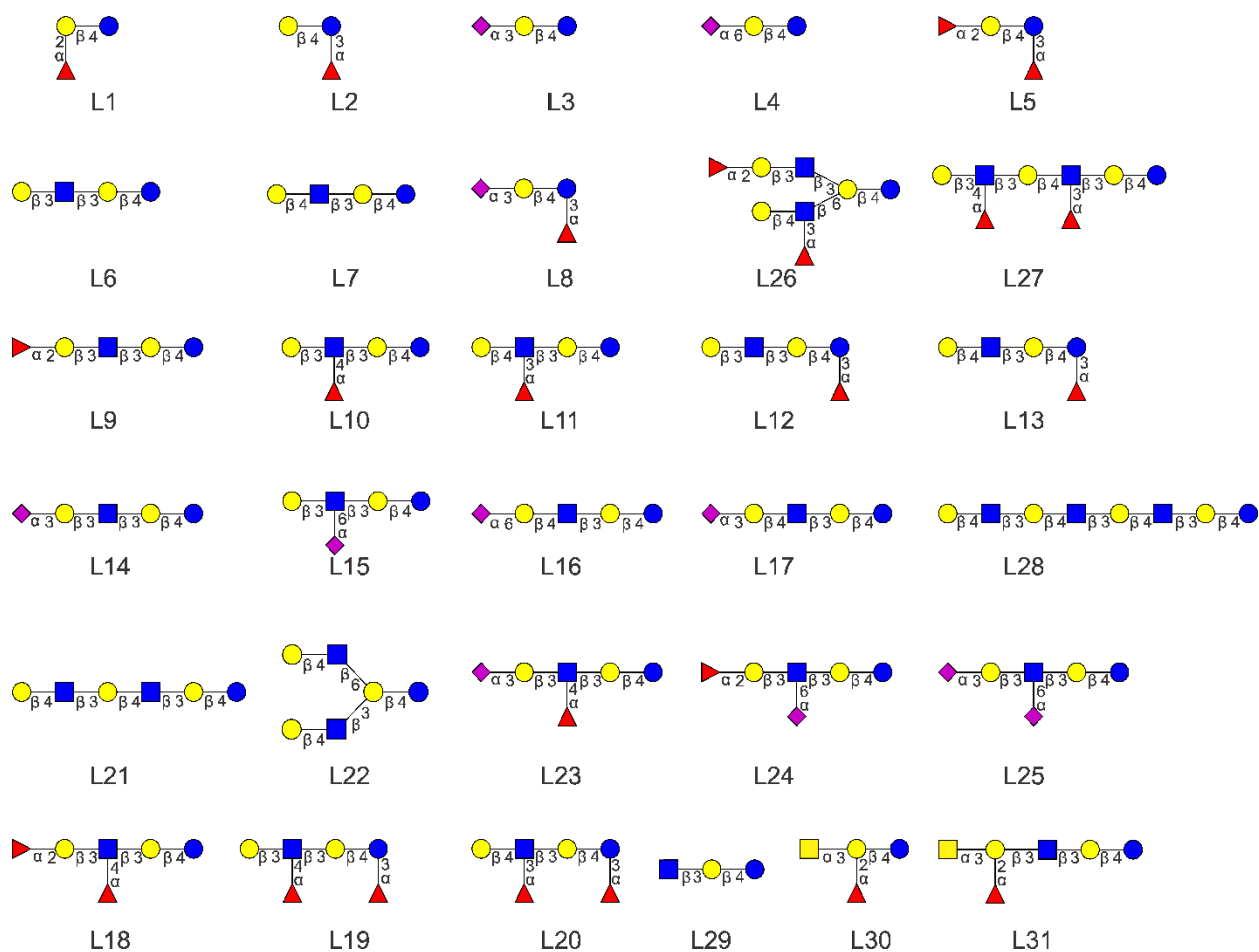
Figure 2.1. Structures of the HMOs (**L1–L31**) used in this study. Monosaccharide key: glucose (⬤), galactose (◯), N-acetylglucosamine (■), sialic acid (◆), fucose (▲).

Table 2.1. List of HMOs (**L1–L31**) used in this study, their MWs, chemical structures, and common names

HMO	MW (Da)	Structure	Common name
L1	488.17	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc	2'-Fucosyllactose
L2	488.17	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	3-Fucosyllactose
L3	633.21	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-Glc	3'-Sialyllactose

L4	633.21	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-Glc	6'-Sialyllactose
L5	634.23	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	Difucosyllactose
L6	707.25	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-tetraose
L7	707.25	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-neotetraose
L8	779.27	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	3'-Sialyl-3'-fucosyllactose
L9	853.31	α- L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-fucopentaose I
L10	853.31	β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-fucopentaose II
L11	853.31	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-fucopentaose III
L12	853.31	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	Lacto-N-neofucopentaose V
L13	853.31	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)[α-L-Fuc-(1→3)]-β-D-Glc	Lacto-N-neofucopentaose
L14	998.34	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyllacto-N-tetraose a
L15	998.34	α-D-Neu5Ac-(2→6)-[β-D-Gal-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- β-D-Glc	Sialyllacto-N-tetraose b
L16	998.34	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyllacto-N-tetraose c
L17	998.34	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyllacto-N-tetraose d
L18	999.36	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-difucohexaose I
L19	999.36	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-Glc	Lacto-N-difucohexaose II
L20	999.36	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α- L-Fuc-(1→3)]-β-D-Glc	Lacto-N-neodifucohexaose
L21	1072.38	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Para Lacto-N-neohexaose
L22	1072.38	β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-Glc	Lacto-N-neohexaose
L23	1144.40	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyl monofucosyll

			acto-N-tetraose
L24	1144.40	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyl-lacto-N-fucopentaose V
L25	1289.44	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Disialyllacto-N-tetraose
L26	1364.50	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc	Difucosyllacto-N-hexaose a
L27	1364.50	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-(1→3)]-β-D-Gal-(1→4)-β-D-Glc	Difucosyl-para-lacto-N-hexaose
L28	1438.29	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal(1→4)-β-D-GlcNAc (1→3)-β-D-Gal-(1→4)-β-D-GlcNAc(1→3)-β-D-Gal(1→4)-β-D-Glc	Lacto-N-neooctaose
L29	545.48	β-D-GlcNAc-(1→3)-β-D-Gal(1→4)-β-D-Glc	Lacto-N-triaose
L30	691.62	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc	Blood group A antigen tetraose type 5
L31	1056.96	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-GlcNAc(1→3)-β-D-Gal-(1→4)-β-D-Glc	Blood group A antigen hexaose type 1

L1 (MW 488.17 Da), **L2** (MW 488.17 Da), **L12** (MW 853.31 Da), **L13** (MW 853.31 Da), **L17** (MW 998.34 Da), **L20** (MW 999.36 Da), **L21** (MW 1072.38 Da), **L23** (MW 1144.40 Da), **L24** (MW 1144.40 Da), **L27** (MW 1364.50 Da) **L28** (MW 1437.36 Da), **L29** (MW 545.50 Da), **L30** (MW 691.62 Da), and **L31** (MW 1056.32 Da) were purchased from Elicityl SA (Crolles, France); **L3** (MW 633.21 Da), **L4** (MW 633.21 Da), **L5** (MW 634.23 Da)**, L6** (MW 707.25 Da), **L7** (MW 707.25 Da), **L8** (MW 779.27 Da), **L9** (MW 853.31 Da), **L10** (MW 853.31 Da), **L11** (MW 853.31 Da), **L14** (MW 998.34 Da), **L15** (MW 998.34 Da), **L16** (MW 998.34 Da)**, L18** (MW 999.36 Da), and **L26** (MW 1364.50 Da) were purchased from IsoSep (Tullinge, Sweden); **L19** (MW 999.36 Da) and **L22** (MW 1072.38 Da) from Dextra (Reading, UK); **L25** (MW 1289.44 Da) was purchased from CarboSynth (Compton, UK). Stock solutions of the HMOs were prepared by

dissolving 0.35–0.50 mg of solid HMO in deionized water to give a final concentration of 10 mM. All stock solutions were stored at -20°C until used.

2.2.3 Mass spectrometry

All experiments were carried out using a Synapt G2 ESI quadrupole-ion mobility separation-time-of-flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK), equipped with a nanoflow ESI (nanoESI) source. Mass spectra were obtained in negative ion mode using cesium iodide (1 mg mL^{-1}) for calibration. The nanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.78 mm i.d.) pulled to ~5 μm outer-diameter using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). Approximately 5 μL of a sample solution, containing 15 μM or 30 μM hGal (15 μM for hGal-3C, 30 μM for hGal-1 or hGal-7) and 3 μM of each HMO (**L1**–**L31**), was loaded into the nanoESI tip. To perform ESI, a platinum wire was inserted into the solution and a voltage of ~-1.0 kV was applied. A cone voltage of 25 V was used, and the source block temperature was maintained at 60 ºC. Trap voltages, ranging from 10 to 80 V, and Transfer voltages, ranging from 10 to 80 V, were used to carry out CID; Argon was used in the Trap and Transfer ion guides at pressures of 2.22 x 10^{-2} and 3.36 x 10^{-2} mbar, respectively. The helium chamber preceding the traveling wave ion mobility separation (TWIMS) device was maintained at 7.72 mbar. The IMS parameters, optimized for each HMO isomer set, were: 2 mL min^{-1} Trap gas flow rate; 150 to 180 mL min^{-1} helium cell gas flow rate; 50 to 90 mL min^{-1} ion mobility gas flow rate; 50 V Trap direct-current bias; 400 to 1000 m s^{-1} ion mobility wave velocity; 15 to 40 V ion mobility wave height. All IMS measurements were carried out using nitrogen as the mobility gas, at a pressure of 3.41 mbar. Data acquisition and processing were carried out using MassLynx (v4.1).

2.3 Results and Discussion

2.3.1 HMO library characterization - IMS arrival times and CID fingerprinting

The HMO library used in the present study includes nine sets of isomers (encompassing a total of 24 HMOs), with each isomer set containing between two and five HMOs. To be able to screen the entire library against a lectin with CaR-ESI-MS simultaneously, all possible ligands, including structural isomers, must be identifiable following their release (as ions) from the protein in the gas phase. As described below, all components of the HMO library can be simultaneously identified by combining measurements of MWs, IMS-ATs, and CID fingerprints.

The IMS measurements were performed initially on individual HMO anions. With the exception of **L25**, which is detected only as the doubly deprotonated species, the HMOs in the library were detected either as singly deprotonated ions alone or as both singly and doubly deprotonated species in the negative ion mode ESI-MS (Figure 2.2). Shown in Figure 2.3 are representative IMS arrival time distributions (ATDs) measured for singly deprotonated (**L1**–**L24** and **L26**–**L31**) and doubly deprotonated (**L25**) HMO ions.
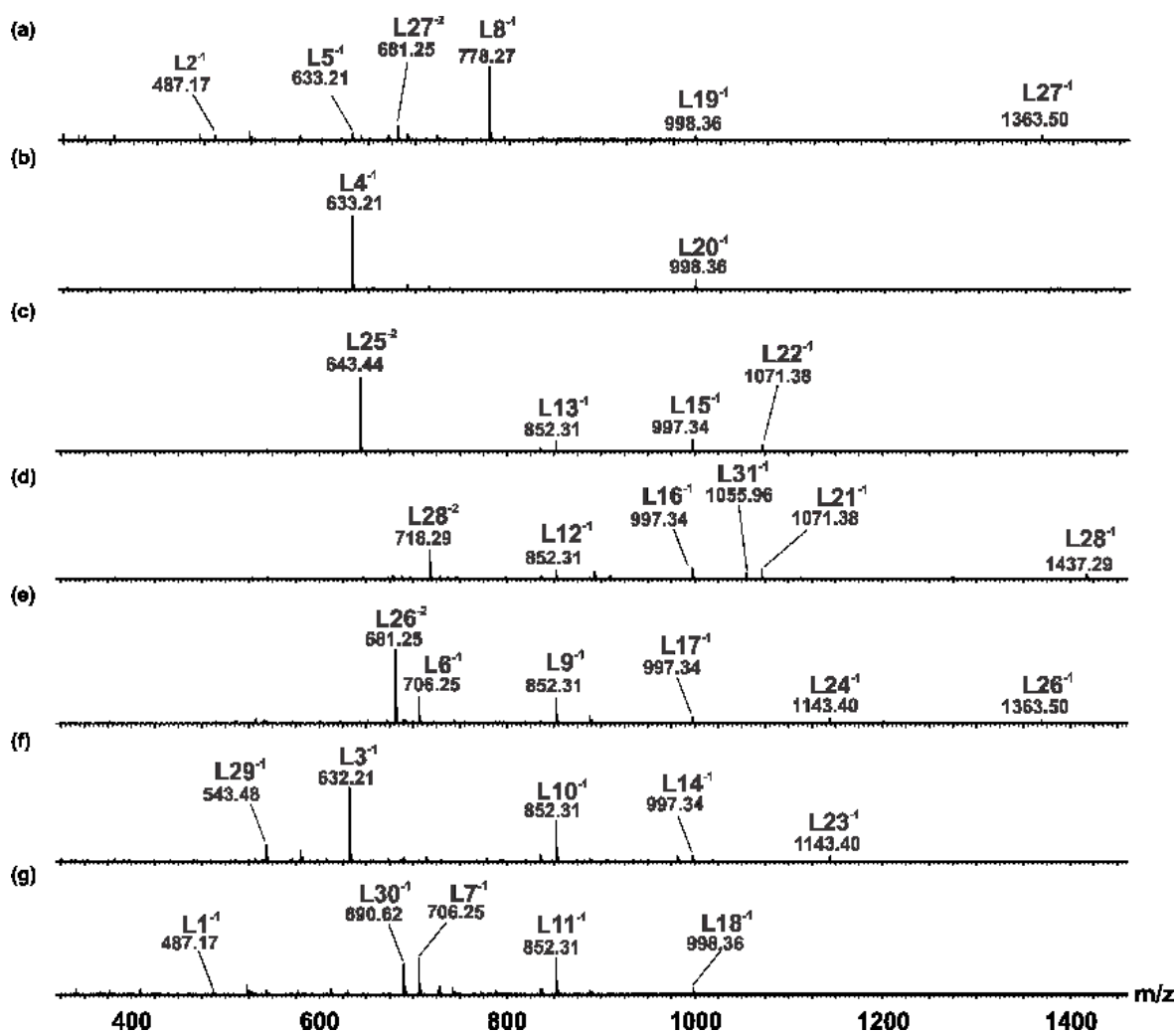
Figure 2.2. Representative ESI mass spectra acquired in the negative ion mode for aqueous ammonium acetate solutions (20 mM, pH 6.8) containing equimolar concentrations (2.5 μM) of (a) **L2**, **L5**, **L8**, **L19,** and **L27**; (b) **L4** and **L20**; (c) **L13**, **L15**, **L22,** and **L25**; (d) **L12**, **L16**, **L21**, **L28,** and **L31**; (e) **L6**, **L9**, **L17**, **L24,** and **L26**; (f) **L3**, **L10**, **L14**, **L23,** and **L29**; and (g) **L1**, **L7**, **L11**, **L18,** and **L30**.
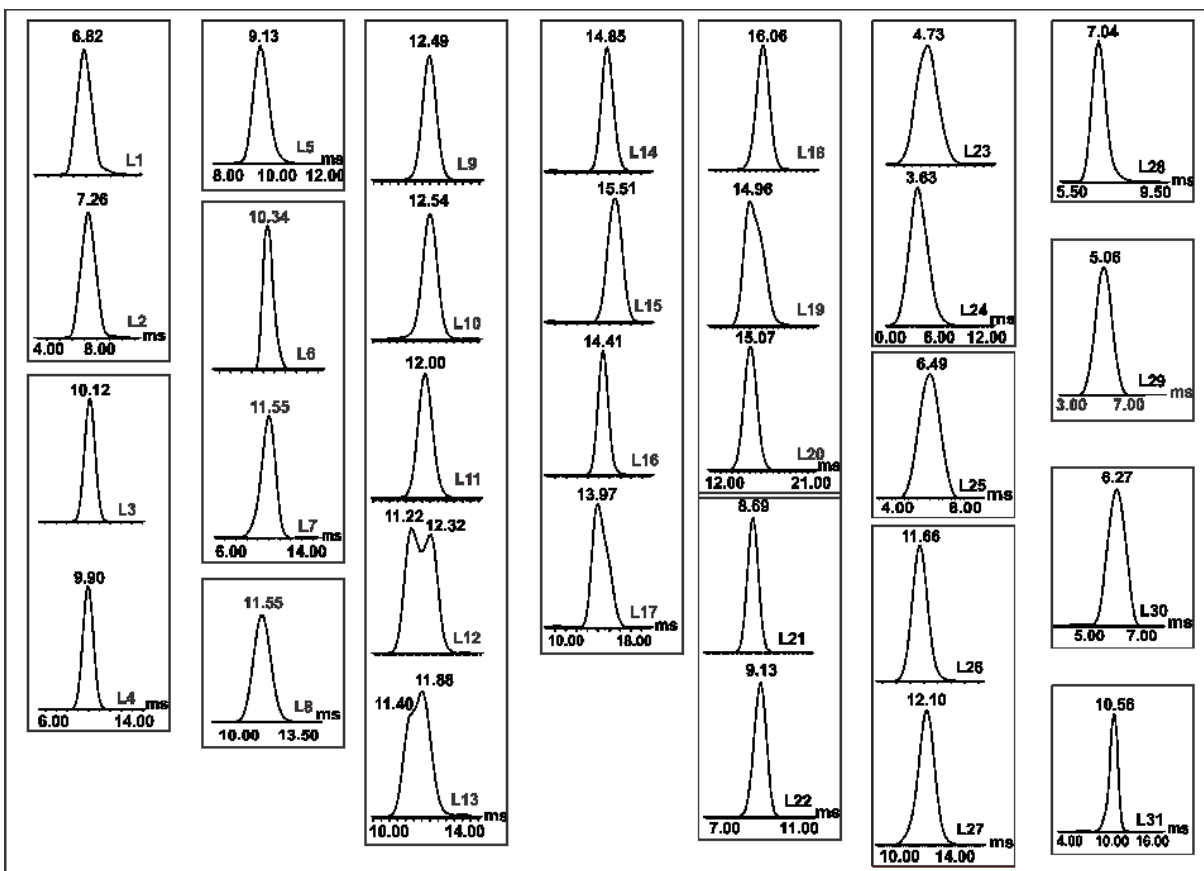
Figure 2.3. IMS-ATDs measured for the singly (**L1**, **L2**, **L3**, **L4**, **L5**, **L6**, **L7**, **L8, L9**, **L10**, **L11**, **L12**, **L13**, **L14**, **L15**, **L16, L17**, **L18**, **L19**, **L20**, **L21**, **L22**, **L23**, **L24, L26**, **L27**, **L28**, **L29**, **L30**, and **L31)** and doubly (**L25**) deprotonated HMO ions.

The IMS parameters were optimized to produce the largest difference in ATs within each isomer set (Table 2.2).

Table 2.2. IMS wave velocities (m/s) and wave heights (V) used to analyze the HMO anions

HMOs	Wave Velocity (m/s)	Wave Height (V)
L1/L2	1000	25
L3/L4	600	15
L5	600	15
L6/L7	1000	25
L8	1000	25
L9-L13	400	16

L14-L17	1000	25
L18-L20	1000	25
L21/22	600	25
L23/24	600	15
L25	875	25
L26/27	1000	35
L28	1000	25
L29	600	40
L30	400	25
L31	800	40

Inspection of the IMS data revealed that, with the exception of the **L12** and **L13** ions, the deprotonated HMO ions exhibit IMS-ATDs that are characterized by a single peak (Figure 2.3), with a full width at half maximum (FWHM) ranging from 0.4 to 1.5 ms. In the case of **L12** and **L13**, the IMS-ATDs exhibit two partially-resolved features, suggestive of the presence of at least two different stable conformations.[44,45] Within a given isomer set, the smallest difference in IMS-ATs is 0.05 ms, while the uncertainty in the IMS-ATs, established from replicated measurements (>10 measurements) performed on different days, is <0.01 ms.

The aforementioned IMS results, which represent the most comprehensive IMS-MS data reported to date for HMOs, demonstrate that each HMO within a given isomer set exhibits a unique IMS-AT.[16,41,46] However, because lectins generally can recognize multiple isomeric oligosaccharides, it was necessary to establish to what extent HMO isomers present as mixtures could be identified by IMS alone. With this in mind, IMS measurements were performed on all possible (47) combinations of HMO isomers (Table 2.3).

Table 2.3. List of all possible combinations of structural isomers in the HMO library

HMOs	**Number of combinations**	**Combinations**

L1/L2	1	**L1+L2**
L3/L4	1	**L3+L4**
L6/L7	1	**L6+L7**
L9-L13	26	**L9+L10, L9+L11, L9+L12, L9+L13, L10+L11, L10+L12, L10+L13, L11+L12, L11+L13, L12+L13, L9+L10+L11, L9+L10+L12, L9+L10+L13, L9+L11+L13, L10+L11+L12, L10+L11+L13, L11+L12+L13, L11+L12+L9, L12+L13+L9, L12+L13+L10, L9+L10+L11+L12, L9+L10+L11+L13, L9+L10+L12+L13, L10+L11+L12+L13, L9+L11+L12+L13, L9+L10+L11+L12+L13**
L14-L17	11	**L14+L15, L14+L16, L14+L17, L15+L16, L15+L17, L16+L17, L14+L15+L16, L14+L15+L17, L14+L15+L17, L15+L16+L17, L14+L15+L16+L17**
L18-L20	4	**L18+L19, L18+L20, L19+L20, L18+L19+L20**
L21/L22	1	**L21+L22**
L23/L24	1	**L23+L24**
L26/L27	1	**L26+L27**

Representative IMS-ATDs produced from equimolar concentration solutions are shown in Figures 2.4–2.6.
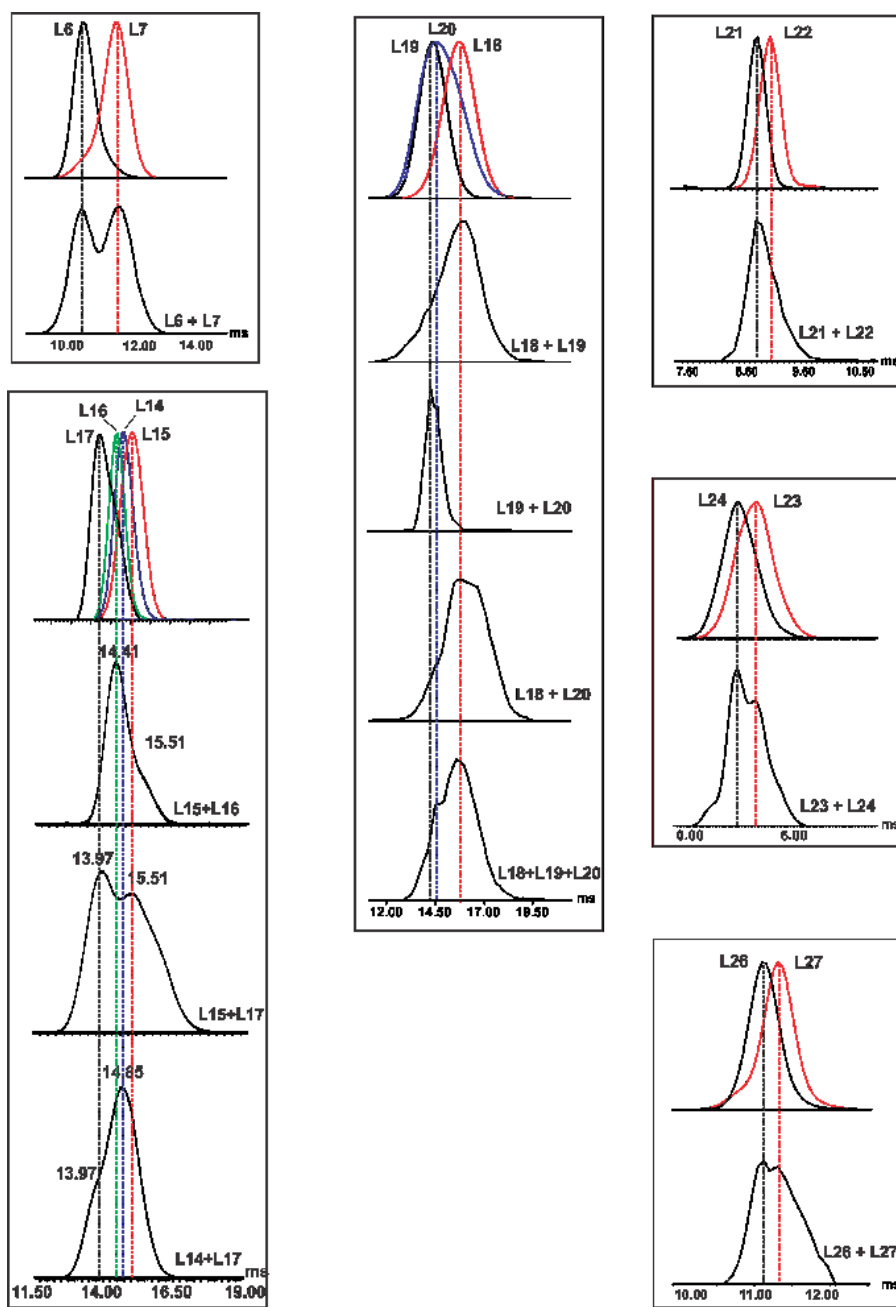
Figure 2.4. IMS-ATDs measured for combinations of singly deprotonated isomeric HMO ions: **L6/L7, L14/L17, L15/L16, L15/L17, L18/L19, L18/L20, L19/L20**, **L18/L19/L20, L21/L22**, **L23**/**L24,** and **L26/L27**. The IMS parameters used for these measurements are listed in Table 2.2.
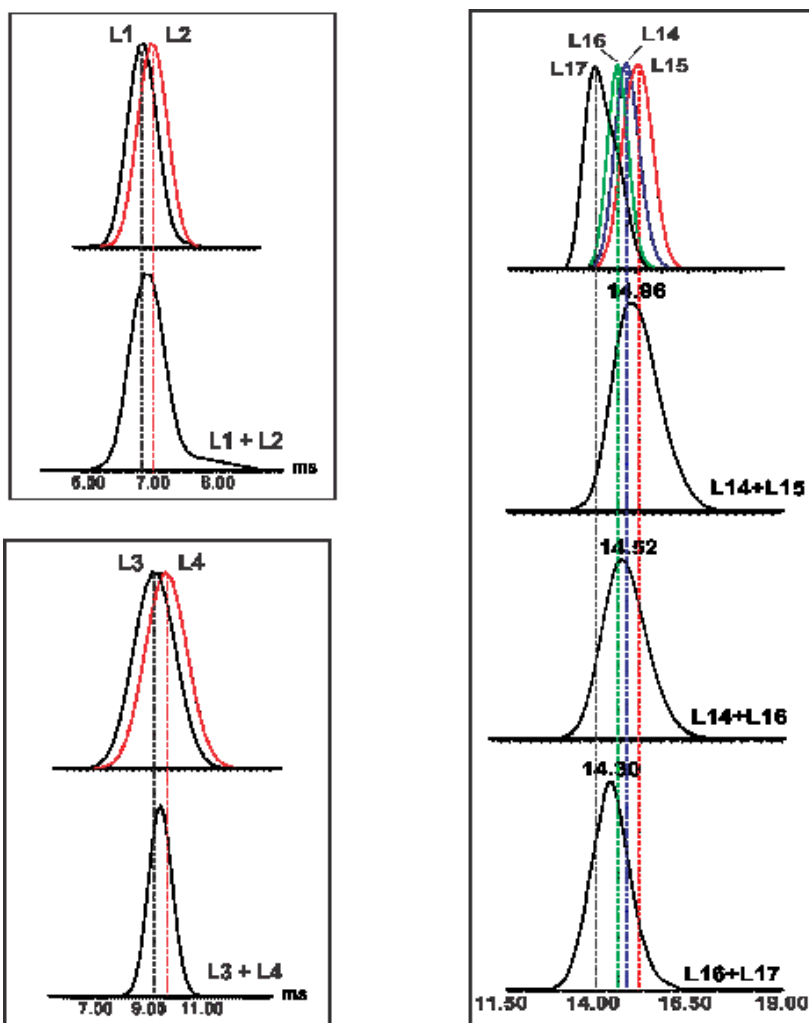
Figure 2.5. IMS-ATDs measured for combinations of singly deprotonated isomeric HMO ions: **L1/L2**, **L3/L4**, **L14/L15**, **L14/L16,** and **L16/L17**. The IMS parameters used for these measurements are listed in Table 2.2.
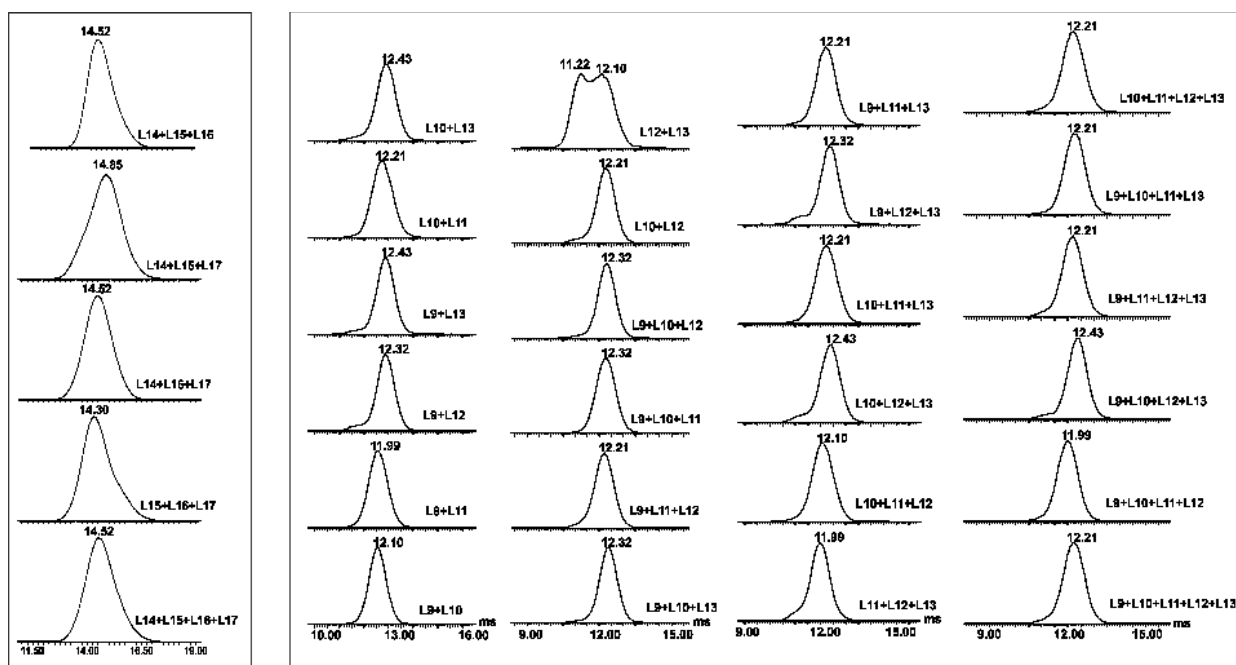
Figure 2.6. IMS-ATDs measured for combinations of singly deprotonated isomeric HMO ions: **L14/L15/L16**, **L14/L15/L17**, **L14/L16/L17**, **L15/L16/L17, L14/L15/L16/L17,** and all possible combination of the **L9**–**L13** isomer set. The IMS parameters used for these measurements are listed in Table 2.2.

For some combinations of isomers (**L6/L7**, **L14/L17**, **L15/L16**, **L15/L17**, **L18/L19**, **L18/L20**, **L19/L20**, **L18/L19/L20**, **L21/L22**, **L23/L24,** and **L26/L27**), the HMOs (when present at similar abundances in the gas phase) can be identified based on the partially resolved IMS-ATDs (Figure 2.4). For example, the IMS-ATD of the **L6/L7** mixture exhibits two features, with maxima (10.34 and 11.55 ms) that match the IMS-ATs measured for **L6** and **L7** individually (Figure 2.3). For some pairs of isomers (**L1/L2**, **L3/L4**, **L14/L15**, **L14/L16,** and **L16/L17**), the measured IMS-ATDs appear as a single feature with an AT corresponding approximately to the average value of the IMS-ATs for the individual isomers (Figure 2.5). For example, the ATs measured for **L3** and **L4**, individually, are 10.12 and 9.90 ms, respectively, while the AT measured when both isomers are present is 10.01 ms. Similar ATs were observed for mixtures of **L3** and **L4** with concentration

ratios between ~1:2 and ~2:1 (Figure 2.7).
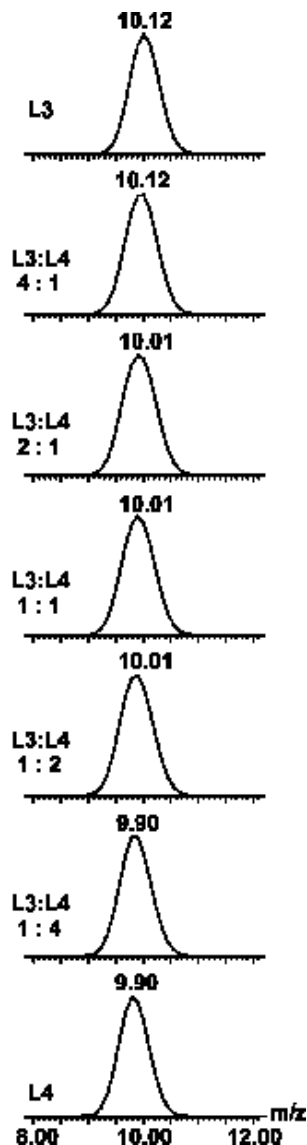
Figure 2.7. IMS-ATDs measured for the singly deprotonated ions of **L3** and **L4,** alone and as a mixture at the molar ratios indicated. The IMS parameters used for these measurements are listed in Table 2.2.

However, at lower or higher concentration ratios, the measured ATs correspond to that of the higher concentration (most abundant) HMO. For the **L14/L15/L16**, **L14/L15/L17**, **L14/L16/L17, L15/L16/L17,** and **L14/L15/L16/L17** combinations, as well as all combinations of **L9–L13**, it

was not possible to identify conclusively which isomers were present from the IMS-ATDs alone (Figure 2.6). For example, an AT of 12.21 ms was measured for nine different isomer combinations (**L10/L11**, **L10/L12**, **L9/11/L13**, **L9/L11/L12**, **L10/L11/L13**, **L10/L11/L12/L13**, **L9/L10/L11/L13**, **L9/L11/L12/L13,** and **L9/L10/L11/L12/L13**, Figure 2.6).

This analysis reveals that, when present at similar abundances in the gas phase, it is possible to identify isomers present in 16 of the 47 possible isomer combinations (Table 2.3) from IMS alone. For the other isomer combinations, a CID fingerprinting step was introduced into the workflow. As illustrated in Figures 2.8–2.11, CID of the deprotonated HMO ions, using optimized collision energies, produces unique fragment ions that serve to distinguish between isomers that could not be identified by IMS.
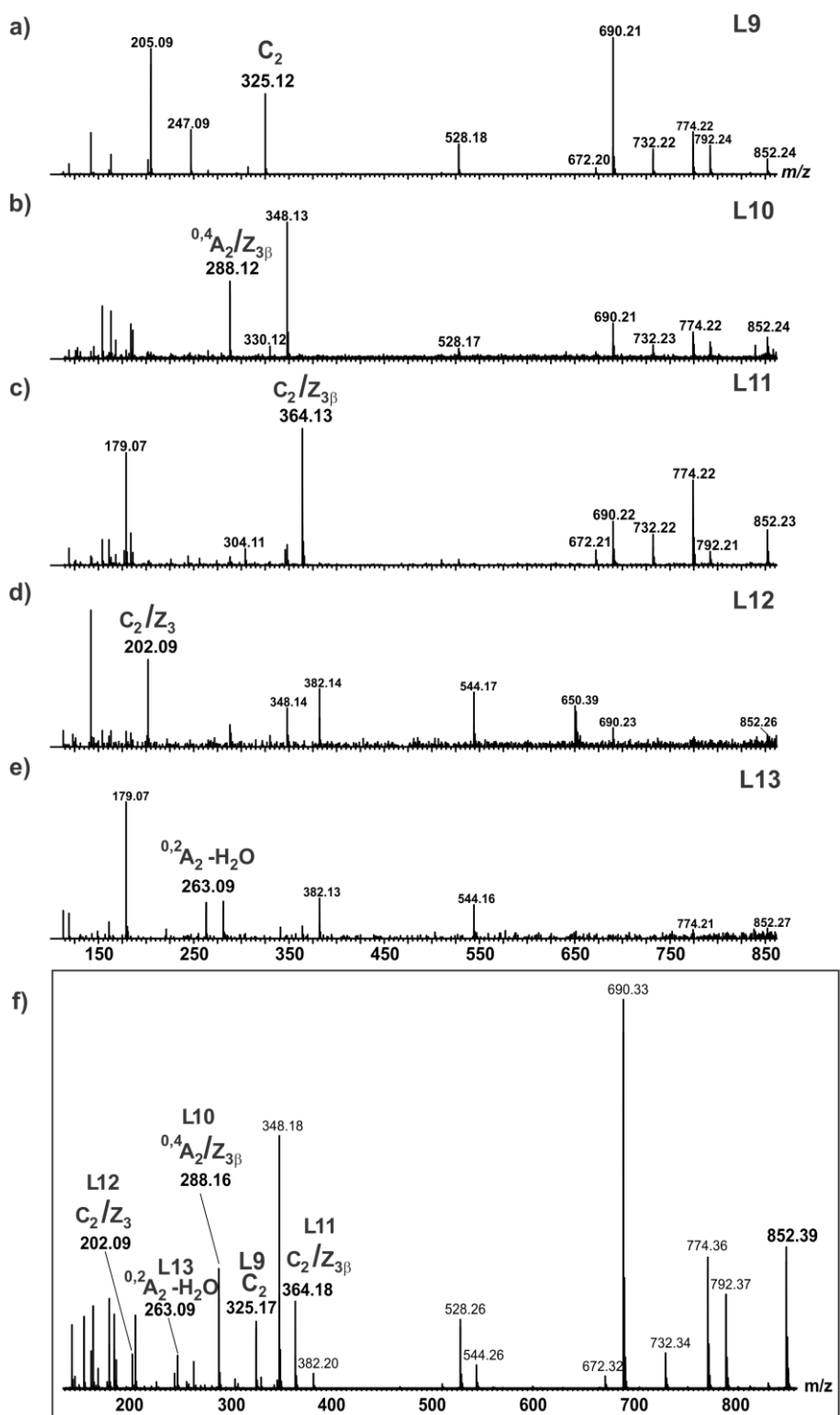
Figure 2.8. CID mass spectra acquired in the Transfer region at 30V for the singly deprotonated anions of (a) **L9,** (b) **L10,** (c) **L11,** (d) **L12,** (e) **L13,** and (f) an equimolar mixture of the **L9-L13** isomers. Fragment ion nomenclature taken from reference 49.
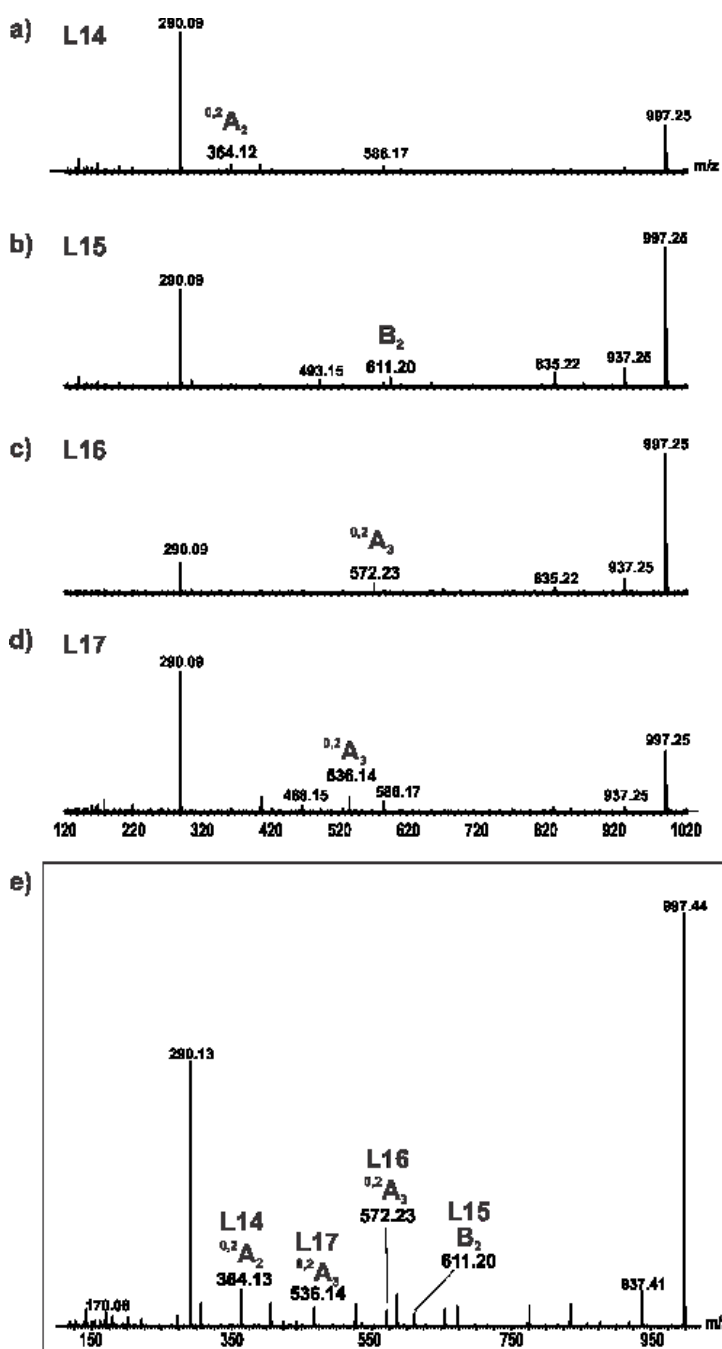
Figure 2.9. CID mass spectra acquired in the Transfer region at 60V for the singly deprotonated anions of (a) **L14,** (b) **L15,** (c) **L16,** (d) **L17,** and (e) an equimolar mixture of the **L14–L17** isomers. Fragment ion nomenclature taken from reference 50.
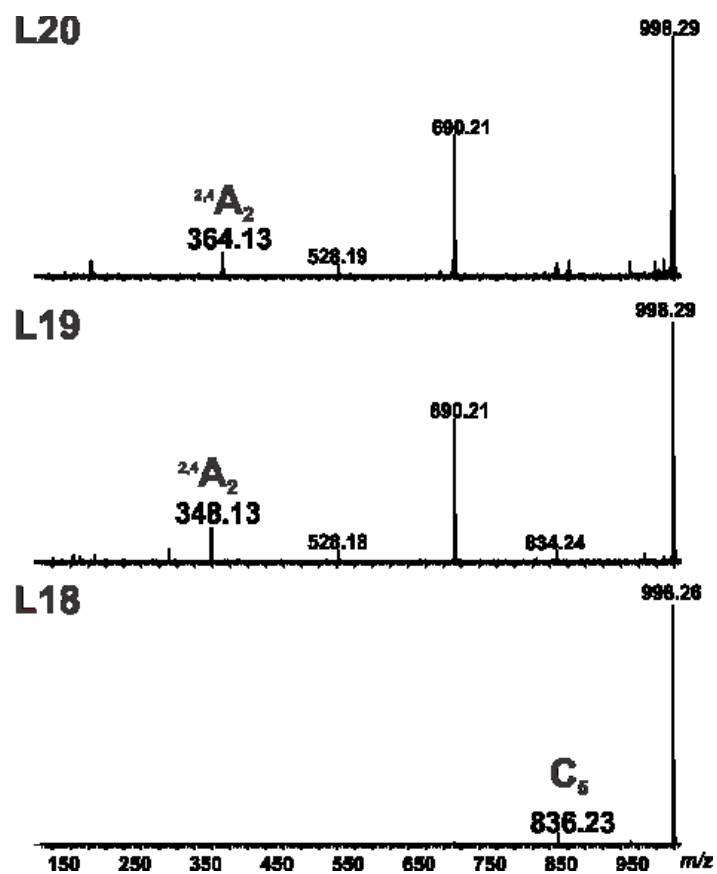
Figure 2.10. CID mass spectra acquired in the Transfer region at 10V for the singly deprotonated anions of **L18**, **L19,** and **L20.** Fragment ion nomenclature taken from references 51–52.
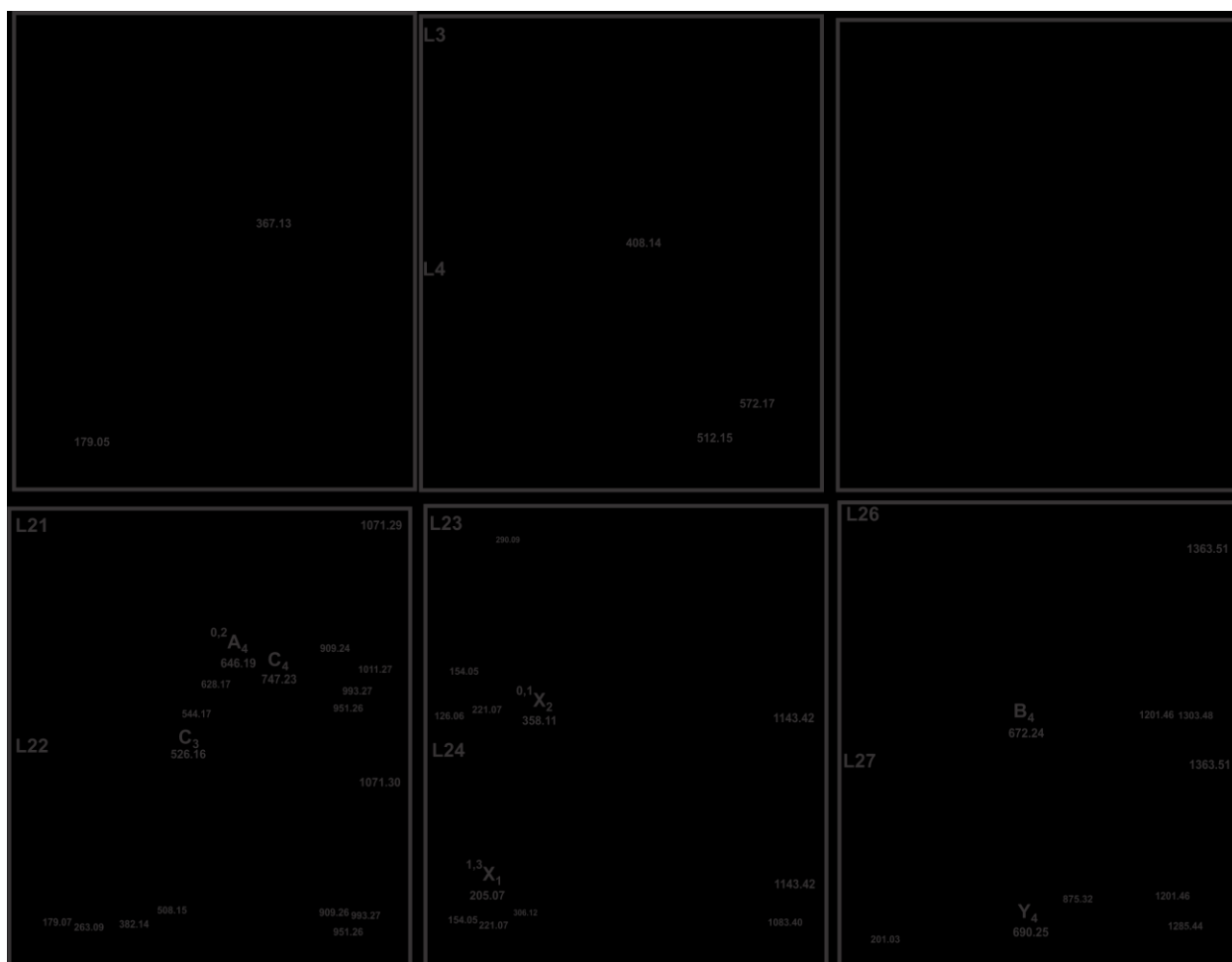
Figure 2.11. CID mass spectra acquired in the Transfer region for the singly deprotonated ions of **L1** and **L2 (**at 20 V), **L3** and **L4** (at 40 V), **L6** and **L7** (at 30 V**), L21** and **L22** (at 50 V), **L23** and **L24** (at 80 V), and **L26** and **L27** (at 30 V). Fragment ion nomenclature taken from references 49–52.

For example, CID (in the Transfer region at 30 V) of the singly deprotonated ions of each HMO in the **L9** – **L13** series produced fragment ions unique to each of the five isomers: C_2 (m/z 325.1) for **L9**, $^{0,4}A_2/Z_3$ (m/z 288.1) for **L10,** C_2/Z_3 (m/z 364.1) for **L11,** C_2/Z_3 (m/z 202.0) for **L12,** and $^{0,2}A_2$–H_2O ions (m/z 263.09) for **L13** (Figures 2.8a–e). These same fragment ions were evident when CID was performed simultaneously on all five isomers **(**Figure 2.8f). Similar results were obtained for the other isomer series (Figures 2.9–2.11).

Taken together, the results of the aforementioned analysis demonstrate that, by combining MWs, IMS-ATs, and CID fingerprints, it is possible to identify all the components of the HMO library simultaneously when present as free oligosaccharides in solution (Table 2.4). To our knowledge, this is the largest library of structurally-related oligosaccharides for which this capability has been demonstrated.

Table 2.4. Overview of the method used for positive identification of each component of the HMO library (**L1-L31**) in the CaR-ESI-MS assay: molecular weight (MW), ion mobility separation (IMS), and collision-induced dissociation (CID) fingerprinting

	Number of HMOs	**HMOs**
MW	7	**L5, L8, L25, L28-L31**
IMS	15	**L1- L4, L6, L7, L18-L24, L26, L27**
CID fingerprinting	9	**L9-L17**

2.3.2 HMO library screening against human lectins

Having established that all thirty-one HMOs (**L1**–**L31**) can be identified from MW, alone or in combination with IMS-ATs and CID fingerprinting, the library was screened against hGal-1, hGal-3C, and hGal-7, and the results were compared with the reported affinities to establish the reliability of the CaR-ESI-MS assay.[31]

hGal-3C: Inspection of a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of hGal-3C (15 μM), P_{ref} (5 μM), and the HMO library (3 μM each) reveals signals corresponding to both free and HMO-bound hGal-3C ions, at charge states -6 and -7 (Figure 2.12a).
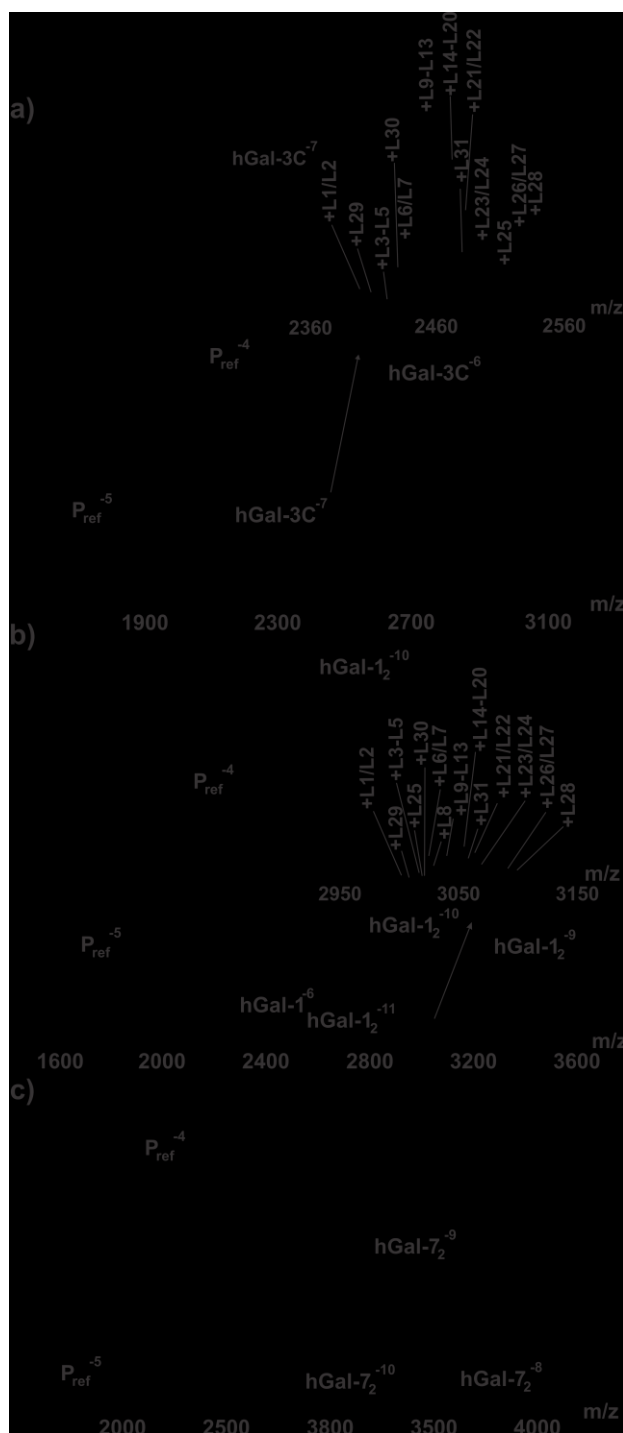
Figure 2.12. Representative ESI mass spectra acquired in the negative ion mode for 20 mM aqueous ammonium acetate solutions (pH 6.8) of P_{ref} (5 µM), HMO library (3 µM each HMO) and (a) hGal-3C (15 µM), (b) hGal-1 (30 µM), or (c) hGal-7 (30 µM). Inset in (a) and (b) shows expanded view of the ions of the (hGal + HMO) complexes.

There is no evidence of hGal-3C bound to two (or more) HMOs. This observation, together with the absence of an ion signal corresponding to P_{ref} bound to any of the HMOs, indicates that, despite the high total concentration of HMOs, non-specific binding to hGal-3C during the ESI process was negligible.[47] Thirteen different HMO MWs were identified from the (hGal-3C + HMO) complexes detected; five of these complexes correspond to HMOs with unique MWs (**L25**, **L28**, **L29**, **L30, and L31**) (Figure 2.12a). It was not possible (using the current instrumentation) to distinguish hGal-3C complexes containing HMOs that differ in MW by only 1 Da (e.g., **L3/L4** (MW 633.2 Da) and **L5** (MW 634.2 Da); **L14–L17 (**MW 998.3 Da), and **L18–L20** (MW 999.3 Da)).

To identify the remaining HMO ligands, all (hGal-3C + HMO) complexes, at -7 charge state, were isolated using the quadrupole mass filter, which was set to pass a wide range of m/z (~200 m/z), and subjected to collisional activation in the Trap region. CID mass spectra were collected using Trap voltages between 10 and 50 V (Figure 2.13).
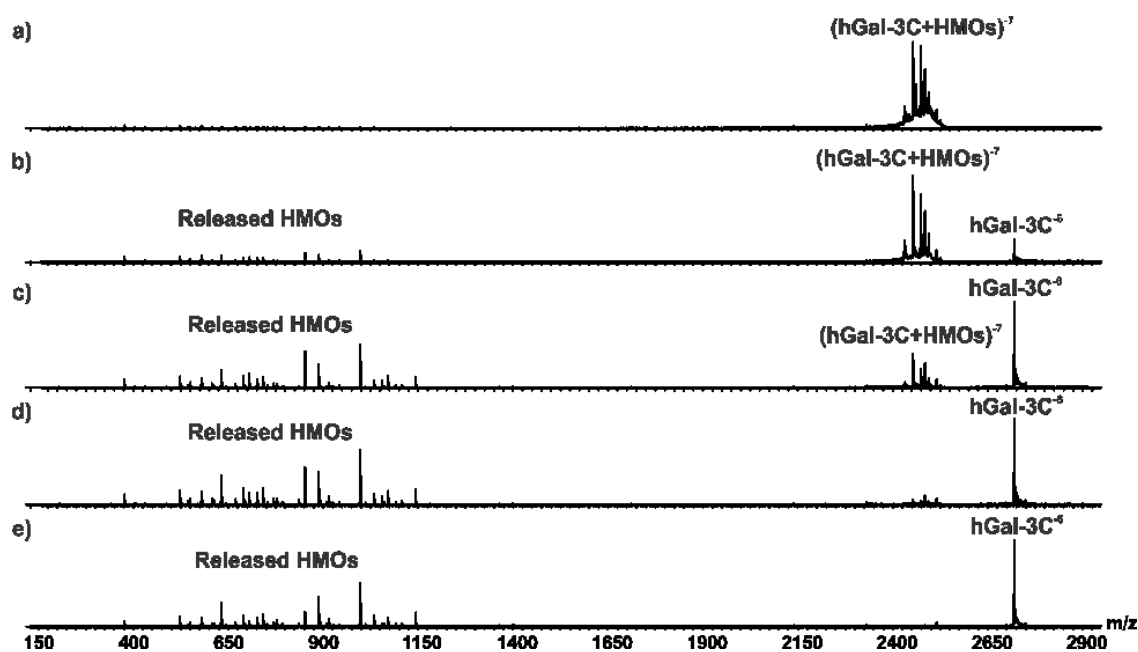
Figure 2.13. CID mass spectra acquired for the (hGal-3C + HMO)$^{7-}$ ions measured at Trap voltages of (a) 10 V, (b) 20 V, (c) 30V, (d) 40, and (e) 50 V.

The optimum Trap voltage for release was selected so as to give the maximum number of MW "hits" but minimize fragmentation of the released HMO ligands. CID carried out at 40 V produced ion signals corresponding to fourteen different HMO MWs, 488.1 Da, 545.5 Da, 633.2 Da, 691.6 Da, 707.2 Da, 853.3 Da, 998.3 Da, 999.3 Da, 1056.3 Da, 1072.3 Da, 1144.4 Da, 1289.4 Da, 1364.5 Da, and 1437.3 Da (Figure 2.14). Five of these MWs correspond to HMOs with unique MWs, *vide supra*; each of the other nine MWs could, in principle, represent between two and five different HMOs.
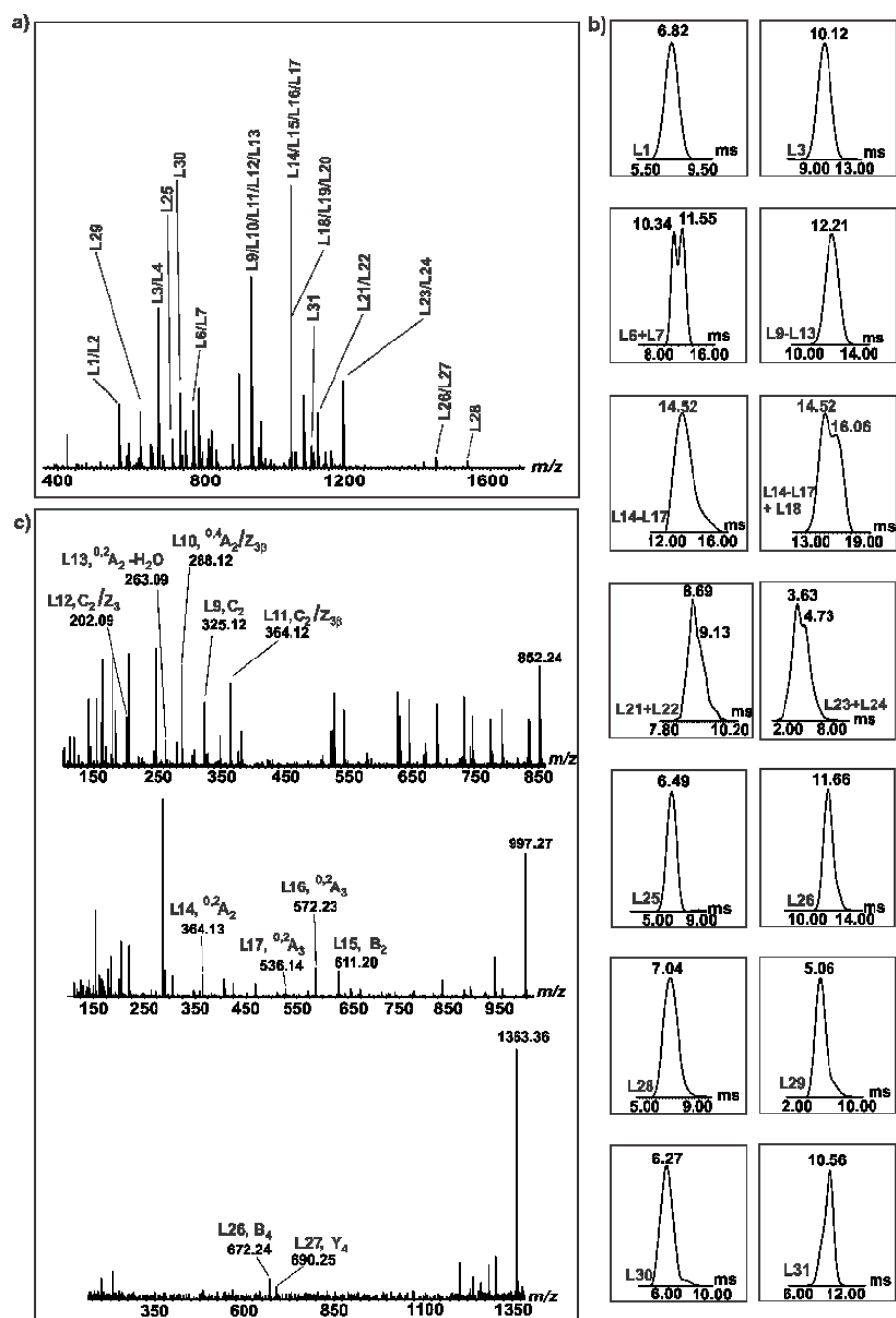
Figure 2.14. (a) Representative CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions produced by ESI from a solution of 20 mM aqueous ammonium acetate solutions (pH 6.8) of hGal-3C (15 μM), P$_{ref}$ (5 μM), and the HMO library (3 μM each), at a Trap voltage 40 V. (b) Representative IMS-ATDs measured for the released HMO anions. The IMS parameters used for these measurements are listed in Table 2.2. (c) Representative CID mass spectra acquired for

released HMO anions with IMS-AT of 12.21 ms (top panel) using a Transfer voltage of 30 V, for ions with IMS-AT of 14.52 ms (middle panel) using a Transfer voltage of 60 V, and for ions with IMS-AT of 11.66 ms (bottom panel) using a Transfer voltage of 30 V.

Comparison of the IMS-ATDs measured for the released HMOs (Figure 2.14) and the free library components (Figures 2.3, 2.4, and 2.5) allowed for the identification of another 10 HMO ligands – **L1**, **L3**, **L6, L7**, **L18**, **L21, L22, L23**, **L24,** and **L26**. The IMS-AT (12.21 ms) measured for the released HMOs with MW 853.3 Da indicated the presence of two or more isomers from the **L9-L13** set. CID fingerprinting of these released HMO ions (IMS-AT of 12.21 ms) produced unique fragment ions arising from **L9** (m/z 325.1), **L10** (m/z 288.1), **L11** (m/z 364.1), **L12** (m/z 202.0), and **L13** (m/z 263.0) (Figure 4c), confirming that all five isomers are specific ligands for hGal-3C. Similarly, CID fingerprinting of the released HMOs with MW 997.3 Da (with the corresponding IMS-AT of 14.52 ms) produced fragment ions indicative of **L14** (m/z 364.1), **L15** (m/z 611.2), **L16** (m/z 572.2), and **L17** (m/z 536.1) being ligands (Figures 2.14c and 2.9). CID fingerprinting, which was performed on all released HMOs in order to provide an additional confidence in ligand assignment, also led to the identification of **L27** as a ligand. **L27** is a low affinity ligand (1.5×10^3 M^{-1}),[31] compared to its structural isomer **L26** (5.8×10^4 M^{-1}),[31] and could not be identified from the IMS-ATD of the released HMO ions (with MW 1364.5). However, CID performed on released HMO ions with IMS-AT of 11.66 ms produced fragment ions indicative of both **L26** (m/z 672.2) and **L27** (m/z 690.2) (Figures 2.14c and 2.11).

The results of CaR-ESI-MS screening of the HMO library against hGal-3C indicate that 25 of the HMOs in the library are ligands for this lectin. Importantly, all 25 are reported to exhibit measurable affinities (ranging from 2×10^3 to 1.3×10^5 M^{-1}) for hGal-3C (Figure 2.15).[31]

Figure 2.15 data table

HMOs	hGal-1 CaR-ESI-MS	hGal-1 K_a	hGal-3C CaR-ESI-MS	hGal-3C K_a	hGal-7 CaR-ESI-MS	hGal-7 K_a
L1	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L2	Binding	$K_a \geq 500$	No Binding	No Binding	No Binding	$K_a \geq 500$
L3	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L4	Binding	$K_a \geq 500$	No Binding	No Binding	Binding	$K_a \geq 500$
L5	Binding	$K_a \geq 500$	No Binding	No Binding	No Binding	No Binding
L6	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$
L7	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L8	Binding	$K_a \geq 500$	Binding	No Binding	No Binding	No Binding
L9	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L10	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^5$	Binding	$K_a \geq 10^4$
L11	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^5$	Binding	$K_a \geq 500$
L12	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L13	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L14	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L15	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L16	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L17	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L18	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L19	Binding	$K_a \geq 500$	No Binding	No Binding	Binding	$K_a \geq 500$
L20	Binding	$K_a \geq 500$	No Binding	No Binding	Binding	$K_a \geq 500$
L21	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L22	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L23	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L24	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L25	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L26	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$
L27	Binding	$K_a \geq 500$	Binding	$K_a \geq 500$	Binding	$K_a \geq 500$
L28	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^5$	Binding	$K_a \geq 500$
L29	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 10^4$
L30	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$
L31	Binding	$K_a \geq 500$	Binding	$K_a \geq 10^4$	Binding	$K_a \geq 500$

Affinities: ● $K_a \geq 10^5$ M^{-1} ● $K_a \geq 10^4$ M^{-1} ○ $K_a \geq 500$ M^{-1} ○ $K_a \leq 500$ M^{-1} ○ No Binding

CaR-ESI-MS: ● Binding ○ No Binding

Figure 2.15. Comparison of the HMO ligands identified by the CaR-ESI-MS assay with the corresponding HMO affinities for hGal-1, hGal-3C, and hGal-7 reported in reference 31.

The successful detection of ligands with affinities that span two orders of magnitude highlights the ability of CaR-ESI-MS to identify both high and low affinity HMO ligands.

hGal-1: Utilizing the same approach, the HMO library was screened against hGal-1, which exists predominantly as a homodimer in solution.[31] A representative ESI mass spectrum acquired

in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of hGal-1 (30 µM), P_{ref} (5 µM), and the HMO library (3 µM each) is shown in Figure 2.12b. Ion signals corresponding to both the free and HMO-bound hGal-1 dimer, at charge states -9 to -11, was detected. Fourteen different HMO MWs were identified from the detected (hGal-1 + HMO) complexes; six of these complexes correspond to HMOs with unique MWs (**L8, L25**, **L28**, **L29, L30,** and **L31**). Non-specific binding during the ESI process was insignificant, as indicated by the absence of ion signal corresponding to P_{ref} bound to any of the HMOs. Using a combination of IMS-ATs and CID fingerprinting, it was found that the other 25 HMOs are also ligands (Figures 2.16–2.18). The finding that all 31 HMOs are ligands for hGal-1 agrees with the reported affinity data (Figure 2.15).[31]
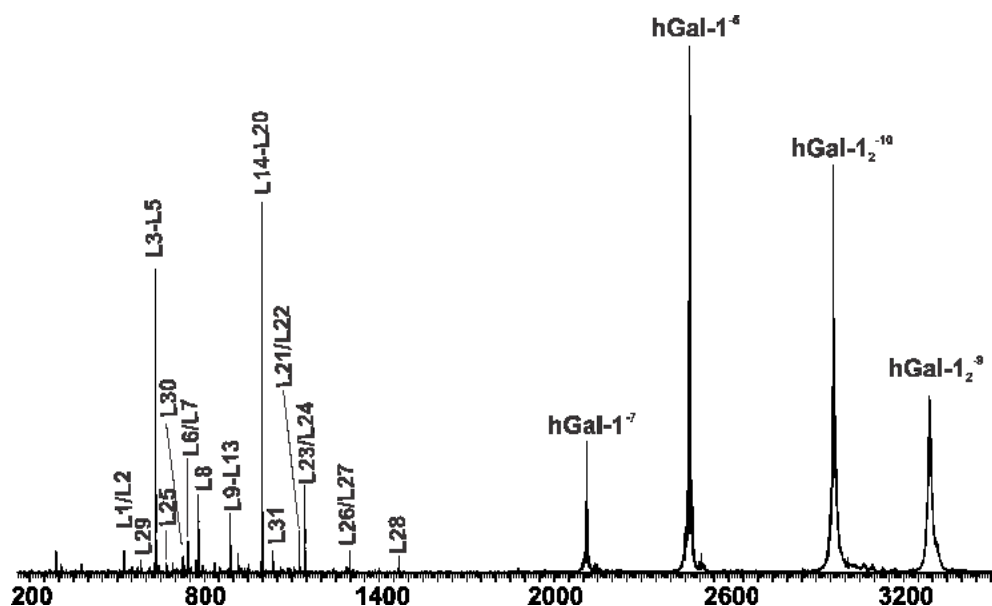
Figure 2.16. CID mass spectrum acquired for the (hGal-1 + HMO)$^{10-}$ ions at a Trap voltage of 80 V.
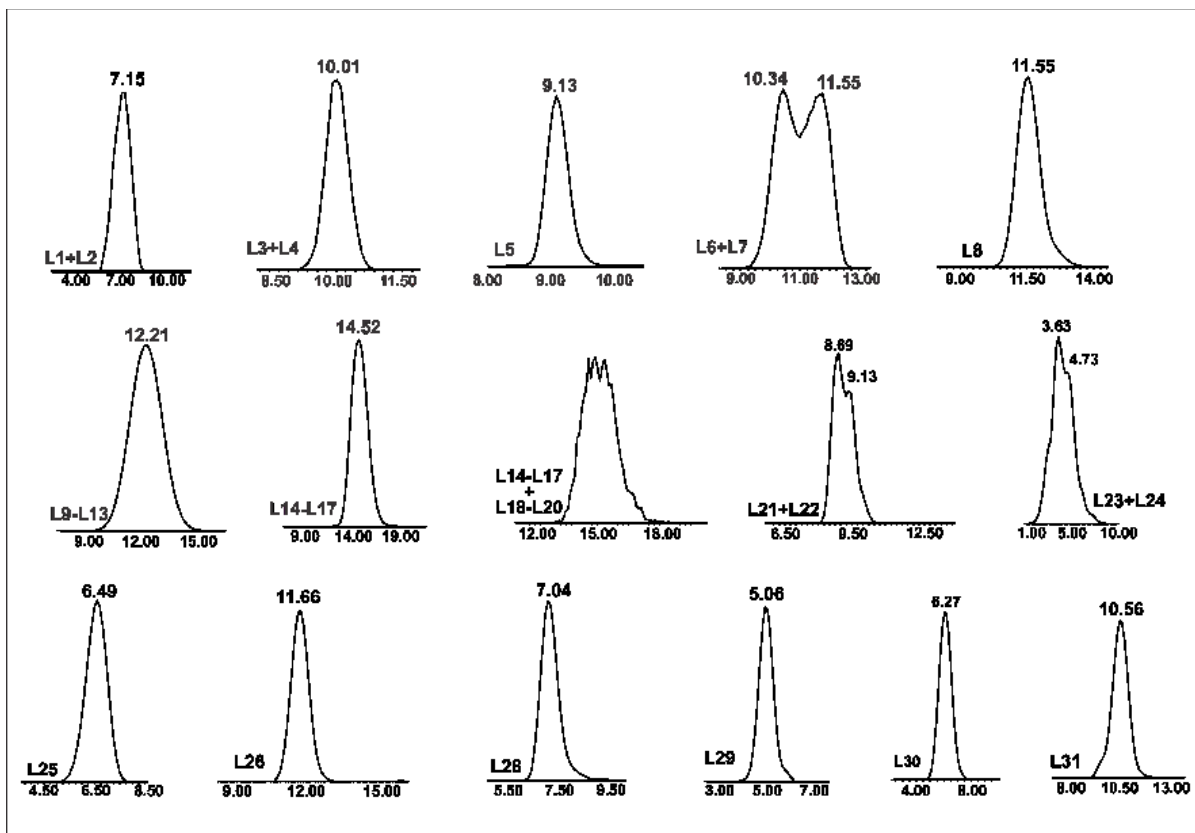
Figure 2.17. IMS-ATDs measured for deprotonated HMO ions following their release from the (hGal-1 + HMO)$^{10-}$ ions at a Trap voltage of 80 V. The IMS parameters used for these measurements are listed in Table 2.2.
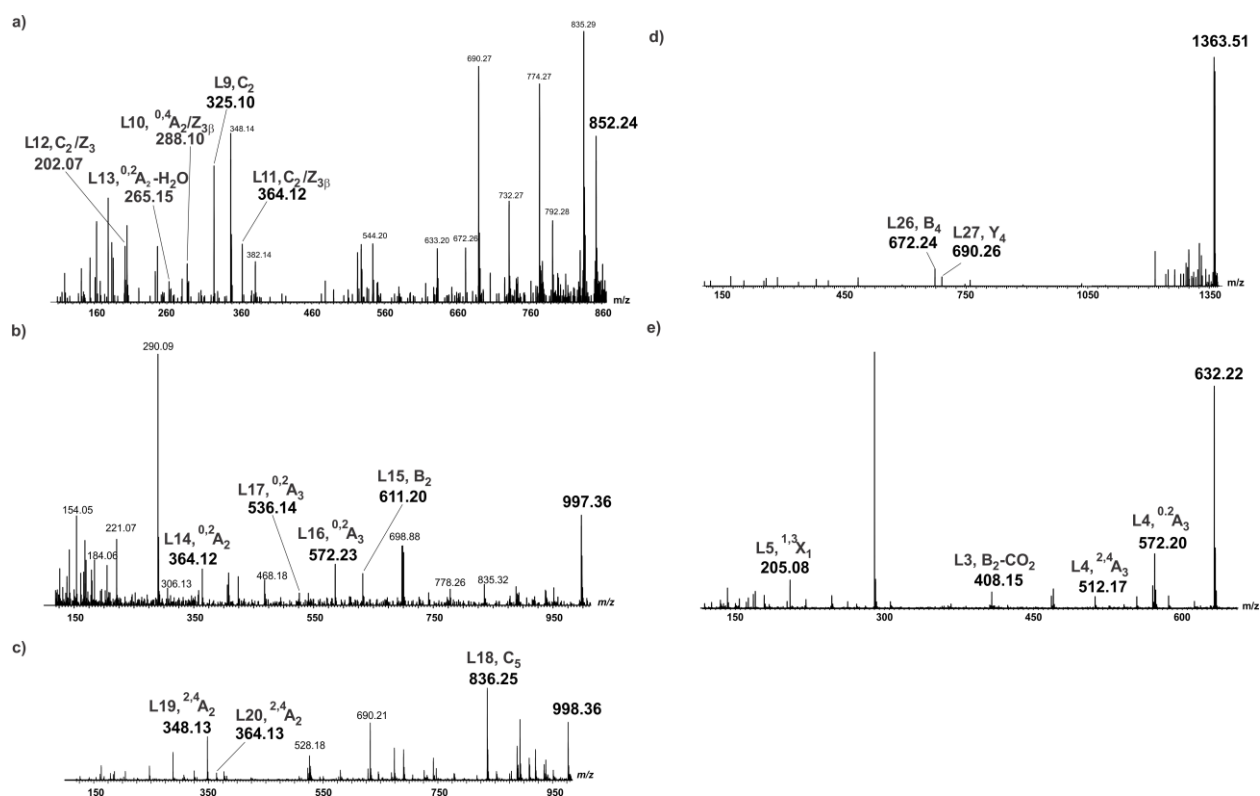
Figure 2.18. CID mass spectra acquired for HMO ions released from $(hGal\text{-}1 + HMO)^{10\text{-}}$. (a) Ions with IMS-AT of 12.21 ms using a Transfer voltage of 30 V, (b) ions with IMS-AT of 14.52 ms using a Transfer voltage of 60 V, (c) ions with IMS-AT of 15.07 ms using a Transfer voltage of 10 V, (d) ions with IMS-AT of 11.66 ms using a Transfer voltage of 30 V, and (e) ions with IMS-AT of 10.01 ms using a Transfer voltage of 40 V.

hGal-7: Shown in Figure 2.12c is a representative negative ion mode ESI mass spectrum acquired for an aqueous ammonium acetate solution (20 mM, pH 6.8) of hGal-7 (30 μM), P_{ref} (5 μM), and the HMO library (3 μM each). Similar to hGal-1, hGal-7 exists predominantly as a homodimer in solution.[31] Ion signals corresponding to both the free and HMO-bound hGal-7 dimer, at charge states -8 to -10, was detected in the mass spectrum. Unlike hGal-3C and hGal-1, the individual HMO-bound hGal-7 dimer ions were not resolved in the mass spectrum, making it impossible to determine the MWs of bound HMO ligands. No P_{ref}-HMO complexes were detected,

55

indicating that non-specific HMO binding was negligible under the experimental conditions used.

Release of the bound ligands resulted in the appearance of 13 different HMO MWs (Figure 2.19).
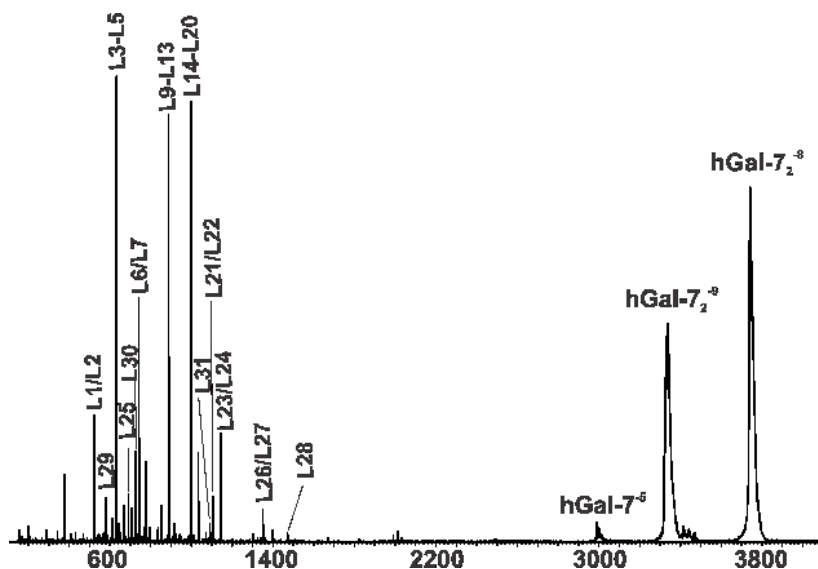
Figure 2.19. CID mass spectrum acquired for all (hGal-7+HMO)$^{-9}$ ions at a Trap voltage of 80 V.
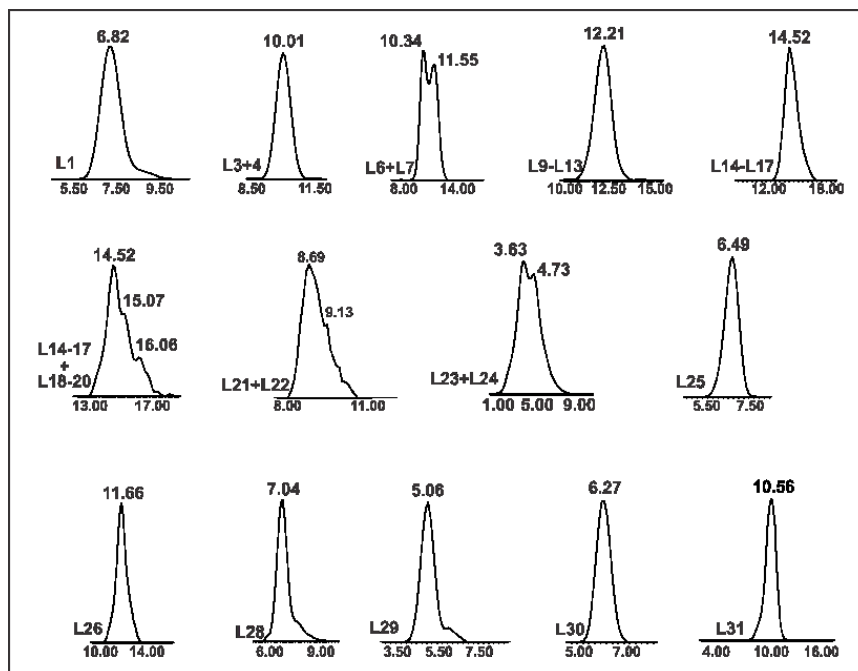
Figure 2.20. IMS-ATDs measured for deprotonated HMO ions following their release from the (hGal-7 + HMO)$^{-9}$ ions at a Trap voltage of 80 V. The IMS parameters used for these measurements are listed in Table 2.2.
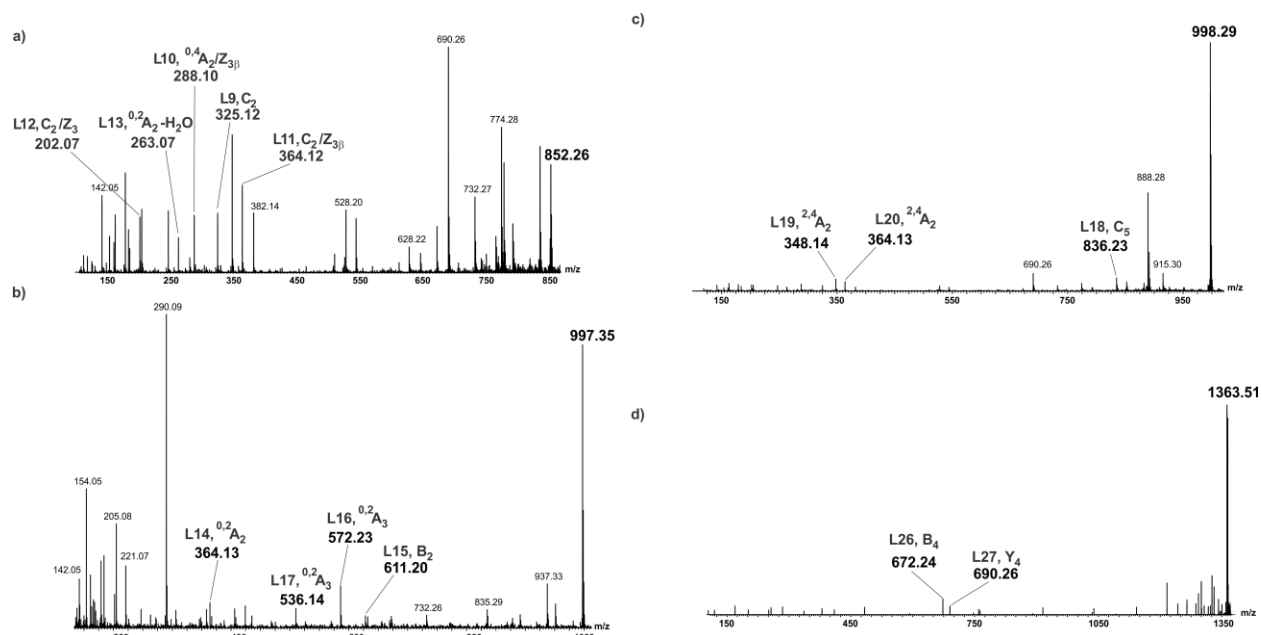
Figure 2.21. CID mass spectra acquired for HMO ions released from $(hGal\text{-}7 + HMO)^{-9}$. (a) Ions with IMS-AT of 12.21 ms using a Transfer voltage of 30 V, (b) ions with IMS-AT of 14.52 ms using a Transfer voltage of 60 V, (c) ions with IMS-AT of 15.07 ms using a Transfer voltage of 10 V, and (d) ions with IMS-AT of 11.66 ms using a Transfer voltage of 30 V.

Based on the IMS-ATs and CID fingerprints measured for the released ions, 28 ligands were identified positively; no binding was detected for **L2, L5,** and **L8** (Figures 2.19–2.21). According to the reported affinity data, all but one (**L8**) of the 31 HMOs are ligands for hGal-7. The two false negatives, **L2** and **L5**, exhibit the lowest affinities (of the the ligands in the library) for hGal-7 (436 and 235 M^{-1}, respectively).[31] Under the solution conditions used, only 0.13% and 0.07% of hGal-1 is expected to be bound to **L2** and **L5**, respectively.

2.3.3 HMO affinity ranking

Because of the different degrees of fragmentation of the HMO anions following their release from the lectins, the relative abundances of released ligands do not reflect the relative abundances of bound ligands in solution accurately.[48] Consequently, it is not possible to establish relative

affinities of the HMO ligands from the relative abundances of released ligands reliably. However, by applying correction factors, based on the extent of fragmentation observed in CID performed on the individual HMO anions, it is possible to identify the majority of the highest affinity HMOs (or HMO isomer sets) for each lectin. This approach is described briefly below.
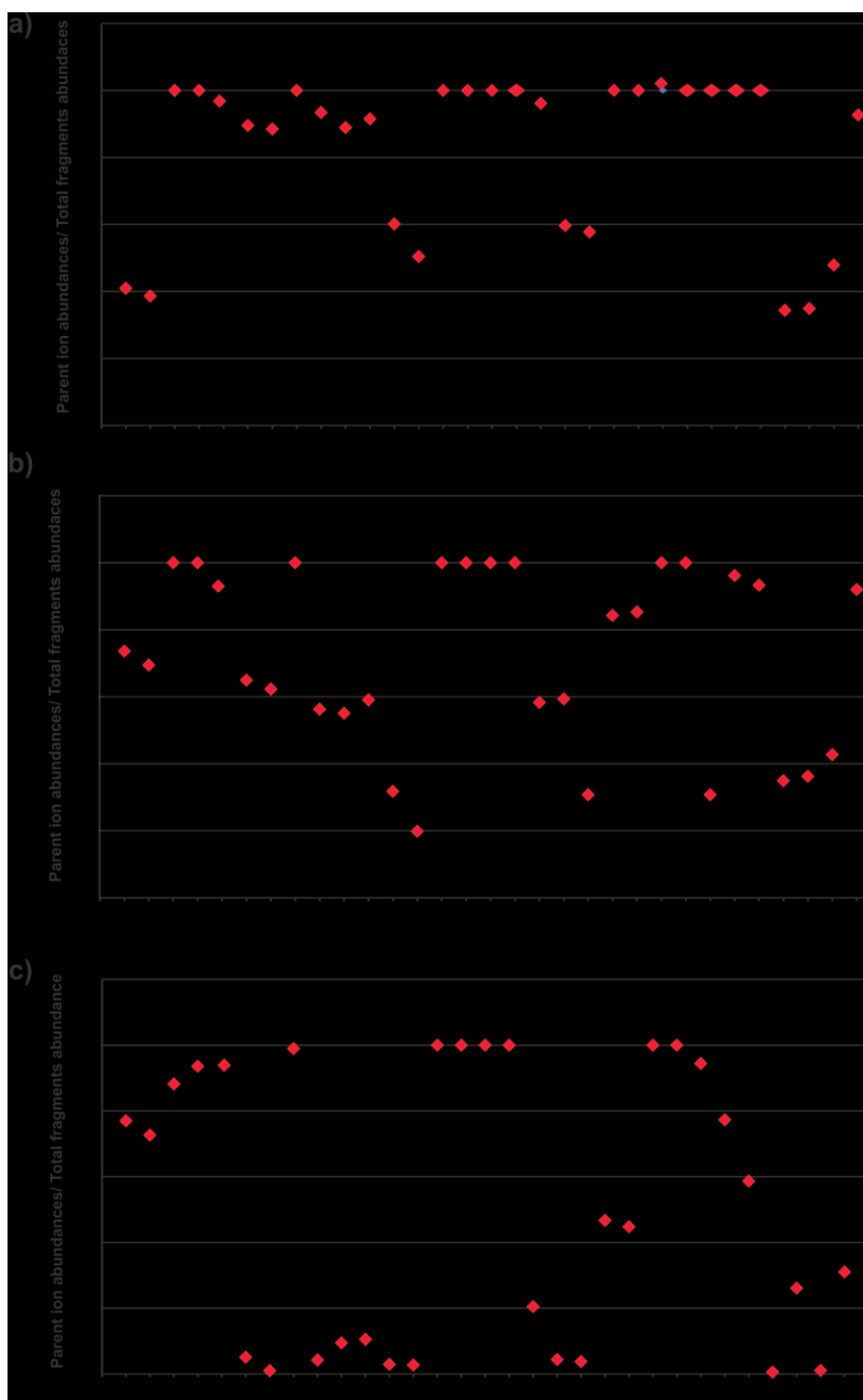
Figure 2.22. Fractional abundances of deprotonated ions of **L1**–**L31** measured using Transfer voltages of (a) 10 V, (b) 20 V, and (c) 30 V.

Plotted in Figure 2.22 is the fraction of each HMO anion that remains un-dissociated when subjected to CID in the Transfer region (at 10 V, 20 V, and 30 V). It can be seen that, at the lowest energy investigated (10 V), 22 of the HMO anions do not undergo any fragmentation, while seven (**L12**, **L13**, **L19**, **L20**, **L28**, **L29,** and **L30**) undergo between 40% and 65% dissociation. At 20 V, the majority (22) of the HMO ions undergo fragmentation, with seven of these (**L12, L13**, **L20, L25**, **L28**, **L29,** and **L30)** undergoing between 50% and 80% dissociation. At 30 V, only seven of the HMO anions remain completely intact and 12 (**L6**, **L7**, **L9, L10**, **L11**, **L12, L13**, **L18**, **L19**, **L20, L28,** and **L30**) undergo >80% dissociation. Based on the extent of fragmentation measured at 10 V, 20 V, and 30 V, a correction factor (which quantitatively accounts for the extent of dissociation) was calculated for each HMO at each voltage (Table 2.5).

Table 2.5. Correction factors, based on the results of CID carried out in the Transfer region at voltages of 10 V, 20 V, and 30 V, to correct the abundances of released HMO anions for the occurrence of fragmentation, and an average correction factor for each isomer set based on correction factors determined at 10 V, 20 V, and 30 V for each isomer within the set

HMO isomer set	HMO	Correction factors			Average energy correction factor	Average correction factor for HMO isomer sets
		10 V	20 V	30 V		
L1/L2	L1	1.3	1.4	2.4	1.7	1.8
	L2	1.4	1.4	2.6	1.8	
L3/L4	L3	1.0	1.0	1.1	1.0	1.0
	L4	1.0	1.0	1.1	1.0	
L5	L5	1.0	1.0	1.1	1.0	1.0
L6/L7	L6	1.1	1.5	19.6	7.4	20.6

	L7	1.1	1.6	98.6	33.8	
L8	**L8**	1.0	1.0	1.0	1.0	**1.0**
L9-L13	**L9**	1.1	1.8	23.7	8.9	**9.0**
	L10	1.1	1.8	10.6	4.5	
	L11	1.1	1.7	9.5	4.1	
	L12	1.7	3.1	34.2	13.0	
	L13	2.0	5.0	36.7	14.6	
L14-L17	**L14**	1.0	1.0	1.0	1.0	**1.0**
	L15	1.0	1.0	1.0	1.0	
	L16	1.0	1.0	1.0	1.0	
	L17	1.0	1.0	1.0	1.0	
L18-L20	**L18**	1.0	1.7	4.9	2.5	**7.2**
	L19	1.7	1.7	22.8	8.7	
	L20	1.7	3.3	26.4	10.5	
L21/L22	**L21**	1.0	1.2	2.1	1.4	**1.5**
	L22	1.0	1.2	2.2	1.5	
L23/L24	**L23**	1.0	1.0	1.0	1.0	**1.0**
	L24	1.0	1.0	1.0	1.0	
L25	**L25**	1.0	3.3	1.1	1.8	**1.8**
L26/L27	**L26**	1.0	1.0	1.3	1.1	**1.2**
	L27	1.0	1.1	1.7	1.3	
L28	**L28**	2.9	2.9	177.4	61.1	**61.1**
L29	**L29**	2.9	2.8	3.8	3.2	**3.2**
L30	**L30**	2.1	2.3	94.2	32.9	**32.9**
L31	**L31**	1.1	1.1	3.2	1.8	**1.8**

There are two practical challenges to correcting the CaR-ESI-MS data for fragmentation of the HMO ions. First, the internal energy distributions of the released (from lectin) HMO anions are not known. In other words, it is not possible to relate the CID voltages used to release the HMO anions to voltages used to fragment the anions (as described above). To account for this uncertainty, an average energy correction factor, one based on the average values measured at 10 V, 20 V, and 30 V, was calculated for each HMO. Secondly, in cases where multiple isomers are ligands, it is not possible to correct the abundances of individual isomers because the relative abundances of the different isomers released from the lectin cannot (necessarily) be ascertained. To deal with this, an average correction factor was calculated for each isomer set, based on the corresponding average energy correction factors. These factors, which are listed in Table 2.5 were used to correct (for fragmentation) the abundances of the released HMO ions (grouped as isomer sets) in the CaR-ESI-MS experiments approximately. Taking L6 (MW 707.25 Da) as an example, at 10 V, the peak intensity ratio of parental peak (m/z 706.25) to total (parental plus all possible fragment peaks) is 0.9. The correction factor for L6 equals to the reversed ratio (1.1). Same calculations were carried for the other two voltages (20 V and 30 V), and the average of three voltages for L6 is 7.4. Since L7 is within the same isomer set as L6, the averaged correction factor of L6 and L7 (20.6) becomes the final correction factor for released ligand with molecular weight 707.25 Da. The normalized relative abundance for MW 707.25 Da (L6 and L7) before correction is 0.036. By multiplying the final correction factor (20.6), the normalized relative abundances after correcting the secondary fragmentation is 0.733, which is the third abundant released molecular weight (Figure 2.23).
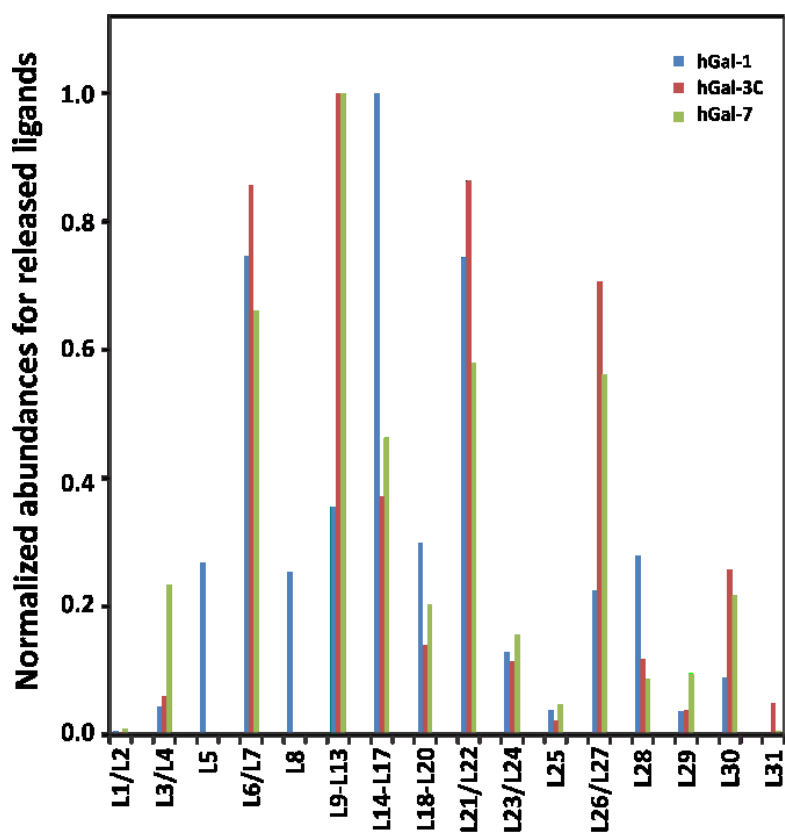
Figure 2.23. Normalized relative abundances (corrected for secondary fragmentation using the factors listed in Table 2.5) of the HMO ligands of hGal-1, hGal-3C, and hGal-7 detected by CaR-ESI-MS. The abundances for all HMOs within an isomer set are added together.

Shown in Figure 2.23 is a plot of the relative abundances of the released HMO ligands from the three hGal proteins, grouped according to MW, following application of these correction factors. For hGal-1, the MWs with the highest ion abundances are 998.3 Da (**L14-L17**), 707.2 Da (**L6/L7**), and 1072.3 Da (**L21/L22**), indicating that one or more HMO within each of these three isomer sets exhibit relatively high affinity for the protein. These results are consistent with the reported HMO affinities–**L6, L22**, **L21,** and **L14** are the highest affinity ligands in this library.[29] For hGal-3C, the HMO MWs with the highest ion abundances are 853.3 Da (**L9-L13**), 1072.3 Da (**L21/L22**), and 707.2 Da (**L6/L7**). This is consistent with **L11**, **L10, L21,** and **L7** being high

affinity (ranked 2nd to 5th highest) ligands for hGal-3C.[31] The relatively low abundance of the highest affinity ligand, **L28**, can be explained by the fact that it is the lone HMO with that MW and is one of the most labile (in the gas phase) HMOs in the library, because of the huge change in the correction factor between 10 V and 30 V, meaning the extend of fragmentation is significant. For hGal-7, the highest corrected ion abundances correspond to HMOs with MWs 853.3 Da (**L9**–**L13**), 707.2 Da (**L6/L7**), 1072.3 Da (**L21/L22**), and 1364.5 Da (**L26/L27**). These findings are consistent with **L26**, **L6, L10,** and **L21** being the 2nd–5th highest affinity ligands in the library for hGal-7.[31] Similar to what was observed for hGal-3C, the highest affinity ligand, **L29**, which has a unique MW in the library and is also somewhat prone to post-release fragmentation, was not among the most abundant HMO MWs.

Taken together, the results of this analysis show that, following correction for post-release fragmentation, the relative abundances of the released ligands allow for the identification of the MWs corresponding to the majority of the highest affinity HMO ligands in the library. Although not quantitative, the HMO MW affinity ranking can serve as a useful guide for follow-up affinity measurements on individual HMOs. Such measurements can be time consuming and expensive due to the relatively low affinities and the high cost of purified HMOs.

2.4 Conclusions

In summary, we have developed a high-throughput, label- and immobilization-free method, based on CaR-ESI-MS, for screening (simultaneously) libraries of free (unmodified) HMOs against lectins *in vitro*. The method requires pure stock solutions of known concentration for each HMO from which CID and IMS measurements are performed prior to analysis of mixtures. Ligand identification is established from the MWs, IMS-ATs, and CID fingerprints of HMO anions released from the target protein in the gas phase. The reliability of the assay was demonstrated by screening a library of 31 HMOs against three hGal proteins. Implemented using an equimolar

concentration library of HMOs, the CaR-ESI-MS assay identified 100% of ligands with reported affinities >500 M^{-1}; no false positives were detected. Although not quantitative, the assay also successfully identified the majority of the highest affinity HMO ligands (or isomer sets that contain the highest affinity ligands) in the library for each of the three hGal. Given the speed (<1 h) and low sample consumption (<5 ng of each HMO and <0.5 μg of protein) of the assay, together with the absence of labeling and immobilization requirements, we expect that the CaR-ESI-MS method will be adopted widely for rapid screening of free HMOs, made from pure stock solutions of known concentration, against lectins. Furthermore, it should be noted that this methodology is not limited necessarily to HMOs and could be extended to libraries of other classes of oligosaccharides constructed from pure stock solutions of known concentration, e.g. histo-blood group antigen oligosaccharides and N-glycans from glycoproteins, for screening.

2.5 Literature cited

1. Newburg, D.S.; Shen, Z.; Warren, C.D. *Adv. Exp. Med. Biol.* **2000**, *478*, 381.

2. Gabrielli, O.; Zampini, L.; Galeazzi, T.; Padella, L.; Santoro, L.; Peila, C.; Giuliani, F.; Bertino, E.; Fabris, C.; Coppa, G.V. *Pediatrics* **2011**, *128*, 1520.

3. Bode, L. *Glycobiology* **2012**, *22*, 1147.

4. Sela, D.A.; Chapman, J.; Adeuya, A.; Kim, J.H.; Chen, F.; Whitehead, T.R.; Lapidus, A.; Rokhsar, D.S.; Lebrilla, C.B.; German, J.B.; Price, N.P.; Tichardson, P.M.; Mills, D.A. *Proc. Natl. Acad. Sci.* **2008**, *105*, 18964.

5. Eiwegger, T.; Stahl, B.; Haidl, P.; Schmitt, J.; Boehm, G.; Dehlink, E; Urbanek, R.; Szepfalusi, Z. *Pediatr. Allergy Immunol.* **2010**, *21*, 1179.

6. Wang, B.; Yu, B.; Karim, M.; Hu, H.; Sun, Y.; McGreevy, P.; Petocz, P.; Held, S.; Brand-Miller, J. *Am. J. Clin. Nutr.* **2007**, *85*, 561.

7. Hickey, R.H. *Int. Dairy J.* **2012**, *22*, 141.

8. Underwood, M.A.; Gaerlan, S.; De Leoz, M.L.A.; Dimapasoc, L.; Kalanetra, K.M.; Lemay, D.G.; German, J.B.; Mills, D.A.; Lebrilla, C.B. *Pediatr. Res.* **2015**, *78*, 670.

9. Newburg, D.S.; Ruiz-Palacios, G.M.; Morrow, A.L. *Annu. Rev. Nutr.* **2005**, 25, 37.

10. Sisk, P.M.; Lovelady, C.A.; Dillard, R.G.; Gruber, K.J.; O'Shea, T.M. *J. Perinatol.* **2007**, *27*, 428.

11. Kobata, A. *Proc. Jpn. Acad. Sci., Ser. B* **2010**, *86*, 731.

12. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. *Annu. Rev. Nutr.* **2000**, *20*, 699.

13. Totten, S.M.; Zivkovic, A.M.; Wu, S.; Ngyuen, U.; Freeman, S.L.; Ruhaak, L.R.; Darboe, M.K.; German, J.B.; Prentice, A.M.; Lebrilla, C.B. *J. Proteome Res.* **2012**, *11*, 6124.

14. Smilowitz, J.T.; Lebrilla, C.B.; Mills, D.A.; German, B.J.; Freeman, S.L. *Annu. Rev. Nutr.* **2014**, *34*, 143.

15. Hong, Q.; Ruhaak, L.R.; Totten, S.M.; Smilowitz, J.T.; German, J.B.; Lebrilla, C.B. *Anal. Chem*. **2014**, *86*, 2640.

16. Xu, G.; Davis, J. C.; Goonatilleke, E.; Smilowitz, J. T.; German, J. B.; Lebrilla, C. B. *J. Nutr.* **2017**, *147*, 117.

17. Ruhaak, L.; Lebrilla, C.B. *Adv. Nutr.* **2012**, *3*, 406S

18. Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W.E. *BBA-Gen. Subjects* 2002, *1572*, 232.

19. Ohtsuka, I.; Sadakane, Y.; Higuchi, M.; Hada, N.; Hada, J.; Kakiuchi, N.; Sakushima, A. *Bioorg. Med. Chem*. **2011**, *19*, 894.

20. Shang, J; Piskarev, V.E.; Xia, M., Huang, P.; Jiang, X.; Likhosherstov, L.M.; Novikova, O.S.; Newburg, D.S.; Ratner, D.M. *Glycobiology*. **2013**, *23*, 1491.

21. Noll, A. J.; Gourdine, J. P.; Yu, Y.; Lasanajak, Y.; Smith, D. F.; Cummings, R. D. *Glycobiology* **2016**, *26*, 655.

22. Ashline, D.J.; Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Prasad, B.V.; Estes, M.K.; Cummings, R.D.; Smith, D.F.; Reinhold, V. N. *Mol. Cell. Proteom.* **2014**, *13*, 2961.

23. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Mickum, M. L.; Ashline, D. J.; Prasad, B. V.; Estes, M. K.; Reinhold, V. N. Cummings, R.D.; Smith, D.F. *Mol. Cell. Proteom.* **2014**, *13*, 2944.

24. Noll, A. J.; Yu, Y.; Lasanajak, Y.; Duska-McEwen, G.; Buck, R. H.; Smith, D. F.; Cummings, R. D. *Biochem. J.* **2016**, *473*, 1343.

25. Grant, O. C.; Smith, H. M.; Firsova, D.; Fadda, E.; Woods, R. J. *Glycobiology* **2014**, *24*, 17.

26. Kilcoyne, M.; Gerlach, J. Q.; Kane, M.; Joshi, L. *Anal. Methods* **2012**, *4*, 2721.

27. Geissner, A.; Seeberger, P.H. *Annu Rev Anal Chem* **2016**, *9*, 223

28. Arthur, C.M.; Rodrigues, L.C.; Baruffi, M.D.; Sullivan, H.C.; Heimburg-Molinaro, J.; Smith, D.F.; Cummings, R.D.; Stowell, S.R. Examining galectin binding specificity using glycan microarrays. In *Galectins: Methods and Protocols*; Stowell, S.R., Cummings, R.D., Eds.; Methods in Molecular Biology 1207; Humana Press: New York, **2015**, 115-131.

29. Paulson, J. C.; Blixt, O.; Collins, B. E. *Nat. Chem. Biol.* **2006**, *2*, 238.

30. He, X. G.; Gerona-Navarro, G.; Jaffrey, S. R. *J. Pharmacol. Exp. Ther*. **2005**, *313*, 1.

31. Shams-Ud-Doha, Km; Kitova, E. N.; Kitov, P., I.; St-Pierre, Y; Klassen, J. S. *Anal. Chem.* **2017**, *89*, 4914.

32. López-Lucendo, M.F.; Solís, D.; André, S.; Hirabayashi, J.; Kasai, K.; Kaltner, H.; Gabius, H.J.; Romero, A. *J. Mol. Biol*. **2004**, *343*, 957.

33. Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U.J.; Akke, M. *J. Am. Chem. Soc*. **2010**, *132*, 14577.

34. Christensen, T.; Toone, E.J. *Methods Enzymol*. **2003**, 362, 486.

35. El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. *Biochemistry* **2012**, *51*, 4244.

36. Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom*. **2012**, *23*, 431.

37. El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2011**, *84*, 50.

38. Kitova, E.N.; El-Hawiet, A.; Schnier, P.D.; Klassen, J.S. *J. Am. Soc. Mass Spectrom*. **2012**, *23*, 431.

39. Hofmann, J.; Stuckmann, A.; Crispin, M.; Harvey, D.J.; Pagel, K.; Struwe, W.B. *Anal. Chem*. **2017**, *89*, 2318.

40. Struwe, W.B.; Baldauf, C.; Hofmann, J.; Rudd, P.M.; Pagel, K. *Chem. Com*. **2016**, *52*, 12353.

41. Williams, J.P.; Grabenauer, M.; Holland, R. J.; Carpenter, C. J.; Wormald, M. R.; Giles, K.; Harvey, D. J.; Bateman, R. H.; Scrivens, J.H.; Bowers, M. T. *Int. J. Mass Spectrom.* **2010**, 298, 119.

42. De Kivit, S.; Kraneveld, A.D.; Garssen, J.; Willemsen, L.E.M. *Eur. J. Pharmacol,* **2011**, *668*, S124.

43. Labrie, M.; Vladoiu, M.; Leclerc, B.G.; Grosset, A.A.; Gaboury, L.; Stagg, J.; St-Pierre, Y. *PLoS One*. **2015**, *10*, e0131307.

44. Li, H.; Bendiak, B.; Siems, W. F.; Gang, D.R.; Hill Jr., H. H. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 2699.

45. Brown, D.J.; Stefan, S.E.; Berden, G.; Steill, J.D.; Oomens, J.; Eyler, J.R.; Bendiak, B. *Carbohyd. Res.* **2011**, *346*, 2469.

46. Fenn, L.,S.; McLean, J. A. *Phys. Chem. Chem. Phys*. **2011**, *13*, 2196.

47. Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010

48. Rezaei Darestani R.; Winter, P.; Kitova, E.N.; Tuszynski, J.A.; Klassen, J.S. *J. Am. Soc. Mass Spectrom*. **2016**, *27*, 876.

49. Zhang, H.; Zhang, S.; Tao, G.; Zhang, Y.; Mulloy, B.; Zhan, X.; Chai, W. *Anal. Chem.* **2013**, *85*, 5940.

50. Wheeler, S.F.; Harvey, D.J. *Anal. Chem.* **2000**, *72*, 5027.

51. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1331.

52. Domon, B.; Costello, C.E. *Glycoconjugate J.* **1988**, *5*, 397.

Chapter 3

Screening Natural Libraries of Human Milk Oligosaccharides Against Lectins Using CaR-ESI-MS[†]

3.1 Introduction

Human milk, in addition to being an essential source of nutrition, provides infants with many important health benefits.[1,2] Human milk contains a variety of active components, including proteins, glycoproteins, and fat globules.[3] Every liter of human milk also contains approximately 5 to 25 g of unconjugated oligosaccharides, known as human milk oligosaccharides (HMOs).[4] Studies have shown that HMOs afford health benefits to breast-fed infants through several mechanisms.[5–7] Of the more than two hundred known HMOs, fewer than fifty are available commercially.[8–12] Given the limited availability of purified HMOs, the use of mixtures of HMO, extracted directly from milk, is an attractive alternative for protein–HMO interaction studies. However, there are few screening technologies that are applied readily to natural libraries. One approach, pioneered by Cummings and co-workers, involves the use of shotgun glycan microarrays, which employ fractions of HMOs, purified from human milk, that are chemically modified and immobilized on a solid surface.[13–15] The use of HMO glycan microarrays allows for the rapid profiling of HMO binding properties of lectins and the discovery of new protein–HMO interactions that might be relevant to human health. However, the derivatization of HMOs at the

---

reducing end glucose is a drawback to this approach since the terminal lactose moiety often is implicated in the binding of HMOs to lectins.[16]

Electrospray ionization mass spectrometry (ESI-MS) is an attractive alternative to glycan microarrays for screening carbohydrate libraries against lectins *in vitro,* and it is particularly well suited for the study of HMO interactions since there is no requirement for labeling or immobilization of the oligosaccharides, which may influence their binding properties.[16] Direct detection and quantification of free and ligand-bound proteins ions by ESI-MS enables the binding stoichiometry and affinity of protein–carbohydrate interactions to be established.[17–21] Moreover, because it is possible to monitor multiple binding equilibria simultaneously, ESI-MS also is amenable to screening carbohydrate libraries.[22] In cases where the protein–carbohydrate complexes cannot be detected or reliably quantified, library screening can be performed using a catch-and-release (CaR)-ESI-MS format, whereby ligands are identified following their release, as ions, from protein–ligand complexes upon collisional activation in the gas phase.[22] The assay is rapid, sensitive and label- and immobilization-free and, although not quantitative, can be used to identify the highest affinity ligands and to guide follow-up quantitative binding measurements.[22]

The CaR-ESI-MS assay has been used previously to screen a variety of defined oligosaccharide libraries, including free HMOs, and shown to identify the highest affinity ligands in libraries containing in excess of 200 different components successfully.[23–29] Here, we describe the application of the assay for screening mixtures of HMOs, extracted from pooled human milk, against lectins to identify specific ligands. To our knowledge, this represents the first demonstration of CaR-ESI-MS for screening natural libraries for protein interactions. The main advantage of using natural HMO libraries, compared to the libraries of purified HMOs, is the larger number and diversity of structures that are present, in particular the inclusion of larger oligosaccharides, which are not commercially available. A disadvantage is that it may not be

possible to assign the structures of all the extracted HMOs unambiguously. However, even in cases where the exact structure is not known, the monosaccharide composition of the HMO ligand(s) can be established readily.

In the present study, the feasibility of using the CaR-ESI-MS assay to screen natural libraries of HMOs was demonstrated using a C-terminal fragment of human galectin-3 (hGal-3C), which contains the carbohydrate recognition domain, as a model HMO-binding lectin. The affinities of 32 free (unmodified) HMOs for hGal-3C were measured recently, and these binding data served to validate the majority of the interactions identified by CaR-ESI-MS for the HMO libraries.[16]

3.2 Experimental Section

3.2.1 Proteins

The recombinant fragment of the C-terminus (residues 107–250) of human galectin-3 (hGal-3C, MW 16,330 Da) was a gift from Prof. C. Cairo (University of Alberta). Lysozyme (MW 14,310 Da), which served as the reference proteins (P_{ref}), was purchased from Sigma-Aldrich Canada (Oakville, Canada). Each protein was dialyzed against 200 mM aqueous ammonium acetate (pH 6.8), concentrated using 10 kDa MW cut-off Amicon Ultra-4 centrifugal filters (Millipore Corp, Bedford, MA), and stored at -20 °C until used.

3.2.2 Human milk oligosaccharides

The structures of the pure HMOs are listed in Table 3.1.

Table 3.1. List of MWs, chemical structures, and common names of purified HMOs (**HMO1–HMO31**)

HMO	MW (Da)	Structure	Common name

HMO1	488.17	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc	2'-Fucosyllactose
HMO2	488.17	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	3-Fucosyllactose
HMO3	633.21	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-Glc	3'-Sialyllactose
HMO4	633.21	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-Glc	6'-Sialyllactose
HMO5	634.23	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	Difucosyllactose
HMO6	707.25	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-tetraose
HMO7	707.25	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-neotetraose
HMO8	779.27	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	3'-Sialyl-3'-fucosyllactose
HMO9	853.31	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-fucopentaose I
HMO10	853.31	β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-fucopentaose II
HMO11	853.31	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-fucopentaose III
HMO12	853.31	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc	Lacto-N-neofucopentaose V
HMO13	853.31	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)[α-L-Fuc-(1→3)]-β-D-Glc	Lacto-N-neofucopentaose
HMO14	998.34	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyllacto-N-tetraose a
HMO15	998.34	α-D-Neu5Ac-(2→6)-[β-D-Gal-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- β-D-Glc	Sialyllacto-N-tetraose b
HMO16	998.34	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyllacto-N-tetraose c
HMO17	998.34	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyllacto-N-tetraose d
HMO18	999.34	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-difucohexaose I
HMO19	999.34	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-Glc	Lacto-N-difucohexaose II

HMO20	999.34	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α- L-Fuc-(1→3)]-β-D-Glc	Lacto-N-neodifucohexaose
HMO21	1072.38	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Para Lacto-N-neohexaose
HMO22	1072.38	β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-neohexaose
HMO23	1144.40	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyl monofucosyllacto-N-tetraose
HMO24	1144.40	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Sialyl-lacto-N-fucopentaose V
HMO25	1289.44	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Disialyllacto-N-tetraose
HMO26	1364.50	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc	Difucosyllacto-N-hexaose a
HMO27	1364.50	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-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Difucosyl-para-lacto-N-hexaose
HMO28	1438.29	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal(1→4)-β-D-GlcNAc(1→3)-β-D-Gal(1→4)-β-D-GlcNAc(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-neooctaose
HMO29	545.48	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc	Lacto-N-triaose
HMO30	691.62	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc	Blood group A antigen tetraose type 5
HMO31	1056.96	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-GlcNAc(1→3)-β-D-Gal(1→4)-β-D-Glc	Blood group A antigen hexaose type 1

HMO1 (MW 488.17 Da), HMO2 (MW 488.17 Da), HMO12 (MW 853.31 Da), HMO13 (MW 853.31 Da), HMO17 (MW 998.34 Da), HMO20 (MW 999.36 Da), HMO21 (MW 1072.38 Da), HMO23 (MW 1144.40 Da), HMO24 (MW 1144.40 Da), HMO27 (MW 1364.50 Da), HMO28 (MW 1438.29 Da), HMO29 (MW 545.48 Da), HMO30 (MW 691.62 Da), and HMO31 (MW 1056.96 Da) were purchased from Elicityl SA (Crolles, France); HMO3 (MW 633.21 Da), HMO4 (MW 633.21 Da), HMO5 (MW 634.23 Da), HMO6 (MW 707.25 Da), HMO7 (MW 707.25 Da),

HMO8 (MW 779.27 Da), **HMO9** (MW 853.31 Da), **HMO10** (MW 853.31 Da), **HMO11** (MW 853.31 Da), **HMO14** (MW 998.34 Da), **HMO15** (MW 998.34 Da), **HMO16** (MW 998.34 Da)**, HMO18** (MW 999.36 Da), and **HMO26** (MW 1364.50 Da) were purchased from IsoSep (Tullinge, Sweden); **HMO19** (MW 999.36 Da) and **HMO22** (MW 1072.38 Da) were purchased from Dextra (Reading, UK); **HMO25** (MW 1289.44 Da) was purchased from CarboSynth (Compton, UK). Stock solutions of each HMO were prepared by dissolving a known mass of the oligosaccharide in ultrafiltered water (Milli-Q, Millipore, Billerica, MA) to give a final concentration of 1 mM. All of the stock solutions were stored at -20 °C until used.

3.2.3 HMO fractions

Fourteen HMO fractions (designated as *Fraction 1 – Fraction 14*) were produced from pooled HMOs (pHMOs) originally isolated from human milk pooled from over 50 different donors with term infants. First, pHMOs were separated by charge using anion exchange chromatography on an anion exchange column QAE Sephadex A-25 (Sigma-Aldrich, St. Louis, USA). The HMO charge fractions were separated further by size using Bio-Gel P-2 Gel (Bio-Rad, Hercules, CA, USA). *Fraction 1-Fraction 14* were weighed separately (*Fraction 1* (1.08 mg), *Fraction 2* (0.95 mg), *Fraction 3* (1.77 mg), *Fraction 4* (1.05 mg), *Fraction 5* (0.95 mg), *Fraction 6* (1.33g), *Fraction 7* (1.00 mg), *Fraction 8* (0.55 mg), *Fraction 9* (0.69 mg), *Fraction 10* (1.20 mg), *Fraction 11* (0.51 mg), *Fraction 12* (0.29 mg), *Fraction 13* (2.11mg), and *Fraction 14* (0.85 mg)) and were dissolved in 40 mL Milli-Q water. Following a 100-fold dilution with Milli-Q water, each of the stock solutions was stored at -20 °C until used. Given that, based on ESI-MS analysis, some of the fractions had similar compositions, *vide infra*, these were pooled to give four new fractions, designated as: *Fr1* (*Fraction 1*), *Fr2* (*Fractions 2–5*), *Fr3* (*Fractions 6–10*), *Fr4* (*Fractions 11–14*).

3.2.4 Mass spectrometry

All CaR-ESI-MS measurements were carried out in negative ion mode using a Synapt G2 ESI quadrupole-ion mobility separation-time-of-flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK), equipped with a nanoflow ESI (nanoESI) source. NanoESI tips were produced in-house from borosilicate capillaries (1.0 mm o.d., 0.78 mm i.d.), pulled to ~5 μm outer-diameter using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). To perform nanoESI, a platinum wire was inserted into the nanoESI tip, and a voltage of -1.0 kV was applied. A Cone voltage of 25 V was used, and the source block temperature was maintained at 50 ºC. For the CaR-ESI-MS measurements, a Trap voltage of 5 V and Transfer voltage between 10 and 80 V were used. Argon was used in Trap and Transfer ion guides at pressures of 2.22 x 10^{-2} and 3.36 x 10^{-2} mbar, respectively. The helium chamber preceding the traveling wave IMS (TWIMS) device was maintained at 7.72 mbar. The IMS parameters, which were optimized for each HMO MW, were: 2 mL min^{-1} Trap gas flow rate; 150 to 180 mL min^{-1} helium cell gas flow rate; 50 to 90 mL min^{-1} ion mobility gas flow rate; 50 V Trap voltage; 400 to 1000 m s^{-1} IMS wave velocity; 15 to 40 V IMS wave height. All IMS measurements were carried out using nitrogen as the mobility gas at a pressure of 3.41 mbar. Data acquisition and processing were carried out using MassLynx (v4.1). The quantitative affinity measurements were carried out in positive ion mode using a Synapt G2S quadrupole-ion mobility separation-time-of-flight (Q-IMS-TOF) mass spectrometer (Waters UK Ltd., Manchester, UK). Details of the instrumental conditions used and data analysis are given as Supporting Information.

3.3 Results and Discussion

A two-step approach was used to identify HMO ligands, present in the fractions, for hGal-3C (Figure 3.1).
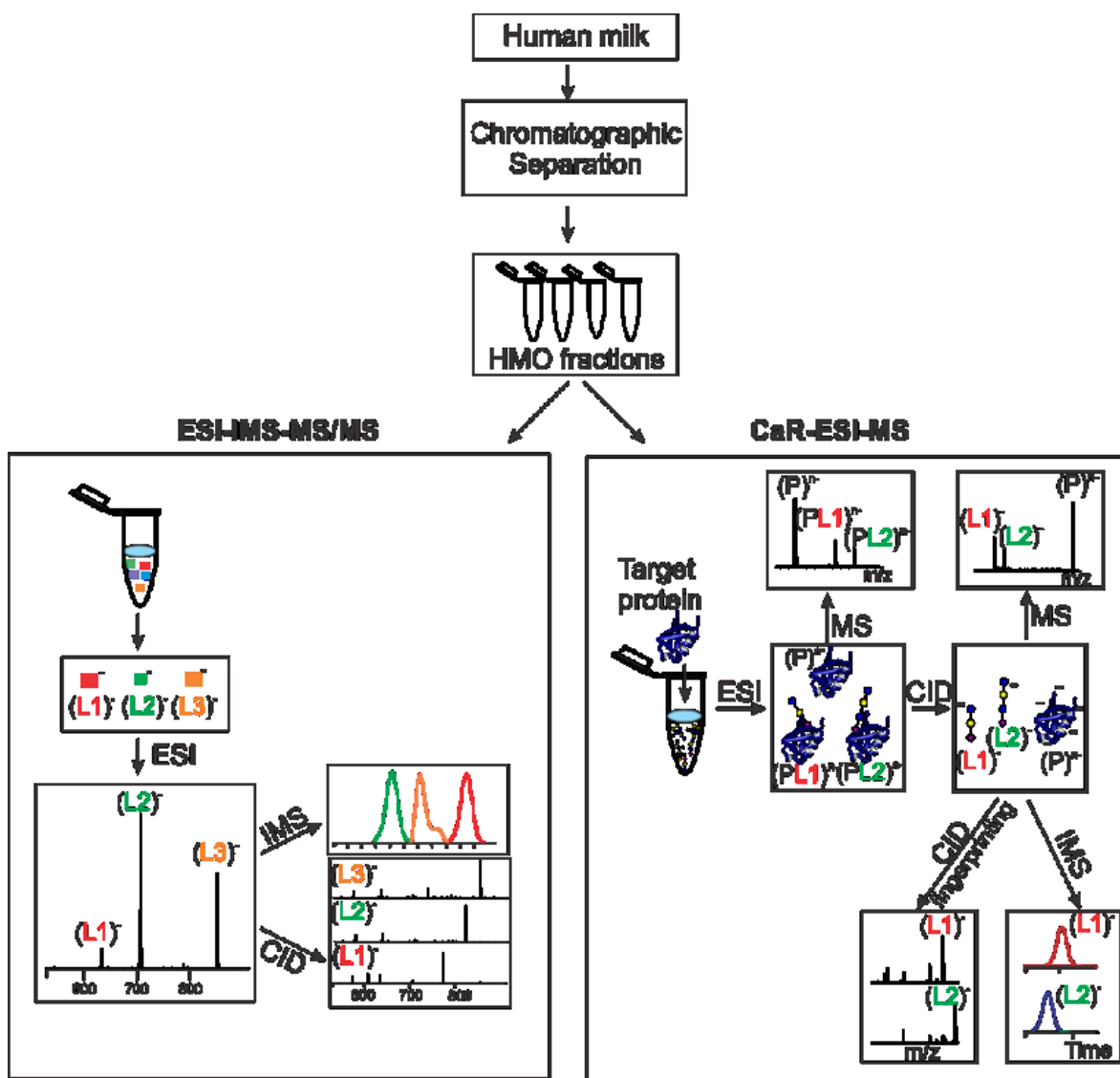
Figure 3.1. Overview of the two-step approach (ESI-IMS-MS/MS and CaR-ESI-MS) for screening HMO mixtures, extracted from human milk, against lectins.

First, each of the fractions was analyzed by ESI-IMS-MS/MS in order to identify, as much as possible, the HMOs present. This was accomplished by comparing the MWs, IMS-ATs, and CID fingerprints of ions detected from the fractions, with the MWs of the most abundant oligosaccharides found in human milk,[8–10] the IMS-ATs, and CID fingerprints recently reported for a library of 31 purified HMOs.[29] In cases where appropriate HMO standards were not available,

possible HMO structures were suggested based on the CID fragmentation data and the structures known to be present in human milk. To our knowledge, the present work represents the first demonstration of using IMS-ATs and CID fingerprinting to identify HMOs in mixtures extracted from human milk. Following the characterization of the fractions, they were screened against the lectin using CaR-ESI-MS and ligands identified from a comparison of the MWs, IMS ATs, and CID fingerprints of the released and free HMOs.

3.3.1 ESI-MS analysis of HMO fractions *Fr1–Fr4*

Figures 3.2–3.7 are representative ESI mass spectra and IMS arrival time distributions (ATDs), acquired in negative ion mode, for aqueous solutions of each of the four fractions.

Figure 3.2. Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of *Fr1* (0.05 μg μL^{-1}).

Figure 3.3. IMS-ATDs of deprotonated HMO ions produced from *Fr1*: (a) m/z 608.14; (b) m/z 632.21; (c) m/z 681.16; (d) m/z 754.18; (e) m/z 790.67; (f) m/z 836.25; (g) m/z 852.31; (h) m/z 863.70; (i) m/z 998.34; (j) m/z 1055.96; (k) m/z 1071.38; (l) m/z 1201.30; (m) m/z 1217.46; and (n) m/z 1363.25. (o) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 852.31.

Figure 3.4. Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of *Fr2* (0.05 μg μL^{-1}).

Figure 3.5. Representative IMS-ATDs acquired in negative ion mode for aqueous ammonium acetate solution (40 mM, pH 6.8) of *Fr2* (0.05 μg μL^{-1}) for deprotonated HMO ions at (a) m/z

608.12; (b) m/z 690.62; (c) m/z 706.25; (d) m/z 836.25; (e) m/z 852.31; and (f) m/z 998.34. (g)

CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z

852.31.

Figure 3.6. (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous

ammonium acetate solution (20 mM, pH 6.8) of *Fr3* (0.05 μg μL^{-1}). IMS-ATDs measured for

deprotonated HMO ions with (b) m/z 341.31, (c) m/z 487.17, and (d) m/z 706.25.

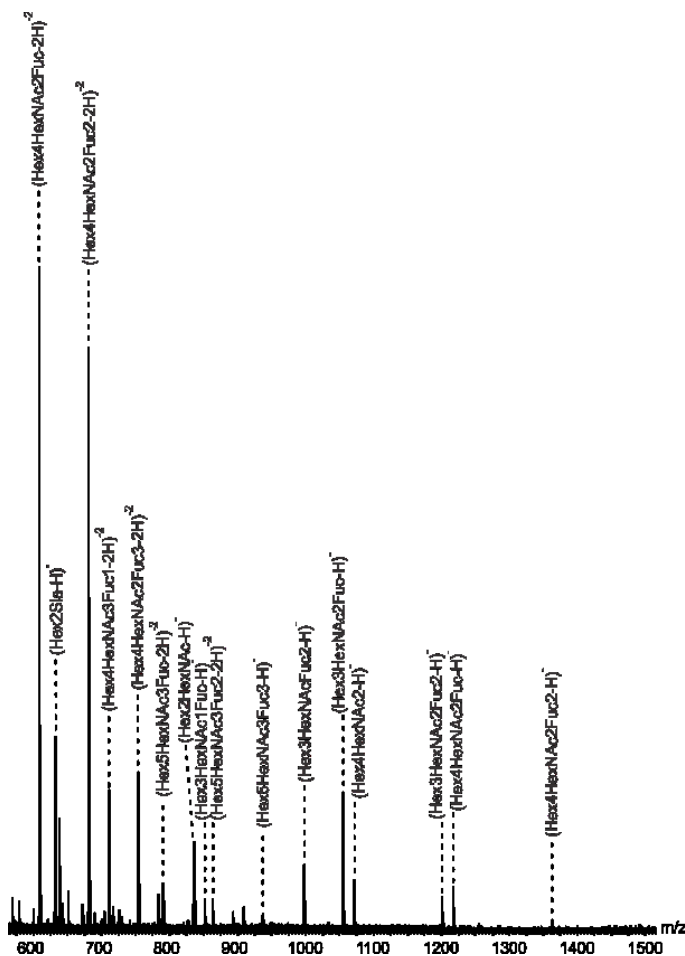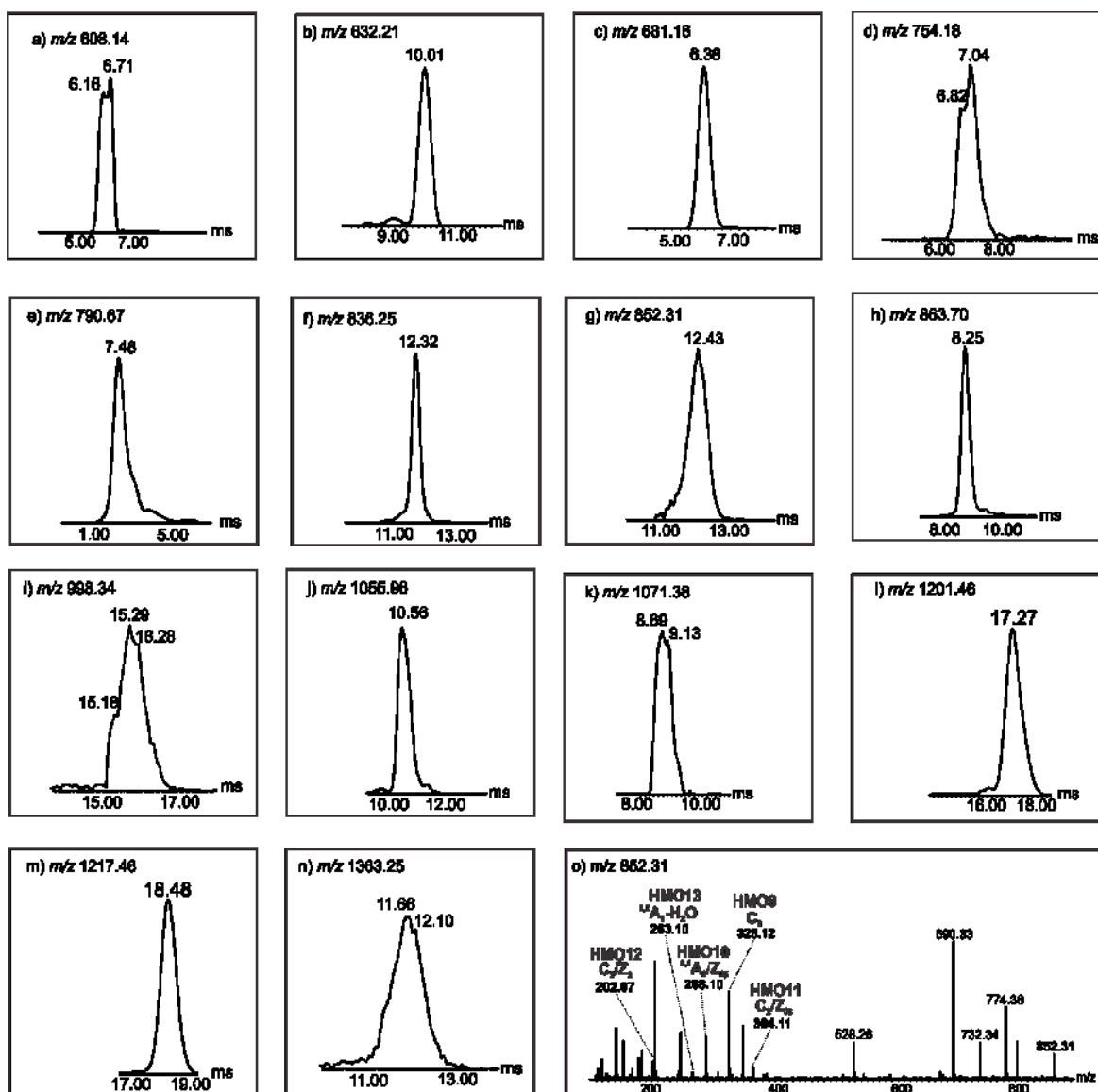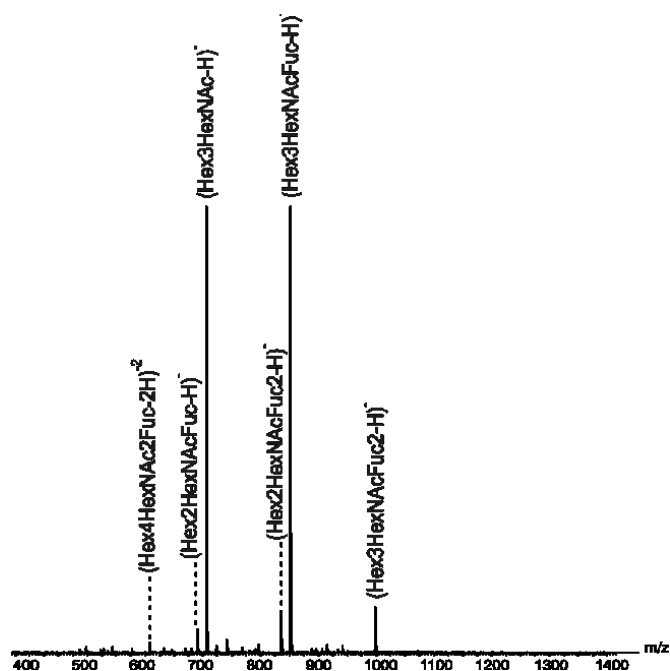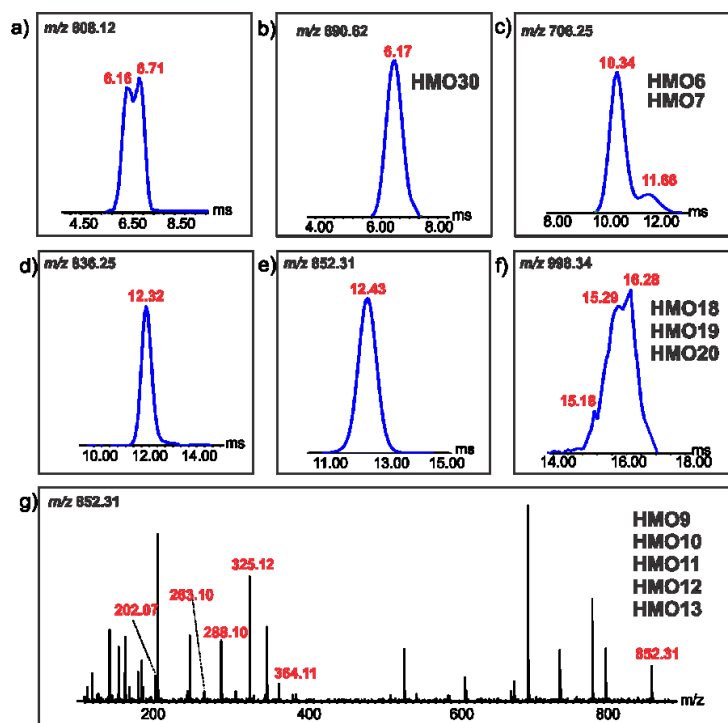Figure 3.7. (a) Representative ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (20 mM, pH 6.8) of *Fr4* (0.05 μg μL⁻¹). IMS-ATDs measured for deprotonated HMO ions with (b) m/z 632.21; (c) m/z 643.44; (d) m/z 753.67; (e) m/z 826.17; (f) m/z 835.20; and (g) m/z 899.18. (i) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 997.34.

Singly and doubly deprotonated ions corresponding to 12 HMO MWs (633.21 Da, 837.25 Da, 853.31 Da, 999.34 Da, 1056.96 Da, 1072.38 Da, 1202.30 Da, 1218.28 Da, 1364.50 Da, 1510.32 Da, 1583.46 Da, and 1729.64 Da) were identified in *Fr1*, six MWs (691.62 Da, 707.25 Da, 837.25 Da, 853.31 Da, 999.34 Da, and 1218.28 Da) in *Fr2*, two in *Fr3* (488.17 Da and 707.25 Da), as well as lactose, and seven in *Fr4* (633.21 Da, 836.20 Da, 998.34 Da, 1289.44 Da, 1509.32 Da, 1655.33 Da, and 1800.44 Da). The monosaccharide compositions of the 21 different HMO MWs, which contain between three and 10 monosaccharide units, were identified, shown in Table 3.2.

Table 3.2. MW and monosaccharide composition (Hex ≡ Glc or Gal; HexNAc ≡ GlcNAc; Fuc ≡ fucose; and Sia ≡ sialic acid) of HMOs identified from ESI-MS analysis of aqueous solutions of *Fr1–Fr4*. The identity of specific HMO structures was based on a comparison of IMS-ATs and CID fingerprints of deprotonated ions produced from *Fr1–Fr4* and those of **HMO1–HMO31**

Fraction	Theoretical MW	Monosaccharide composition	Confirmed/Putative HMO Structures
Fr1	633.21	Hex2Sia	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO3**) α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-Glc (**HMO4**)
	837.25	Hex2HexNAcFuc2	α-L-Fuc-(1→3/4)-β-D-GlcNAc-(1→3/6)-[α-L-Fuc-(1→2/3)]-β-D-Gal-(1→4)-β-D-Glc [a]
	853.31	Hex3HexNAcFuc	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO9**) β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO10**) β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO11**) β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc (**HMO12**) β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)[α-L-Fuc-(1→3)]-β-D-Glc (**HMO13**)
	999.34	Hex3HexNAcFuc2	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO18**) 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-Glc (**HMO19**) β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc (**HMO20**)
	1056.96	Hex3HexNAc2Fuc	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-GlcNAc(1→3)-β-D-Gal(1→4)-β-D-Glc (**HMO31**)

1072.38	Hex4HexNAc2	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO21**) β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)]-β-D-Glc-(1→4)-Glc (**HMO22**)
1202.30	Hex3HexNAc2Fuc2	β-D-GlcNAc-(1→3/4)-[L-Fuc-(1→2)]-β-D-Gal-(1→3)-[L-Fuc-(1→4)]-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc [a]
1218.28	Hex4HexNAc2Fuc	β-D-Gal-(1→3/4)-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→3/4)-[L-Fuc-(1→3/4)]-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc [a]
1364.50	Hex4HexNAc2Fuc2	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc (**HMO26**) 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-(1→3)]-β-D-Gal-(1→4)-β-D-Glc (**HMO27**)
1510.32	Hex4HexNAc2Fuc3	α-L-Fuc-(1→2)-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]-β-D-Gal-(1→4)-β-D-Glc [a]
1583.46	Hex5HexNAc3Fuc	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc [a] β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc [a]
1729.64	Hex5HexNAc3Fuc2	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→3)-β-D-GlcNAc]-β-D-Gal-(1→4)-β-D-Glc [a] 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-(1→6)-[β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc [a]

			β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc [a]
			β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc [a]
	691.62	Hex2HexNAcFuc	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc (**HMO30**)
	707.25	Hex3HexNAc	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO6**) β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO7**)
	837.25	Hex2HexNAcFuc2	β-D-GlcNAc-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-Gal-(1→4)-β-D-Glc[a] α-D-Neu5Ac-(2→6)-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc[a]
Fr2	853.31	Hex3HexNAcFuc	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO9**) β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO10**) β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO11**) β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc (**HMO12**) β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)[α-L-Fuc-(1→3)]-β-D-Glc (**HMO13**)
	999.34	Hex3HexNAcFuc2	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO18**) 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-Glc (**HMO19**) β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α- L-Fuc-(1→3)]-β-D-Glc (**HMO20**)
	1218.28	Hex4HexNAc2Fuc	β-D-Gal-(1→3/4)-β-D-GlcNAc (1→3/6)-β-D-Gal-(1→3/4)-[L-Fuc-

			(1→3/4)]-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc
Fr3	342.31	Hex2	β-D-Gal-(1→4)-β-D-Glc (**Lactose**)
	488.17	Hex2Fuc	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc (**HMO1**) β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc (**HMO2**)
	707.25	Hex3HexNAc	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO6**) β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO7**)
Fr4	633.21	Hex2Sia	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO3**) α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-Glc (**HMO4**)
	836.20	Hex2HexNAcSia	β-D-GlcNAc-(1→3/4)-[α-D-Neu5Ac-(2→6)]-β-D-Gal-(1→4)-β-D-Glc [a] α-D-Neu5Ac-(2→6)-)-β-D-GlcNAc-(1→3/4)-β-D-Gal-(1→4)-β-D-Glc [a]
	998.34	Hex3HexNAcSia	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO14**) α-D-Neu5Ac-(2→6)-[β-D-Gal-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- β-D-Glc (**HMO15**) α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO16**) α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO17**)
	1289.44	Hex3HexNAcSia2	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO25**)
	1509.32	Hex4HexNAc2FucSia	Neu5Ac-(2→6)-β-D-Gal-(1→3/6)-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→3)-[L-Fuc-(1→3/4)]-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc [a]

1655.33	Hex4HexNAc2Fuc2Sia	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-[α-L-Fuc-(1→2)-β-D-Gal-(1→4)-α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-β-D-Gal-(1→4)-β-D-Glc [a] α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-[α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)]-β-D-GlcNAc-(1→3)-]-β-D-Gal-(1→4)-β-D-Glc [a]
1800.44	Hex4HexNAc2FucSia2	α-D-Neu5Ac-(2→6)-[α-D-Neu5Ac-(2→6) -β-D-Gal-(1→3)-]-β-D-GlcNAc-(1→4)-[α-L-Fuc-(1→2)-]-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc [a]

a. Possible structures based on CID results and identified HMOs reported in references 13 and 14.

Some of the MWs were found in more than one fraction (e.g. 707.25 was detected in both *Fr2* and *Fr3*). Of the detected MWs, 14 correspond to neutral HMOs (12 of these are fucosylated), with the remaining seven corresponding to acidic HMOs (three are fucosylated).

The deprotonated ions (singly or doubly charged) associated with each HMO MW were subjected to IMS analysis and CID fingerprinting, and the results compared to those measured for **HMO1–HMO31**.[29] As an example, the IMS and CID data acquired for the HMOs in *Fr2* are shown in Figure 3.5; the corresponding results obtained for *Fr1*, *Fr3,* and *Fr4* are shown in Figures 3.3, 3.6, and 3.7.

Inspection of Figure 3.5 reveals that the IMS-ATDs measured for deprotonated ions corresponding to MWs of 691.62 Da, 837.25 Da, and 853.31 Da exhibit single features, with ATs of 6.17 ms (Figure 3.5b), 12.32 ms (Figure 3.5d), and 12.43 ms (Figure 3.5e), respectively. For the MWs 707.25 Da, 999.34 Da, and 1218.28 Da, their IMS-ATDs consisted of two and three features, respectively, with ATs of 10.34 Da and 11.66 Da (for 707.16 Da); 15.18 ms, 15.29 ms, and 16.28 ms (for 999.34 Da); and 6.16 ms and 6.71 ms (for 1218.28 Da) (Figures 3.5c, f, and a, respectively).

The IMS-ATD measured for 853.31 Da has a broad distribution (FWHM 1.2 ms) centred at 12.43 ms, consistent with the presence of multiple isomers (Figure 3.5e).[29]

Four of the measured MWs (691.62 Da, 707.25 Da, 999.34 Da, and 853.31 Da) present in *Fr2* coincide with those of HMOs in the 31 component library (**HMO1–HMO31**). Comparison of the IMS-ATDs reveals that the ion corresponding to MW 691.62 Da, which has an AT of 6.17 ms, matches that of **HMO30** (Figure 3.5b). The two features (with ATs of 10.34 ms and 11.66 ms) observed for the ions corresponding to MW 707.25 Da match the IMS-ATs measured for **HMO6** and **HMO7** (Figure 3.5c). The three partially resolved features (with ATs of 15.18 ms, 15.29 ms, and 16.28 ms) measured for the ions with MW 999.34 Da match those of **HMO18**, **HMO19,** and **HMO20**, respectively (Figure 3.5f). Additionally, the CID mass spectrum acquired for deprotonated ions of MW 853.31 Da revealed the presence of unique HMO fragments arising from five different HMOs, C_2 (m/z 325.12) from **HMO9**, $^{0,4}A_2/Z_{3\beta}$ (m/z 288.10) from **HMO10,** $C_2/Z_{3\beta}$ (m/z 364.11) from **HMO11,** C_2/Z_3 (m/z 202.07) from **HMO12**, and $^{0,2}A_2–H_2O$ (m/z 263.10) from **HMO13** (Figure 3.5g).

Following the same approach, the other three fractions were analyzed, and the HMOs identified are listed in Table 3.2. Together, the four fractions are found to contain at least 35 different HMOs, corresponding to 21 different MWs. The structures of 25 HMOs (corresponding to 11 different MWs) were established from the measured IMS-ATs and CID fingerprints and comparison with those of purified HMOs. Of these, 18 are neutral HMOs (of which 14 are fucosylated) and seven are acidic HMOs. For the other MWs, only monosaccharide compositions could be determined. However, it was possible to suggest HMO structures based on previously identified HMOs having the same monosaccharide compositions.[8,9] In certain cases, it was possible to reduce the number of structures based on an analysis of the fragment ions produced by CID (Figures 3.8–3.17).

Figure 3.8. CID mass spectrum acquired for deprotonated HMO ions at m/z 835.20. Fragmentation scheme shown for (b) β-GlcNAc-(1→3)-[α-Neu5Ac-(2→6)]-β-Gal-(1→4)-β-Glc and (c) α-Neu5Ac-(2→6)-β-GlcNAc-(1→3/6)-β-Gal-(1→4)-β-Glc.

Figure 3.9. CID mass spectrum acquired for deprotonated HMO ions at m/z 836.25. (b) Fragmentation scheme shown for α-L-Fuc-(1→3)-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc.
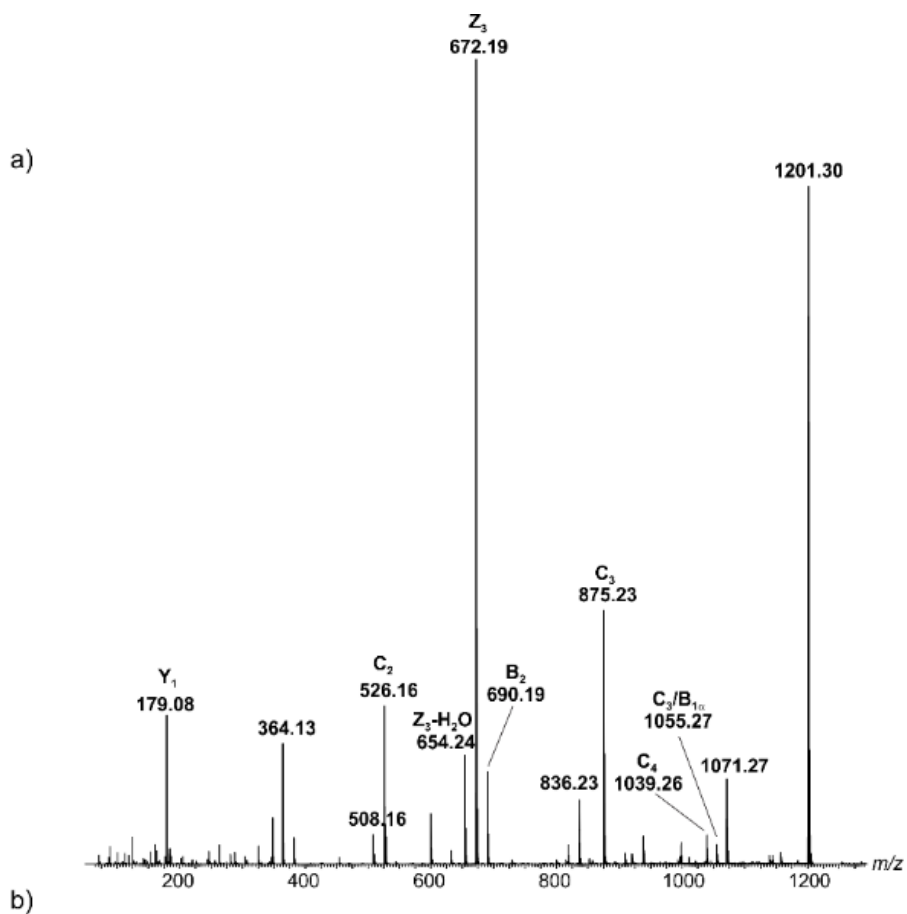
Figure 3.10. CID mass spectrum acquired for deprotonated HMO ions at m/z 1201.30. (b) Fragmentation scheme shown for β-D-GlcNAc-(1→3)-[L-Fuc-(1→3)]-β-D-Gal-(1→4)-[L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-β-D-Gal-(1→4)-β-D-Glc.

Figure 3.11. CID mass spectrum acquired for deprotonated HMO ions at m/z 1217.28. (b) Fragmentation scheme shown for β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-β-D-Gal-(1→4)-β-D-Glc.

Figure 3.12. CID mass spectrum acquired for deprotonated HMO ions at m/z 1508.32. (b) Fragmentation scheme shown for α-D-Neu5Ac-(2→6)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-β-D-Gal-(1→4)-β-D-Glc.

Figure 3.13. CID mass spectrum acquired for deprotonated HMO ions at m/z 1509.32. (b) Fragmentation scheme shown for α-L-Fuc-(1→3)-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-(1→6)-]-β-D-Gal-(1→4)-β-D-Glc.
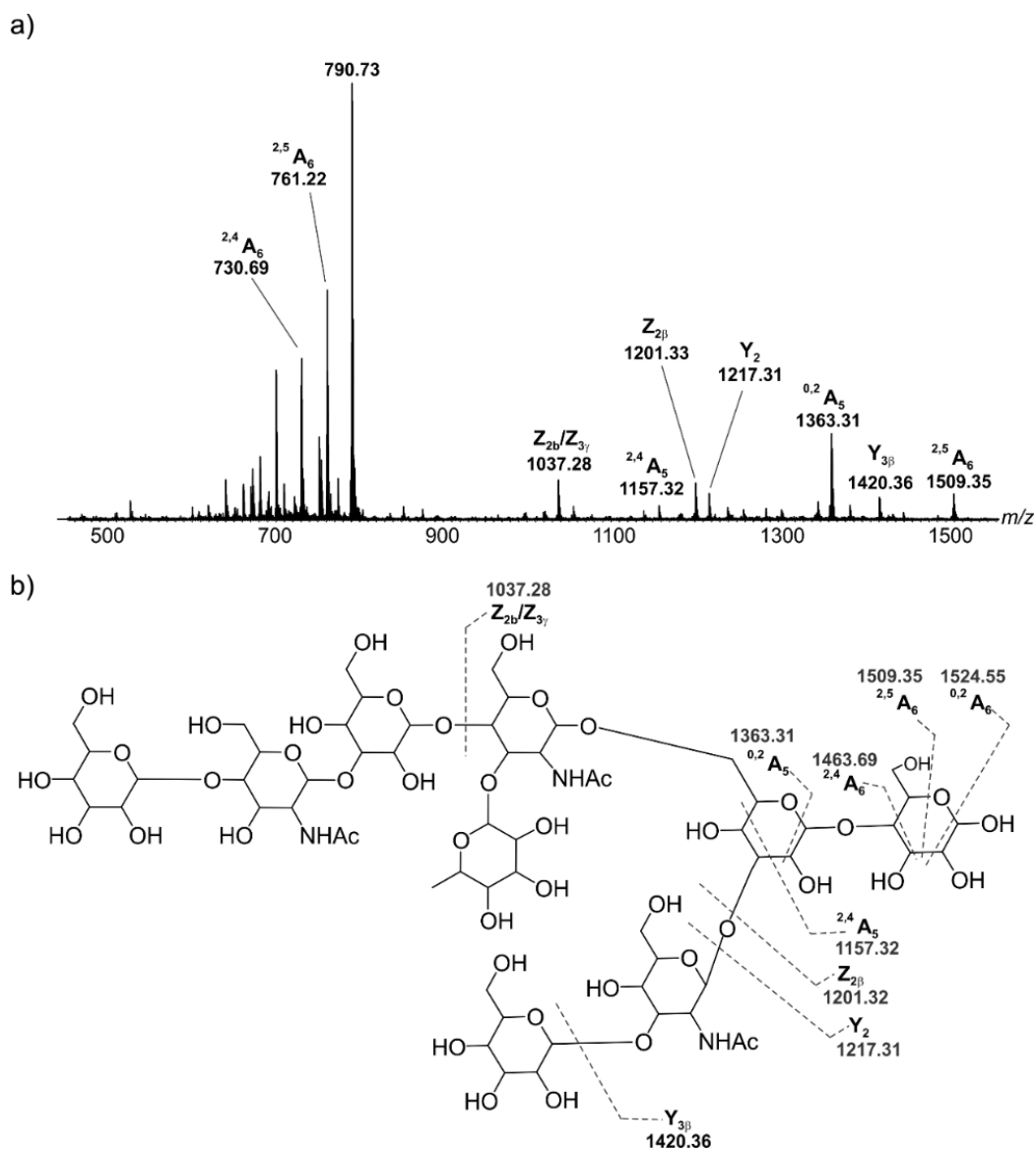
Figure 3.14. CID mass spectrum acquired for deprotonated HMO ions at m/z 790.73. (b) Fragmentation scheme shown for β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-]-β-D-Gal-(1→4)-β-D-Glc.
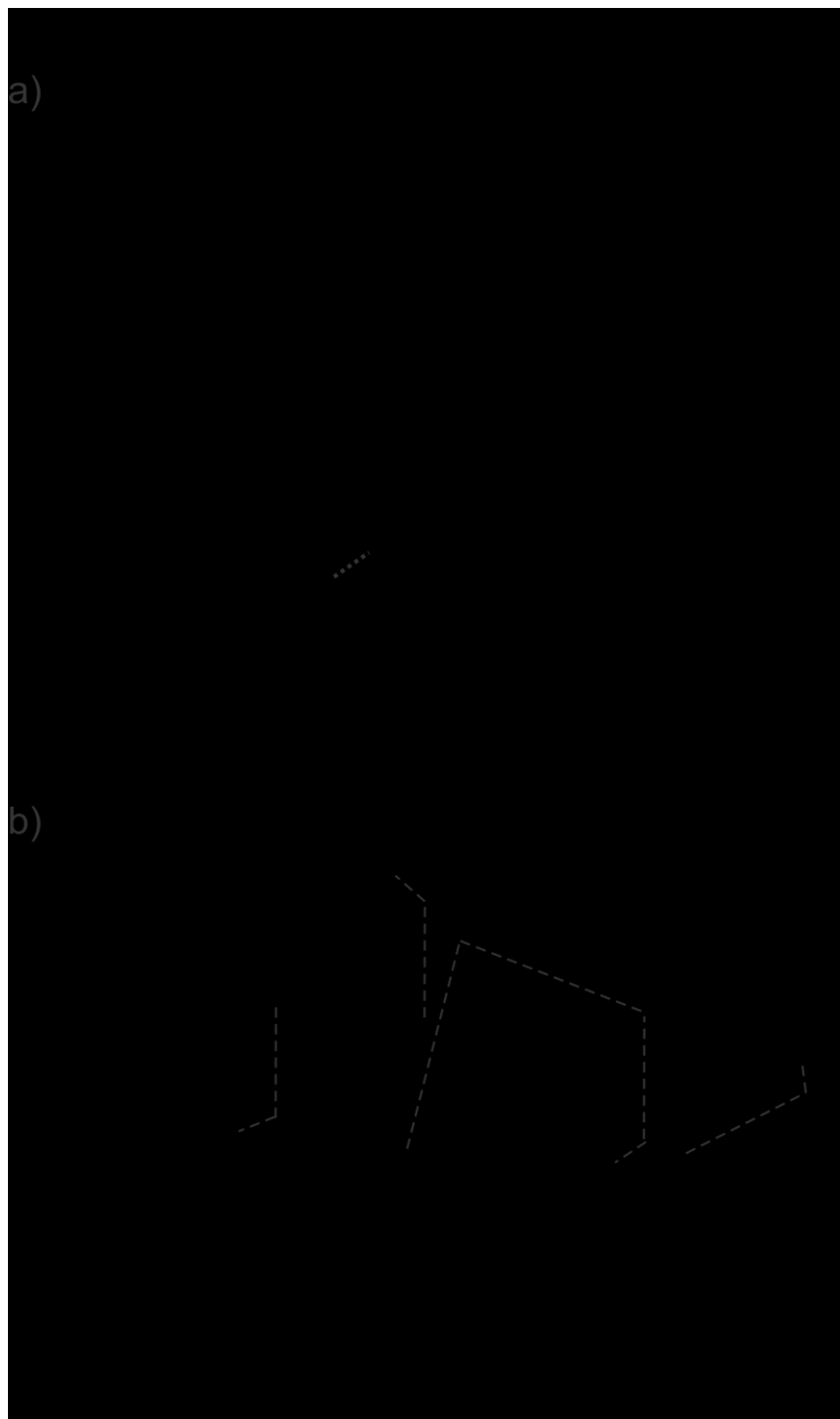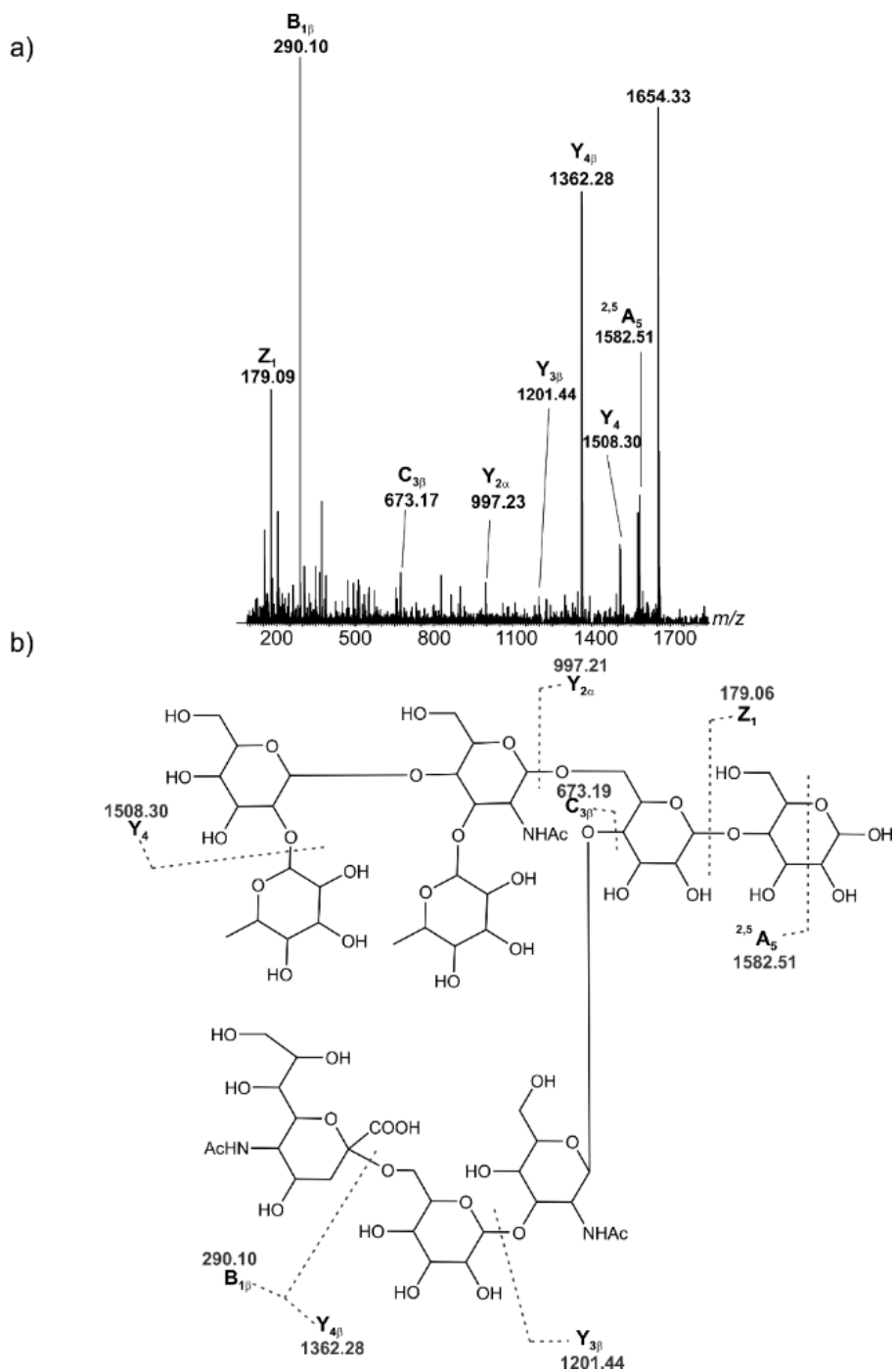
Figure 3.15. CID mass spectrum acquired for deprotonated HMO ions at m/z 1654.33. (b) Fragmentation scheme shown for L-Fuc-(1→2)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)-]-β-D-GlcNAc-(1→6)-[α-D-Neu5Ac-(2→6)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→4)-]-β-D-Gal-(1→4)-β-D-Glc.
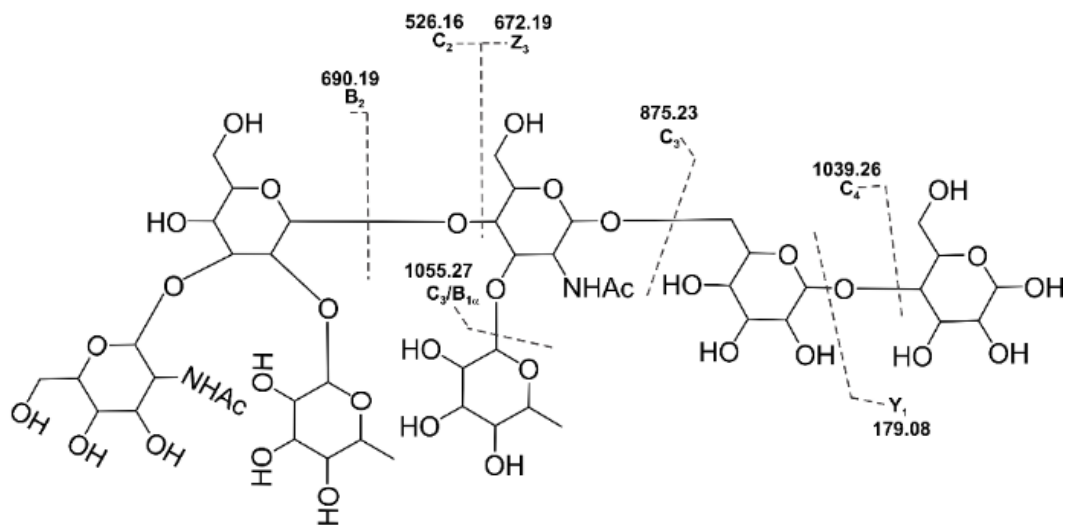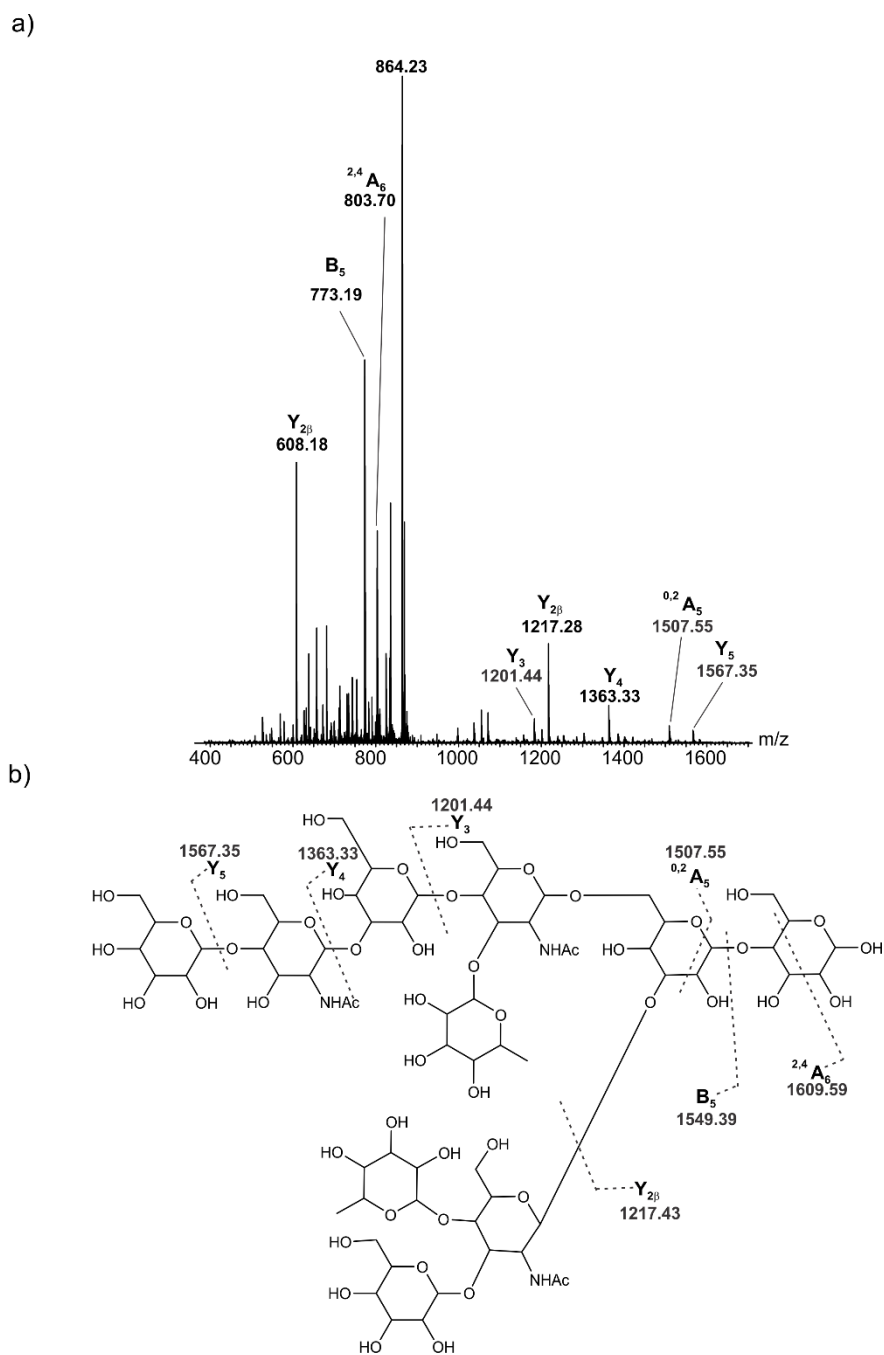
Figure 3.16. CID mass spectrum acquired for deprotonated HMO ions at m/z 864.23. (b) Fragmentation scheme shown for β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→3)-[α-L-Fuc-(1→4)-]-β-D-GlcNAc-(1→3)-]-β-D-Gal-(1→4)-β-D-Glc.
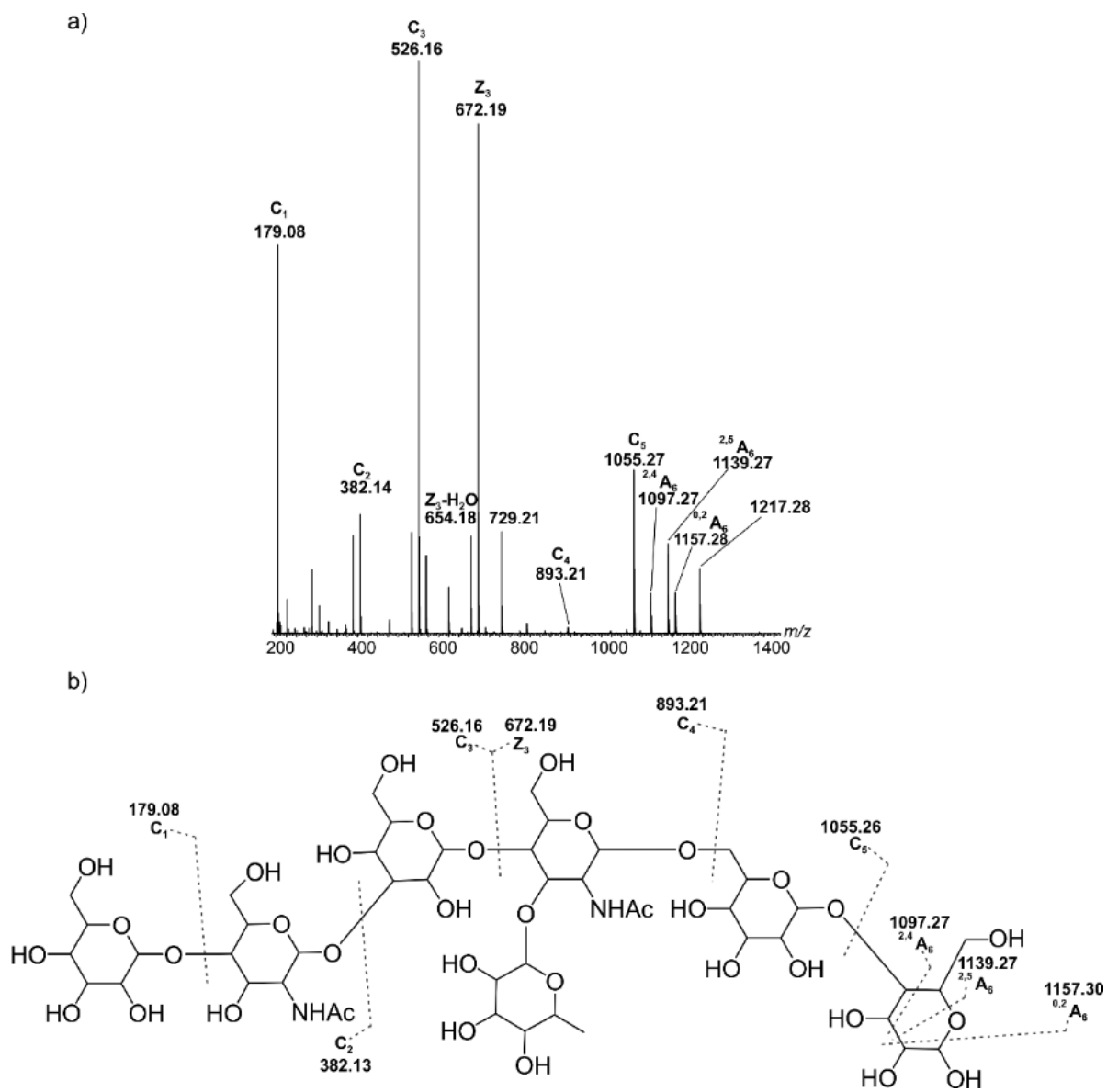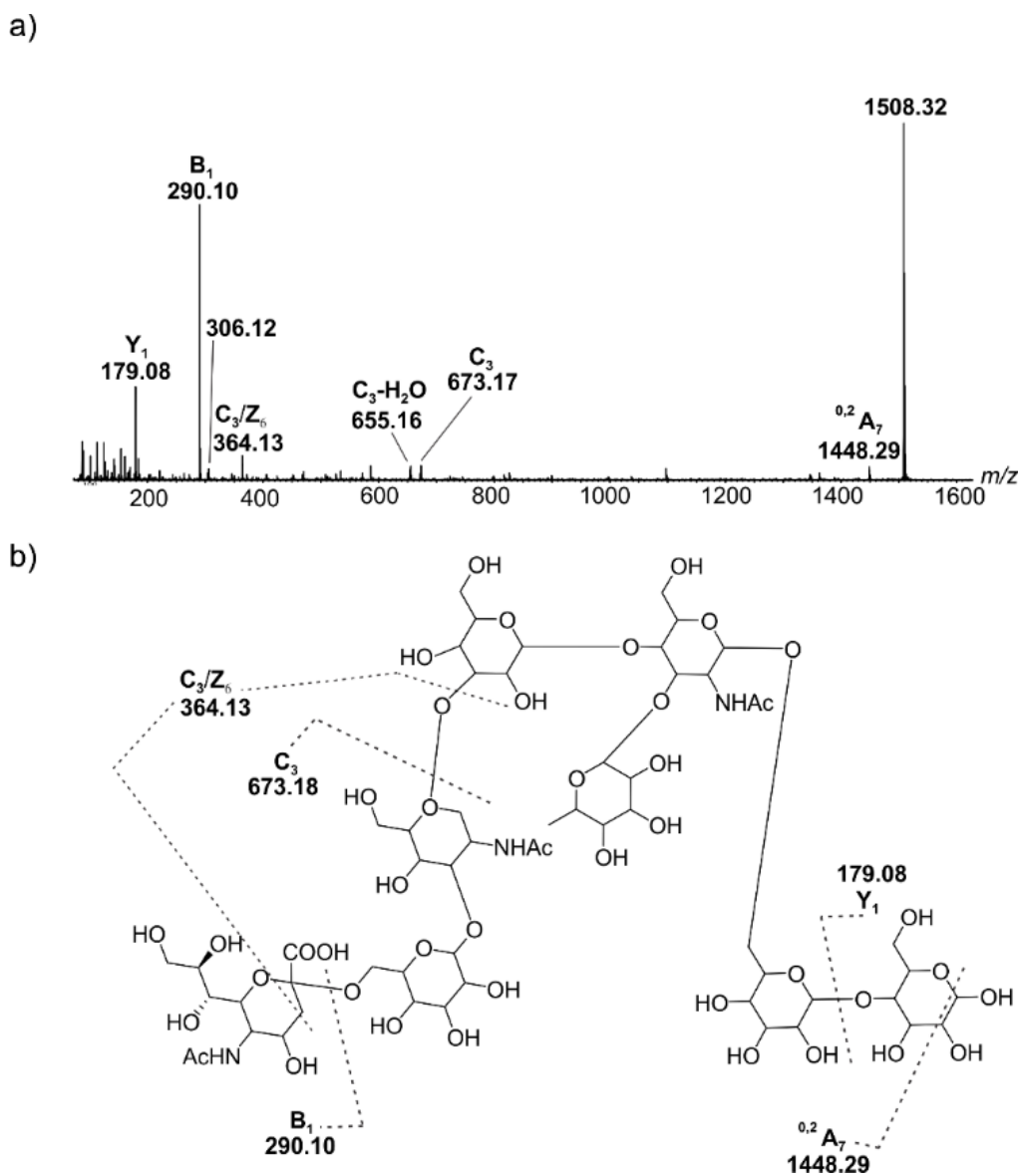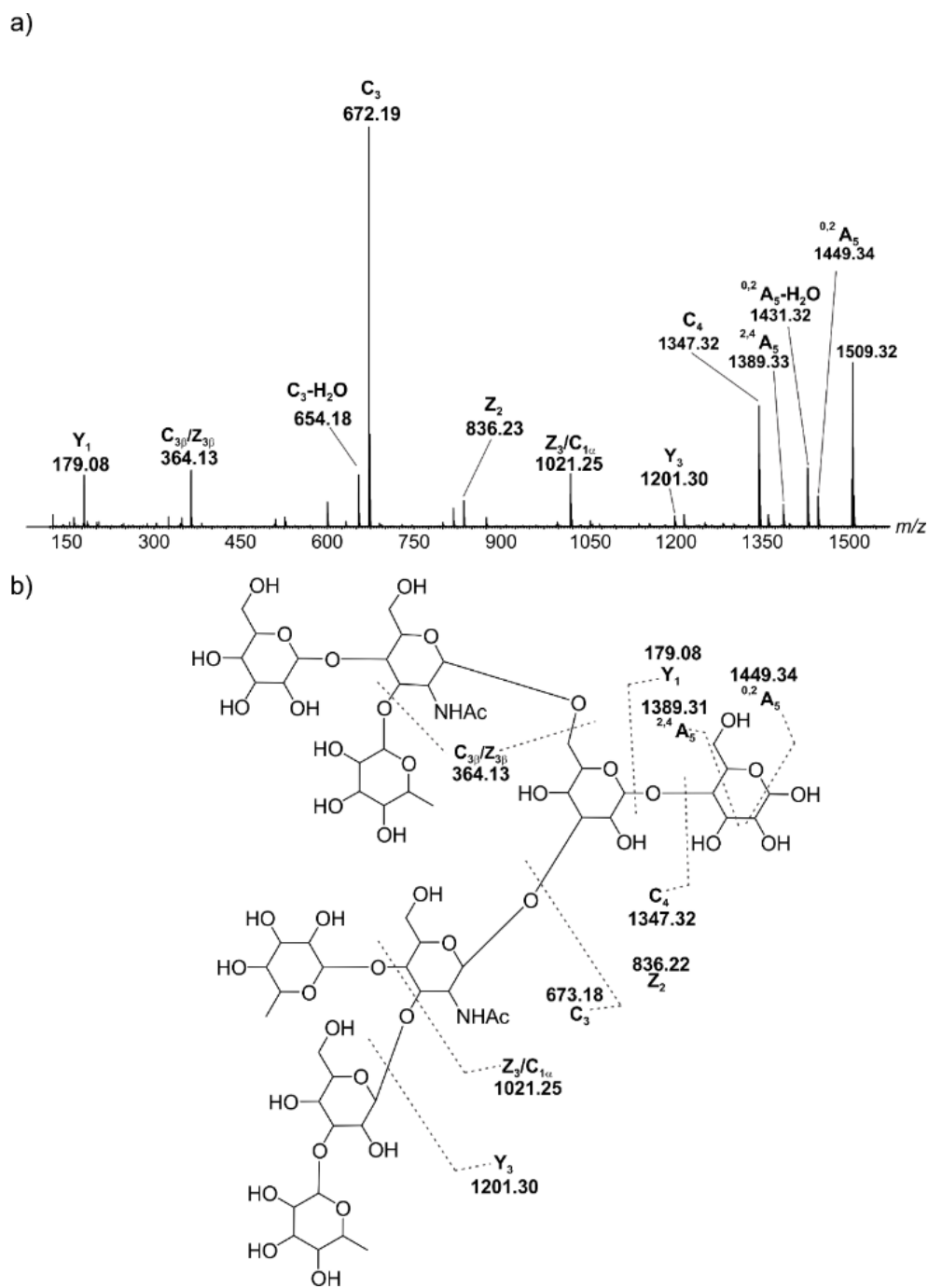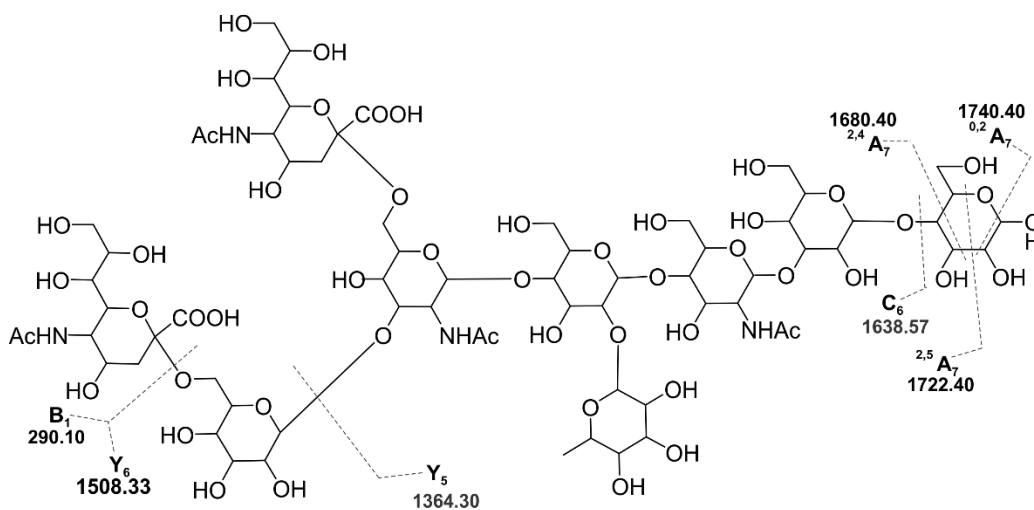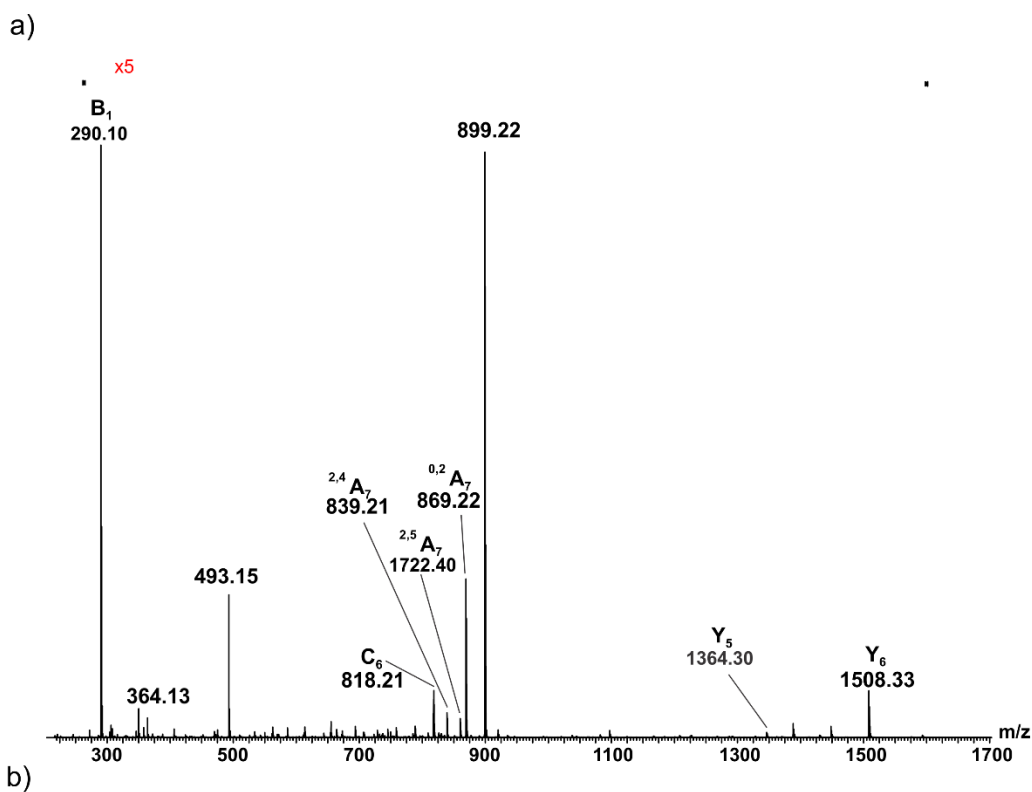
Figure 3.17. CID mass spectrum acquired for deprotonated HMO ions at m/z 899.22. (b) Fragmentation scheme shown for α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-[α-D-Neu5Ac-(2→6)-]-β-D-GlcNAc-(1→4)-[L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc.

For example, the CID of deprotonated HMOs with MW 836.25 Da (Hex2HexNAcSia) produced fragment ions at m/z 306.14, m/z 470.14, and m/z 493.15 (Figure 3.8), which are indicative of α2→6 linked sialic acid, α-Neu5Ac-(2→6)-β-Gal, and α-Neu5Ac-(2→6)-β-GlcNAc, respectively.[30] These data, taken together with the previously identified HMOs with Hex2HexNAcSia composition, suggest the presence of both β- GlcNAc-(1→3)-[α-Neu5Ac-(2→6)]-β-Gal-(1→4)-β-Glc and α-Neu5Ac-(2→6)-β-GlcNAc-(1→3/6)-β-Gal-(1→4)-β-Glc. It is worth mentioning that those are possible structures, but the exact structure could only be obtained by the integration of the fragmentation data with other tools as exoglycosidases

3.3.2 Screening HMO Fractions against hGal-3C

Having established the HMO compositions of *Fr1–Fr4*, each fraction was screened against hGal-3C. Shown in Figure 3.18a is a representative ESI mass spectrum acquired in the negative ion mode for aqueous ammonium acetate solutions (40 mM, pH 6.8) of hGal-3C (15 μM), P$_{ref}$ (5 μM), and *Fr2* (0.05 μg μL^{-1}). Signals corresponding to hGal-3C bound to HMOs with five different MWs (691.62 Da, 707.25 Da, 836.25 Da, 853.31 Da, and 999.34 Da) were detected.

Figure 3.18. (a) Representative ESI mass spectrum acquired in the negative ion mode for 40 mM aqueous ammonium acetate solutions (pH 6.8) of P_{ref} (5 μM), *Fr2* (0.05 μg μL^{-1}), and hGal-3C (15 μM), (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at a Trap voltage of 40 V showing the released HMOs ligands; IMS-ATDs of (c) m/z 690.62; (d) m/z 706.25; (e) m/z 836.25; (f) m/z 852.31; and (g) m/z 998.34. (h) CID mass spectrum acquired for released HMO anions with IMS-AT of 12.43 ms using a Transfer voltage of 30 V.

100

The absence of signals corresponding to HMO-bound P_{ref} ions indicates that nonspecific HMO–hGal-3C binding during the ESI process was negligible.[21] From the MWs of the detected hGal-3C-HMO complexes, it was possible to identify only one of the HMO ligands (**HMO30**, MW 691.62 Da); each of the other detected MWs could, in principle, correspond to multiple HMOs.

To identify the other HMO ligands, the (hGal-3C + HMO) complexes at the -7 charge state, the most abundant charge state detected, were isolated using the quadrupole mass filter (set to pass a range of ~200 m/z ions), and subjected to collisional activation in the Trap region. The selected Trap voltage (40 V) allowed for the efficient release of the HMOs from hGal-3C without causing significant secondary fragmentation. Signals corresponding to deprotonated HMO ions (as well as chloride adducts) of the five different MWs were detected (Figure 3.18b). From comparison of the IMS-ATDs and CID fingerprints of the released HMO ions with available data for purified HMOs of the same MW, nine HMO ligands (**HMO6**, **HMO7**, **HMO9**, **HMO10**, **HMO11**, **HMO12**, **HMO13**, **HMO18,** and **HMO30**) were identified positively (Figure 3.18c–h) in addition to the HMO with MW 836.25 Da.

Similar analysis of the CaR-ESI-MS data acquired for the other three fractions with hGal-3C was performed (Figures 3.19–3.21).

Figure 3.19. (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (20 mM, pH 6.8) of P_{ref} (5 μM), *Fr1* (0.05 μg μL^{-1}), and hGal-3C (15 μM). (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at Trap voltage 40 V. IMS-ATDs of released HMO ions at (c) m/z 632.21, (d) m/z 836.25, (e) m/z 852.31, (f) m/z 998.34, (g) m/z 1055.32, (h) m/z 1071.38, (i) m/z 1217.46, and (j) m/z 1363.25. (k) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 852.31.

Figure 3.20. (a) Representative ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (20 mM, pH 6.8) of P_{ref} (5 μM), *Fr3* (0.05 μg μL^{-1}), and hGal-3C (15 μM), (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at Trap voltage 40 V. IMS-ATDs of released HMO ions at (c) m/z 487.17 and (d) m/z 706.25.
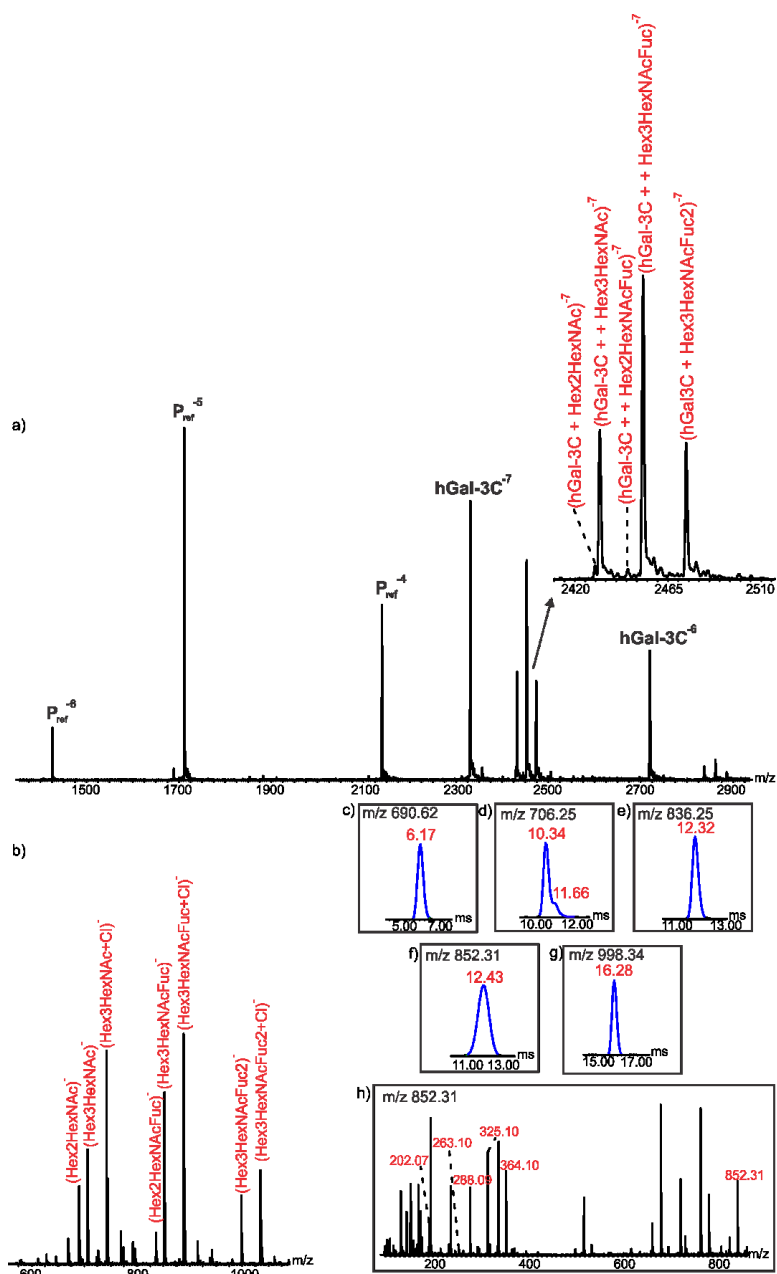
Figure 3.21. (a) Representative ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (20 mM, pH 6.8) of P_{ref} (5 µM), *Fr4* (0.05 µg µL^{-1}), and hGal-3C (15 µM). (b) CID mass spectrum acquired for all (hGal-3C + HMO)$^{7-}$ ions at Trap voltage 40 V. IMS-ATDs of released HMO ions at (c) m/z 632.21, (d) m/z 643.44, (e) m/z 753.67, (f) m/z 826.17, (g) m/z 899.18, and (h) m/z 997.34. (i) CID mass spectrum acquired in the Transfer region at 30 V for deprotonated HMO ions at m/z 997.34.

Taken together, the CaR-ESI-MS data obtained for the four fractions revealed HMO ligands corresponding to 17 different MWs (Table 3.3). The structures of 21 HMO ligands

(corresponding to eleven different MWs) were established from IMS-ATs and CID fingerprints. For the remaining six MWs, only monosaccharide compositions (and their corresponding putative structures) could be determined (Table 3.3). Notably, all of the 21 HMO ligands for which the structures were conclusively identified (**HMO1**, **HMO3**, **HMO6, HMO7, HMO9 - HMO17, HMO18**, **HMO21**, **HMO22, HMO25 - HMO27, HMO30,** and **HMO31**) were shown previously to bind to hGal-3C with measurable affinity.[16]

Table 3.3. MW and monosaccharide composition (Hex ≡ Glc or Gal; HexNAc ≡ GlcNAc; Fuc ≡ fucose; and Sia ≡ sialic acid) of HMO ligands of hGal-3C identified from *Fr1 – Fr4* using CaR-ESI-MS. The identity of specific HMO structures was based on a comparison of IMS-ATs and CID fingerprints of deprotonated ligand ions released from hGal-3C and those of **HMO1** – **HMO31**

MW (Da)	Monosaccharide composition	Confirmed/Putative HMO Structures
488.17	Hex2Fuc	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc (**HMO1**)
633.21	Hex2Sia	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO3**)
691.62	Hex4HexNAcFuc	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc (**HMO30**)
707.25	Hex3HexNAc	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO6**) β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO7**)
836.20	Hex2HexNAcSia	β-D-GlcNAc-(1→3/4)-[α-D-Neu5Ac-(2→6)]-β-D-Gal-(1→4)-β-D-Glc [a] α-D-Neu5Ac-(2→6)-β-D-GlcNAc-(1→3/4)-β-D-Gal-(1→4)-β-D-Glc [a]

837.25	Hex2HexNAcFuc2	α-L-Fuc-(1→3/4)-β-D-GlcNAc-(1→3/6)-[α-L-Fuc-(1→2/3)]-β-D-Gal-(1→4)-β-D-Glc [a]	
853.31	Hex3HexNAcFuc	α- L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO9**)	
		β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO10**)	
		β-D-Gal-(1→4)-[α-L-Fuc- (1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO11**)	
		β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-Glc (**HMO12**)	
		β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)[α-L-Fuc-(1→3)]-β-D-Glc (**HMO13**)	
998.34	Hex3HexNAcSia	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO14**)	
		α-D-Neu5Ac-(2→6)-[β-D-Gal-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- β-D-Glc (**HMO15**)	
		α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO16**)	
		α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO17**)	
999.34	Hex3HexNAcFuc2	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO18**)	
1056.32	Hex3HexNAc2Fuc	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-GlcNAc(1→3)-β-D-Gal(1→4)-β-D-Glc (**HMO31**)	
1072.38	Hex4HexNAc2	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO21**)	

		β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)]-β-D-Glc-(1→4)-Glc (**HMO22**)	
1218.28	Hex4HexNAc2Fuc	β-D-Gal-(1→3/4)-β-D-GlcNAc(1→3/6)-β-D-Gal-(1→3/4)-[L-Fuc-(1→3/4)]-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc [a]	
1289.44	Hex3HexNAcSia2	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-D-Neu5Ac-(2→6)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc (**HMO25**)	
1364.50	Hex4HexNAc2Fuc2	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-β-D-Glc (**HMO26**) 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-(1→3)]-β-D-Gal-(1→4)-β-D-Glc (**HMO27**)	
1509.32	Hex4HexNAc2FucSia	Neu5Ac-(2→6)-β-D-Gal-(1→3/6)-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→3)-[L-Fuc-(1→3/4)]-β-D-GlcNAc-(1→3/6)-β-D-Gal-(1→4)-β-D-Glc [a]	
1656.33	Hex4HexNAc2Fuc4	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-[α-L-Fuc-(1→2)-β-D-Gal-(1→4)-α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-β-D-Gal-(1→4)-β-D-Glc [a]	
		α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-[α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)]-β-D-GlcNAc-(1→3)-]-β-D-Gal-(1→4)-β-D-Glc [a]	
1800.44	Hex4HexNAc2FucSia2	α-D-Neu5Ac-(2→6)-[α-D-Neu5Ac-(2→6) -β-D-Gal-(1→3)-]-β-D-GlcNAc-(1→4)-[α-L-Fuc-(1→2)-]-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc [a]	

a. Possible structures based on CID results and identified HMOs reported in references 13 and 14.

3.4 Conclusions

This work describes the first application of CaR-ESI-MS for screening natural libraries of HMOs, derived from pooled breast milk, against target protein. A total of 21 different HMO MWs were identified in the four fractions used in the present study. The structures of 25 HMOs, corresponding to 11 different MWs, were identified based on their IMS-ATs and CID fingerprints.

To our knowledge, this is the first report describing the use of IMS-ATs and CID fingerprinting for identifying HMOs in mixtures. For the other MWs, monosaccharide compositions and, in some cases, possible structures were established. Implementation of the assay was demonstrated using hGal-3C, which served as a model HMO binding lectin. The assay revealed HMO ligands corresponding to 17 different MWs. From a comparison of IMS-ATs and CID fingerprints measured for the released HMO ligands and for a library of 31 pure HMOs, the structures of 21 HMO ligands were identified. Each of these glycans was shown previously to bind to hGal-3C. The presence of HMO ligands at six other MWs also was ascertained; however, the exact structures of these ligands could not be established conclusively.

3.5 Literature cited

1. Lara-Villoslada, F.; Olivares, M.; Sierra, S.; Rodríguez, J.M.; Boza, J.; Xaus, J. *Br J Nutr.* **2007**, *98*, S96.

2. Ballard, O.; Morrow, A.L. *Pediatr Clin North Am.*, **2013**, *60*, 49.

3. Martín, R.; Langa, S.; Reviriego, C.; Jimínez, E.; Marín, M.L.; Xaus, J.; Fernández, L.; Rodríguez, J.M. *J Pediatr.* **2003**, *143*, 754.

4. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. *Annu Rev Nutr.* **2000**, *20*, 699.

5. Kulinich, A.; Liu, L. *Carbohydr Res.* **2016**, *432*, 62.

6. Smilowitz, J.T.; Lebrilla, C.B.; Mills, D.A.; German, J.B.; Freeman, S.L. *Annu. Rev. Nutr.* **2014**, *34*, 143.

7. Bode, L. *Glycobiology* **2012**, *22*, 1147.

8. Wu, S.; Tao, N.; German, J.B.; Grimm, R.; Lebrilla, C. *J. Proteome Res.* **2010**, *9*, 4138.

9. Wu, S.; Grimm, R.; German, J.B.; Lebrilla, C.B. *J. Proteome Res.* **2011**, *10*, 856.

10. Finke, B.; Stahl, B.; Pfenninger, A.; Karas, M.; Daniel, H.; Sawatzki, G. *Anal. Chem.* **1999**, *71*, 3755.

11. Stahl, B.; Thurl, S.; Zeng, J.; Karas, M.; Hillenkamp, F.; Steup, M.; Sawatzki, G. *Anal. Biochem.* **1994**, *223*, 218.

12. Domon, B.; Costello, C. *Glycoconj. J.* **1988**, *5*, 397.

13. Xuezheng, S.; Lasanajak, Y.; Xia, B.; Heimburg-Molinaro, J.; Rhea, J.M.; Ju, H.; Zhao, C.; Molinaro, R.J.; Cummings, R.D.; Smith, D.F. *Nat. Methods* **2011**, *8*, 85.

14. Noll, A.J.; Gourdine, J.P.; Yu, Y.; Lasanajak, Y.; Smith, D.F.; Cummings, R.D. *Glycobiology* **2016**, *26*, 655.

25. Noll, A.J.; Yu, Y.; Lasanajak, Y.; Duska-McEwen, G.; Buck, R.H.; Smith, D.F.; Cummings, R.D. *Biochem. J.* **2016**, *473*, 1343.

16. Shams-Ud-Doha, K.; Kitova, E.N.; Kitov, P.; St-Pierre, Y.; Klassen, J.S. *Anal Chem.* **2017**, *89*,4914.

17. Kitova, E.N.; El-Hawiet, A.; Schnier, P.D.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.

18. Lin, H.; Kitova, E.N.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 104.

19. Liu, L.; Kitova, E.N.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 310.

20. Wang, W.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2003**, *75*, 4945.

21. Sun, J.; Kitova, E.N.; Wang, W.; Klassen, J.S. *Anal. Chem.* **2006**, *78*, 3010.

22. El-Hawiet, A.; Shoemaker, G.K.; Daneshfar, R.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2011**, *84*, 50.

23. Leney, A.C.; Fan, X.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2014**, *86*, 5271.

24. El-Hawiet, A.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2013**, *85*, 7637.

25. Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E.N.; Li, C.; Jia, F.; Cairo, C.W.; Klassen, J.S. *Anal. Chem.* **2012**, *84*, 7618.

26. Han, L.; Kitova, E.N.; Tan, M.; Jiang, X.; Klassen, J.S. *J. Am. Chem. Soc.* **2014**, *25*, 111.

27. Han, L.; Tan, M.; Xia, M.; Kitova, E.N.; Jiang, X.; Klassen, J.S. *J. Am. Chem. Soc.* **2014**, *136*, 12631.

28. Kitova, E.N.; El-Hawiet, A.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1908.

29. El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E.N.; St-Pierre, Y.; Klassen, J.S. *Anal. Chem.* **2017**, *89*, 8713.

30. Wheeler, S.F.; Harvey, D.J. *Anal. Chem.* **2000**, *72*, 5027.

Chapter 4

Identifying Interactions between Human Milk Oligosaccharides and VP8*

Domain of Rotavirus

4.1 Introduction

Rotavirus is one of the leading causes of gastrointestinal infection in children. Each year, over 500,000 deaths are reported due to severe diarrheal diseases caused by rotaviruses.[1] Human rotaviruses have a wide diversity of genotypes, which are due to mutation, genetic reassortment, or other reasons. The genotypes of the rotavirus strains are determined by the outer capsid proteins. Among the 11 double-stranded RNA segments in rotavirus, two are responsible for coding outer capsid proteins, VP7 and VP4.[2] Similar to influenza, the naming system of rotavirus is based on these two proteins; VP7 decides G serotypes, and VP4 determines P serotypes. Nowadays, 27 G genotypes and 37 P genotypes have been discovered, and their combinations lead to over 70 rotavirus genotypes.[3] The strains involved in this study are G1P[8] (Wa), G2P[4] (DS1), and G10P[11] (N155). While the first two strains are spread widely around the world and are accountable for most rotavirus infections in children,[5] the third one mainly attacks newborns, therefore, it is a neonate-specific strain.[6] Previous research showed that the VP8* domain of the VP4 spike protein is responsible for cell attachment, and it is critical for initiating the infection.[7] Figure 4.1 shows the structures of VP8* from the three strains used in this study.

VP8*-G10P[11]   VP8*-G1P[8]   VP8*-G2P[4]
PDB: 4YFW   PDB: 2I2S   PDB: 2AEN
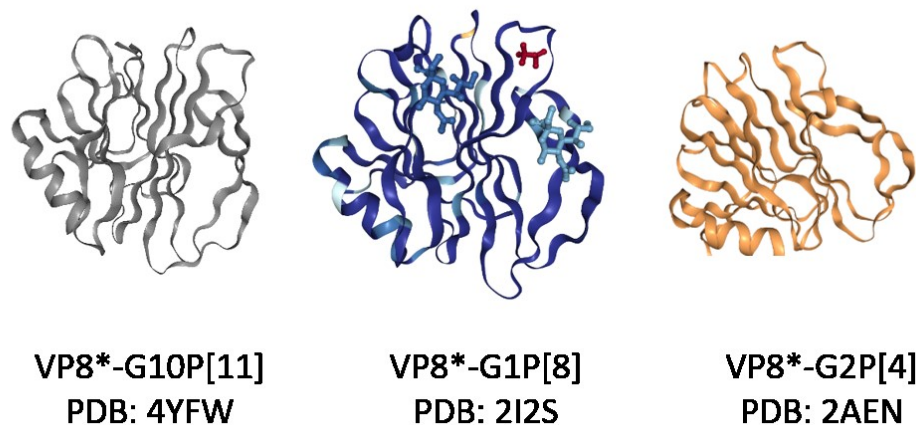
Figure 4.1. Crystal structures of spike protein VP8* domain of three rotavirus strains, figures are extracted from NGL viewer.[4]

Similar to many other viruses, rotavirus invasion starts with an interaction with intestinal epithelial glycan.[11] The common belief is that HMOs prevent viral attachment by acting like a receptor analog.[12] Studies have shown that HMOs played important roles in blocking cell attachment from noroviruses.[13,14] Thus, it is reasonable to believe that HMOs protect infants from rotavirus infection in a similar way as they protect from norovirus. In this study, the interactions of HMOs and other oligosaccharides with VP8* from different strains are characterized. We were hoping that by screening HMOs against three strains of rotavirus, the question of whether there are unified or distinguished defensive mechanisms of HMOs among different human rotavirus strains can be answered.

Previous studies classified rotavirus into sialidase-sensitive and insensitive types.[15] It commonly was believed that most animal strains are sialidase-sensitive, for example, porcine rotavirus CRW-8 and monkey rotavirus RRV.[15] A few animal strains, plus all human strains, belong to the sialidase-insensitive type, such as human rotavirus Wa and DS-1.[16,17] It was believed that the cell attachment process of animal rotavirus depends on sialic acid, but human strains do not require the presence of sialic acid.[17,18] Although there have been a few studies on human

rotavirus attachment mechanisms, the binding motif for VP8* between different P genotypes remain unclear. In this study, we characterized the interactions between VP8* protein from three human rotavirus strains (G1[8], G2P[4], and G10P[11]), and seven HMO libraries composed of 35 purified HMOs that contain H and A type HBGAs, in addition to GM1 and A-type HBGA hexaose type 2, by direct ESI-MS assay (Figure 4.2). Among the three strains, clear differences in binding preference were discovered.
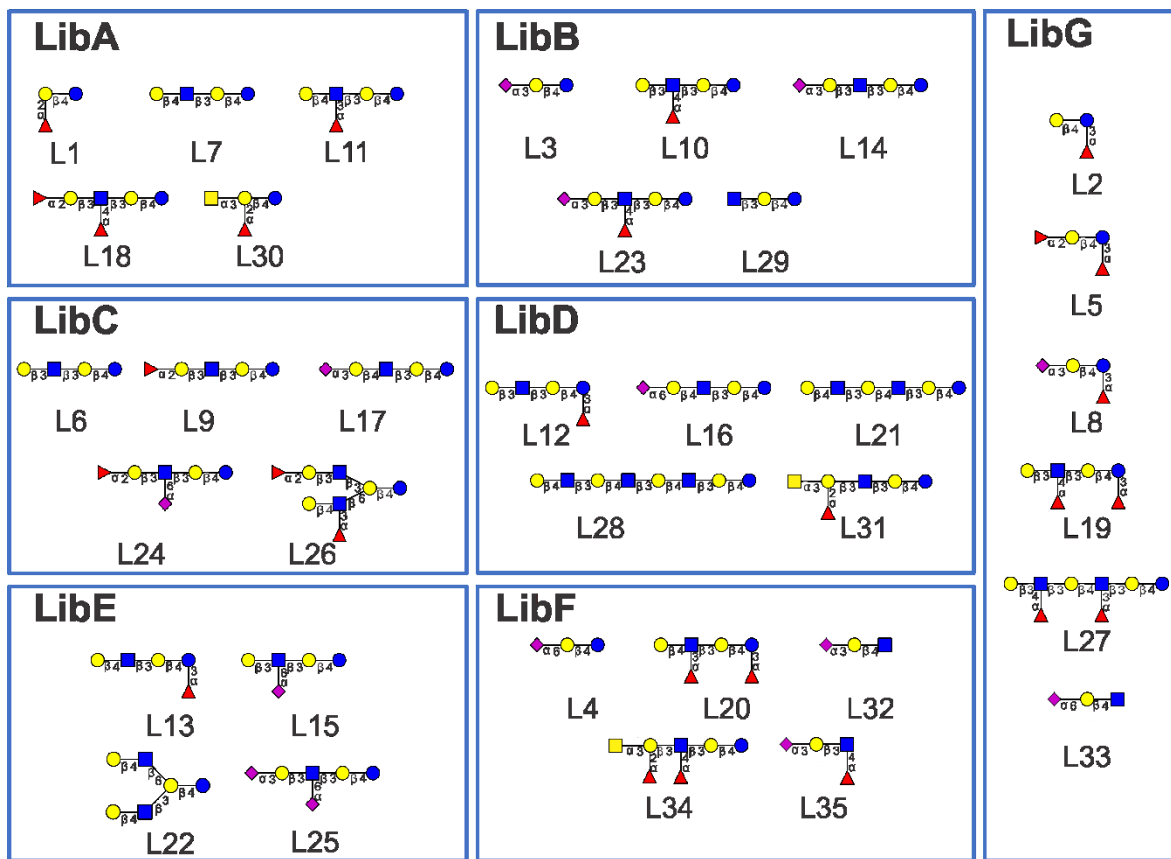
Figure 4.2. The composition of seven HMO libraries A–G.

4.2 Experimental and Methods

4.2.1 Proteins

VP8* domain from three rotavirus strains, G1P[8], G2P[4], and G10P[11], were provided by Dr. B. V. Venkatar Prasad (Baylor College of Medicine). Single chain variable fragment of mAb Se

155-4 (ScFv), as the reference protein for G10P[11], was produced in house using recombinant technology as described elsewhere.[19] Cytochrome c (Cyt C), as the reference protein for G1P[8] and G2P[4], was purchased from Sigma-Aldrich Canada (Oakville, Canada). All proteins were dialyzed against 200 mM aqueous ammonium acetate (pH 6.8) using 10 kDa MW cutoff Amicon Ultra-4 centrifugal filters (Millipore Corp, Bedford, MA). The final concentrations of stock solutions of the proteins were: VP8*-P[4]: 229 µM, VP8*-P[8]: 394 µM, VP8*-P[11]: 214 µM. The protein stock solutions were stored at -20 °C until used.

4.2.2 Human milk oligosaccharides and other oligosaccharides

In total, 35 HMOs (L1–L35) were used in this study. The information of L1–L31 were described in Chapter 2 and listed in table 2.1 The structures of L32–L35 are listed in table 4.1.

Table 4.1. Molecular weight, structure and common name of HMOs included in the libraries

HMO	Monoisotopic MW (Da)	Structure	Common name
L32	674.24	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-β-D-GlcNAc	3'-Sialyl-N-acetyllactosamine
L33	674.24	α-D-Neu5Ac-(2→6)-β-D-Gal-(1→4)-β-D-GlcNAc	6'-a-Sialyl-N-acetyllactosamine
L34	1202.44	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc(1→3)-β-D-Gal-(1→4)-β-D-Glc	A-heptasaccharide
L35	820.30	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc	3'-Sialyl Lewis A

Blood group A antigen (A-type HBGA) hexaose type 2 (MW 1056.39 Da) and GM1 (MW 998.34 Da) were purchased from Elicityl SA (Crolles, France); L32, L33, and L34 were purchased from Dextra (Reading, UK); and HMO35 were purchased from CarboSynth (Compton, UK). Stock

solutions of each oligosaccharide were prepared by dissolving a known mass of the oligosaccharide in ultrafiltered water (Milli-Q, Millipore, Billerica, MA) to give a final concentration of 1 mM. All stock solutions were stored at -20 °C until used.

4.2.3 Mass spectrometry

The settings and parameters for mass spectrometry were described in Chapter 2, Section 2.2.3.

4.2.4 Determination of Ligand Affinities from ESI-MS Data

For a protein with one binding site and with the presence of multiple ligands (L_1, L_2, …L_x), the apparent binding constant ($K_{a, Lx}$) of an individual ligand can be obtained from Eq 4.1:

$$K_{a,L_x} = \frac{R_x}{[L_x]_0 - \frac{[P]_0 R_x}{(1 + R_1 + R_2 + \cdots R_x)}} \tag{4.1}$$

where $[L]_0$ and $[P]_0$ are the initial concentrations of L and P, respectively. The R_x term represents the concentration ratio of ligand-bound to free protein (i.e., $[PL_x]/[P]$), and it can be determined from the abundances (Ab) of the gaseous ligand-bound and free P ions, summed over all charge states (n), Eq 4.2:

$$R_x = \frac{Ab\ (PL_x)}{Ab\ (P)} = \frac{\sum_n Ab(PL_x)^{n+}}{\sum_n Ab(P)^{n+}} \tag{4.2}$$

If the ligand binding sites are not equivalent or in cases of cooperative binding (positive or negative), the relationship given by Eq 4.1 will not be valid.

4.3 Result and Discussion

4.3.1 Rotavirus VP8*-P[11]

Figure 4.3 is a typical spectrum of VP8*-P[11] (MW 18674 Da) screened against HMO LibG. The enlarged section showed that there are six protein–HMO complexes, and the ligands can be determined through their molecular weights (L2, MW 488.17 Da, L5, MW 634.23 Da, L8, MW 779.27 Da, L19, MW 999.34 Da, L27, MW 1364.50 Da, and L33, MW 674.60 Da). After correction by the reference protein method,[17] only four ligands (L2, L8, L27, and L33) showed

affinities higher than 10^3 M^{-1}, whereas K_a values for L5 and L19 are 750 M^{-1} and 921 M^{-1}, respectively.
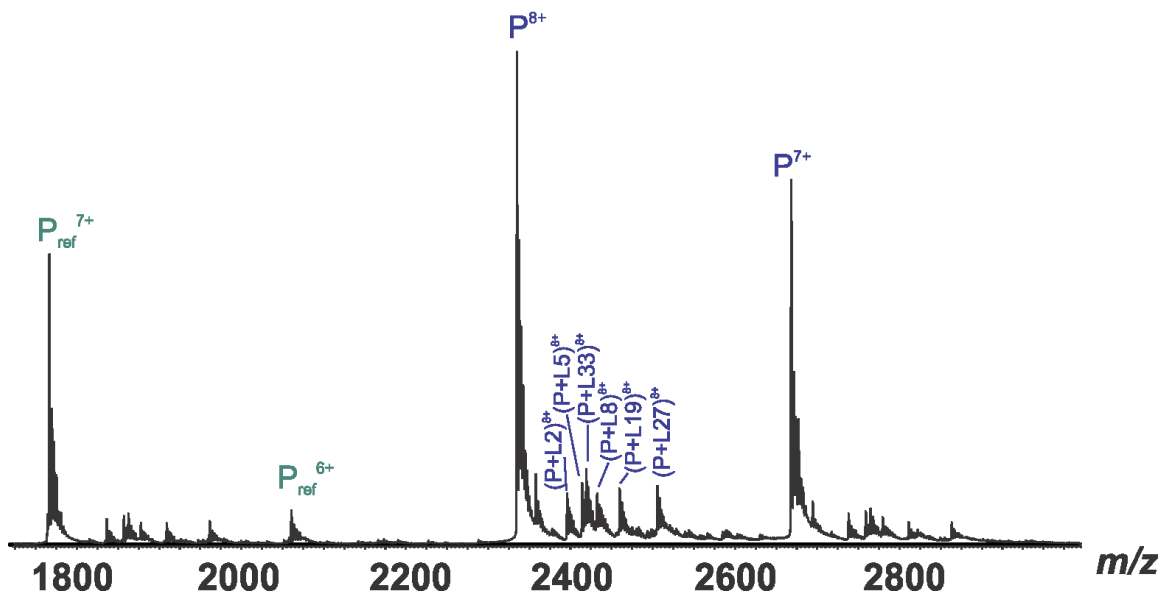
Figure 4.3. A ESI-MS spectrum of a solution containing 100 mM ammonium acetate (pH 6.8), P_{ref} (0.6 μM), HMO LibG (20 μM), and VP8*-P[11] (4.3 μM) in the positive ion mode. PL complexes are labeled in the region of 8+ charge state.

The same approach was applied to the other HMO libraries. Figure 4.4 is the spectrum of ammonium acetate solutions containing VP8*-P[11], P_{ref}, and HMO LibA–LabF. In general, no strong binding ($K_a > 10^4$ M^{-1}) was detected for screening VP8*-P[11] to oligosaccharide in this study.
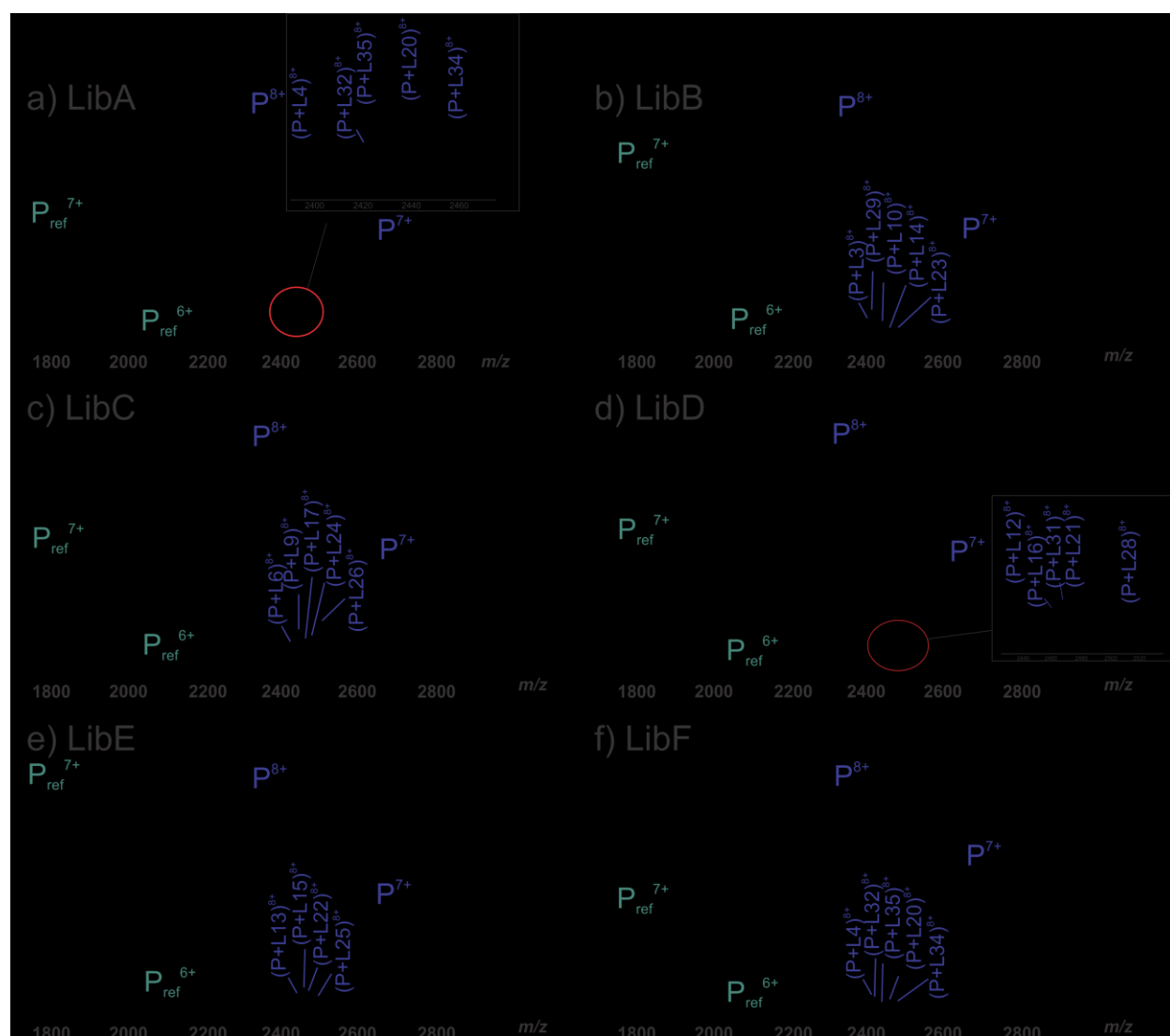
Figure 4.4. ESI-MS spectra of solutions containing 100 mM ammonium acetate (pH 6.8), P_{ref} (0.6 μM), VP8*-P[11] (4.3 μM) and a) LibA (20 μM), b) LibB (20 μM), c) LibC (20 μM), d) LibD (20 μM), e) LibE (20 μM), and f) LibF (20 μM), in the positive ion mode. PL complexes are labeled in the region of 8+ charge state.

The binding constants (K_a) were obtained from three concentrations of each library, with all experiments carried out in triplicate. Figure 4.5 and Table 4.2 show the summary of the K_a values of all 35 HMOs after nonspecific binding correction, and it suggests that 17 binders have a weak interaction to VP8*-P[11] (10^3 M^{-1} <K_a<10^4 M^{-1}).
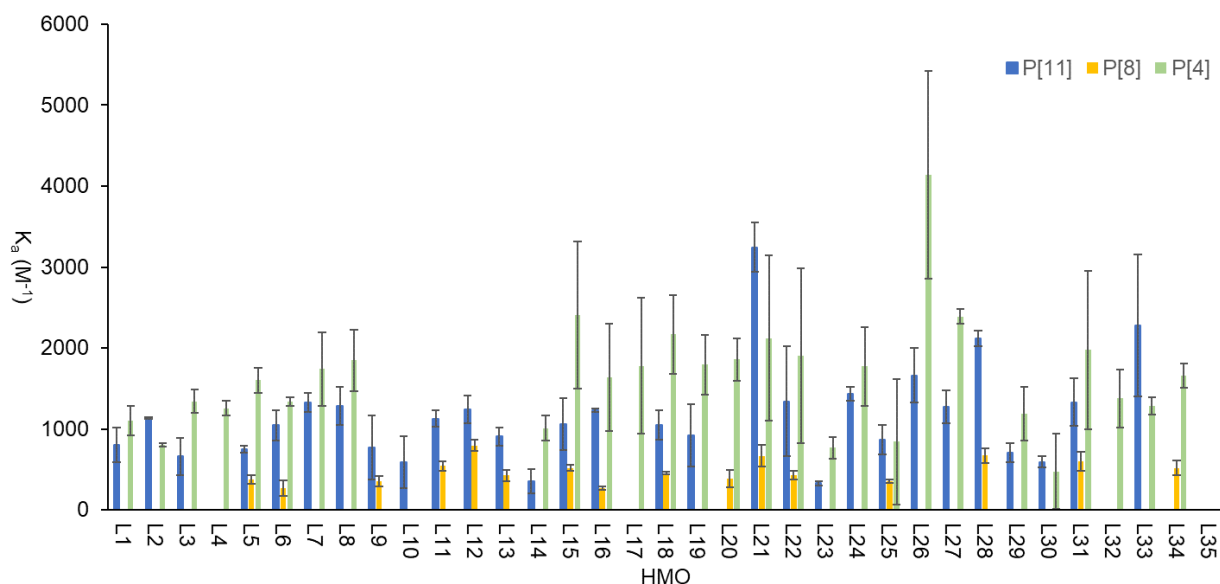
Figure 4.5. A summary of K_a values of 35 HMO ligand (L1–L35) to VP8* of three human rotavirus strains, P[11] (blue), P[8] (orange), and P[4] (green).

Table 4.2. A summary of K_a values of 35 HMO ligand (L1–L35) to VP8* of three human rotavirus strains: VP8*-P[11], VP8*-P[8], and VP8*-P[4]

HMO	VP8*-P[11]		VP8*-P[8]		VP8*-P[4]	
	Ka (M^{-1})	STDEVA	Ka (M^{-1})	STDEVA	Ka (M^{-1})	STDEVA
L1	803	217	0	0	1106	181
L2	1134	7	0	0	798	21
L3	661	228	0	0	1343	147
L4	0	0	0	0	1256	89
L5	750	47	374	56	1603	154
L6	1047	186	271	95	1336	56
L7	1329	117	0	0	1739	455
L8	1285	238	n/a	n/a	1846	381
L9	771	396	352	66	n/a	n/a
L10	586	322	0	0	n/a	n/a
L11	1128	103	545	58	n/a	n/a
L12	1244	168	797	71	n/a	n/a
L13	905	111	428	70	n/a	n/a
L14	356	147	0	0	1011	156
L15	1063	322	517	40	2405	906
L16	1228	23	271	22	1639	662

118

L17	0	0	0	0	1778	839
L18	1047	182	456	20	2169	485
L19	921	385	0	0	1796	368
L20	0	0	385	110	1856	262
L21	3241	304	670	137	2122	1022
L22	1343	675	430	56	1903	1082
L23	330	27	0	0	766	136
L24	1434	82	0	0	1772	488
L25	868	179	355	21	845	776
L26	1664	335	n/a	n/a	4142	1284
L27	1278	203	n/a	n/a	2389	93
L28	2118	95	671	91	n/a	n/a
L29	704	120	0	0	1185	333
L30	593	67	0	0	477	465
L31	1330	294	599	113	1974	976
L32	0	0	0	0	1377	357
L33	2277	877	0	0	1286	106
L34	0	0	519	88	1662	151
L35	0	0	n/a	n/a	n/a	n/a

From the top five binders (L21, L33, L28, L26, and L24) of VP8*-P[11], it appears that type 2 LacNAc (Galβ1-4GlcNAc) (L21, L28, and L26) and α2-6 sialic acid (L33 and L24) are preferred by G10P[11] (Figure 4.6).
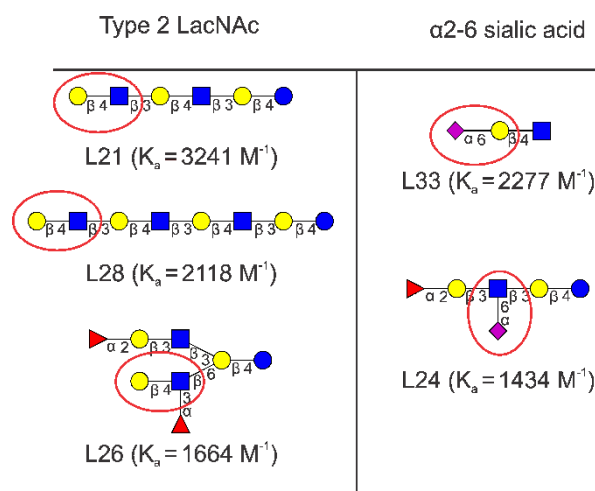
Figure 4.6. HMO structures of the top five binders to VP8*-P[11] from the HMO libraries.

Based on the shotgun glycan microarray (SGM) data reported by Smith and colleagues,[20] P[11] from the human strain has interactions with both type 1 and 2 LacNAc, whereas bovine P[11] showed binding to type 2 LacNAc only. Results from ESI-MS showed VP8*-P[11] has interactions with HMOs that contain either type 1 or 2 LacNAc, but it binds type 2 LacNAc in higher affinity. The sequence alignment of VP8* proteins from both strains showed great similarity (93.5%). Therefore, it is reasonable to assume that the neonatal-specific strain maintains some features of the bovine strain, therefore, it prefers type 2 LacNAc over type 1.

It is believed widely that the cell entry process of human rotaviruses does not depend on sialic acid, however, our results suggested otherwise. In over 35 purified HMOs, 12 of them have one or more sialic acid. The HMOs with an α2-6 link terminal sialic acid (L15, L16, and L33) showed higher binding affinity (average 1.5×10^3 M^{-1}) than the HMOs with an α2-3 link (L3, L8, L14, L17, L23, and L32), with an average K_a of 0.4×10^3 M^{-1}, except for L4. A group of scientists from Australia reported that the internal sialic acid from GM1 played an important role in the interaction with VP8*-P[8], suggesting that the presence of sialic acid is not completely irrelevant to the interaction between human rotaviruses and their ligands.[21] However, to the best of our knowledge, no study has been done to confirm the interactions between VP8*-P[11] and GM1 or other sialic acid containing oligosaccharides. In addition to the HMO libraries, we screened GM1 against VP8*-P[11] using a direct ESI-MS assay. For a series of ligand concentrations, results showed that the K_a of VP8*-P[11] to GM1 is higher than all the HMOs in the libraries, which is $3.8 \times 10^3 \pm 634$ M^{-1}. There is evidence showing that the glycan binding site of VP8*-P[11] is different from P[8]. With the affinity of VP8*-P[11] to GM1, it is possible that P[11] has more than one binding site for glycan ligands compared to other strains, yet it may be unable to bind two ligands at the same time.

To investigate the binding pattern of VP8*-P[11] to HMOs further, we compared its interactions to histo-blood group antigens (HBGAs), which are common ligands for bacterial and viral protein and some present in human milk. The protein showed very low affinity ($<10^3$ M^{-1}) to H-type HGBAs in this library, as suggested by the K_a values of L1 ($0.8 \times 10^3 \pm 217$ M^{-1}) and L9 ($0.8 \times 10^3 \pm 396$ M^{-1}). However, the bindings of VP8*-P[11] to A-type HBGAs are slightly higher than to H-type HBGAs since the K_a values for L30, L31, and A-type HBGA hexoase type 2 are $0.6 \times 10^3 \pm 67$ M^{-1}, $1.3 \times 10^3 \pm 294$ M^{-1}, and $1.8 \times 10^3 \pm 187$ M^{-1}, respectively (Figure 4.5 and Table 4.2). However, there is no clear preference between H or A type HBGAs and VP8*-P[11]. Since the presence of HGBA in human milk depends highly on the genetic profile of the mother, the similarity of interactions between VP8*-P[11] and both types of HGBAs supports a clinical report that breastfed infants with a different blood type mother react to rotavirus similarly.[22]

Among the 14 HMOs that presented both in the ESI-MS and SGM results, eight of them have affinities higher than 10^3 M^{-1} suggested by ESI-MS, whereas only one (L21) were discovered by SGM with high affinity (Figure 4.7).[20] The lack of ability of detecting weak HMO binders results false negatives and leads to an incorrect conclusion of the true binding motif of rotavirus VP8* protein. Nevertheless, both assays agreed that L21 (pLNnH) has a higher affinity to VP8*-P[11] than other available HMOs.
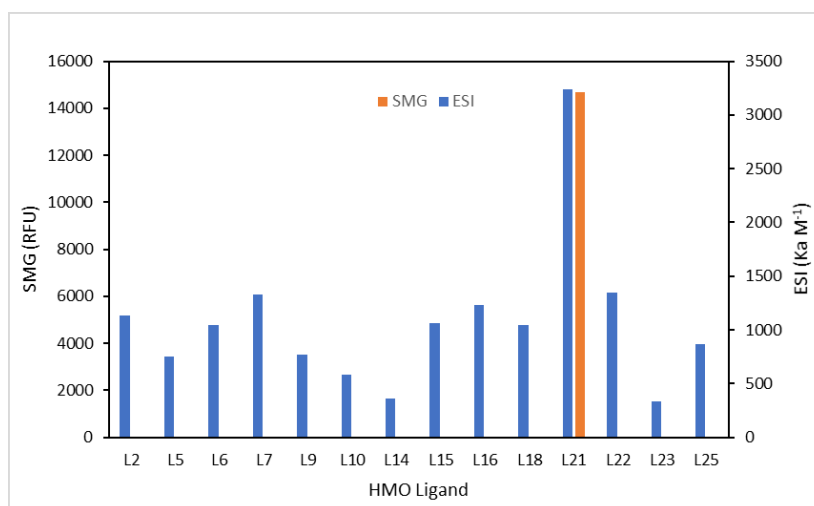
Figure 4.7. Comparison of binding affinity results from ESI-MS (blue) and SGM (orange).

4.3.2 Rotavirus VP8*-P[8]

VP8*-P[8] showed a uniform very low affinity ($K_a < 10^3$ M^{-1}) to 31 available HMOs (Figure 4.4).

Figure 4.8 shows spectra of VP8*-P[8] to HMO libraries A-F.
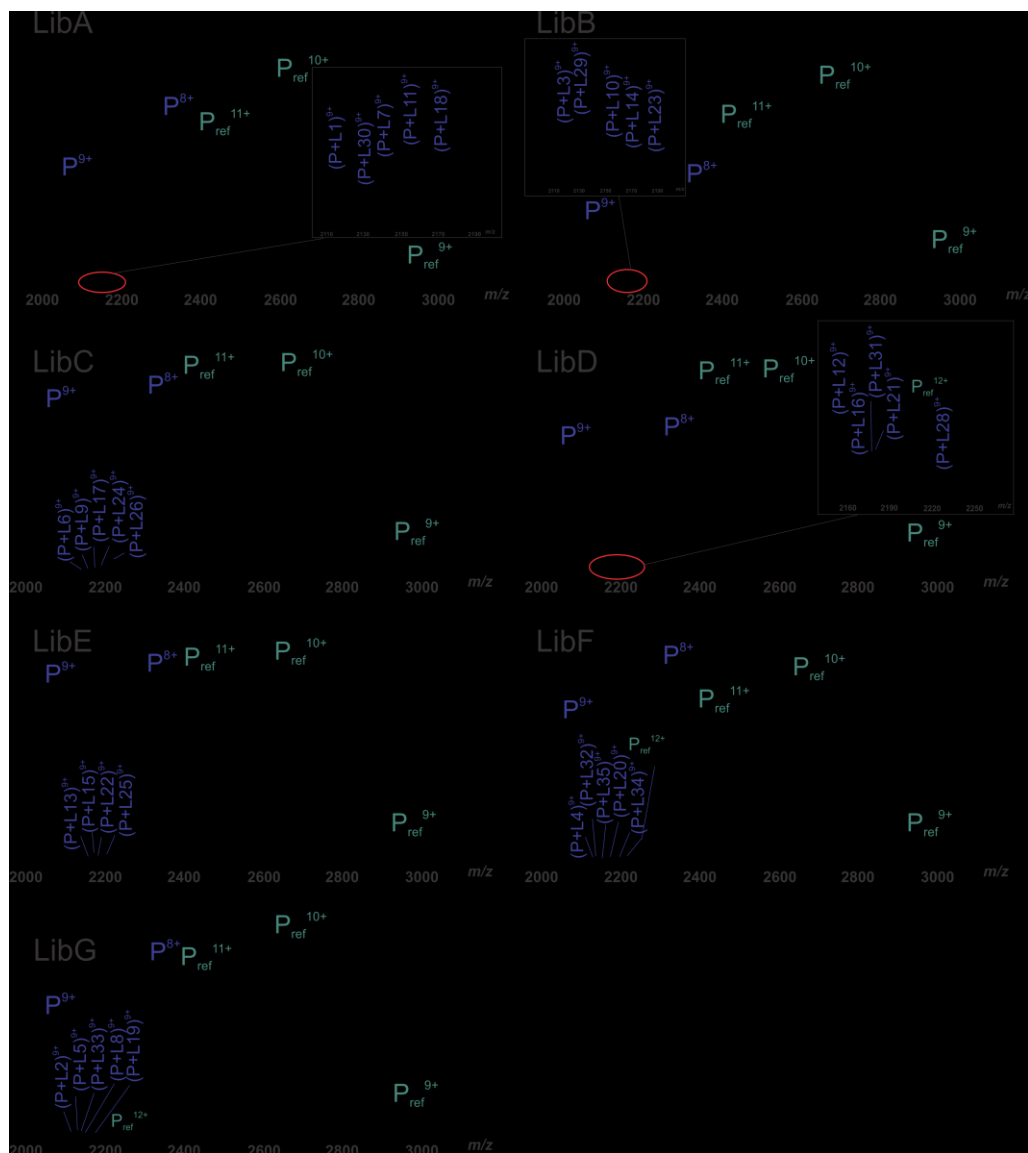
Figure 4.8. ESI-MS spectra of solutions containing 100 mM ammonium acetate (pH 6.8), P_{ref} (0.6 µM), VP8*-P[8] (3.9 µM) with LibA (25 µM), LibB (25 µM), LibC (25 µM), LibD (25 µM), LibE (25 µM), LibF (25 µM), and LibG (25 µM), in the positive ion mode. PL complexes are labeled in the region of 9+ charge state.

The binding target for G1P[8] remained a mystery although its crystal structures were discovered in 2007.[23] Some researchers believed that H-type HBGAs are potential binding partners and that A-type HBGAs are not, but other studies suggest otherwise.[24,25] Some experiments were done with the HBGAs present in milk and saliva, where interferences may exist. ESI-MS results showed very limited binding to both types of HBGAs, with an average K_a of 0.2×10^3 M^{-1}. In addition, NMR data coupled with cell culture assay suggested that P[8] has interaction with GM1a but not with GD1.[26] From our ESI-MS results, the K_a value obtained from a concentration titration of VP8*-P[8] to GM1a is almost zero (Figure 4.9).
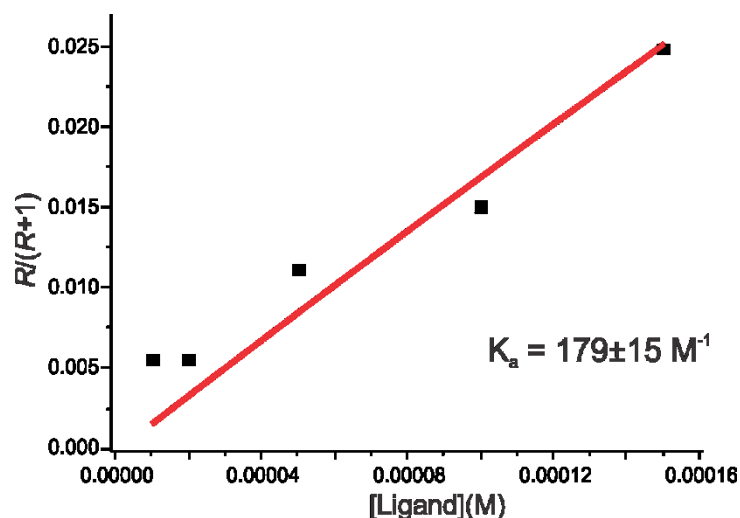
Figure 4.9. A titration of concentrations of GM1a to VP8*-P[8].

Several factors may be the reason for the differences between the results, including the absence of co-binders and the absence of a full-length of the spike protein.[24,27] Therefore, more studies are required to draw conclusions on the true affinity of VP8*-P[8] to oligosaccharides.

4.3.2 Rotavirus VP8*-P[4]

VP8*-P[4] showed a wide variety of binding preferences from the HMO libraries. Out of 28 available HMOs, 24 are found to have $K_a > 10^3$ M^{-1} to this protein. Figure 4.10 shows the spectra of VP8*-P[4] to seven HMO libraries.
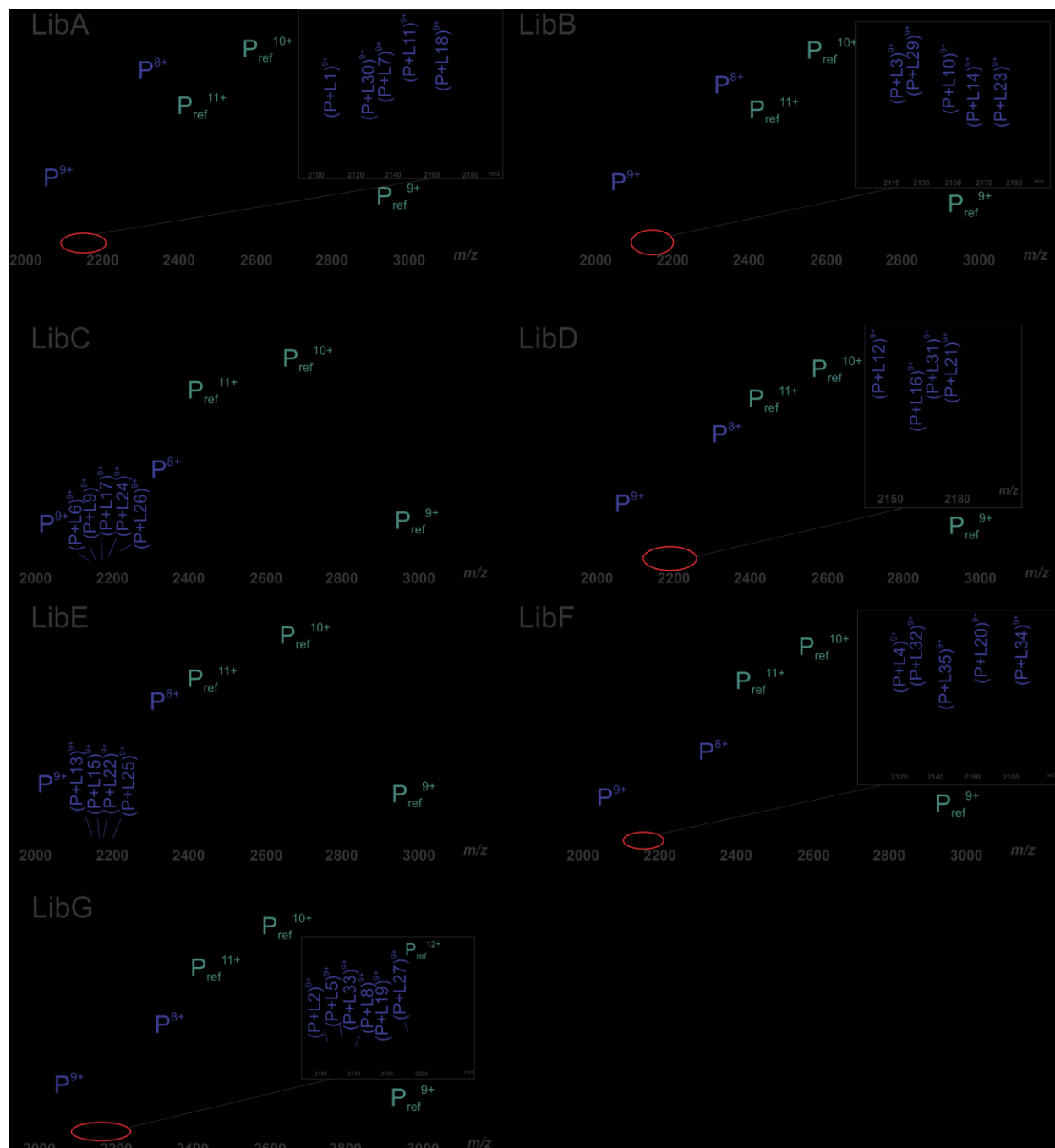
Figure 4.10. ESI-MS spectra of solutions containing 100 mM ammonium acetate (pH 6.8), P_{ref} (1.6 μM), VP8*-P[4] (4.6μM) with LibA (15 μM), LibB (15 μM), LibC (15 μM), LibD (15 μM),

LibE (15 μM), LibF (15 μM), and LibG (15 μM), in the positive ion mode. PL complexes are labeled in the region of 9+ charge state.

From the data presented in Figure 4.4, three of the top five binders (L26, L15, L27, L18, and L21) have two fucoses, indicating that fucose could be important in the interactions. The behavior of this protein to HBGAs is unclear due to the protein–ligand complex peak in 8+ charge stage was overlapping partially with the reference protein peak in 11+ charge stage. To overcome this problem, further studies were suggested in Chapter 5.

4.4 Conclusions

The binding patterns of different human rotavirus strains to HMOs were investigated. Three sialidase-insensitive strains, G2P[8], G1P[4], and G10P[11], were used in this study. Our data revealed that the binding preference of the VP8* domain from their spike proteins, which is responsible for cell attachment, to HMOs and some common oligosaccharides are very different.

The neonatal-specific rotavirus strain, G10P[11], targets newborns only. Although the native receptor(s) of G10P[11] remain(s) unclear, some studies suggested that type 1 and 2 LacNAc are the potential binding motifs. Direct ESI-MS results showed that type 2 LacNAc and α2-6 link sialic acid are preferred by this strain. The interaction of VP8*-G10P[11] to GM1 was measured, and it was higher than all HMOs; this could be an indication of the important role of gangliosides played in rotavirus cell attachment.

G1P[8] and G2P[4] are known as the strains responsible for the majority of global infections. However, they showed different preference in binding to HMOs. G1P[8] has almost no binding to HMOs used in this study, with uniform low affinities ($K_a < 10^3$ M^{-1}). Possible reasons could be the absence of the co-binder of the protein or the full-length protein. It is also possible

that HMOs have no effect in blocking the viral attachment of this strain. On the other hand, G2P[4] showed various affinities to different structures, from $K_a < 10^3$ M^{-1} to $K_a < 10^4$ M^{-1}.

Taken together, data from the direct ESI-MS assay revealed that the affinities of HMOs to different rotavirus strains varies. This could be an indication of different protection effects of human milk to different rotaviruses or of different intrusion mechanisms for each strain. The behavior of the viruses is more diverse than we expected. By understanding the differences between the interaction patterns, new therapeutic strategies for targeting a specific strain can be developed. Therefore, future studies should be carried out to complete the picture.

4.5 Literature cited

1. Parashar, U.D. *Emerg. Infect. Dis.* **2003,** *9*, 565.

2. Estes, M.K.; Cohen, J. *Microbiol. Rev.* **1989,** *53*, 410.

3. Theuns, S.; Heylen, E.; Zeller, M. *J. Virol.* **2015**, *89*, 1043.

4. Rose, A.S.; Bradley, A.R.; Valasatava, Y.; Duarte, J.M.; Prlić, A.; Rose, P.W. ACM Proceedings of the 21st International Conference on Web3D Technology (Web3D '16), **2016** 185.

5. Agocs, M.M.; Serhan, F.; Yen, F, C. *MMWR Morb Mortal Wkly Rep* **2014**, *3*, 634.

6. Banerjee, I.; Gladstone, B.P.; Le Fevre, A.M. *J. Infect. Dis.* **2007**, *195*, 625.

7. Arias, C.F.; Isa, P.; Guerrero, C.A.; Méndez, E.; Zárate, S.; López, T.; Espinosa, R.; Romero, P.; López S. *Med. Res. Arch.* **2002**, *33*, 356.

8. Jiang, V.; Jiang, B.; Tate, J.; Parashar, U.D.; Patel, M.M. *Hum. Vaccines*. **2010**, *6*, 532.

9. Armah, G.E.; Sow, S.O.; Breiman, R.F. *Lancet,* **2010**, *376*, 606.

10. Madhi, S.A.; Cunliffe, N.A.; Steele, D. *N. Engl. J. Med.* **2010**, *362*, 289.

11. Ramani, S.; Hu, L.; Prasad, B.V.V.; Estes, M.K. *Cell Mol. Gastroenterol Hepatol.* **2016**, *2*, 263.

12. Bode, L. *Early Hum. Dev.* **2015**, *91*, 619.

13. Ling, H.; Kitova, E.N.; Tan, M.; Jiang, X.; Klassen, J.S. *J. Am. Soc. Mass Spectrom*. **2014**, *25*, 111.

14. Shang, J.; Piskarev, V.E.; Xia, M.; Huang, P.; Jiang, X.; Likhosherstov, L.M.; Ratner, D.M. *Glycobiology*, **2013**, *23*, 1491.

15. Ciarelt, M.; Estes, M.K.; *J. gen. virol.* **1999,** *80*, 943.

16. Ciarlet, M. *J. virol.* **2002**, *76*, 4087.

17. Fukudome, K.; Yoshie, O.; Konno, T. *Virology,* **1989**, *172*, 196.

18. Guo, C.; Nakagomi, O.; Mochizuki, M.; Ishida, H.; Kiso, M.; Ohta, Y.; Suzuki, T.; Miyamoto, D.; Hidari, K.; Suzuki. Y. *J. Biochem.* **1999** *126*, 683.

19. Zdanov, A.; Li, Y.; Bundle, D.R.; Deng, S.J.; Mackenzie, C.R.; Narang, S.A.; Young, N.M.; Cygler, M. *Proc. Natl. Acad. Sci.* **1994**, *91*, 6423.

20. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Michkum, M.L.; Ashline, D.J.; Prasad, B.V.V.; Estes, M.K.; Reinhold, V.N.; Cummings, R.D.; Smith, D.F *Mol. Cell Proteomics* **2014,** *13*, 2944.

21. Haselhorst, T. *Nat. Chem. Biol.* **2008**, *5*, 91.

22. Das, S.; Sahoo, G.C.; Das, P.; Singh, U.K.; Jaiswal, A.K.; Singh, P.; Kumar, R. *PLoS One.* **2016**, *11,* e0146243.

23. Blanchard, H.; Yu, X.; Coulson, B.S.; von Itzstein, M. *J. Mol. Biol.* **2007**, *267*, 1215.

24. Huang, P.; Xia, M.; Tan, M,; Zhong, W,; Wei, C.; Wang, L.; Morrow, A.; Jiang, X. *J. Virol.* **2012**, *86*, 4833.

25. Bohm, R.; Fleming, F.E.; Maggioni, A.; Dang, V.T.; Holloway, G.; Coulson, B.S.; von Itzstein, M.; Haselhorst, T. *Nat. Comm.* **2015**, *5*, 5907.

26. Haselhorse, T.; Fleming, F.E.; Dyason, J.C.; Hartnell, R.D; Yu, X.; Holloway, G.; Santegoets, K.; Kiefel, M.J.; Blanchard, H.; Coulson, B.S.; von Itzstein, M. *Nat. Chem. Biol.* **2009**, *5*, 91.

27. Lopez, S.; Aria, C.F. *Trends Microbiol.* **2004**, *12*, 271.

Chapter 5

Conclusions and Future Work

The work in this thesis includes the development of mass spectrometry methods to screen carbohydrate-binding proteins to human milk oligosaccharides (HMOs). By using the direct and CaR-ESI-MS assays, the interactions between HMOs and several important lectins were studied.

5.1 Conclusions

Chapter 2 demonstrated a high-throughput and label- and immobilization-free method, using CaR-ESI-MS assay, for screening the interactions between an HMO library and lectins. This method was based on the MW, IMS-AT and CID fingerprint information extracted from purified HMOs. First, information of three dimensions (MW as the first, IMS-AT as the second, and CID fingerprint as the third) were obtained from 31 purified HMOs. The main challenge here was to distinguish the IMS-ATs from the isomer sets with three or more structures. Next, lectins were screened against a pooled library composed of all purified HMOs with equimolar concentrations. Human galectin 1, human galectin 7, and a C-terminal fragment of human galectin 3 (hGal-1, hGal-7, and hGal-3C) were used as target lectins. The HMO binders in the library were released from the target proteins in the gas phase, and their identification information was compared with the one obtained from the individual purified HMOs. Results from a CaR-ESI-MS assay showed that this method was able to identify the ligands with >500 M^{-1} affinities with no false positives. With fragmentation correction, relative affinities of the release ligands, identified as molecular weight sets, were obtained; this gives an affinity rank of the molecular weights. Although establishing the identification information took a long time, it can be used to screen all proteins. To screen one lectin to the library requires less than one hour and with limited sample consumption (<5 ng of each HMO and <0.5 μg of each lectin). The work in this Chapter laid a solid pathway

for screening proteins against carbohydrate libraries not limited to HMOs, with a rapid and label-free method, and it provides structural information of the ligands.

Chapter 3 showed an application of the method developed in Chapter 2 by screening natural HMO libraries. For the natural libraries, the composition is unknown, and the concentration of each HMO is not available, therefore, it is a more challenging environment for screening. Prior to screen against the target protein, the composition of the natural HMO libraries (from pooled human milk fractions) needed to be established. In total, 21 different HMO MWs were discovered in four milk fractions. Within the 21 MWs, 11 of them were confirmed with 25 structures by comparing their IMS-ATs and CID fingerprints with the purified HMOs that were obtained from Chapter 2. For those MWs with no purified standards to compare with, the monosaccharide composition and some possible structures were suggested.

Once the composition of the fractions was established, hGal-3C, a well studied lectin, served as the model protein for this project. We discovered 17 MWs that bind to hGal-3C from the milk fractions. With identification through IMS-ATs and CID fingerprints established from Chapter 2, 21 HMO structures were identified. The results were confirmed by the results from Chapter 2 that purified HMOs with the same structure were binders to hGal-3C. This is the first demonstration of using a reliable, fast, and label free mass spectrometry method to study the interaction between proteins and natural libraries. Although it cannot provide quantitative information, it has potential for semi-quantitative measurements by comparing the abundance ratio (R) difference from different protein concentrations.[1]

Chapter 4 investigated the binding patterns of different human rotavirus strains to HMOs. Although three strains involved in the study, G2P[8], G1P[4], and G10P[11], were classified as sialidase-insenetive strains, the binding preference of the VP8* domain from their spike proteins,

which are responsible for cell attachment to HMOs and some common oligosaccharides, they are very different.

G1P[8], known as the strain that is responsible for the majority of global infections, showed very low affinities ($K_a < 10^3$ M^{-1}) to all HMOs used in this study. Several explanations can be given for this phenomenon. One could be the absence of the co-binder of the protein, another could be the absence of the full-length protein. Alternatively, it is possible that HMOs have no effect in blocking the viral attachment from this strain. G2P[4] has a very similar sequence to G1P[8], however, unlike G1P[8] with almost no binding to the HMOs, G2P[4] showed various affinities to different structures, from $K_a < 10^3$ M^{-1} to $K_a < 10^4$ M^{-1}.

The neonatal-specific rotavirus strain, G10P[11], has a different target patient age range. So far, the native receptor(s) of G10P[11] remains unclear. The results showed that type 2 LacNAc and α2-6 link sialic acid could be its potential binding partners. In addition, the relative high affinity of VP8*-G10P[11] to GM1 suggested that gangliosides, which are mainly embedded in the cell surface, may play a critical role in rotavirus cell attachment.

In summary, the affinities of HMOs to different rotavirus stains varies. This could be an indication of the different protection effects of human milk on different rotaviruses, or it could reveal different intrusion mechanisms of each strain. The behavior of the viruses is more diverse than we expected. By understanding the differences between the interaction patterns, new therapeutic strategies for targeting a specific strain can be developed. Therefore, future studies should be carried out to complete the picture.

5.2 Future Work

5.2.1 Additional screening of rotavirus VP8* to more oligosaccharides and human milk fractions

Following Chapter 4, some possible extensions are suggested to complete the project. First, due to the complex peak overlapping issue, measurements of the interactions between H-type HBGAs in the HMO libraries and VP8*-P[4] were not available. The binding to alternative oligosaccharides, such as H-type HBGA tetraose, should be measured. Second, there is evidence that the binding site of VP8* is similar to human galectin 3 (hGal-3) and that the conformation of bound Lacto-N-tetraose (LNT) to VP8*-P[11] is similar to the LNT bound to hGal-3. Thus, it is possible that the binding of galectin to LNT promotes the interaction between LNT and VP8*-P[11]. Although there is no evidence showing that the cell entry process of rotavirus depends on the presence of galectin 3, it is worth trying to understand the relationship between the interaction of hGal-3 and VP8* with HMOs. A proposed experiment is to increase the concentration of hGal-3 in a solution containing a fixed amount of LNT and VP8*-P[11] protein, to see whether the addition of hGal-3 promotes the affinity between LNT and VP8*-P[11].

As described in Chapter 2 and 3, a CaR-ESI-MS assay was applied to screen natural libraries against proteins. Using the same method by screening human milk fractions to rotavirus VP8* protein from different strains, a more complete picture can be obtained because the higher molecular weight HMOs in the milk fractions are not available commercially. In addition, adding linkage specific glycosidases can be used to obtain structural information of HMO binders with VP8*. For example, α-1,2-fucosidase could be added to the human milk fractions and the affinity of L24 to VP8*-P[11] monitored. If only α2-6 sialic acid were responsible for the binding, the removal of α1-2 fucose should not change the binding constant, or the change should be very limited.

5.2.2 Discovery of a secondary binding site of BabA

Previous work showed the possibility of having multiple binding sites for blood group antigen-binding adhesion (BabA) from *Helicobacter pylori (H. pylori)*.[2] Although crystal structures

suggested that fucose is critical in the interaction between BabA and oligosaccharides, A. El-Hawiet et. al. found that acidic (3'-Sialyllactose) and non-fucosylated (LNT) HMOs bind to the target protein, with the reported ligand (Lacto-N-difucohexaose II) occupying the known binding site.

By having access to mutated BabA with a modified binding site, this hypothesis could be examined. Using a direct ESI-MS assay with a series of the ligand concentrations, the binding constant of ligands to mutated BabA can be measured. If there is a secondary binding site present on BabA, the K_a value of the mutated BabA to Lacto-N-difucohexaose II should be reduced significantly, whereas the affinities of 3'-Sialyllactose and LNT should be similar to the affinities of the wild-type BabA.

5.2.3 Investigation of interactions between LabA from *H. pylori* to HMOs

LacDiNAc-specific binding Adhesin (LabA) is a newly discovered adhesion from *H. pylori*.[3] Its potential binding partner is LacDiNAc (GalNAcβ1-4GlcNAc), a disaccharide motif found in gastric mucin. A good understanding of the binding mechanisms of LabA to its binder(s) is critical for developing inhibitors that help against *H. pylori* infections. The binding constant of LabA to the LacDiNAc analogue (4-methoxyphenyl 4-O-(2-acetamido-2-deoxy-β-D-galactopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranoside, 530.52 Da, purchased from CarboSynth (Compton, UK)) was measured. Figure 5.1 shows a mass spectrum of an aqueous ammonium acetate solution of LabA 26695, Cytochrome C (Cyt c) served as P_{ref}, and the LacDiNAc analogue.
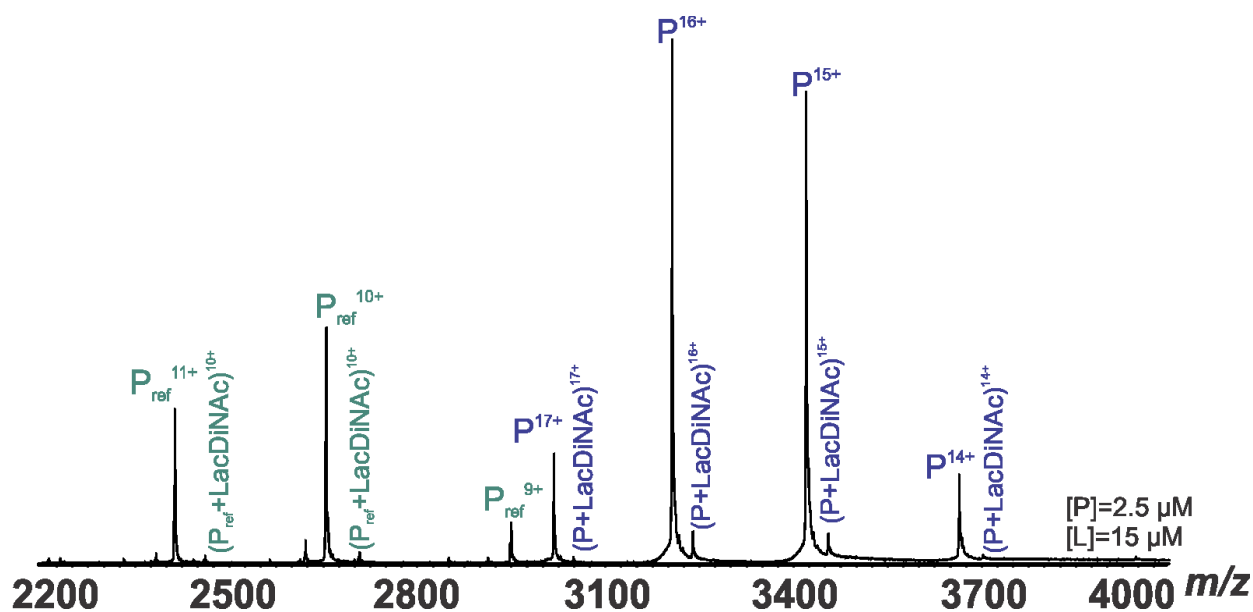
Figure 5.1. A ESI mass spectrum acquired in the positive ion mode for 100 mM aqueous ammonium acetate solutions (pH 6.8, 25 °C) of LabA 26695 (2.5 µM), P_{ref} (3.2 µM), and the LacDiNAc analogue (15 µM).

Only one isoform of LabA was detected, with a molecular weight of 51 277 Da. After correction for nonspecific binding, the affinity of LabA to the LacDiNAc analogue was found to be $3 \times 10^3 \pm 252$ M^{-1}. However, since the analogue has a modification on the reducing end, it could be an interference on the binding constant. On the other hand, chitotriose, chitotetraose, and chitohexaose have the same motif as LacDiNAc, with no modification on the reducing end (Figure 5.2). Therefore, the affinities of LabA to these structures were measured.
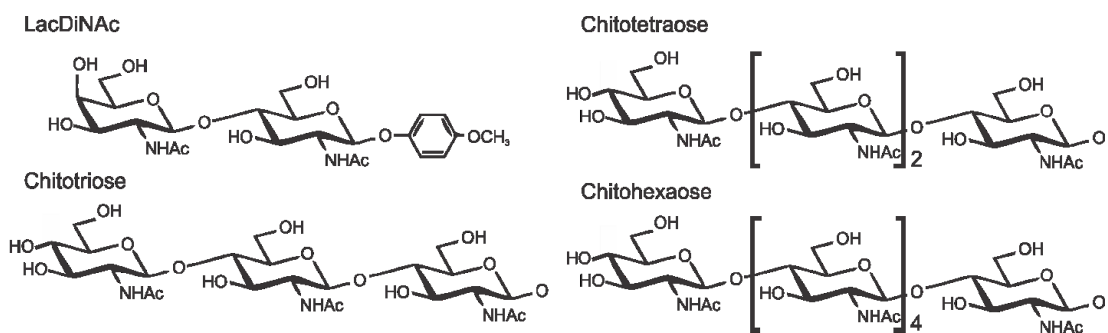
Figure 5.2. Structures of LacDiNAc analogue, chitotriose, chitotetraose, and chitohexaose.
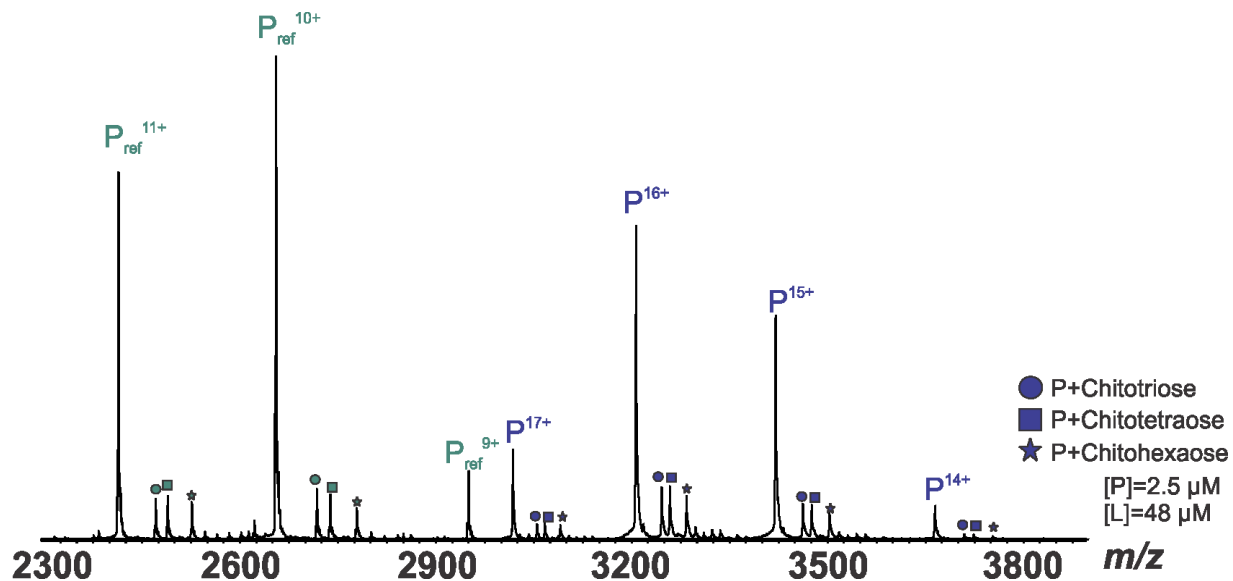
Figure 5.3. A ESI mass spectrum acquired in the positive ion mode for 100 mM aqueous ammonium acetate solutions (pH 6.8, 25 °C) of LabA 26695 (2.5 µM), P_{ref} (3.2 µM), chitotriose (48 µM), chitotetraose (48 µM), and chitohexaose (48 µM).

Shown in Figure 5.3 is a mass spectrum obtained from an aqueous ammonium acetate solution with LabA 26695, P_{ref}, and three chito-oligosaccharides. The affinities for the three chito-oligosaccharides were found to range from 1×10^3 M^{-1} to 2×10^3 M^{-1} (Figure 5.4), suggesting that the modification on the reducing end did not change the binding constant significantly.
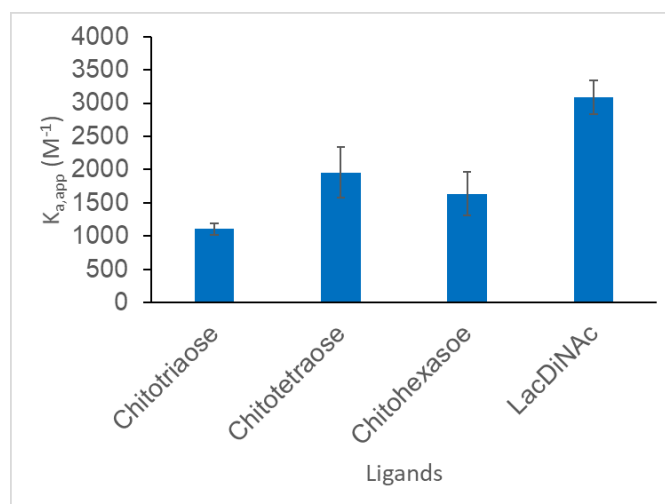
Figure 5.4. A summary of the binding constant ($K_{a,app}$ (M^{-1})) of LabA 26695 to four ligands.

Direct ESI-MS assay results showed that the affinity of LabA 26695 to its potential binding motif is relatively weak ($K_a < 10^4$ M^{-1}). Therefore, it is well worth screening the HMO library to the protein, using the methods developed in Chapter 2 and 3, to obtain more information on the binding pattern of LabA.

5.2.4 Enhancement of studying low affinity interactions using multiple-temperature mass spectrometry

R. Daneshfar et al. developed a variable-temperature device that is compatible with nano-ESI-MS.[4] This method can be beneficial for studying low affinity interactions, which are a shared feature for most carbohydrate–protein interactions, in two ways. First, the lower the temperature is, the higher the binding constant will be for a model system. This means that a higher free protein and protein–complex peak ratio (R) can be obtained from a spectrum, therefore, a more reliable result. Second, from results at different temperatures, the binding constant at 37 °C can be determined; this is physiological temperature for the most interactions we studied. Preliminary experiments have been done on rotavirus VP8*-P[11] to LNnT at 16–0 °C. However, there was

not much change in K_a values between the room temperature and <4 °C. A possible reason for this could be that the energy change in this particular interaction is too low to be observed during the temperature change. An alternative model system with relative high binding affinity (<10^3 M^{-1}), such as norovirus to some of the HMOs, should be tested.

5.3 Literature cited

1. Elena, E.K.; El-Hawiet, A.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1908.

2. El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K. Kitova, E.N.; Kitov, P.I; Bode, L.; Hage, N.; Falcone, F. H.; Klassen, J.S. *Analyst*, **2018**, *143*, 536.

3. Rossez, Y.; Gosset, P.; Bonneca, I.G.; Magalhaes, A.; Ecobichon, C.; Reis, C.A.; Cieniewski-Bernard, C.; Curt, M.J.C.; Leonard, R.; Maes, E.; Sperandio, B.; Slomianny, C.; Sansonetti, P.J.; Michalski, J.; Robbe-Masselot, C. *J. Infect. Dis.* **2014**, *210*, 1286.

4. Daneshfar, R.; Kitova, E.N.; Klassen, J.S *J. Am. Chem. Soc.* **2004**, *126*, 4786.

Bibbiography

Chapter 1

1. World Health Organization, Global strategy for infant and young child feeding. Geneva, Switzerland: World Health Organization and UNICEF, **2003**.

2. WHO Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality. Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. Lancet 2000, *355*, 451.

3. Horta, B.L.; Victora, C.G. Short-term effects of breastfeeding: a systematic review of the benefits of breastfeeding on diarrhea and pneumonia mortality. Geneva: World Health Organization, **2013.**

4. Horta, B.L.; de Mola, C.L.; Victora, C.G. *Acta Paediatr Suppl* **2015**, *104*, 30.

5. Bode L., *Glycobiology* **2012**, *22*, 1147.

6. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. *Annu. Rev. Nutr.* **2000**, *20*, 699.

7. Kumazaki, T.; Yoshiba, A. *Proc. Natl. Acad. Sci.* **1984**, *81*, 4193.

8. Johnson, P.H.; Watkins, W.M.; *Glycoconj. J.* **1992**, *9*, 241.

9. György, P.N.R.; Rose, C.S. *Arch Biochem. Biophys.* **1954**, *48*, 193.

10. Morelli, L.; *J Nutr.* **2008**, *138*, 1791S.

11. Angeloni, S.; Ridet, J.L.; Kusy, N.; Gao, H.; Crevoisier, F.; Guinchard, S. *Glycobiology* **2005**, *15*, 31.

12. Bode, L. *J. Nutr.* **2006**, *136*, 2127.

13. Wang, B.; McVeagh, P.; Petocz, P.; Brand-Miller, J. *Am. J. Clin. Nutr.* **2003**, *78*, 1024.

14. Kramer, M.S.; Aboud, F.; Mironova, E.; the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group *Arch. Gen. Psychiatry.* **2008**, *65*, 578.

15. Bode, L. *Nutr. Rev.* **2009**, 67, Suppl 2, S183.

16. Coppa, G.V.; Zampini, L.; Galeazzi, T.; Facinelli, B.; Ferrante, L.; Capretti, R. *Pediatr. Res.* **2006**, *59*, 377.

17. Weichert, S. *J. Virol.* **2016**, *90*, 4843.

18. Wu, S.; Tao, N.; German, J.B.; Grimm, R.; Lebrilla, C.B. *J. Proteome Res.* **2010**, *9*, 4138.

19. Pierce, M.M.; Raman, C.S.; Nall, B.T. *Methods* **1999**, *19*, 213.

20. Perret, S.; Sabin, C.; Dumon, C. *Biochem. J.* **2005**, *389*, 325.

21. Utsuno, K.; Uludag, H. *Biophys. J.* **2010**, *99*, 201.

22. Koromyslova, A.; Tripathi, S.; Morozov, V.; Schroten, H.; Hansman, G.S. *Virology* **2017**, *508*, 81.

23. Satoh, A.; Kojima, K.; Koyama, T.; Ogawa, H.; Matsumoto, I. *Anal. Biochem.* **1998**, *260*, 96.

24. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Mickum, M.L.; Ashline, D.J.; Prasad, B.V.; Estes, M.K.; Reinhold, V.N.; Cummings, R.D.; Smith, D.F. *Mol. Cell Proteomics* **2014**, *13*, 2944.

25. Song, X.; Lasanajak, Y.; Xia, B.; Smith, D.F.; Cummings, R.D. *ACS Chem. Biol.* **2009**, *4*, 741.

26. Ashline, D.J.; Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Prasad, V.; Estes, M.K.; Cummings, R.D.; Smith, D.F.; Reinhold, V.N. *Mol. Cell Proteomics* **2014**, *13*, 2961.

27. Shams-Ud-Doha, K.; Kitova, E.N.; Kitov, P.I.; St-Pierre, Y.; Klassen, J.S. *Anal. Chem.* **2017**, *89*, 4914.

28. Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818.

29. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605.

30. Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M. *Science* **1989**, *246*, 64.

31. Karas, M.; Bahr, U.; Dulcks, T. *J. Anal. Chem.* **2000**, *366*, 669.

32. Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300.

33. Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A.

34. Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804.

35. Kebarle, P.; Verkerk, U.H. *Mass Spectrom. Rev.* **2009**, *28*, 898.

36. Fenn, J.B. *Angew Chem Int Ed Engl.* **2003**, *42*, 3871.

37. Wu, X.; Oleschuk, R.D.; Cann, N.M. *The Analyst* **2012**, *137*, 4150.

38. Jecklin, M. C.; Touboul, D.; Jain, R.; Toole, E. N.; Tallarico, J.; Drueckes, P.; Ramage, P.; Zenobi, R. *Anal. Chem.* **2009**, *81*, 408.

39. Karas, M.; Bahr, U.; Dulcks, T.; Fresenius J. *Anal. Chem.* **2000**, *366*, 669

40. Konermann, L.; Ahadi, E.; Rodriguez, A.D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2.

41. Iavarone, A.T.; Williams, E.R. *J. Am. Chem. Soc.* **2003**, *125*, 2319.

42. Dole, M.; Mark, L.L.; Hines, R.L.; Mobley, R.C.; Ferguson, L.D.; Alice, M.B. *J. Chem. Phys.* **1968**, *49*, 2240.

43. Batey, J.H. *Vacuum* **2014**, *101*, 410.

44. De Hoffmann, E.; Stroobant, V. Mass Spectrometry Principles and Applications 3th Ed. John Wiley & Sons.: New York, **2007**.

45. Quadrupole mass analyzers: theoretical and practical considerations. In Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics.

46. Scarff, C.A.; Patel, V.J.; Thalassinos, K.; Scrivens, J.H. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 625.

47. Giles, K.; Pringle, S.D.; Worthington, K.R.; Little, D.; Wildgoose, J.L.; Bateman, R.H. *Rapid Comm. Mass Spectrom.* **2004**, *18*, 2401.

48. Wu, L.; Vogt, F.G. *J. Pharm. Biomed. Anal.* **2012**, *69*, 133.

49. Zandkarimi, F.; Wickramasekara, S.; Morré, J.; Stevens, J.F.; Maier, C.S. Electrospray Ionization Traveling Wave Ion Mobility Spectrometry Mass Spectrometry for the Analysis of Plant Phenolics: An Approach for Separation of Regioisomers. D.R. Gang Ed.; Springer Cham: Switzerland **2013**. 21.

50. McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 599.

51. Wiley, W.C.; McLaren, L.H. *Rev. Sci. Instrum.* **1955**, *26*, 1150.

52. Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrom. Rev.* **2000**, *19*, 65.

53. Mamyrin, B.A.; Karataev, V.I.; Shmikk, D.V.; Zagulin, V.A. *Sov. Phys. JETP.* **1973**, *37*, 45.

54. Kitova, E.N.; El-Hawiet, A.; Schnier, P.D.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.

55. El-Hawiet, A.; Shoemaker, G.K.; Daneshfar, R.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2012**, *84*, 50.

56. Wang, W.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2005**, *77*, 3060.

57. Sun, J.; Kitova, E.N.; Wang, W.; Klassen, J.S. *Anal. Chem.* **2006**, *78*, 3010.

58. Jecklin, M.C.; Touboul, D.; Bovet, C.; Wortmann, A.; Zenobi, R. *J. Am Soc. Mass Spectrom.* **2008**, *19*, 332.

59. Yu, Y.; Kirkup, C.E.; Pi, N.; Leary, J.A. *J. Am Soc. Mass Spectrom.* **2004**, *15*, 1400.

60. Sun, J.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2007**, *79*, 416.

61. Bagal, D.; Kitova, E.N.; Liu, L.; El-Hawiet, A.; Schnier, P.D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801.

Chapter 2

1. Newburg, D.S.; Shen, Z.; Warren, C.D. *Adv. Exp. Med. Biol.* **2000**, *478*, 381.

2. Gabrielli, O.; Zampini, L.; Galeazzi, T.; Padella, L.; Santoro, L.; Peila, C.; Giuliani, F.; Bertino,

E.; Fabris, C.; Coppa, G.V. *Pediatrics* **2011**, *128*, 1520.

3. Bode, L. *Glycobiology* **2012**, *22*, 1147.

4. Sela, D.A.; Chapman, J.; Adeuya, A.; Kim, J.H.; Chen, F.; Whitehead, T.R.; Lapidus, A.; Rokhsar, D.S.; Lebrilla, C.B.; German, J.B.; Price, N.P.; Tichardson, P.M.; Mills, D.A. *Proc. Natl. Acad. Sci.* **2008**, *105*, 18964.

5. Eiwegger, T.; Stahl, B.; Haidl, P.; Schmitt, J.; Boehm, G.; Dehlink, E; Urbanek, R.; Szepfalusi, Z. *Pediatr. Allergy Immunol.* **2010**, *21*, 1179.

6. Wang, B.; Yu, B.; Karim, M.; Hu, H.; Sun, Y.; McGreevy, P.; Petocz, P.; Held, S.; Brand-Miller, J. *Am. J. Clin. Nutr.* **2007**, *85*, 561.

7. Hickey, R.H. *Int. Dairy J.* **2012**, *22*, 141.

8. Underwood, M.A.; Gaerlan, S.; De Leoz, M.L.A.; Dimapasoc, L.; Kalanetra, K.M.; Lemay, D.G.; German, J.B.; Mills, D.A.; Lebrilla, C.B. *Pediatr. Res.* **2015**, *78*, 670.

9. Newburg, D.S.; Ruiz-Palacios, G.M.; Morrow, A.L. *Annu. Rev. Nutr.* **2005**, 25, 37.

10. Sisk, P.M.; Lovelady, C.A.; Dillard, R.G.; Gruber, K.J.; O'Shea, T.M. *J. Perinatol.* **2007**, *27*, 428.

11. Kobata, A. *Proc. Jpn. Acad. Sci., Ser. B* **2010**, *86*, 731.

12. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. *Annu. Rev. Nutr.* **2000**, *20*, 699.

13. Totten, S.M.; Zivkovic, A.M.; Wu, S.; Ngyuen, U.; Freeman, S.L.; Ruhaak, L.R.; Darboe, M.K.; German, J.B.; Prentice, A.M.; Lebrilla, C.B. *J. Proteome Res.* **2012**, *11*, 6124.

14. Smilowitz, J.T.; Lebrilla, C.B.; Mills, D.A.; German, B.J.; Freeman, S.L. *Annu. Rev. Nutr.* **2014**, *34*, 143.

15. Hong, Q.; Ruhaak, L.R.; Totten, S.M.; Smilowitz, J.T.; German, J.B.; Lebrilla, C.B. *Anal. Chem*. **2014**, *86*, 2640.

16. Xu, G.; Davis, J. C.; Goonatilleke, E.; Smilowitz, J. T.; German, J. B.; Lebrilla, C. B. *J. Nutr.*

2017, *147*, 117.

17. Ruhaak, L.; Lebrilla, C.B. *Adv. Nutr.* **2012**, *3*, 406S

18. Hirabayashi, J.; Hashidate, T.; Arata, Y.; Nishi, N.; Nakamura, T.; Hirashima, M.; Urashima, T.; Oka, T.; Futai, M.; Muller, W.E. *BBA-Gen. Subjects* 2002, *1572*, 232.

19. Ohtsuka, I.; Sadakane, Y.; Higuchi, M.; Hada, N.; Hada, J.; Kakiuchi, N.; Sakushima, A. *Bioorg. Med. Chem*. **2011**, *19*, 894.

20. Shang, J; Piskarev, V.E.; Xia, M., Huang, P.; Jiang, X.; Likhosherstov, L.M.; Novikova, O.S.; Newburg, D.S.; Ratner, D.M. *Glycobiology*. **2013**, *23*, 1491.

21. Noll, A. J.; Gourdine, J. P.; Yu, Y.; Lasanajak, Y.; Smith, D. F.; Cummings, R. D. *Glycobiology* **2016**, *26*, 655.

22. Ashline, D.J.; Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Prasad, B.V.; Estes, M.K.; Cummings, R.D.; Smith, D.F.; Reinhold, V. N. *Mol. Cell. Proteom.* **2014**, *13*, 2961.

23. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Mickum, M. L.; Ashline, D. J.; Prasad, B. V.; Estes, M. K.; Reinhold, V. N. Cummings, R.D.; Smith, D.F. *Mol. Cell. Proteom.* **2014**, *13*, 2944.

24. Noll, A. J.; Yu, Y.; Lasanajak, Y.; Duska-McEwen, G.; Buck, R. H.; Smith, D. F.; Cummings, R. D. *Biochem. J.* **2016**, *473*, 1343.

25. Grant, O. C.; Smith, H. M.; Firsova, D.; Fadda, E.; Woods, R. J. *Glycobiology* **2014**, *24*, 17.

26. Kilcoyne, M.; Gerlach, J. Q.; Kane, M.; Joshi, L. *Anal. Methods* **2012**, *4*, 2721.

27. Geissner, A.; Seeberger, P.H. *Annu Rev Anal Chem* **2016**, *9*, 223

28. Arthur, C.M.; Rodrigues, L.C.; Baruffi, M.D.; Sullivan, H.C.; Heimburg-Molinaro, J.; Smith, D.F.; Cummings, R.D.; Stowell, S.R. Examining galectin binding specificity using glycan microarrays. In *Galectins: Methods and Protocols*; Stowell, S.R., Cummings, R.D., Eds.; Methods in Molecular Biology 1207; Humana Press: New York, **2015**, 115-131.

29. Paulson, J. C.; Blixt, O.; Collins, B. E. *Nat. Chem. Biol.* **2006**, *2*, 238.

30. He, X. G.; Gerona-Navarro, G.; Jaffrey, S. R. *J. Pharmacol. Exp. Ther*. **2005**, *313*, 1.

31. Shams-Ud-Doha, Km; Kitova, E. N.; Kitov, P., I.; St-Pierre, Y; Klassen, J. S. *Anal. Chem.* **2017**, *89*, 4914.

32. López-Lucendo, M.F.; Solís, D.; André, S.; Hirabayashi, J.; Kasai, K.; Kaltner, H.; Gabius, H.J.; Romero, A. *J. Mol. Biol*. **2004**, *343*, 957.

33. Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U.J.; Akke, M. *J. Am. Chem. Soc*. **2010**, *132*, 14577.

34. Christensen, T.; Toone, E.J. *Methods Enzymol*. **2003**, 362, 486.

35. El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. *Biochemistry* **2012**, *51*, 4244.

36. Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass Spectrom*. **2012**, *23*, 431.

37. El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2011**, *84*, 50.

38. Kitova, E.N.; El-Hawiet, A.; Schnier, P.D.; Klassen, J.S. *J. Am. Soc. Mass Spectrom*. **2012**, *23*, 431.

39. Hofmann, J.; Stuckmann, A.; Crispin, M.; Harvey, D.J.; Pagel, K.; Struwe, W.B. *Anal. Chem*. **2017**, *89*, 2318.

40. Struwe, W.B.; Baldauf, C.; Hofmann, J.; Rudd, P.M.; Pagel, K. *Chem. Com*. **2016**, *52*, 12353.

41. Williams, J.P.; Grabenauer, M.; Holland, R. J.; Carpenter, C. J.; Wormald, M. R.; Giles, K.; Harvey, D. J.; Bateman, R. H.; Scrivens, J.H.; Bowers, M. T. *Int. J. Mass Spectrom.* **2010**, 298, 119.

42. De Kivit, S.; Kraneveld, A.D.; Garssen, J.; Willemsen, L.E.M. *Eur. J. Pharmacol,* **2011**, *668*, S124.

43. Labrie, M.; Vladoiu, M.; Leclerc, B.G.; Grosset, A.A.; Gaboury, L.; Stagg, J.; St-Pierre, Y. *PLoS One*. **2015**, *10*, e0131307.

44. Li, H.; Bendiak, B.; Siems, W. F.; Gang, D.R.; Hill Jr., H. H. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 2699.

45. Brown, D.J.; Stefan, S.E.; Berden, G.; Steill, J.D.; Oomens, J.; Eyler, J.R.; Bendiak, B. *Carbohyd. Res.* **2011**, *346*, 2469.

46. Fenn, L.,S.; McLean, J. A. *Phys. Chem. Chem. Phys*. **2011**, *13*, 2196.

47. Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010

48. Rezaei Darestani R.; Winter, P.; Kitova, E.N.; Tuszynski, J.A.; Klassen, J.S. *J. Am. Soc. Mass Spectrom*. **2016**, *27*, 876.

49. Zhang, H.; Zhang, S.; Tao, G.; Zhang, Y.; Mulloy, B.; Zhan, X.; Chai, W. *Anal. Chem.* **2013**, *85*, 5940.

50. Wheeler, S.F.; Harvey, D.J. *Anal. Chem.* **2000**, *72*, 5027.

51. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1331.

52. Domon, B.; Costello, C.E. *Glycoconjugate J.* **1988**, *5*, 397.

Chapter 3

1. Lara-Villoslada, F.; Olivares, M.; Sierra, S.; Rodríguez, J.M.; Boza, J.; Xaus, J. *Br J Nutr.* **2007**, *98*, S96.

2. Ballard, O.; Morrow, A.L. *Pediatr Clin North Am.*, **2013**, *60*, 49.

3. Martín, R.; Langa, S.; Reviriego, C.; Jiménez, E.; Marín, M.L.; Xaus, J.; Fernández, L.; Rodríguez, J.M. *J Pediatr.* **2003**, *143*, 754.

4. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. *Annu Rev Nutr.* **2000**, *20*, 699.

5. Kulinich, A.; Liu, L. *Carbohydr Res.* **2016**, *432*, 62.

6. Smilowitz, J.T.; Lebrilla, C.B.; Mills, D.A.; German, J.B.; Freeman, S.L. *Annu. Rev. Nutr.* **2014**, *34*, 143.

7. Bode, L. *Glycobiology* **2012**, *22*, 1147.

8. Wu, S.; Tao, N.; German, J.B.; Grimm, R.; Lebrilla, C. *J. Proteome Res.* **2010**, *9*, 4138.

9. Wu, S.; Grimm, R.; German, J.B.; Lebrilla, C.B. *J. Proteome Res.* **2011**, *10*, 856.

10. Finke, B.; Stahl, B.; Pfenninger, A.; Karas, M.; Daniel, H.; Sawatzki, G. *Anal. Chem.* **1999**, *71*, 3755.

11. Stahl, B.; Thurl, S.; Zeng, J.; Karas, M.; Hillenkamp, F.; Steup, M.; Sawatzki, G. *Anal. Biochem.* **1994**, *223*, 218.

12. Domon, B.; Costello, C. *Glycoconj. J.* **1988**, *5*, 397.

13. Xuezheng, S.; Lasanajak, Y.; Xia, B.; Heimburg-Molinaro, J.; Rhea, J.M.; Ju, H.; Zhao, C.; Molinaro, R.J.; Cummings, R.D.; Smith, D.F. *Nat. Methods* **2011**, *8*, 85.

14. Noll, A.J.; Gourdine, J.P.; Yu, Y.; Lasanajak, Y.; Smith, D.F.; Cummings, R.D. *Glycobiology* **2016**, *26*, 655.

25. Noll, A.J.; Yu, Y.; Lasanajak, Y.; Duska-McEwen, G.; Buck, R.H.; Smith, D.F.; Cummings, R.D. *Biochem. J.* **2016**, *473*, 1343.

16. Shams-Ud-Doha, K.; Kitova, E.N.; Kitov, P.; St-Pierre, Y.; Klassen, J.S. *Anal Chem.* **2017**, *89*, 4914.

17. Kitova, E.N.; El-Hawiet, A.; Schnier, P.D.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431.

18. Lin, H.; Kitova, E.N.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 104.

19. Liu, L.; Kitova, E.N.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 310.

20. Wang, W.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2003**, *75*, 4945.

21. Sun, J.; Kitova, E.N.; Wang, W.; Klassen, J.S. *Anal. Chem.* **2006**, *78*, 3010.

22. El-Hawiet, A.; Shoemaker, G.K.; Daneshfar, R.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2011**, *84*, 50.

23. Leney, A.C.; Fan, X.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2014**, *86*, 5271.

24. El-Hawiet, A.; Kitova, E.N.; Klassen, J.S. *Anal. Chem.* **2013**, *85*, 7637.

25. Zhang, Y.; Liu, L.; Daneshfar, R.; Kitova, E.N.; Li, C.; Jia, F.; Cairo, C.W.; Klassen, J.S. *Anal. Chem.* **2012**, *84*, 7618.

26. Han, L.; Kitova, E.N.; Tan, M.; Jiang, X.; Klassen, J.S. *J. Am. Chem. Soc.* **2014**, *25*, 111.

27. Han, L.; Tan, M.; Xia, M.; Kitova, E.N.; Jiang, X.; Klassen, J.S. *J. Am. Chem. Soc.* **2014**, *136*, 12631.

28. Kitova, E.N.; El-Hawiet, A.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1908.

29. El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K.; Kitova, E.N.; St-Pierre, Y.; Klassen, J.S. *Anal. Chem.* **2017**, *89*, 8713.

30. Wheeler, S.F.; Harvey, D.J. *Anal. Chem.* **2000**, *72*, 5027.

Chapter 4

1. Parashar, U.D. *Emerg. Infect. Dis.* **2003,** *9*, 565.

2. Estes, M.K.; Cohen, J. *Microbiol. Rev.* **1989,** *53*, 410.

3. Theuns, S.; Heylen, E.; Zeller, M. *J. Virol.* **2015**, *89,* 1043.

4. Rose, A.S.; Bradley, A.R.; Valasatava, Y.; Duarte, J.M.; Prlić, A.; Rose, P.W. ACM Proceedings of the 21st International Conference on Web3D Technology (Web3D '16), **2016** 185.

5. Agocs, M.M.; Serhan, F.; Yen, F, C. *MMWR Morb Mortal Wkly Rep* **2014**, *3*, 634.

6. Banerjee, I.; Gladstone, B.P.; Le Fevre, A.M. *J. Infect. Dis.* **2007**, *195*, 625.

7. Arias, C.F.; Isa, P.; Guerrero, C.A.; Méndez, E.; Zárate, S.; López, T.; Espinosa, R.; Romero, P.; López S. *Med. Res. Arch.* **2002**, *33*, 356.

8. Jiang, V.; Jiang, B.; Tate, J.; Parashar, U.D.; Patel, M.M. *Hum. Vaccines*. **2010**, *6*, 532.

9. Armah, G.E.; Sow, S.O.; Breiman, R.F. *Lancet,* **2010**, *376*, 606.

10. Madhi, S.A.; Cunliffe, N.A.; Steele, D. *N. Engl. J. Med.* **2010**, *362*, 289.

11. Ramani, S.; Hu, L.; Prasad, B.V.V.; Estes, M.K. *Cell Mol. Gastroenterol Hepatol.* **2016**, *2*, 263.

12. Bode, L. *Early Hum. Dev.* **2015**, *91*, 619.

13. Ling, H.; Kitova, E.N.; Tan, M.; Jiang, X.; Klassen, J.S. *J. Am. Soc. Mass Spectrom*. **2014**, *25*, 111.

14. Shang, J.; Piskarev, V.E.; Xia, M.; Huang, P.; Jiang, X.; Likhosherstov, L.M.; Ratner, D.M. *Glycobiology*, **2013**, *23*, 1491.

15. Ciarelt, M.; Estes, M.K.; *J. gen. virol.* **1999,** *80*, 943.

16. Ciarlet, M. *J. virol.* **2002**, *76*, 4087.

17. Fukudome, K.; Yoshie, O.; Konno, T. *Virology,* **1989**, *172*, 196.

18. Guo, C.; Nakagomi, O.; Mochizuki, M.; Ishida, H.; Kiso, M.; Ohta, Y.; Suzuki, T.; Miyamoto, D.; Hidari, K.; Suzuki. Y. *J. Biochem.* **1999** *126*, 683.

19. Zdanov, A.; Li, Y.; Bundle, D.R.; Deng, S.J.; Mackenzie, C.R.; Narang, S.A.; Young, N.M.; Cygler, M. *Proc. Natl. Acad. Sci*. **1994**, *91*, 6423.

20. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Michkum, M.L.; Ashline, D.J.; Prasad, B.V.V.; Estes, M.K.; Reinhold, V.N.; Cummings, R.D.; Smith, D.F *Mol. Cell Proteomics* **2014,** *13*, 2944.

21. Haselhorst, T. *Nat. Chem. Biol.* **2008**, *5*, 91.

22. Das, S.; Sahoo, G.C.; Das, P.; Singh, U.K.; Jaiswal, A.K.; Singh, P.; Kumar, R. *PLoS One.* **2016**, *11,* e0146243.

23. Blanchard, H.; Yu, X.; Coulson, B.S.; von Itzstein, M. *J. Mol. Biol.* **2007**, *267*, 1215.

24. Huang, P.; Xia, M.; Tan, M,; Zhong, W,; Wei, C.; Wang, L.; Morrow, A.; Jiang, X. *J. Virol.* **2012**, *86*, 4833.

25. Bohm, R.; Fleming, F.E.; Maggioni, A.; Dang, V.T.; Holloway, G.; Coulson, B.S.; von Itzstein, M.; Haselhorst, T. *Nat. Comm.* **2015**, *5*, 5907.

26. Haselhorse, T.; Fleming, F.E.; Dyason, J.C.; Hartnell, R.D; Yu, X.; Holloway, G.; Santegoets, K.; Kiefel, M.J.; Blanchard, H.; Coulson, B.S.; von Itzstein, M. *Nat. Chem. Biol.* **2009**, *5*, 91.

27. Lopez, S.; Aria, C.F. *Trends Microbiol.* **2004**, *12*, 271.

Chapter 5

1. Elena, E.K.; El-Hawiet, A.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1908.

2. El-Hawiet, A.; Chen, Y.; Shams-Ud-Doha, K. Kitova, E.N.; Kitov, P.I; Bode, L.; Hage, N.; Falcone, F. H.; Klassen, J.S. *Analyst*, **2018**, *143*, 536.

3. Rossez, Y.; Gosset, P.; Bonneca, I.G.; Magalhaes, A.; Ecobichon, C.; Reis, C.A.; Cieniewski-Bernard, C.; Curt, M.J.C.; Leonard, R.; Maes, E.; Sperandio, B.; Slomianny, C.; Sansonetti, P.J.; Michalski, J.; Robbe-Masselot, C. *J. Infect. Dis.* **2014**, *210*, 1286.

4. Daneshfar, R.; Kitova, E.N.; Klassen, J.S *J. Am. Chem. Soc.* **2004**, *126*, 4786.