University of Alberta

Characterization of Glycosyltransferases and Glycosidases using Electrospray Mass Spectrometry

by

Naoto Soya

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

Fall 2011
Edmonton, Alberta

Examining Committee

John S. Klassen, Chemistry

Christopher W. Cairo, Chemistry

X. Chris Le, Chemistry

Michael J. Serpe, Chemistry

Christine M. Szymanski, Biological Sciences

Thomas Peters, University of Lüebeck

Abstract

Carbohydrates are one of the major classes of organic compounds found in nature and living organisms, and the diverse roles of carbohydrates are crucial for most life forms. The transformation and degradation of glycans are effectively regulated by carbohydrate-modifying enzymes. This thesis describes the characterization of carbohydrate-modifying enzymes, glycosyltransferases and glycosidases, using electrospray mass spectrometry (ES-MS) to gain a better understanding of these enzymes.

The ES-MS binding assay was applied to quantify the affinities of the human blood group synthesizing glycosyltransferases (GTA and GTB) for their H-antigen substrate in the absence and presence of bound uridine 5′-diphosphate (UDP) and divalent metal cofactor Mn^{2+}. The presence of UDP and Mn^{2+} in the binding site had a marked influence on the association constant (K_a), enthalpy (ΔH_a) and entropy (ΔS_a) for the association of H-antigen to GTs. Moreover, the interactions between GT and nucleotide-sugar donor were investigated. Our results revealed that Mn^{2+} enhances the affinities of donors by 20 ~ 100 times. However, donors undergo enzyme-catalyzed hydrolysis in the presence of Mn^{2+} resulting in monosaccharide and UDP.

The catalytic mechanism of GTA and GTB was also investigated using ES-MS. To trap the glycosyl-enzyme intermediates in their enzymatic reaction, GT mutants, in which the putative catalytic nucleophile Glu^{303} was replaced with Cys, were utilized. The formation of intermediates was observed by incubation of GT

mutants with donor substrates. Tandem MS analysis confirmed Cys^{303} as the site of glycosylation. Incubation of the purified intermediates with H-antigen resulted in the decrease of intermediates and the formation of the trisaccharide products. Our results suggest that the GT mutants could operate by a double displacement mechanism.

The rate of substrate cleavage by the human neuraminidase 3 (NEU3) was measured using ES-MS. The kinetic analysis using synthetic substrates revealed that NEU3 activity depended upon the hydrophobicity of the aglycone. In addition, the substrates with incorporated azide groups in the Neu5Ac residue at either *C*9 or the *N*5-Ac position were cleaved by NEU3. However, the incorporation of larger aryl groups was tolerated only at *C*9, but not at *N*5-Ac.

Acknowledgements

I am truly grateful to the following people who enriched my academic life at the University of Alberta. Without their assistance and understanding, this thesis would never have been completed.

First, I would like to express my gratitude to my supervisor, Prof. John Klassen. I really appreciate providing me the opportunity for research as well as his guidance, encouragement and kind supports to my work. I also would like to thank my supervisor committee members, Prof. Christopher Cairo, Prof. Chris Le, Prof. Michael Serpe and Prof. Christine Szymanski for their thoughtful and helpful comments into my research. Special thanks must go to Prof. Thomas Peters form University of Lüebeck who made an effort to participate in my final defense despite his many other commitments.

I was fortunate to be surrounded by many excellent colleagues and collaborators. My gratitude also goes to my collaborator, Prof. Monica Palcic in Carlsberg Laboratory for her kind supports and helpful discussion. I thank Blake Zheng for his kind help and teaching biological techniques and tools. Moreover, I would like to acknowledge the support and friendship from the current and past members of the Klassen group: Dr. Elena Kitova, Dr. Rambod Daneshfar, Dr. Glen Shoemaker, Dr. Jiangxiao Sun, Igor Sinelnikov, Messele Fentabil, Lan Liu, Nian Sun, Jacquelyn Jhingree, Amr El-Hawiet, Lu Deng, Yixuan Zhang, Hong Lin, Ling Han and Justin Powers.

It is my pleasure to thank my friends at the University of Alberta who made

my life in Edmonton so much better: Dr. Akihiro Imamura, Dr. Yasuharu Watanabe, Dr. Fumie Sunahori, Dr. Eri Adams, Hidenori Mizuno, Satoshi Takebayashi, Tomomi Matsumoto and Hiyo Koizumi.

Finally, I would like to thank my parents, my wife and my daughters who have always been a source of encouragement and motivation for this PhD study.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

4MU-NA	2′-(4-methylumbelliferyl-α-D-N-acetyl-neuraminic acid
AC	Alternating current
ACN	Acetonitrile
Ala or A	Alanine
Arg or R	Arginine
CAZy	Carbohydrate-active enzyme
Cer	Ceramide
CID	Collision-induced dissociation
CRM	Charge residue model
Cys or C	Cysteine
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EC	Enzyme commission
EDTA	Ethylenediaminetetraacetic acid
ES	Electrospray
FAC-MS	Frontal affinity chromatography-mass spectrometry
FTICR	Fourier-transform ion cyclotron resonance
FWHM	Full width at half maximum
Fuc	Fucose
Gal	Galactose
GalNAc	N-Acetylgalactosaminie

GH	Glycosyl hydrolase (or glycosidase)
Glc	Glucose
Gly or G	Glycine
GT	Glycosyltransferase
GTA	α-(1→3)-N-Acetylgalactosaminyltransferase
GTB	α-(1→3)-Galactosyltransferase
IEM	Ion evaporation model
ITC	Isothermal titration microcalorimetry
IS	Internal standard
K_a	Association constant
K_d	Dissociation constant
K_m	Michaelis constant
k_{rel}	Relative reaction rate
Leu or L	Leucine
MALDI	Matrix-assisted laser desorption ionization
MBP	Maltose-binding protein
MeOH	Methanol
Met or M	Methionine
MS	Mass Spectrometry and Mass spectrometer
MS/MS	Tandem mass spectrometry
MW	Molecular weight
MWCO	Molecular weight cut-off
m/z	Mass-to-charge ratio

nanoES	Nano-flow electrospray
NEU	Neuraminidase
NEU3	Neuraminidase 3
Neu5Ac	N-Acetylneuraminic acid
NMR	Nuclear magnetic resonance spectroscopy
OTase	Oligosaccharyltransferase
P_{ref}	Reference protein
RF	Response factor
rf	Radio frequency
scFv	Single chain variable fragment
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser or S	Serine
SPR	Surface plasmon resonance
S/N	Signal-to-noise ratio
UDP	Uridine 5′-diphosphate
UDP-Gal	Uridine 5′-diphosphogalactose
UDP-GalNAc	Uridine 5′-diphospho-N-acetylgalactosamine
UDP-Glc	Uridine 5′-diphosphoglucose
UDP-GlcNAc	Uridine 5′-diphospho-N-acetylglucosamine
UPLC	Ultra-performance liquid chromatography
V_{max}	Maximum rate at saturating substrate concentration
ΔG_a	Gibbs free energy of association
ΔH_a	Enthalpy of association

ΔS_a	Entropy of association

Chapter 1

Characterization of Glycosyltransferases and Glycosidases using Electrospray Mass Spectrometry

1.1 Introduction

Carbohydrates, or saccharides, are one of the major classes of organic compounds found in nature and living organisms.[1] Although carbohydrates perform diverse roles in living organisms, their main roles are as a source of energy in energy storage such as amylose, starch and glycogen, or as a structural component.[1] The role of carbohydrates as an energy source has been extensively studied, and their metabolic pathways in living organisms is well understood.[1] Cellulose[2,3] and chitin,[4] cellular structural components in plant cell walls and animal exoskeletons respectively, are also well known biopolymers made up carbohydrates. In addition to their established roles, however, carbohydrates covalently attached to proteins or lipids have been a focus of recent research due to their significant role in biological development and activities. Glycosylation of proteins is one of the most common post-translational modifications,[5,6] and it plays an important role in protein solubility, structural conformation and interactions with other molecules.[7] Moreover, glycoproteins and glycolipids embedded in the membrane bilayer have numerous biological roles in cell-cell

communications, signaling, bacterial and viral infection, adhesion, inflammation as well as normal and pathological cellular developments.[8–18]

The diverse roles of carbohydrates are crucial for most life forms. Cells in different tissues or organs synthesize different oligosaccharides, or glycans, of various sequence and structure. The synthesis of these glycans will also vary depending on their differentiation state or their physiological environment. The transformation and degradation of glycans in all living organisms are effectively regulated by carbohydrate-modifying enzymes. Carbohydrate-modifying enzymes such as glycosyltransferases (GTs: EC.2.4.x.x) and glycosidases (or glycosyl hydrolase, GHs: EC.3.2.1.x) are a large family of enzymes that are encoded by about 1–2 % of the coding regions of all genomes sequenced to date,[15] and they are classified into families based on their amino acid sequence similarities[19], which can be found in the CAZy (Carbohydrate-Active enZyme) databases at http://www.cazy.org. GTs catalyze the transfer of a monosaccharide from an activated donor, such as a sugar-nucleotide, to a specific acceptor substrate.[20–22] They are responsible for the biosynthesis of the specific glycan structures that are covalently attached to proteins, peptides, lipids, nucleic acids and natural products forming complex biopolymers such as cellulose,[2,3] chitin,[4] starch,[23,24] as well as therapeutically important glycosylated natural product antibiotics[25,26] and anti-cancer agents.[27,28] Oligosaccharyltransferases (OTases: EC 2.4.1.119), which is a type of GTs, catalyze the transfer of glycans from lipid carriers, such as dolichol phosphate and undecaprenol pyrophosphate, to the asparagine residues or serine/ threonine residues of acceptor proteins.[29–31] This reaction catalyzed by OTases is

a central step to biosynthesize N-linked and O-linked glycoproteins.[29–31] To account for the vast numbers of structures found on glycoproteins and glycolipids, it is estimated that more than two hundred GTs exist, with each one forming a unique glycosidic linkage.[32] In contrast, GHs catalyze the cleavage of specific glycosidic linkages between carbohydrate residues of glycans.[33–35] Like the formation of glycosidic bond by GTs, enzyme-catalyzed selective hydrolysis of carbohydrates is also a biologically widespread process and crucial for the modification of glycans as well as degradation of biopolymers such as starch or glycogens to monosaccharides for energy metabolism.[33–36] GHs have also been known to be involved in pathogenesis mechanisms of various diseases,[37,38] anti-bacterial defense[39] and processes in the biosynthesis of N-linked glycoproteins.[40,41]

Altered and aberrant glycosylation or deglycosylation often results in diseases and disorders such as cancers, tumorigenesis, inflammation and other disorders.[42–49] In some cases, these unusual glycosylation or deglycosylation are due to the mutation of certain GTs or GHs caused by genetic disorders.[50,51] Therefore, GTs and GHs have been recognized as attractive drug targets, and inhibitors that reduce specific glycosylations or deglycosylations might be new therapeutics.[47,52–58] In fact, therapeutic GT or GH inhibitors can be found in the treatments against various infections and diseases.[59–62] Moreover, GTs and GHs have been recognized as ideal tools to synthesize oligosaccharides, vaccine and natural sugar products.[22,63–70]

This thesis focuses on the characterization of GTs and GHs using electrospray ionization mass spectrometry (ES-MS) in order to gain a better understanding of these enzymes. In particular, the present work deals with the human blood group synthesizing GTs (GTA and GTB) as well as the membrane-associated human neuraminidase 3 (NEU3). GTA and GTB catalyze the final step in the biosynthesis of ABO(H) blood group A- and B-antigens by transferring a monosaccharide, N-acetylgalactosamine (GalNAc) and galactose (Gal), to the disaccharide acceptor H-antigen, respectively[71,72] (Figure 1.1). A divalent metal cofactor is also required for the proper enzymatic reaction. An amazing feature of these two enzymes is that despite only differing in four of 354 amino acids (Arg/Gly 176, Gly/Ser 235, Leu/Met 266, and Gly/Ala 268), they possess different donor specificities.[71] Although these highly homologous enzymes have been extensively studied by various analytical technique such as X-ray crystallography,[73–80] nuclear magnetic resonance (NMR) spectroscopy[81–84] and radiochemical kinetic assays,[85–87] there remains outstanding questions. For instance, the native oligomeric forms of the recombinant soluble fragments of the GTs are unclear.[75] Moreover, information regarding the binding of nucleotide-sugar donors to GTA and GTB has been limited since attempts to soak or co-crystallize these enzymes with donor molecules have proven difficult.[73,78] The most challenging work in the case of GTA and GTB is in elucidating their catalytic mechanism, which remains unclear despite years of exhaustive studies.[20] Although support for two suggested mechanisms (the double-displacement

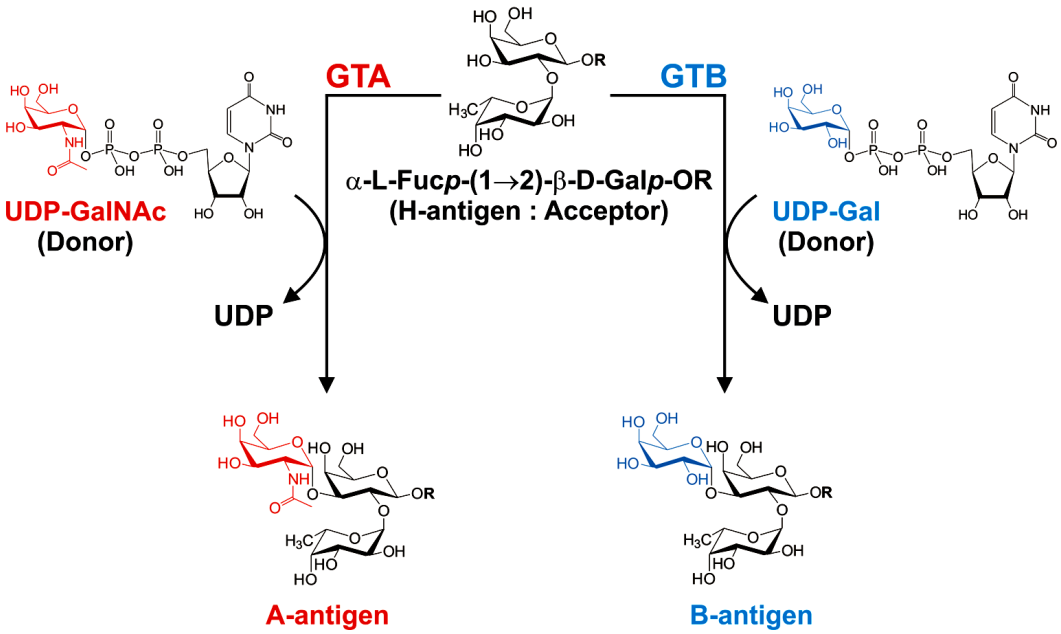

Figure 1.1 The enzymatic reactions catalyzed by the human blood group A and B glycosyltransferases (GTA and GTB). Shown here, the monosaccharide is transferred from a nucleotide donor to the disaccharide H-antigen, α-L-Fuc*p*-(1→2)-β-D-Gal*p*-OR (where R may be a lipid, protein or oligosaccharide), resulting in the trisaccharide A- or B-antigens, α-D-GalNAc*p*-(1→3)-[α-L-Fuc*p*-(1→2)]-β-D-Gal*p*-OR and α-D-Gal*p*-(1→3)-[α- L-Fuc*p*-(1→2)]-β-D-Gal*p*-OR, respectively.

mechanism[88,89] and the S_Ni-like mechanism[90]) exist, there is no direct evidence for both mechanisms.

The human neuraminidases (NEU) consist of a family of four isoforms (NEU1 – NEU4), which differ in their subcellular localization, roles and enzymatic properties including substrate specificity.[91] Among these isoforms, NEU3 is known as a plasma membrane-associated sialidase and catalyzes the cleavage of terminal sialic acid from ganglioside substrates such as GM3, GD1a, GD1b and GT1b[92,93] (Figure 1.2). Recently, NEU3 has gained interest due to its role in diseases and cell surface functions.[94–97] However, despite its apparent importance in disease and membrane structure, to date the substrate specificity of NEU3 has not been well defined at the molecular level. In fact, recent work has primarily examined the enzymatic activity of NEU3 using natural ganglioside substrates. A detailed study of the NEU3 activity using synthetic glycolipid substrates analogues could aid in a better understanding of the substrate recognition by NEU3, and thus the design of new therapeutical inhibitors.

In the present work, we seek to develop a more complete understanding of GTA, GTB and NEU3 using ES-MS assays. In the following sections, a brief introduction of ES-MS, the *direct* ES-MS binding assay, ES-MS kinetic assay, peptide sequencing using tandem MS (MS/MS) utilized in present studies will be discussed.

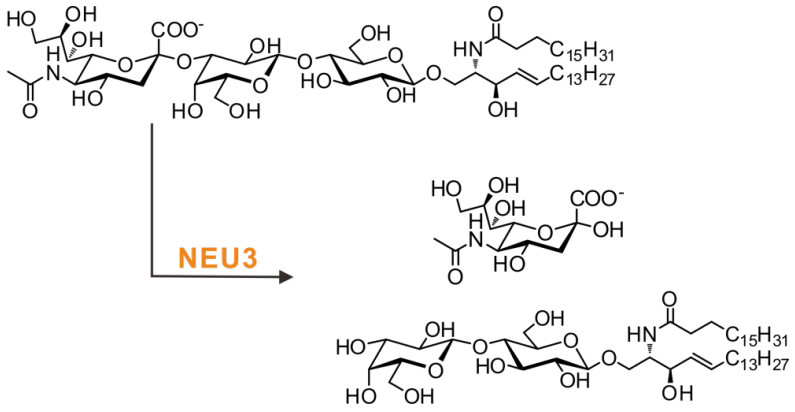

Figure 1.2 The enzymatic reactions catalyzed by the human neuraminidase 3 (NEU3). Shown here, the terminal sialic acid is cleaved from GM3, resulting in α-D-Neu5Ac and β-D-Gal*p*-(1→4)-β-Glc*p*-(1→1)-Cer.

1.2 Mass Spectrometry

A mass spectrometer is an analytical tool to separate ions according to their mass-to-charge ratio (m/z), detect them as electronic signals and produce mass spectrum. The instrument principally can be divided into three parts, an ion source to generate gaseous ions, a mass analyzer to separate ions and a detector. There exist a number of ionization techniques and analyzers for MS,[98] and the suitable combination of ionization technique and an analyzer rely on the nature of analytes and the type of information required.

The utility of mass spectrometry for the analysis of biological molecules has grown enormously in the past decades due to, in particular, the emergence of two "soft" ionization techniques, matrix-assisted laser desorption/ionization (MALDI)[99,100] and electrospray ionization (ES).[101–103] In MALDI, analytes are co-crystalized with a chemical matrix, which protects analytes from being destroyed by laser irradiation as well as facilitates their vaporization and ionization. They are deposited under vacuum, and the matrix sublimates and carries the analyte molecules into the gas phase upon the exposure to pulsed laser irradiation.[99,100] In ES, analytes are sprayed directly from solutions by applying a strong electric field gradient to the solutions under atmospheric pressure. While both ionization techniques allows the generation of intact macromolecular ions, ES is usually utilized for mass spectrometric studies of noncovalent protein complexes formed in solution.[104] There are also a number of different types of mass analyzers including magnetic sector, quadrupole, ion trap, time of flight, Fourier-transform ion cyclotron resonance (FTICR) and most recently orbitrap.[98] In the present

study, nano-flow ES (nanoES) combined with FTICR-MS was used to study noncovalent protein-ligand complexes; these topic are discussed further in the following sections.

1.2.1 Electrospray Ionization

The mechanism of the ES processes involves the production of charged droplets from electrolytes dissolved in a solvent. The small droplets are formed by applying a strong electric field to a metal capillary or a metal wire, such as platinum (Pt), inserted inside the glass tip containing the solution under atmospheric pressure. The subsequent shrinkage of the droplets by solvent evaporation accompanied with droplet fissions leads to the formation of fine, highly charged droplets from which the gas phase ions are produced.[101–103] Shown in Figure 1.3 is a diagram describing the ES processes involved in the formation of gas phase ions. In positive-ion mode, a high positive voltage (3 ~ 6 kV) is applied to the ES emitter, typically a stainless steel capillary. The electric field induces charge separation in the solution, whereby the positively charged electrolyte drifts towards the liquid surface leading to the formation of a liquid cone referred to as a Taylor cone. At a sufficiently high electric field, the liquid cone becomes unstable and emits a thin liquid filament which subsequently breaks up into small positively charged droplets. With solvent evaporation, these droplets start to shrink into the smaller droplets. As the charge density on the droplet surface increases to near the Rayleigh limit, the point at which the Coulombic repulsion of the surface charges is equal to the surface tension of the

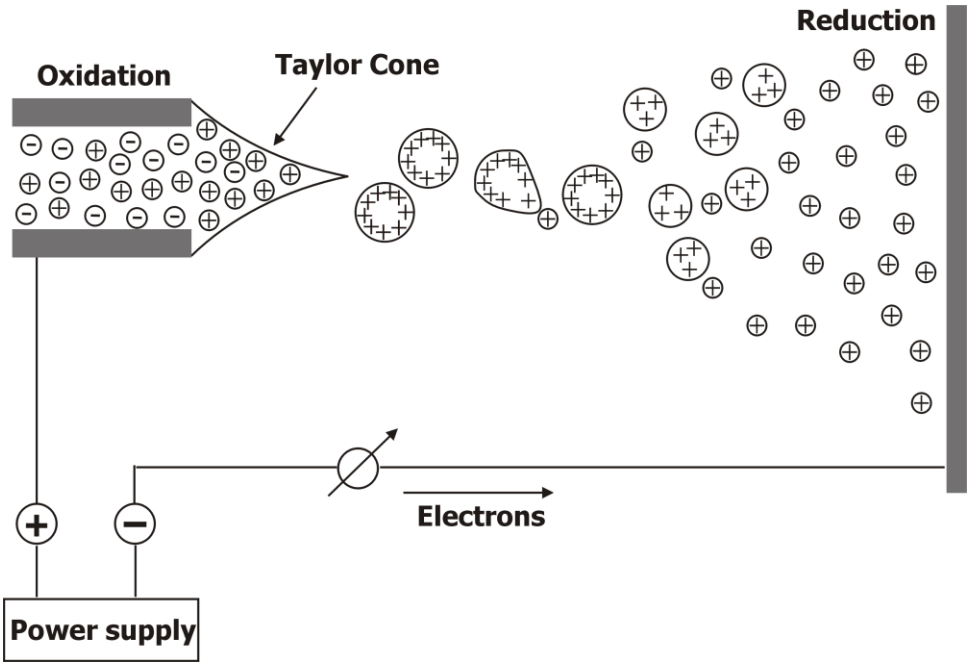

Figure 1.3 Schematic diagram of ES performed in positive-ion mode and the processes that lead to the formation of gaseous ions.

droplets, these droplets undergo Rayleigh fission, eventually forming small highly charged offspring droplets.

Production of gaseous ions from small offspring droplets follows one of two mechanisms: the ion evaporation model (IEM) and the charge residue model (CRM). The IEM[105,106] assumes that ion emission occurs directly from very small, highly charged droplets. This model is considered to operate for small molecules.[102] In contrast, gaseous ions of macromolecules, such as proteins and non-covalent protein-ligand complexes, produced by ES are generally thought to form by the CRM.[102,107] According to this model, the droplets undergo multiple evaporation/fission events, and finally fine droplets containing a single analyte molecule is produced.[108] Further solvent evaporation from these droplets results in multiply charged, gaseous ions of macromolecules.

In the present work, the low flow variant of ES, nanoES,[109–111] was utilized. The mechanism of nanoES is considered to be similar as ES, except that nanoES operates at lower solution flow rates (typically ~ 20 nL/min), lower voltage (~ 1 kV in positive-ion mode) and produces correspondingly smaller droplets than conventional ES. In our group, nanoES tips were prepared by pulling glass capillaries (borosilicate tube) to ~5 μm o.d. at one end using a pipette puller. The advantages of nanoES include the lower sample consumption (typically 2 ~ 3 μL) and smaller amount (pico-gram level) of required protein.[109–111] This is critical for the work presented here since both protein and ligand are either expensive or difficult to synthesize. In addition, the smaller initial droplets with a diameter less than 200 nm generated by nanoES, which are

100 to 1000 times smaller than those produced by conventional ES,[110,112] readily allows the transfer of proteins and protein-ligand complexes from buffered aqueous solutions to the gas phase. In addition, the shorter lifetimes of the nanoES droplets due to the smaller initial droplet size are likely advantageous for preserving the original solution composition throughout the formation of gaseous ions. Furthermore, nanoES can minimize the extent of nonspecific binding; *vide infra*, from ES processes due to there being fewer analyte molecules in each droplet.[113]

1.2.2 Fourier-transform Ion Cyclotron Resonance Mass Spectrometry

Ions generated by nanoES are separated based on *m/z* and detected by an FTICR-MS. FTICR-MS achieves the highest resolving power and mass accuracy among available MS instruments. High mass resolution is critical for the quantitative analysis of protein-ligand binding affinities to adequately resolve protein ions from protein-ligand complex ions with similar *m/z* values. The general operating principles of FTICR-MS are described in many reviews and books,[98,114–116] and, therefore, only a brief overview is given here.

The principle of the FTICR-MS is based on ion cyclotron motion, which arises from the circular motion of charged particles in magnetic field with a frequency that is related to their mass and the number of charge. Figure 1.4 illustrates the cyclotron motion of a positively and negatively charged ion subjected to a static magnetic field. The cyclotron frequency, ω_c, is described in *eq* 1.1:

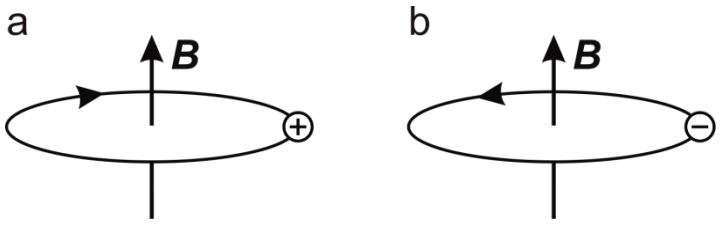

Figure 1.4 Cyclotron motion of a) a positive ion and b) a negative ion in magnetic field, *B*.

$$\omega_c = \frac{zeB}{m} \tag{1.1}$$

where z is the number of charge on the ion, e is the elementary charge, B is the magnetic field strength, and m is the mass of the ion. A notable feature of *eq* 1.1 is that ω_c is inversely proportional to the *m/z* of a give ion. Moreover, all ions of a given *m/z* rotate at the same frequency, independent of their velocities. The ultra-high resolution achieved by FTICR-MS is a direct result of insensitivity of the cyclotron frequency to the kinetic energy distribution.

Figure 1.5 describes how a mass spectrum is generated from the ion cyclotron motion. In order to produce a signal for the ions trapped in the ICR cell, a packet of ions with given *m/z* needs to be excited by applying on oscillating electromagnetic field such as provided by an AC signal generator. If the frequency of the applied field matches the ω_c of the ions, they will absorb energy and thus increase their orbital radius of the circular trajectory, but keeping a constant cyclotron frequency. Shown in Figure 1.5 is the spiral trajectory of excited ions with the same *m/z* at their cyclotron frequency. As the coherently orbiting ions pass a second, opposing pair of electrodes (detection plates), which are parallel to the magnetic axis, they induce an alternating current (image current) on the plates (Fig. 1.5). The packet of ions gain enough cyclotron energy to move out near the detection plates, but the ions absorbing the energy will not achieve their orbital radius that exceeds the dimensions of the ICR cell. The amplitude of this image current is proportional to the number of ions in the ICR cell while the frequency of the alternating current matches the ω_c of the ions.

A Fourier-transform converts the detected image current from the time

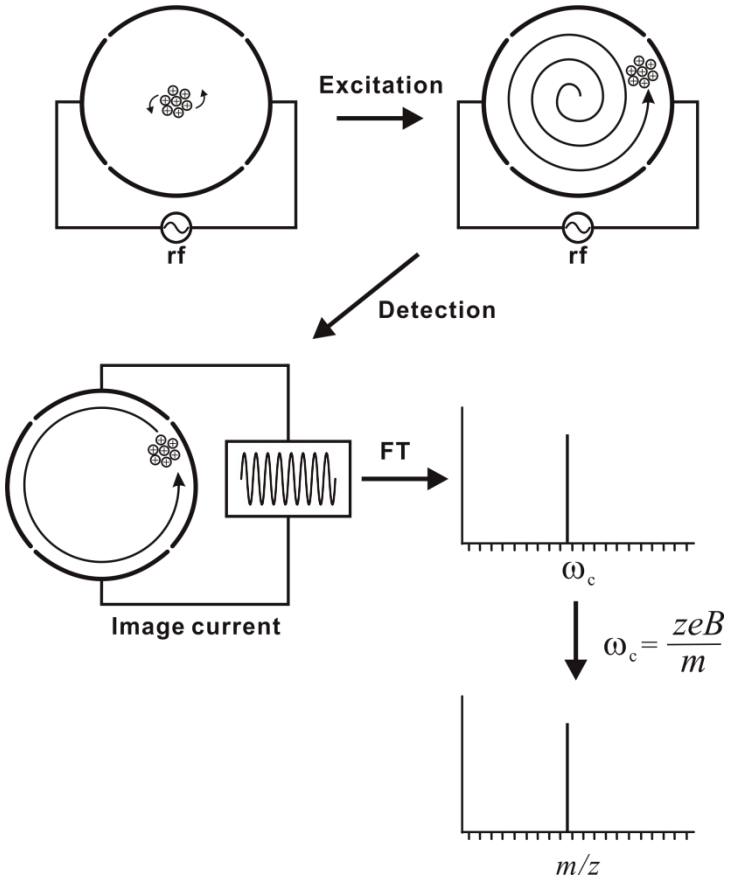

Figure 1.5 Illustration of excitation, image current detection and the production of mass spectrum by FTICR-MS.

domain into the frequency domain, and a mass spectrum can be obtained using equation 1.1 (Figure 1.5). The resolving power of FTICR-MS can routinely reach hundreds of thousands, typically measured as the full width at half maximum (FWHM).[98,116] The resolving power is proportional to the magnetic field strength (with higher magnetic field having higher resolution), and the acquisition time. The acquisition time is the duration of the detection phase, which is determined by the dataset size and the frequency of sampling. Longer acquisition time (larger dataset size) results in higher resolution. Therefore, high vacuum (10^{-10} mbar) is necessary in the cell region of FTICR-MS, to avoid the collision with residual gases and the deactivation of the ions.

Shown in Figure 1.6a is a schematic diagram of the Bruker Apex-II nanoES-FTICR-MS used in the present work. Briefly, fine droplets produced by nanoES are introduced into the mass spectrometer through a heated metal capillary, which facilitates the desolvation of the droplets. Resulting gaseous ions are transmitted through a skimmer and accumulated in the hexapole for a certain time period to enhance the signal-to-noise (S/N) ratio. After accumulation, ions are introduced into the ICR cell for detection. The typical base pressure for the instrument is ~5×10^{-10} mbar, maintained by the differential pumping system.

The mass spectrometer used in the enzyme kinetics of NEU3 is a Bruker Apex-Qe nanoES-Qh-FTICR-MS (Figure 1.6b). Apex-Qe is a hybrid quadrupole/FTICR mass spectrometer, in which two mass analyzers are combined. The first quadrupole acts as a mass filter to select and isolate targeted analyte efficiently for tandem MS (MS/MS) analysis. In the present work, however, the

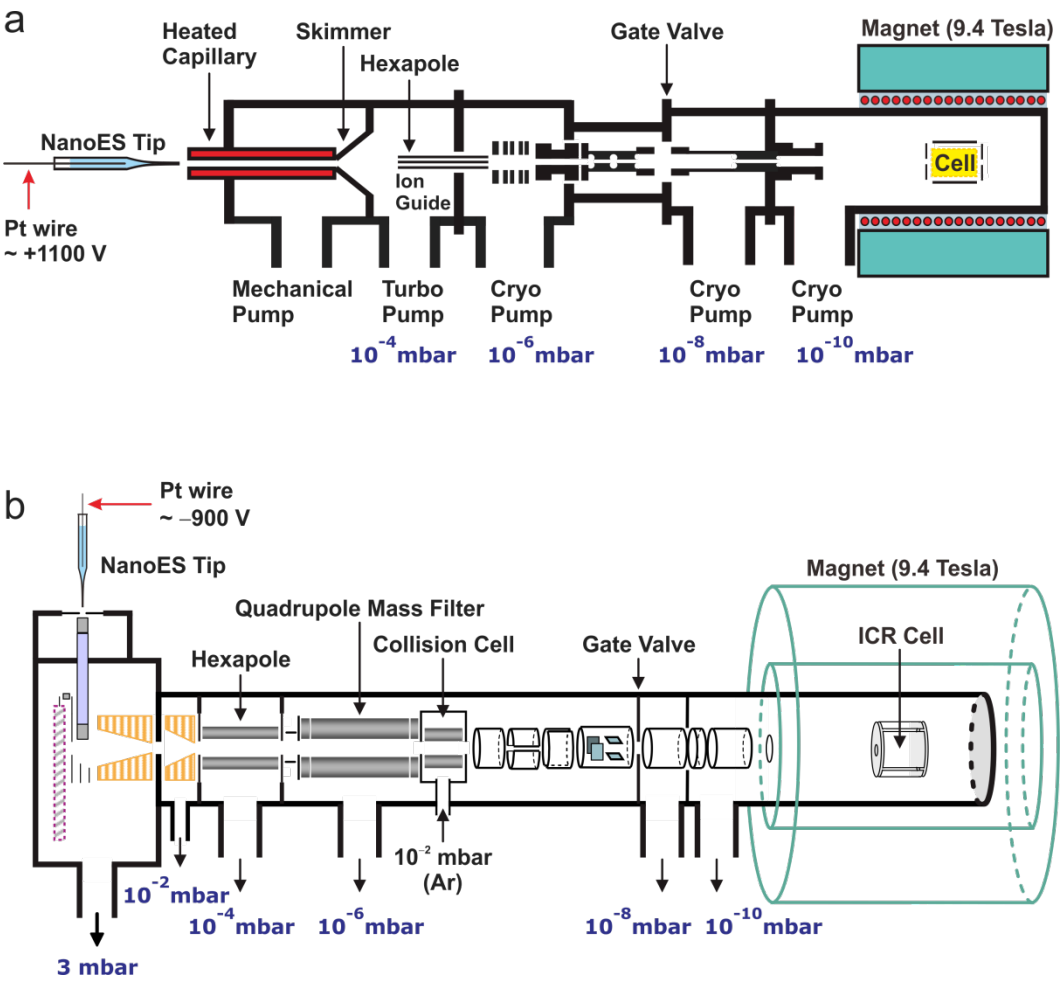

Figure 1.6 Schematic diagrams of (a) the Bruker Apex-II nanoES-FTICR-MS and (b) the Bruker Apex-Qe nanoES-Qh-FTICR-MS used in this study.

quadrupole was operated in radio frequency (rf)-only mode for the transmission of all ions. The fine droplets produced by nanoES are introduced into the mass spectrometer through a glass sampling capillary (i.d. 0.5 mm). Nitrogen gas at 90 °C is used as drying gases to facilitate the desolvation of the droplets. Resulting gaseous ions are transmitted through the funnels and skimmers. The ions are stored electrodynamically in the first hexapole followed by further accumulation in a second hexapole collision cell. Following accumulation, the ions are transferred from the hexapole collision cell, and through a series of ion optics, introduced into the ICR cell for detection.

1.3 *Direct* ES-MS Binding Assay

An important first step in an enzymatic reaction is the formation of enzyme-substrate complexes. It is, thus, worthwhile to characterize their interactions with substrates and to understand their substrate specificity or recognition mechanism at the molecular level. A better understanding of these interactions will also aid the design of new therapeutic enzyme inhibitors. Two methods that have greatly enhanced our understanding of these interactions, allowing the atomic resolution of protein-ligand complexes, are X-ray crystallography and NMR spectroscopy.[117–120] The first protein crystal structure, sperm whale myoglobin, was solved in 1958 by sir John Cowdery Kendrew.[121] Since then, protein crystallography has been used extensively in the determination of protein structure as well as the structure of protein-ligand complexes at the atomic level. To this day, X-ray crystallography remains the gold standard in the

structural characterization of protein complexes. According to the protein data bank[122] (http://www.rcsb.org/pdb/home/home.do), as of June 2011 over 73,503 protein structures have been determined using X-ray crystallography. However, not all biological molecules are easy to crystallize, and there is also the concern that the interactions in the crystalline state could be different from those in solution.[118] NMR spectroscopy is an important alternative to X-ray crystallography and provides precise three-dimensional structure details and binding interactions of protein-ligand complexes in solution with atomic resolution. However, two disadvantages of NMR include the large sample size requirements (milli-gram level) and a limited mass range (up to 40 kDa, although the structure of a 90 kDa protein was recently reported).[123] Moreover, both of these techniques are very labor intensive including the careful determination of experimental conditions and extensive data processing.

Surface plasmon resonance (SPR) is a widely used technique to determine the affinity and stoichiometry of protein-ligand complexes as well as the association and dissociation kinetics of their interactions.[124,125] It can also provide real time information regarding the binding process; thus, the dynamic aspects of interactions can be studied. Binding measurements by SPR require the immobilization of proteins to the sensor surface coated by gold film. The ligands are then flowed past the surface at a constant concentration. This is followed by a dissociation phase, in which buffer solution without the ligands is passed over the surface. Binding of ligands in solution to surface-immobilized proteins alters the refractive index of the sensor surface, which is monitored in real time. The

corresponding signal-versus-time graph contains information regarding the amount of bound ligands, their affinity for the protein and the association and dissociation rate of protein-ligand complexes.[124] SPR can measure interactions with a dissociation constant (K_d) of μM to sub nM and requires a very small amount of sample. However, one major concern with this technique is that the immobilization process may affect the nature of the interactions being investigated. For example, the protein may lose its native conformation, or the accessibility of the binding site may be compromised. Moreover, any adsorption of a ligand to the surface will produce a signal change; thus, nonspecific binding is a concern.[124]

Isothermal titration microcalorimetry (ITC) is one of the most widely used methods to quantify protein-ligand interactions and is the only technique that provides a direct measurement of the enthalpy of association.[126,127] In the ITC experiment, one reactant is gradually added to a solution containing the other reactant. The heat change involved in each addition is monitored by the calorimeter. The analysis of the energetic process provides information on biomolecular interaction, protein denaturation and enzyme inhibition. However, the large amount of sample consumed (mg) and low throughput (5–10 measurements/day) are major limitations for ITC.

Recently, ES-MS has emerged as a promising tool for quantifying protein-ligand interactions in solution, and MS-based binding assays have been developed by many groups.[104,128–130] The *direct* ES-MS binding assay, which has been extensively developed by the Klassen group, is based on the direct detection of

free and ligand-bound protein ions by ES-MS.[131] The association constant (K_a) for a given protein-ligand complex is determined from the ratio (R) of the total abundance (Ab) of bound and unbound protein ions (e.g., PL^{n+}, P^{n+}) measured in the gas phase by ES-MS for solutions of known initial concentrations of protein ($[P]_o$) and ligand ($[L]_o$), *eqs* 1.2 and 1.3.

$$K_a = \frac{R}{[\mathrm{L}]_o - \frac{R}{1+R}[\mathrm{P}]_o} \tag{1.2}$$

$$\frac{[\mathrm{PL}]_{eq}}{[\mathrm{P}]_{eq}} = \frac{\sum_n Ab(\mathrm{PL}^{n+})}{\sum_n Ab(\mathrm{P}^{n+})} = R \tag{1.3}$$

The technique boasts a number of strengths, including its simplicity (no labeling or immobilization required), speed (measurements can usually be completed within a few seconds), and the unique ability to provide direct insight into stoichiometry and to study multiple binding equilibria simultaneously. Additionally, when performed using nanoES, the ES-MS assay affords high sensitivity, normally consuming picomoles or less of analyte per analysis. The ES-MS assay has been applied to a number of protein-ligand complexes,[132–138] and recently a new ES-MS approach for quantifying labile protein-ligand interactions was developed by the Klassen group.[139] In many instances, the K_a values agree well with those obtained by more established analytical methods, such as ITC and SPR. However, despite these many successful examples, there remain issues regarding the reliability of the binding data obtained by ES-MS.

The successful implementation of the direct ES-MS technique for the determination of reliable values of K_a requires that the equilibrium ratio of bound-

to-unbound protein initially present in solution be preserved both throughout the ES process and in the gas phase. Any physical or chemical process that alters this ratio will lead to incorrect K_a values. Below, the solution and gas phase processes that are known to affect the affinity measurements are summarized.

1.3.1 Non-uniform Response Factors

The abundance of each protein species measured by ES-MS is related to the solution concentration by its response factor (*RF*),[140,141] which collectively accounts for the ionization, ion transmission and detection efficiencies:

$$\frac{[\mathrm{PL}]_{\mathrm{eq}}}{[\mathrm{P}]_{\mathrm{eq}}} = \frac{RF_{PL}Ab(\mathrm{PL})}{RF_{P}Ab(\mathrm{P})} = RF(\frac{Ab(\mathrm{PL})}{Ab(\mathrm{P})}) \qquad (1.4)$$

Underlying the derivation of the above expression for K_a is the assumption of uniform *RF* values for the unbound and ligand-bound protein ions; therefore $RF_P = RF_{PL}$ and $RF_{PL}/RF_P = RF = 1$.[140,141] If this assumption is valid, the abundance ratio can be directly used as the concentration ratio. This assumption is likely valid only in cases where the ligand is small compared to the protein, such that the size and surface properties of the free and ligand-bound protein are similar.[140,141] This is the case for the enzyme-substrate complexes analyzed in the present study, as the addition of a small substrate (< 1000 Da) to a large protein (> 60 kDa) has little effect on its response factor.

1.3.2 pH Change during Electrospray

Electrochemical reactions, which occur at the electrode in the ES tip, can alter the composition of the solution.[142] In aqueous solution the dominant

electrochemical reactions occurring at a chemically inert electrode are oxidation (positive ion mode) and reduction (negative ion mode) of H_2O leading to the production of H_3O^+ (*eq* 1.5a) and OH^- (*eq* 1.5b), respectively.

$$2H_2O \rightarrow O_2 + 4H^+ + 4e^- \quad (1.5a)$$

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad (1.5b)$$

At low solution flow rates (<100 nL/min), the resulting pH changes can be large, >1 pH unit after 30 min of spraying.[142] Given that binding affinities of protein-ligand complexes can be sensitive to pH, such changes may result in error into the affinity measurements. Thus, the use of ES solutions with 5 ~ 10 mM ammonium acetate buffer or short spraying times (< 10 min)[142] are recommended to minimize errors in K_a introduced by pH changes.

1.3.3 In-source Dissociation

The collision of resident gaseous molecules in the ion source can give rise to collision-induced dissociation (CID), or in-source dissociation (ISD), of the protein-ligand complexes. The dissociation of protein-ligand complexes by ISD can alter the relative abundance of bound and unbound protein ions, resulting in false-negative K_a values. Especially in the atmospheric pressure ionization technique, such as ES, ISD must be concerned for the MS binding assay. The influence of ISD depends on the configuration of the ion source and the stability of the complexes in the gas-phase environment. Collisional heating of the gaseous complex ions by resident gases may occur at various stages during the ion sampling process, such as within the heated metal sampling capillary (if used), in

the nozzle (or orifice)-skimmer region, and during accumulation of ions within hexapoles in FTICR-MS. In cases where the gaseous complexes are susceptible to ISD, low temperatures in sampling capillary, low potentials across lens elements and short accumulation times in hexapoles are essential to determine reliable K_a values. However, there are usually compromises between the use of gentle ion source conditions and sensitivity. Thus, a balance must be found between minimizing dissociation and obtaining sufficient sensitivity. Where sufficiently gentle sampling conditions are not feasible, the extent of ISD may be reduced through the addition of small organic molecules, such as imidazole,[143] to the ES solution or the introduction of gas phase additive, such as SF_6.[144] From the research reported by the Klassen group and the Loo group, protein-carbohydrate and protein-nucleotide diphosphate complexes are quite stable in gas-phase environment.[145,146] However, the accumulation times in hexapole and the voltages applied to the ion focusing lens throughout the MS analysis were minimized to avoid potential gas-phase dissociation of the complexes.

1.3.4 Nonspecific Binding

One of the largest challenges in developing ES-MS based binding assays is the occurrence of nonspecific binding during the ES processes. These are complexes that are not present in bulk solution but form in the ES droplets, due to concentration effects.[145,147] The formation of such false-positives during the ES process is undesirable as it obscures the true binding stoichiometry and introduces error into the measured K_a value. Generally, the formation of nonspecific protein-

ligand complexes can be minimized by limiting the initial concentration of ligand. For very weak ligand interactions (K_a <10^4 M^{-1}), however, high initial concentrations of ligand are required to produce detectable levels of complex.[148] In such cases, nonspecific binding is often unavoidable. In the present work, the reference protein method, which was developed by the Klassen group, is utilized to correct the ES mass spectra for the occurrence of nonspecific protein-ligand binding.[148] Briefly, this method involves the addition of a reference protein (P_{ref}), which exhibits no specific binding with the ligands of interest, to the ES solution. The occurrence of nonspecific protein-ligand binding is monitored by the appearance of ions corresponding to nonspecific complexes of P_{ref} and ligand in the mass spectrum. It was shown previously that the extent of nonspecific binding of neutral and acidic carbohydrates is independent of the size and structure of the protein.[145,148] As a result, the distribution of ligands bound nonspecifically to P_{ref} can be used to correct the measured intensities of the free and ligand-bound protein ions for the occurrence nonspecific ligand binding, leading to more reliable values of K_a.

1.4 ES-MS Kinetic Assay

In enzyme kinetics, the reaction rate catalyzed by enzymes is measured by monitoring the formation of enzymatic products or the degradation of substrates, and the factors that influence the reaction rate, such as temperature, pH, the interaction between enzyme and substrate as well as the presence of inhibitors, are examined. Widely used techniques to measure enzyme kinetics are spectroscopic

and radiochemical assays.[149,150] In spectroscopic assays, a synthetic substrate with a chromophore is used to elucidate enzymatic activity. The rate of the enzymatic reaction is monitored by measuring a change in emission (in fluorometric assay) or absorption (in spectrophotometric assay) of light. In a radiochemical assay, a radiolabeled substrate is utilized, and the release of the radioactive label from the substrates results in a change in radiation intensities, which is monitored and related to product formation. Although these assays are well established and widely employed, there are several limitations. The critical drawback in both spectroscopic and radiochemical assays is that these assays depend on chromogenic or radioactive substrates, respectively. When the availability of chromogenic and radioactive-labeled substrates is limited, they must be prepared by labor-intensive, multistep synthesis. Moreover, the modification by chromogenic agents can alter enzyme kinetics in a spectroscopic assay.[151] A radioactive-labeled substrate is more preferred due to the fact that it possesses similar recognition properties to an original substrate, and they usually exhibit identical enzymatic kinetics. However, the use of radioisotopes requires careful operation and generates hazardous radioactive waste.

A number of groups have employed ES-MS to study enzyme kinetics for different enzymes.[152–158] As mentioned above, the considerable advantages of ES-MS are the low consumption of enzyme and synthetic substrates, which can both be quite precious, as well as the capability to detect virtually any products and/or substrates without chemical modifications. However, one of the major issues for the application of ES-MS to enzyme kinetics is that the correlation between the

abundance of ions in the mass spectra and concentration in electrosprayed solution must be known to quantify the concentration of enzymatic products. As already mentioned, the abundance of gas-phase ions measured by ES-MS is related to the analyte concentration by its response factor, and the difference in the response factors originates from the difference in ES ionization efficiency, transmission efficiency in the mass analyzer and detector efficiency.[140,141] In cases where response factors differ, the relative abundance of the gas phase ions measured by ES-MS will not reflect the concentrations of the individual species in solution.[140,141] Moreover, the response factors can differ in each measurement, which results in varied signal intensities for even same concentration of analytes. Therefore, the response factors influence the accuracy and precision in quantification of analytes. In the following section, a calibration curve method is introduced, which can correlate the solution concentration of a given analyte to its intensity in the mass spectrum.

1.4.1 Calibration Curve

To implement a quantitative kinetic assay with ES-MS, an internal standard (IS), which has similar chemical properties and functionalities to the enzymatic product, is used to cancel out any effects relating to response factors.[159–161] The concentration of enzymatic products can be quantified using calibration curves. The signal intensity of the product and an internal standard (I_P and I_{IS}, respectively) are used to determine an intensity ratio (I_P/I_{IS}). Calibration curves are generated by plotting I_P/I_{IS} against a known concentration ratio of

product and internal standard ([P]/[IS]) to provide a slope (*m*) and intercept (*Int*). The experimental concentration of the product, [P], is then estimated using the following equation:[161]

$$[\mathrm{P}] = \left(\frac{I_{\mathrm{P}}}{I_{\mathrm{IS}}} - Int \right) \times \frac{[\mathrm{IS}]}{m} \tag{1.7}$$

1.4.2 Enzyme Kinetics

In general, the time course of product formation in an enzymatic reaction is initially linear, but the rate of product formation starts to decline at longer incubation times (Figure 1.7a). The factors that reduce the enzyme activity include substrate depletion, product inhibition, change in pH and ionic strength as well as denaturation of the enzyme.[149,150] However, these effects should not be significant at the early time periods, and thus initial rates (v_0) are used for the analysis in enzyme kinetics to avoid these complexities. To measure v_0, enzyme assays are carried out at the early time periods in which less than 10–20% of the total substrate consumption has occurred. The v_0 values can be determined by plotting the concentration of the product against the reaction time and drawing the initial linear portion of the progress curve (Figure 1.7a).

When v_0 of the enzymatic reaction is measured over a range of substrate concentration, a plot of v_0 against substrate concentration displays saturation kinetics as shown in Figure 1.7b. Saturation occurs since the number of active sites in enzymes is limited, and they are fully occupied by excess substrate. In addition, reaction rate as a function of substrate concentration often follows the Michaelis-Menten equation, *eq* 1.8:

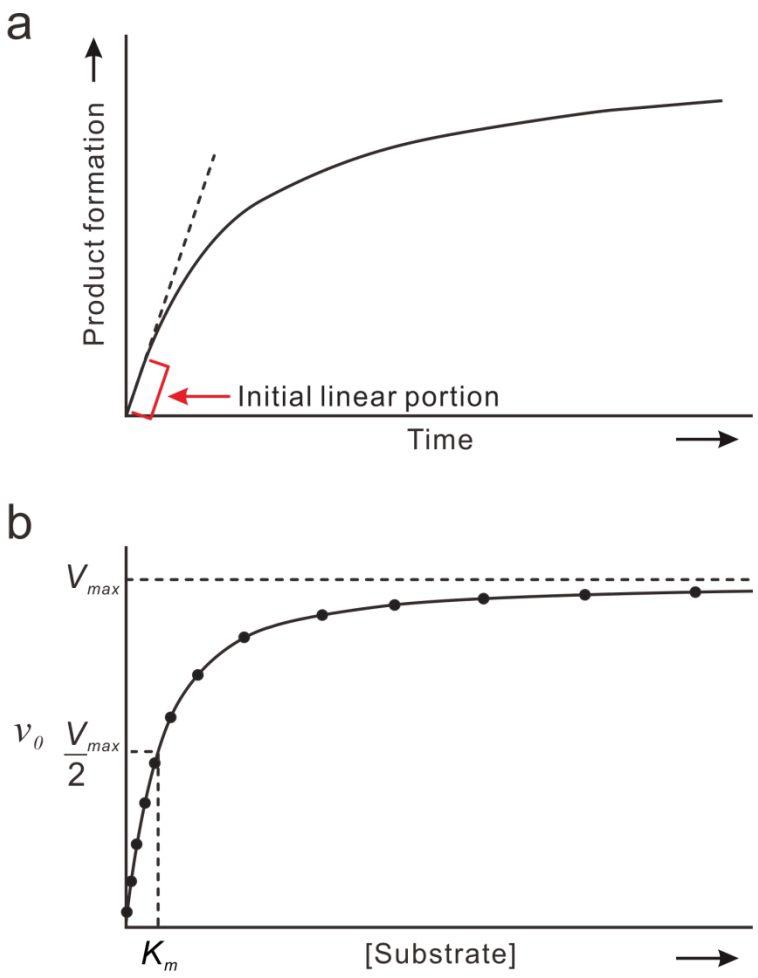

Figure 1.7 Schematic diagrams of (a) a representative progress curve of enzymatic reaction and (b) a representative Michaelis-Menten plot.

$$v_0 = \frac{V_{max}[\mathrm{S}]}{K_m + [\mathrm{S}]} \tag{1.8}$$

where [S] is substrate concentration, V_{max} is the maximum rate at saturating substrate concentration and K_m is Michaelis constant representing the concentration of substrates at which v_0 is one half of the value of V_{max} (Figure 1.7b).

1.5 Peptide Sequencing using Tandem Mass Spectrometry

Proteins and peptides are built up from 20 different amino acids covalently bound together by peptide bonds. Each amino acid has a different side chain, giving their various properties such as acidity, basicity, hydrophilicity or hydrophobicity. The characteristics, biological function and conformation of the proteins and peptides are determined by their amino acids compositions.

Determination of the amino acid sequence of proteins and peptides is essential to study their higher order structures and properties. One of conventional and well-established methodologies to determine the amino acid sequence of peptides is the Edman degradation.[162–165] This method employs a series of chemical reactions. In the first step of the Edman degradation, the polypeptide is reacted with phenylisothiocyanate, which reacts with the N-terminal amino acid group. Acidification then catalyzes intramolecular cyclization and selectively cleaves the N-terminal amino acid as a phenylthiohydantoin. The cleaved amino acid derivative is separated and identified by chromatography or electrophoresis. The remaining polypeptide is isolated, and the Edman degradation is repeated for the identification of next N-terminal amino acid group. For protein sequencing,

proteins are digested chemically or enzymatically to peptides, and the sequence of the resulting peptide fragments are determined by the Edman degradation. In modern approach of the Edman degradation, these procedures are fully automated,[163–165] and an automated sequencer is commercially available. There are, however, several limitations in the Edman degradation. If the N-terminal of polypeptides is chemically blocked such as acetylation, formylation, or pyroglutamyl formation, the Edman reaction does not proceed. These modifications of the N-terminal residue in proteins are frequently observed in biological samples, although the blocking may also occur artificially during the handling of the samples.[165] Moreover, the Edman degradation is sensitive for chemical contaminants containing an amine group, which may react with phenylisothiocyanate.[165] Furthermore, the Edman degradation is difficult to apply to peptide mixtures, which frequently result in erroneous sequencing.[165]

Recently, MS/MS-based techniques has been a powerful methodology for protein and peptide sequencing.[166,167] Compared to the individual analysis of an amino acid residue at N-terminal in the Edman degradation, the measurements of MS/MS are much faster due to the simultaneous detection of peptide fragment ions. In addition, peptide mixture can be analyzed by MS/MS due to the isolation of the peptide, *vide infra*. Moreover, MS/MS-based methodologies for the determination of covalent modification sites in peptides and proteins, such as glycosylation, phosphorylation or sulfation, has also been developed by many research.[168–171]

The MS/MS experiments consist of 1) the isolation of ions, 2) the fragmentation of isolated ions and 3) the analysis of resulting product ions. Although there are a couple of fragmentation techniques, for instance infrared multiphoton dissociation,[172] electron-capture dissociation[173] or, more recently, electron transfer dissociation,[174] the most widely used and established technique for peptide sequencing is CID.[98,166] It was discussed already that CID can result in the dissociation of protein-ligand complexes, but CID is also utilized as a technique for the dissociation of gaseous ions in MS. In CID, the fragmentation of the selectively isolated peptide ions is achieved by energetic collisions with the inert gases, such as nitrogen or argon. The fraction of the translational kinetic energy in the ions is converted into internal vibrational energy by the collisions, bringing the ions into an excited state. Frequently, CID is carried out using hybrid tandem instruments, which are composed of the first MS for isolation of ions, a collision cell where excitation and dissociation of the peptide ions take place by CID and the second MS to analyze and detect resulting product ions. The peptide sequence is determined based on interpretation of peptide fragmentation spectra.

Most frequently, the product ions formed by the cleavage along the peptide backbone are utilized for peptide sequencing. They are assigned based on their cleavage sites (Figure 1.8). Series of a- and x-ions form when dissociation occurs between the α-carbon and the carbonyl group. The fragmentation involving the peptide bonds produces b- and y-ions. The ions generated by the dissociation between the α-carbon and the amino group are labeled as c- and z-ions. Product ions retaining the charge on the N-terminal of the peptide are represented by a-, b-

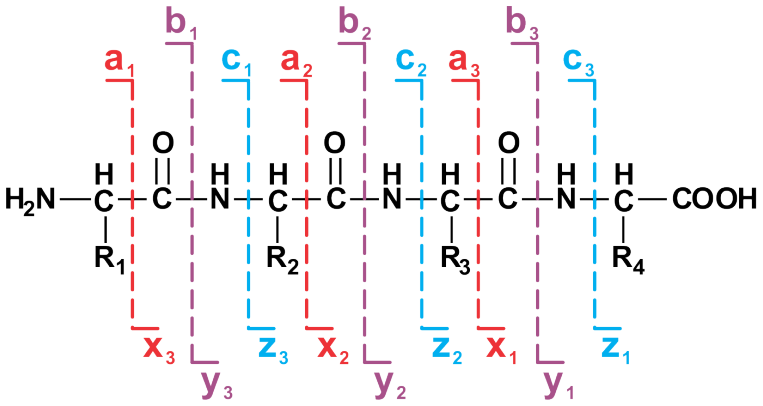

Figure 1.8 Main fragmentation paths of peptides in CID-MS/MS.

or c-ions, whereas ions containing the charge on the C-terminal of the peptide are assigned with x-, y- or z-ions. Although the types of product ions observed in MS/MS experiments depend on various parameters including the amino acid composition, the size of peptide, collision gases used for CID, the charge state of the selected ion and time scale of the instrument, series of a-, b- and y-ions are preferentially produced in CID-MS/MS.[175] Amino acid residues can be determined by analyzing mass difference of product ions in the same type (e.g. mass difference between b_n and b_{n-1} ions).

1.6 Organization of the Thesis

Using the ES-MS assays described above, this thesis presents the characterization of GTA and GTB as well as NEU3. This thesis can be divided into three sections. The first section, chapter 2 and 3, describes the binding properties of GTA and GTB, especially interactions between GTs and their substrates and products. In the second section, chapter 4, the catalytic mechanism of retaining glycosyltransferases is examined using mutant GTA and GTB. The third section, chapter 5, presents the enzyme kinetics of NEU3 for its natural and synthetic substrates to examine the substrate recognition by NEU3.

In Chapter 2, the affinities of GTA and GTB for their common disaccharide acceptor substrate, H-antigen, in the absence and presence of bound UDP and Mn^{2+} were determined using temperature-controlled ES-MS.[176] The presence of bound UDP and Mn^{2+} in the donor binding site had a marked influence on the thermodynamic parameters for association of acceptor to GTA

and GTB. Both the enthalpy and entropy of association (ΔH_a, ΔS_a) decreased significantly. However, the free energy of association (ΔG_a) was unchanged at physiological temperature. The differences in the ΔH_a and ΔS_a values determined in the presence and absence of bound UDP were attributed to structural changes in the GTs induced by the simultaneous binding of H-antigen and UDP.

In Chapter 3, the *direct* ES-MS binding assay was applied to quantify the interaction between GT and nucleotide-sugar donor, donor and acceptor analogs as well as trisaccharide products for the first time. The first comparative thermodynamic study of GTA and GTB provides new insights into these GTs and their interactions with substrate and product. Our study using ES-MS proved that recombinant soluble fragments of GTs exist as homodimers in aqueous solution at physiological pH. Our results also revealed that a divalent metal cofactor, Mn^{2+}, remarkably influences the affinities of GT-donor complexes. In the absence of Mn^{2+}, neither GTA nor GTB exhibited any appreciable affinity for donors, UDP-GalNAc and UDP-Gal. Upon introduction of Mn^{2+}, however, the affinities of both donors to GTs enhanced by 20 ~ 100 times. Moreover, donors undergo enzyme-catalyzed hydrolysis in the presence of GTs and Mn^{2+} resulting in the formation of monosaccharide and UDP. GTA and GTB bind the donor analogues UDP-GlcNAc, UDP-Glc with affinities similar those measured for UDP-Gal and UDP-GalNAc (GTB only), suggesting that the native donors and donor analogues bind to the GTA and GTB through similar interactions. Binding of GTA and GTB with their respective enzymatic products, the trisaccharide A- and B-antigens was also investigated for the first time.

In Chapter 4, direct detection of covalent glycosyl-enzyme intermediates for mutants of GTA and GTB by ES-MS is reported. To trap the intermediates, mutants of GTA or GTB, in which the putative catalytic nucleophile Glu^{303} was replaced with Cys (i.e. GTA_{E303C} and GTB_{E303C}), was utilized. The formation of covalent glycosyl-enzyme intermediates was observed by incubation of GT mutants with their respective donor substrates in the presence Mn^{2+}. Tandem MS analysis of tryptic peptides containing Cys^{303} using CID confirmed Cys^{303} as the site of glycosylation. Incubation of the isolated glycosyl-enzyme intermediates with the disaccharide H-antigen results in the decrease of intermediates and the formation of the corresponding enzymatic trisaccharide products. Our results suggest that the GTA_{E303C} and GTB_{E303C} mutants could operate by a double displacement mechanism.

In Chapter 5, the rate of substrate cleavage by NEU3 was measured using an ES-MS kinetic assay to obtain relative reaction rates (k_{rel}). NEU3 substrate activity was directly dependent upon the hydrophobicity of the aglycone. In addition, the substrates with incorporated azide groups in the Neu5Ac residue at either *C*9 or the *N*5-Ac position were substrates, and in the case of the *N*5-azidoacetyl derivative the activity was superior to GM3. However, the incorporation of larger aryl groups was tolerated only at *C*9, but not at *N*5-Ac. Although our results do not exclude the possibility of the influence of other elements in substrates, a two-site model for enzyme specificity, requiring interaction at both the Neu5Ac residue and the hydrophobic aglycone was proposed.

1.7 Literature Cited

(1) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry 5th Ed.* **2008**, New York: W.H. Freeman.

(2) Somerville, C. *Annu. Rev. Cell Dev. Biol.* **2006**, *22*, 53–78.

(3) Mutwil, M.; Debolt, S.; Persson, S. *Curr. Opin, Plant Biol.* **2008**, *11*, 252–257.

(4) Merzendorfer, H. *J. Comp. Physiol. B* **2006**, *176*, 1–15.

(5) Blom, N.; Sicheritz-Pontén, T.; Gupta, R.; Gammeltoft, S.; Brunak, S. *Proteomics* **2004**, 4, 1633–1649.

(6) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto G. J. Jr. *Angew. Chem. Int. Ed*. **2005**, *44*, 7342–7372.

(7) Roth, J.; Zuber, C; Park, S.; Jang, I.; Lee, Y.; Kysela, K. G.; Fourn, V. L.; Santimaria, R.; Guhl, B.; Cho, J. W. *Mol. Cells*. **2010**, *30*, 497–506.

(8) Varki, A. *Glycoboiology* **1993**, *3*, 97–130.

(9) Simons, K.; Ikonen, E. *Nature* **1997**, *387*, 569–572.

(10) Dennis, J. W.; Granovsky, M.; Warren, C. E. *Bioessays* **1999**, *21*, 412–421.

(11) Rudd, P. M.; Wornald, M. R.; Stanfield, R. L.; Huang, M.; Mattsson, N.; Speir, J. A.; DiGennaro, J. A.; Fetrow, J. S.; Dwek, R. A.; Wilson, I. A. *J. Mol. Biol.* **1999**, *293*, 351–366.

(12) Rudd, P. M.; Elliott, T., Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370−2376.

(13) Lowe, J. B.; Marth, J. D. *Annu. Rev. Biochem*. **2003**, *72*, 643−691.

(14) Szymanski, C. M.; Wren, B. W. *Nat. Rev. Microbiol.* **2005**, *3*, 225−237.

(15) Ohtsubo, K.; Marth, J. D. *Cell* **2006**, *126*, 855−867.

(16) Vigerust, D. J.; Shepherd, V. L. *Trends Microbiol.* **2007**, *15*, 211−218.

(17) van Meer, G.; Voelker, D. R.; Feigenson, G. W. *Nat. Rev. Mol. Cell Biol*. **2008**, *9*, 112−124.

(18) Wennekes, T.; van den Berg, R. J. B. H. N.; Boot, R. G.; van der Marel G. A.; Overkleeft, H. S.; Aerts, J. M. F. G. *Angew. Chem. Int*. *Ed.* **2009**, *48*, 8848−8869.

(19) Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. *Nucleic Acids Rev*. **2009**, *37*, D233−D238.

(20) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem*. **2008**, *77*, 521−555.

(21) Wagner, G. K.; Pesnot, T. *ChemBioChem* **2010**, *11*, 1939−1949.

(22) Palcic, M. M. *Curr. Opin. Chem. Biol*. **2011**, *15*, 226−233.

(23) Ball, S. G.; Morell, M. K. *Annu. Rev. Plant Biol.* **2003**, *54*, 207−233.

(24) Orzechowski, S. *Acta Biochim. Pol*. **2008**, *55*, 435−445.

(25) Walsh, C.; Freel Meyers, C. L.; Losey, H. C. *J. Med. Chem*. **2003**, *46*, 3425−3436.

(26) Křen, V.; Řezanka, T. *FEMS Microbiol*. *Rev.* **2008**, *32*, 858−889.

(27) Griffith, B. R.; Langenhan, J. M.; Thorson, J. S. *Curr. Opin. Biotechnol*. **2005**, *16*, 622–630.

(28) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I. K.; Li, L.; Thorson, J. S. *Sicence* **2006**, *313*, 1291–1294.

(29) Dempski, R. E. Jr; Imperiali, B. *Curr. Opin. Chem. Biol.* **2002**, *6*, 844–850.

(30) Faridmoayer, A.; Fentabil, M. A.; Mills, D. C.; Klassen, J. S.; Feldman, M. F. *J. Bacteriol.* **2007**, *189*, 8088–8098.

(31) Ielmini, M. V.; Feldman, M. F. *Glycobiology* **2011**, *21*, 734–742.

(32) Taniguchi, N.; Honke, K.; Fukuda, M. *Handbook of Glycosyltransferases and Related Genes.* **2002**, Tokyo: Springer.

(33) Davies, G.; Henrissat, B. *Structure* **1995**, *3*, 853–859.

(34) Zechel, D. L.; Withers, S. G. *Acc. Chem. Res.* **2000**, *33*, 11–18.

(35) Bojarová, P.; Křen, V. *Trends Biotechnol*. **2009**, *27*, 199–209.

(36) Marín-Navarro, J.; Polaina, J. *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1267–1273.

(37) Liu, C.; Eichelberger, M. C.; Compans, R. W.; Air, G. M. *J. Virol.* **1995**, *69*, 1099–1106.

(38) Gamblin, S. J.; Skehel, J. J. *J. Biol. Chem*. **2010**, *285*, 28403–28409.

(39) Fischetti, V. *Trends Microbiol*. **2005**, *13*, 491–496.

(40) Helenius, A.; Aebi, M. *Annu. Rev. Biochem.* **2004**, *73*, 1019–1049.

(41) Spiro, R. G. *Cell Mol. Life Sci*. **2004**, *61*, 1025–1041.

(42) Couldrey, C.; Green, J. E. *Breast Cancer Res*. **2000**, *2*, 321–323.

(43) Aebi, M.; Hennet, T. *Trends Cell Biol*. **2001**, *11*, 136–141.

(44) Schachter, H. *Cell Mol. Life Sci*. **2001**, *58*, 1085–1104.

(45) Hakomori, S. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 10231–10233.

(46) Lowe, J. B. *Curr. Opin. Cell Biol.* **2003**, *15*, 531–538.

(47) Dube, D. H.; Bertozzi, C. R. *Nat. Rev*. *Drug Discovery* **2005**, *4*, 477–488.

(48) Cazet, A.; Julien, S.; Bobowski, M.; Burchell, J.; Delannoy, P. *Breast Cancer Res.* **2010**, *12*, 204–216.

(49) Drake, P. M.; Cho, W.; Li, B.; Prakobphol, A.; Johansen, E.; Anderson, N. L.; Regnier, F. E.; Gibson, B. W.; Fisher, S. J. *Clin. Chem*. **2010**, *56*, 223–236.

(50) Lowe, J. B.; Marth, J. D. *Annu. Rev. Biochem*. **2003**, *72*, 643–691.

(51) Freeze, H. H. *Nat. Rev. Genet*. **2006**, *7*, 537–551.

(52) Asano, N. *Glycobiology* **2003**, *13*, 93R–104R.

(53) Osborn, H. M. I.; Evans, P. G.; Gemmell, N.; Osborne, S. D. *J. Pharm. Pharmacol*. **2004**, *56*, 691–702.

(54) Jillian. R. B.; Crawford, B. E.; Esko, J. *Crit. Rev. Biochem. Mol. Biol*. **2007**, *42*, 481–515.

(55) Rempel, B.; Withers, S. G. *Glycobiology* **2008**, *18*, 570–586.

(56) Hinou, H.; Nishimura, S. *Curr. Top. Med. Chem*. **2009**, *9*, 106–116.

(57) Izumi, M.; Yuasa, H.; Hashimoto, H. *Curr. Top. Med. Chem*. **2009**, *9*, 87–105.

(58) Hosoguchi, K.; Maeda, T.; Furukawa, J.; Shinohara, Y.; Hinou, H.; Sekiguchi, M.; Togame, H.; Takemoto, H.; Kondo, H.; Nishimura S. *J. Med. Chem*. **2010**, *53*, 5607−5619.

(59) Reece, P. A. *J. Med. Virol.* **2007**, *79*, 1577−1586.

(60) Kajimoto, T.; Nobe, M. *Curr. Top. Med. Chem.* **2009**, *9*, 13−33.

(61) Ogawa, S.; Kanto, M. *Curr. Top. Med. Chem.* **2009**, *9*, 58−75.

(62) Jackson, R. J.; Cooper, K. L.; Tappenden, P.; Rees, A.; Simpson, E. L.; Read, R. C.; Nicholson, K. G. *J. Infect.* **2011**, *62*, 14−25.

(63) Murata, T.; Usui, T. *Trends. Glycosci. Glycotechnol.* **2000**, *12*, 161−174.

(64) Ajisaka, K.; Yamatomo, Y. *Trends. Glycosci. Glycotechnol.* **2002**, *14*, 1−11.

(65) Jakeman, D. L.; Withers, S. G. *Trends. Glycosci. Glycotechnol.* **2002**, *14*, 13−25.

(66) Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E. *AAPS J.* **2006**, *8*, E190−E195.

(67) Williams, G. J.; Zhang, C.; Thorson, J. S. *Nat. Chem. Biol.* **2007**, *3*, 657−662.

(68) Gantt, R. W.; Goff, R. D.; Williams, G. J.; Thorson, J. S. *Angew. Chem. Int. Ed.* **2008**, *47*, 8889−8892.

(69) Boltje, T. J.; Buskas, T.; Boons, G. *Nat. Chem.* **2009**, *1*, 611−622.

(70) Homann, A.; Seibel, J. *Nat. Prod. Rep.* **2009**, *26*, 1555−1571.

(71) Yamamoto, F.; Clausen, H.; White, T.; Marken, J.; Hakomori S. *Nature* **1990**, *345*, 229−233.

(72) Palcic, M. M.; Seto, N. O. L.; Hindsgaul, O. *Transf. Med.* **2001**, *11*, 315–323.

(73) Patenaude, S. I.; Seto, N. O. L.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol.* **2002**, *9*, 685–690.

(74) Nguyen, H. P.; Seto, N. O. L.; Cai, Y.; Leinala, E. K.; Borisova, S. N.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem.* **2003**, *278*, 49191–49195.

(75) Lee, H. J.; Barry, C. H.; Borisova, S. N.; Seto, N. O. L.; Zheng, R. X. B.; Blancher, A.; Evans, S. V.; Palcic, M. M. *J. Biol. Chem.* **2005**, *280*, 525–529.

(76) Letts, J. A.; Rose, N. L.; Fang, Y. R.; Barry, C. H.; Borisova, S. N.; Seto, N. O. L.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem.* **2006**, *281*, 3625–3632.

(77) Persson, M.; Letts, J. A.; Hosseini-Maaf, B.; Borisova, S. N.; Palcic, M. M.; Evans, S. V.; Olsson, M. L. *J. Biol. Chem.* **2007**, *282*, 9564–9570.

(78) Alfaro, J. A.; Zheng, R. B.; Persson, M.; Letts, J. A.; Polakowski, R.; Bai, Y.; Borisova, S. N.; Seto, N. O. L.; Lowary, T.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem.* **2008**, *283*, 10097–10108.

(79) Pesnot, T.; Jørgensen, R.; Palcic, M. M.; Wagner, G. K. *Nat. Chem. Biol.* **2010**, *6*, 321–323.

(80) Schuman, B.; Persson, M.; Landry, R. C.; Polakowski, R.; Weadge, J. T.; Seto, N. O. L.; Borisova, S. N.; Palcic, M. M.; Evans, S. V. *J. Mol. Biol.* **2010**, *402*, 399–411.

(81) Angulo, J.; Langpap, B.; Blume, A.; Beit, T.; Meyer, B.; Krishna, N. R.; Peters, H.; Palcic, M. M.; Peters, T. *J. Am. Chem. Soc.* **2006**, *128*, 13529−13538.

(82) Blume, A.; Angulo, J.; Beit, T.; Peters, H.; Benie, A. J.; Palcic, M. M.; Peters, T. *J. Biol. Chem.* **2006**, *281*, 32728−32740.

(83) Sindhuwinata, N.; Munoz, E.; Munoz, F. J.; Palcic, M. M.; Peters, H.; Peters, T. *Glycobiology* **2010**, *20*, 718−723.

(84) Rademacher, C.; Landström, J.; Sindhuwinata, N.; Palcic, M. M.; Widmalm, G.; Peters, T. *Glycoconj. J.* **2010**, *27*, 349−358.

(85) Seto, N. O. L.; Palcic, M. M.; Compston, C. A.; Li, H.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem.* **1997**, *272*, 14133−14138.

(86) Seto, N. O. L.; Compston, C. A.; Szpacenko, A.; Palcic, M. M. *Carbohydr. Res.* **2000**, *324*, 161−169.

(87) Marcus, S. L.; Polakowski, R.; Seto, N. O. L.; Leinala, E.; Borisova, S.; Blancher, A.; Roubinet, F.; Evans, S. V.; Palcic, M. M. *J. Biol. Chem.* **2003**, *278*, 12403−12405.

(88) Lairson, L. L.; Chiu, C. P. C.; Ly, H. D.; He, S.; Wakarchuk, W. W.; Strynadka, N. C. J.; Withers, S. G. *J. Biol. Chem.* **2004**, *279*, 28339−28344.

(89) Monegal, A.; Planas, A. *J. Am. Chem. Soc.* **2006**, *128*, 16030−16031.

(90) Errey, J. C.; Lee, S. S.; Gibson, R. P.; Fleites, C. M.; Barry, C. S.; Jung, P. M. J.; O'Sullivan, A. C.; Davis, B. G.; Davies, G. J. *Angew. Chem. Int. Ed.* **2010**, *49*, 1234−1237.

(91) Miyagi, T. *Trends Glycosci. Glycotechnol.* **2010**, *22*, 162–172.

(92) Miyagi, T.; Wada, T.; Iwamatsu, A.; Hata, K.; Yoshikawa, Y.; Tokuyama, S.; Sawada, M. *J. Biol. Chem.* **1999**, *274*, 5004–5011.

(93) Ha, K. T.; Lee, Y. C.; Cho, S. H.; Kim, J. K.; Kim, C. H. *Mol. Cells* **2004**, *17*, 267–273.

(94) Kakugawa, Y.; Wada, T.; Yamaguchi, K.; Yamanami, H.; Ouchi, K.; Sato, I.; Miyagi, T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10718–10723.

(95) Kato, K.; Shiga, K.; Yamaguchi, K.; Hata, K.; Kobayashi, T.; Miyazaki, K.; Saijo, S.; Miyagi, T. *Biolchem. J.* **2006**, *394*, 647–656.

(96) Shiozaki, K.; Yamaguchi, K.; Sato, I.; Miyagi, T. *Cancer. Sci.* **2009**, *100*, 588–594.

(97) Miyagi, T.; Wada, T.; Yamaguchi, K.; Hata, K.; Shiozaki, K. *J. Biochem.* **2008**, *144*, 279–285.

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

(99) Zenobi, R.; Knochenmuss, R. *Mass Spectrom. Rev.* **1997**, *17*, 337–366.

(100) Karas, M.; Kruger, R. *Chem. Rev.* **2003**, *103*, 427–439.

(101) Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A–986A.

(102) Kebarle, P. *J. Mass Spectrom.* **2000**, *35*, 804–817.

(103) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2001**, *20*, 362–387.

(104) Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1–23.

(105) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287–2294.

(106) Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, *71*, 4451–4463.

(107) de la Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93−104.

(108) Dole, M.; Mack, L. L.; Hines, R. L. *J. Chem. Phys*. **1968**, *49*, 2240−2249.

(109) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1−8.

(110) Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Fotsis, T.; Mann, M. *Nature* **1996**, *379*, 466–469.

(111) Karas, M.; Bahr, U.; Dülcks, T. *Fresenius J. Anal. Chem.* **2000**, *366*, 669−676.

(112) Jecklin, M. C.; Touboul, D.; Bovet, C.; Wortmann, A.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 332−343.

(113) Juraschek, R.; Dülcks, T.; Karas, M. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300−308.

(114) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass. Spectrom. Rev.* **1998**, *17*, 1−35.

(115) Barrow, M. P.; Burkitt, W. I.; Derrick, P. J. *Analyst* **2005**, *130*, 18−28.

(116) Marshall, A. G.; Hendrickson, C. L. *Annu. Rev. Anal. Chem.* **2008**, *1*, 579−599.

(117) Yee, A.; Pardee, K.; Christendat, D.; Savchenko, A.; Edwards, A. M.; Arrowsmith, C. H. *Acc. Chem. Res.* **2003**, *36*, 183−189.

(118) Staunton, D.; Owen, J.; Campbell, I. D. *Acc. Chem. Res.* **2003**, *36*, 207−214.

(119) Bonvin, A. M. J. J.; Boelens, R.; Kaptein, R. *Curr. Opin. Chem. Biol.* **2005**, *9*, 501−508.

(120) Angulo, J.; Rademacher, C.; Biet, T.; Benie, A. J.; Blume, A.; Peters, H.; Palcic, M.; Parra, F.; Peters, T. *Methods Enzymol.* **2006**, *416*, 12–30.

(121) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662–666.

(122) Berman,, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(123) Wishart, D. *Curr. Pharm. Biotechnol.* **2005**, *6*, 105–120.

(124) Schuck, P. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 541–566.

(125) Philips, K. S.; Cheng, Q. *Anal. Bioanal. Chem.* **2007**, *387*, 1831–1840.

(126) Bundle, D. R.; Sigurskjold, B. W. *Methods Enzymol.* **1994**, *247*, 288–305.

(127) Ababou, A.; Ladbury, J. E. *J. Mol. Recognit.* **2006**, *19*, 79–89.

(128) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175–186.

(129) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1–27.

(130) Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev.* **2004**, *23*, 368–389.

(131) Wang, W.; Kitova, E. N.; Klassen, J. S. *Methods Enzymol.* **2003**, *362*, 376–397.

(132) Rademacher, C.; Shoemaker, G. K.; Kim, H.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T.; Lowary, T. L. *J. Am. Chem. Soc.* **2007**, *129*, 10489–10502.

(133) Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587–92.

(134) Dingle, T.; Wee, S.; Mulvey, G. L.; Greco, A.; Kitova, E. N.; Sun, J.; Lin, S.; Klassen, J. S.; Palcic, M. M.; Ng, K. K.; Armstrong, G. D. *Glycobiology* **2008**, *18*, 698–706.

(135) Hyland, R. M.; Sun, J.; Griener, T. P.; Mulvey, G. L.; Klassen, J. S.; Donnenberg, M. S.; Armstrong, G. D. *Cell. Microbiol.* **2008**, *10*, 177–187.

(136) Soya, N.; Shoemaker, G. K.; Monica, M. M.; Klassen, J. S. *Glycobiology* **2009**, *19*, 1224–1234.

(137) Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2010**, *82*, 2170–2174.

(138) Liu, L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 310–318.

(139) El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1893–1899.

(140) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; Pauw, E. D. *J. Mass Spectrom.* **2003**, *38*, 491–501.

(141) Gabelica, V.; Rosu, F.; Pauw, E. D. *Anal. Chem.* **2009**, *81*, 6708–6715.

(142) Van Berkel, G. J.; Asano, K. G.; Schnier, P. D. *J. Am. Soc. Mass Spectrom.* **2000**, *12*, 853–862.

(143) Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 416–425.

(144) Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801–7806.

(145) Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2005**, *77*, 3060–3071.

(146) Yin, S.; Xie, Y.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1199–1208.

(147) Pramanik, B. N.; Bartner, P. L.; Mirza, U. A.; Liu, Y. H.; Ganguly, A. K. *J. Mass Spectrom.* **1998**, *33*, 911–920.

(148) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010–3018.

(149) Eisenthal, R.; Danson, M. J. *Enzyme Assays: Practical Approach 2nd Ed.* **2002**, New York: Oxford University Press.

(150) Harris, T. K.; Keshwani, M. M. *Methods Enzymol.* **2009**, *463*, 57–71.

(151) Wallenfels, K. *Methods Enzymol.* **1962**, *5*, 212–219.

(152) Zechel, D. L.; Konermann, L.; Withers, S. G.; Douglas, D. J. *Biochemistry* **1998**, *37*, 7664–7669.

(153) Bothner, B.; Chavez, R.; Wei, J.; Strupp, C.; Phung, Q.; Schneemann, A.; Siuzdak, G. *J. Biol. Chem.* **2000**, *275*, 13455–13459.

(154) Ge, X.; Sirich, T. L.; Beyer, M. K.; Desaire, H.; Leary, J. A. *Anal. Chem.* **2001**, *73*, 5078–5082.

(155) Gao, H.; Petzold, C. J.; Leavell, M. D.; Leary, J. A. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 916–924.

(156) Gao, H.; Leary, J. A. *Anal. Biochem.* **2004**, *329*, 269–275.

(157) Pi, N.; Hoang, M. B.; Gao, H.; Mougous, J. D.; Bertozzi, C. R.; Leary, J. A. *Anal. Biochem.* **2005**, *341*, 94–104.

(158) Danan, L. M.; Yu, Z. H.; Ludden, P. J.; Jia, W. T.; Moore, K. L.; Leary, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1633–1642.

(159) Hsieh, F. Y. L.; Tong, X.; Wachs, T.; Ganem, B.; Henion, J. *Anal. Biochem.* **1995**, *229*, 20–25.

(160) Pi, N.; Meyers, C. L. F.; Pacholec, M.; Walsh, C. T.; Leary, J. A. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 10036–10041.

(161) Pi, N.; Yu, Y.; Mougous, J. D.; Leary, J. A. *Protein Sci.* **2004**, *13*, 903–912.

(162) Edman, P. *Acta Chem. Scand.* **1950**, 283–293.

(163) Hewick, R. M.; Hunkapillar M. W.; Hood, L. E.; Dreyer, W. J. *J. Biol. Chem.* **1981**, *256*, 7990–7997.

(164) Gooley, A. A.; Ou, K.; Russell, J.; Wilkins, M. R.; Sanchez, J. C.; Hochstrasser, D. F.; Williams, K. L. *Electrophoresis* **1997**, *18*, 1068–1072.

(165) Smith, J. B. *Peptide Sequencing by Edman Degradation. Encyclopedia of Life Sciences*, **2001**, New York: John Wiley & Sons.

(166) Steen, H.; Mann, M. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 699–711.

(167) Wysocki, V. H.; Resing, K. A.; Zhang, Q.; Cheng, G. *Methods* **2005**, *35*, 211–222.

(168) Garcia, B. A.; Shabanowitz, J.; Hunt, D. F. *Methods* **2005**, *35*, 256–264.

(169) Kaji, H.; Yamauchi, Y.; Takahashi, N.; Isobe, T. *Nat. Protoc.* **2006**, *1*, 3019–3027.

(170) Yu, Y.; Hoffhines, A. J.; Moore, K. L.; Leary, J. A. *Nat. Methods* **2007**, *4*, 583–588.

(171) Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G. *Nat. Methods* **2007**, *4*, 798–806.

(172) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809–2815.

(173) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265–3266.

(174) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. USA*. **2004**, *101*, 9528–9533.

(175) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2005**, *24*, 508–548.

(176) Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc*. **2004**, *126*, 4786–4787.

Chapter 2

Temperature-dependent Cooperativity in Donor-acceptor Substrate Binding to the Human Blood Group Glycosyltransferases[†]

2.1 Introduction

Glycosyltransferases represent an enormous family of enzymes that catalyze the synthesis of glycosidic linkages by the stereo- and regiospecific transfer of saccharides from activated sugar-nucleotide donors to specific acceptors.[1,2] We have identified the highly homologous blood group A and B synthesizing enzymes as models for the study of retaining glycosyltransferase reactions. Remarkably, despite differing in only 4 out of 354 amino acids, these enzymes possess altered donor specificity. Human α-(1→3)-N-acetylgalactosaminyltransferase (GTA) and α-(1→3)-galactosyltransferase (GTB), catalyze the transfer of GalNAc or Gal from the sugar-nucleotide donors, uridine 5′-diphosphate-GalNAc (UDP-GalNAc) and uridine 5′-diphosphate-Gal (UDP-Gal), respectively, to α-L-Fuc*p*-(1→2)-β-D-Gal*p*-OR acceptors (where R is a glycoprotein or glycolipid, Figure 1.1). Although the catalytic mechanisms of GTA and GTB have been probed,[3,4] there remain a number of outstanding

[†] A version of this chapter has been published: Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587–592.

questions. Notably, an ordered catalytic mechanism (where donor binding precedes acceptor binding), which may imply cooperative substrate binding, has been suggested based on kinetic and NMR data.[5,6] Direct evidence for or against cooperative binding, which may provide further insight into the precise catalytic mechanism of these enzymes, is currently lacking. Establishing whether donor and acceptor binding to GTA and GTB operate in a cooperative fashion represents a significant experimental challenge. The substrate affinities are relatively low ($<10^5$ M^{-1}), *vide infra*, which leads to a distribution of protein species in solution at typical concentrations. The situation is further complicated by the presence of a divalent metal ion cofactor, thought to be manganese (Mn^{2+}), which is also required for optimal activity of these enzymes. As a result, direct insight into binding stoichiometry is necessary in order to accurately quantify the substrate affinities.

There are a number of established analytical techniques used to quantify protein-carbohydrate interactions *in vitro*, each with particular strengths and weaknesses. Isothermal titration microcalorimetry (ITC) is probably the most widely used method and is the only technique that provides a direct measure of the enthalpy of association (ΔH_a).[7] However, the ITC method requires large (mg) and often prohibitive quantities of both protein and ligand for each analysis. Other commonly used methods include surface plasmon resonance (SPR), which affords high sensitivity and can be used to evaluate the on-off rate constants,[8] and frontal affinity chromatography-mass spectrometry (FAC-MS), which is capable of rapidly screening target proteins against libraries of carbohydrate ligands.[9,10]

However, these methods require the immobilization of either the protein (FAC-MS) or the ligand (SPR). Furthermore, none of the above techniques provide direct information regarding binding stoichiometry and are, therefore, of limited use for studying multi-substrate enzyme systems.

Mass spectrometry (MS), combined with electrospray ionization (ES), has emerged as a powerful tool for quantifying the binding stoichiometry and affinity for protein-ligand complexes in solution.[11–15] The strengths of the ES-MS technique are simplicity (no labeling or immobilization required), speed (measurements can usually be completed within a few seconds), and specificity (the unique ability to provide direct insight into stoichiometry and to study multiple binding equilibria simultaneously). When performed using nano-flow ES (nanoES), the ES-MS assay affords high sensitivity, normally consuming picomoles or less of protein and ligand per analysis. Additionally, enthalpies (ΔH_a) and entropies (ΔS_a) of association can be estimated from the temperature dependence of the K_a values determined using temperature-controlled ES-MS.[16]

Here, we describe the results of the first thermodynamic study, performed using the *direct* ES-MS binding assay, of the synthetic H-disaccharide acceptor, α-L-Fuc*p*-(1→2)-β-D-Gal*p*-O$(CH_2)_7CH_3$ (**1**), binding to GTA and GTB in the absence and presence of UDP (**2**) and a metal ion cofactor, Mn^{2+}. Affinities for **1** were measured over a range of solution temperatures and the corresponding thermodynamic parameters (ΔG_a, ΔH_a and ΔS_a) were established. Importantly, it is shown that occupancy of the donor site dramatically decreases the ΔH_a and ΔS_a

parameters for acceptor binding. However, at physiological temperature, ΔG_a is independent of donor site occupancy.

2.2 Materials and Methods

2.2.1 Sample Preparation

The ES-MS measurements were carried out using recombinant, soluble fragments (amino acids 63–354) of the human blood group glycosyltransferases, GTA (monomer MW 34,483 Da) and GTB (monomer MW 34,519 Da), consisting of the full C-terminal and catalytic domains as well as a truncated N-terminal domain.[4] The enzymes were expressed in *Escherichia coli* and purified using procedures described previously.[4,17] The recombinant single chain variable fragment (scFv) of the monoclonal antibody Se155-4 (MW 26,539 Da) was also produced using established protocols.[18] ScFv was used as reference protein in the binding measurements to account for possible non-specific contributions to the mass spectra from the non-specific attachment of **1** to the GTs during the ES process. All protein solutions were exchanged into an aqueous 50 mM ammonium acetate (pH 7) solution except for scFv, which was exchanged directly into Milli-Q water, using an Amicon microconcentrator with a molecular weight cutoff (MWCO) of 10 kDa. The GTB, GTA, and scFv concentrations were determined by lyophilizing a known volume of the protein solution and measuring the corresponding mass of the protein. The synthetic ligand, α-L-Fuc*p*-(1→2)-β-D-Gal*p*-O$(CH_2)_7CH_3$ (**1**),[19] was obtained from Dr. Ole Hindsgaul (Carlsberg Laboratory), and UDP (**2**) was purchased from Sigma-Aldrich, Canada. Ligand

stock solutions were prepared by dissolving the ligand into Milli-Q water; these were stored at –20 °C. The ES solutions were prepared from the stock solutions and used for MS analysis immediately after preparation.

2.2.2 Mass Spectrometry

All experimental measurements were performed using an Apex II 9.4 Tesla FTICR-MS instrument (Bruker, Billerica, MA) equipped with a temperature-controlled nanoflow electrospray (nanoES) ion source.[16] NanoES tips, with an outer diameter of ~5 μm, were pulled from borosilicate tubes (1 mm o.d., 0.68 mm i.d.) using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). A platinum wire inserted into one end of the nanoES tip was used to establish electrical contact with the nanoES solution. A potential of +600 to +800 V was applied to the platinum wire in the nanoES tip in order to spray the solution. The tip was positioned 1–2 mm from a stainless steel sampling capillary using a microelectrode holder. The typical solution flow rate was ~20 nL min^{-1}. Charged droplets and solvated ions emitted by the nanoES tip were introduced into the vacuum chamber of the mass spectrometer through a heated stainless steel sampling capillary (0.43 mm i.d.) maintained at an external temperature of 66 °C. The gaseous ions sampled by the capillary (+48 V) were transmitted through a skimmer (+4 V) and accumulated for 5 s in an rf hexapole (+600 V p-p). The ions were subsequently ejected from the hexapole and injected at –2700 V into the bore of the superconducting magnet, decelerated and introduced into the ion cell. The typical base pressure for the instrument was ~5 × 10^{-10} mbar. Data

acquisition was performed using the XMASS software (version 5.0). The time-domain signals, consisting of the sum of 50 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier-transformation.

2.2.3 Determination of K_a Values using the *direct* ES-MS Binding Assay

The GTA_2 and GTB_2 homodimers (≡ GT_2) possess two acceptor (**1**) binding sites. The relevant reactions and equilibrium expressions are given below:

$$GT_2 + \mathbf{1} \rightleftharpoons (GT_2 + \mathbf{1}) \qquad K_{a,1} = \frac{[GT_2 + \mathbf{1}]}{[GT_2][\mathbf{1}]} \tag{2.1a}$$

$$(GT_2 + \mathbf{1}) + \mathbf{1} \rightleftharpoons (GT_2 + 2(\mathbf{1})) \qquad K_{a,2} = \frac{[GT_2 + 2(\mathbf{1})]}{[GT_2 + \mathbf{1}][\mathbf{1}]} \tag{2.1b}$$

The equilibrium concentrations $[GT_2]$, $[GT_2 + \mathbf{1}]$ and $[GT_2 + 2(\mathbf{1})]$ can be determined from the relative abundance of the corresponding ions observed in the mass spectrum and the equations of mass balance given below:

$$[GT_2]_o = [GT_2] + [GT_2 + \mathbf{1}] + [GT_2 + 2(\mathbf{1})] \tag{2.2a}$$

$$[\mathbf{1}]_o = [\mathbf{1}] + [GT_2 + \mathbf{1}] + 2[GT_2 + 2(\mathbf{1})] \tag{2.2b}$$

After substituting the equilibrium concentrations with the corresponding ion intensity ratios, the K_a values can be calculated from *eq* 2.3 and 2.4:

$$K_{a,1} = \frac{R_1}{[\mathbf{1}]_o - \frac{(R_1 + 2R_2)[GT_2]_o}{1 + R_1 + R_2}} \tag{2.3a}$$

$$K_{a,2} = \frac{R_2}{R_1([\mathbf{1}]_o - \frac{(R_1 + 2R_2)[GT_2]_o}{1 + R_1 + R_2})} \tag{2.3b}$$

where the ratios R_1 (= [GT_2 + **1**]/[GT_2]) and R_2 (= [GT_2 + 2(**1**)]/[GT_2]) are calculated using *eq* 2.4:

$$R_1 = \frac{\sum_{n} Ab(\mathrm{GT_2} + \mathrm{L})^{n+}}{\sum_{n} Ab(\mathrm{GT_2})^{n+}} \tag{2.4a}$$

$$R_2 = \frac{\sum_{n} Ab(\mathrm{GT_2} + 2(\mathrm{L}))^{n+}}{\sum_{n} Ab(\mathrm{GT_2})^{n+}} \tag{2.4b}$$

The values of $K_{a,1}$ and $K_{a,2}$ measured for GTA and GTB binding to **1** will differ due to the presence of two, equivalent acceptor binding sites.[20] The values are related by statistical factors related to the number of equivalent free and occupied binding sites. The intrinsic K_a (the affinity of each binding site) can be determined for each and any of the ligand binding reactions:

$$K_a = \frac{K_{a,1}}{2} \tag{2.5a}$$

$$K_a = 2 \times K_{a,2} \tag{2.5b}$$

Any non-specific binding between GT_2 and **1** during the ES process will influence the intensities measured for the GT_2^{n+}, $(GT_2 + \mathbf{1})^{n+}$, and $(GT_2 + 2(\mathbf{1}))^{n+}$ ions.[15] Recently, a method to quantitatively account for the contribution of nonspecific complexes to the nanoES mass spectra was developed in our laboratory.[21] Briefly, this method involves the addition of a reference protein (P_{ref}), which exhibits no specific affinity for the ligands of interest, to the ES solution. The distribution of ligands bound nonspecifically to P_{ref} can be used to correct the measured intensities of the P^{n+} and specific PL^{n+} ions for the occurrence of

nonspecific ligand binding. A complete description of the correction process is given elsewhere.[21]

The enthalpies and entropies of association (ΔH_a and ΔS_a) for GTA and GTB binding to **1** were estimated from the temperature dependence of the intrinsic K_a, which was modeled using the linear form of the van't Hoff equation:

$$\ln K_a = -\frac{\Delta H_a}{R_g T} + \frac{\Delta S_a}{R_g} \tag{2.6}$$

where R_g is the ideal gas constant and T is the temperature.

2.3 Results and Discussion

2.3.1 Effect of UDP Binding on the Acceptor Affinity

The affinities of GTA and GTB for **1** were measured over a range of solution temperatures (10–40°C). Shown in Figures 2.1a and 2.2a are illustrative ES mass spectra measured for solutions of GTA (7 μM) with **1** (30 μM), and GTB (8 μM) with **1** (40 μM), respectively, at 24 °C and pH 7. The metal cofactor, Mn^{2+}, was also present in excess (100 μM). The recombinant forms of GTA and GTB used in this study exist exclusively as non-covalently bound homodimers (i.e., GTA_2, GTB_2) in aqueous solutions at neutral pH.[22] According to the ES-MS data, at the concentrations investigated, GTA_2 and GTB_2 each bind to 0–2 molecules of **1**. Importantly there is no evidence for the attachment of Mn^{2+} to either enzyme, which suggests the GTs possess very low affinities for Mn^{2+}, a result that is consistent with related retaining glycosyltransferases.[23] The distributions of bound **1**, as determined from the mass spectra, reveal that each homodimer possesses

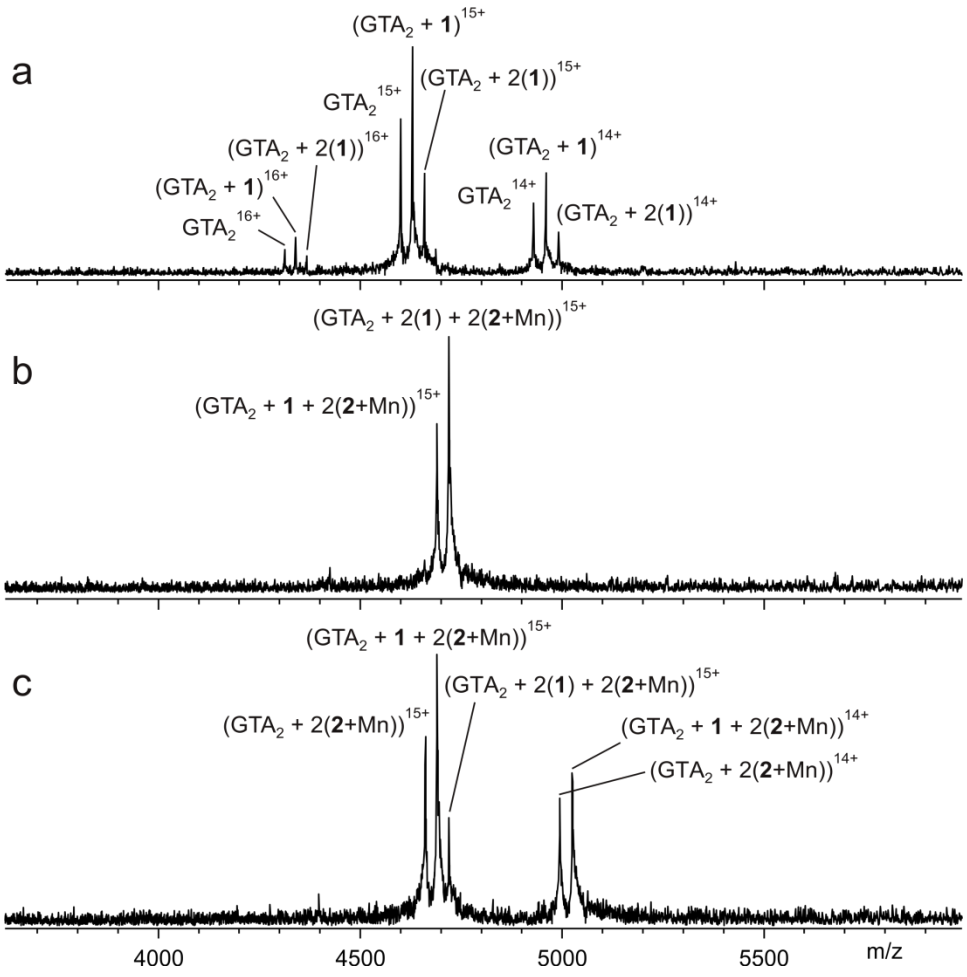

Figure 2.1 Illustrative nanoES mass spectra of solutions (pH 7, 24 °C) containing 7 μM GTA_2 and 100 μM Mn^{2+}. (a) 30 μM **1**, (b) 30 μM **1** and 50 μM **2**, or (c) 10 μM **1** and 50 μM **2**.

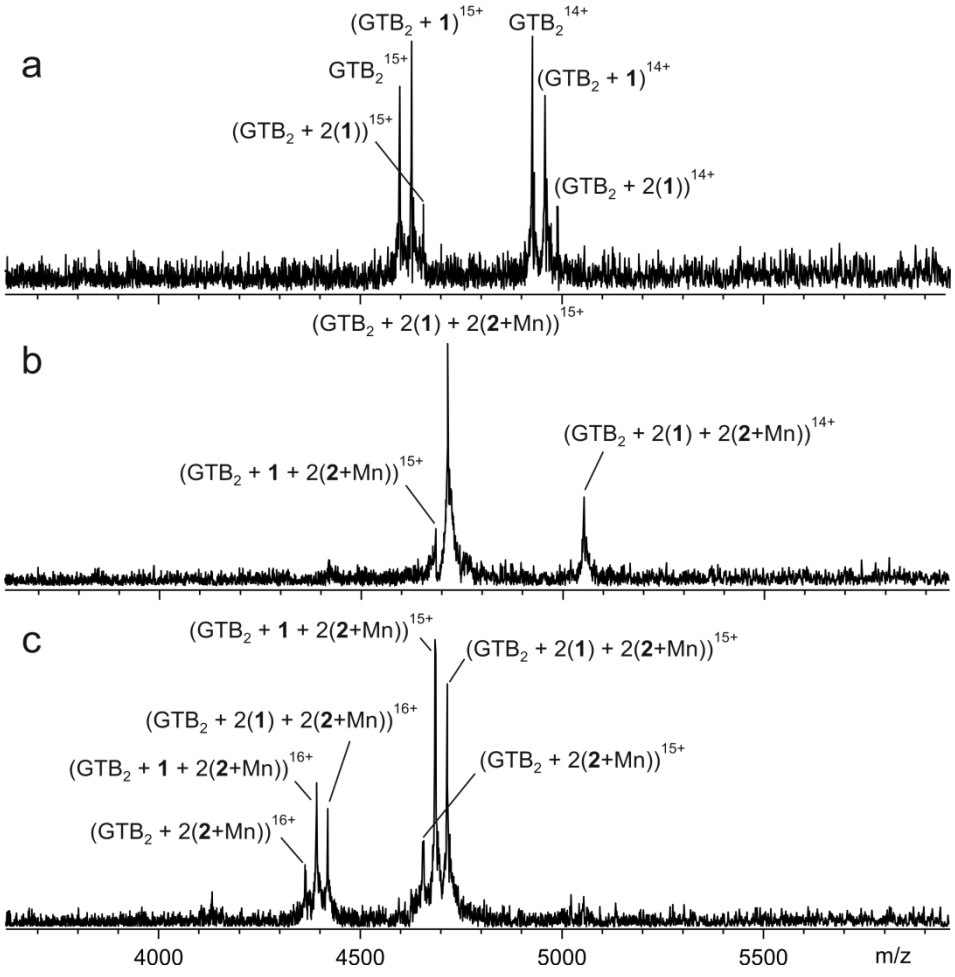

Figure 2.2 Illustrative nanoES mass spectra of solutions (pH 7, 24 °C) containing 8 μM GTB_2 and 100 μM Mn^{2+}. (a) 40 μM **1**, (b) 40 μM **1** and 50 μM **2**, or (c) 20 μM **1** and 50 μM **2**.

two equivalent binding sites for **1**, with intrinsic affinities (K_a) of $(3.2 \pm 0.3) \times 10^4$ M^{-1} (GTA) and $(1.7 \pm 0.3) \times 10^4$ M^{-1} (GTB) at 24 °C.

To establish whether the presence of bound donor influences the affinities of the GTs for **1**, binding measurements were performed in the presence of **2**. The native donors could not be used because of the rapid changes in donor and acceptor substrate concentrations resulting from enzymatic reaction.[22] Shown in Figures 2.1b and 2.2b are illustrative ES mass spectra measured for solutions of GTA (7 μM) with **1** (30 μM), and GTB (8 μM) with **1** (40 μM), respectively, with the addition of **2** (50 μM) and Mn^{2+} (100 μM) at 24 °C and pH 7. Importantly, the distributions of bound **1** change markedly upon addition of **2** to the solutions. At the concentrations investigated, the active site of each enzyme is fully occupied with both **2** and Mn^{2+}. Under these conditions, the affinities of GTA and GTB for **1** are significantly enhanced, and the (GTA_2 + 2(**1**) + 2(**2** + Mn)) and (GTB_2 + 2(**1**) + 2(**2** + Mn)) species dominate in solution. Reduction in the concentration of **1** allowed the K_a values to be determined for the sequential addition of **1** to (GTA_2 + 2(**2** + Mn) and (GTB_2 + 2(**2** + Mn), (Figures 2.1c and 2.2c). The presence of **2** and Mn^{2+} in the active site results in a 5-fold increase in the intrinsic affinity of **1** for GTA ($(1.6 \pm 0.3) \times 10^5$ M^{-1}) and a more substantial 9-fold increase for GTB ($(1.5 \pm 0.3) \times 10^5$ M^{-1}).

An increase in acceptor affinity in the presence of donor has been previously observed in ITC studies of a closely related bovine α-(1→3)-galactosyltransferase binding to its N-acetyllactosamine acceptor.[24] In this case, it was shown that acceptor binding was undetectable in the absence of native donor.

It should be noted, however, that due to limitations of the ITC technique for studying weak interactions ($K_a \leq 10^4$ M^{-1}), it was not possible to establish the binding stoichiometry from these experiments.

2.3.2 Temperature-dependence of Donor-Acceptor Binding Cooperativity

Affinity measurements were performed at varying temperatures and the intrinsic K_a values are reported in the form of van't Hoff plots (Figure 2.3). Importantly, the van't Hoff plots (which exhibit excellent linearity, $r^2 \approx 0.99$) determined in the absence and presence of **2** are markedly different for both GTs. In the absence of **2**, the ΔH_a and ΔS_a values for acceptor binding are: –8.5 ± 0.3 kcal mol^{-1} and –8.0 ± 0.9 cal mol^{-1} K^{-1}, respectively, for GTA; –5.8 ± 0.1 kcal mol^{-1} and –0.3 ± 0.4 cal mol^{-1} K^{-1}, respectively, for GTB (Table 2.1). At higher temperatures, the donor sites are no longer fully occupied. For example, at 37 °C, distributions of bound **2** (and Mn^{2+}) and **1** are observed (Figure 2.4). The determination of K_a is not possible under these conditions because the ions corresponding to the attachment of **2** to the GTs are not fully resolved from those corresponding to the attachment of **1**. Consequently, the K_a values shown in Figure 2.4 are restricted to temperatures where full donor site occupancy was achieved. In the presence of bound **2**, the corresponding ΔH_a and ΔS_a values are: –37 ± 6 kcal mol^{-1} and –99 ± 2 cal mol^{-1} K^{-1}, respectively, for GTA; –36 ± 2 kcal mol^{-1} and –96 ± 5 cal mol^{-1} K^{-1}, respectively, for GTB. The similarities in the ΔH_a and ΔS_a values for the GTs suggest that the nature of the binding interactions with **1** is identical for GTA and GTB when **2** is bound. Importantly, for each GT,

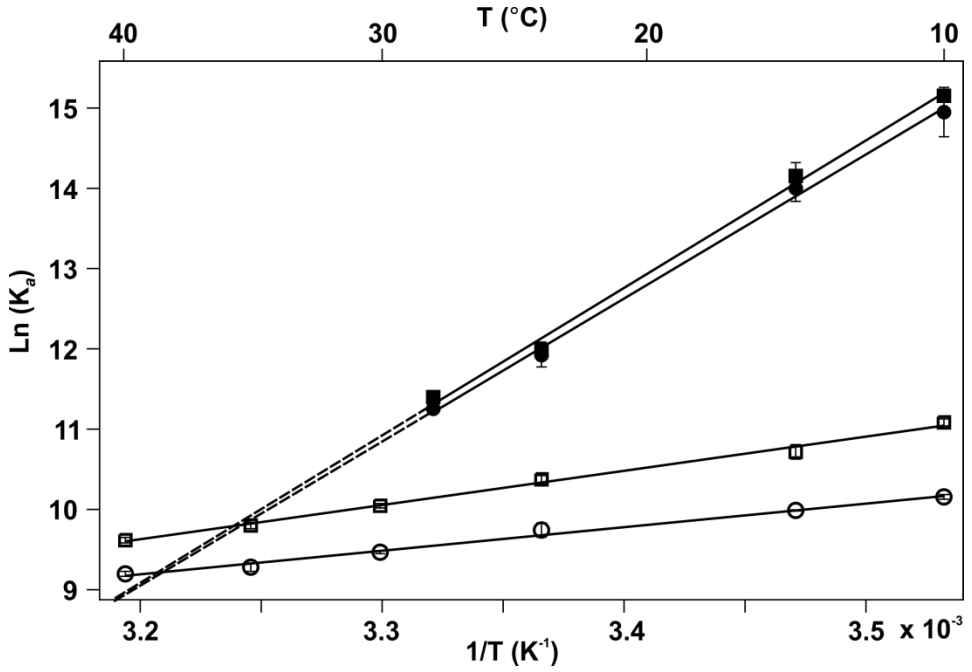

Figure 2.3 Van't Hoff plots for the association of **1** with GTA_2 (□), (GTA_2 + 2(**2** + Mn)) (■), GTB_2 (○) and (GTB_2 + 2(**2** + Mn)) (●). All points represent five replicates and error is shown as the standard deviation.

Table 2.1 Thermodynamic parameters (ΔH_a and ΔS_a) for the association of **1** with GTA or GTB in the presence and absence of bound **2** (and Mn^{2+}) in aqueous solutions at pH 7.[a,b,c]

Enzyme	ΔH_a (kcal mol^{-1})	ΔS_a (cal mol^{-1} K^{-1})
GTA	-8.5 ± 0.3	-8.0 ± 0.9
GTA + **2** + Mn^{2+}	-37 ± 6	-99 ± 2
GTB	-5.8 ± 0.1	-0.3 ± 0.4
GTB + **2** + Mn^{2+}	-36 ± 2	-96 ± 5

a. All solutions contained 100 μM $Mn(C_2H_4O_2)_2$.

b. The reported ΔH_a and ΔS_a represent intrinsic values, in which the statistical factors arising from the two equivalent binding sites for **1** have been accounted for.

c. The reported errors are one standard deviation.

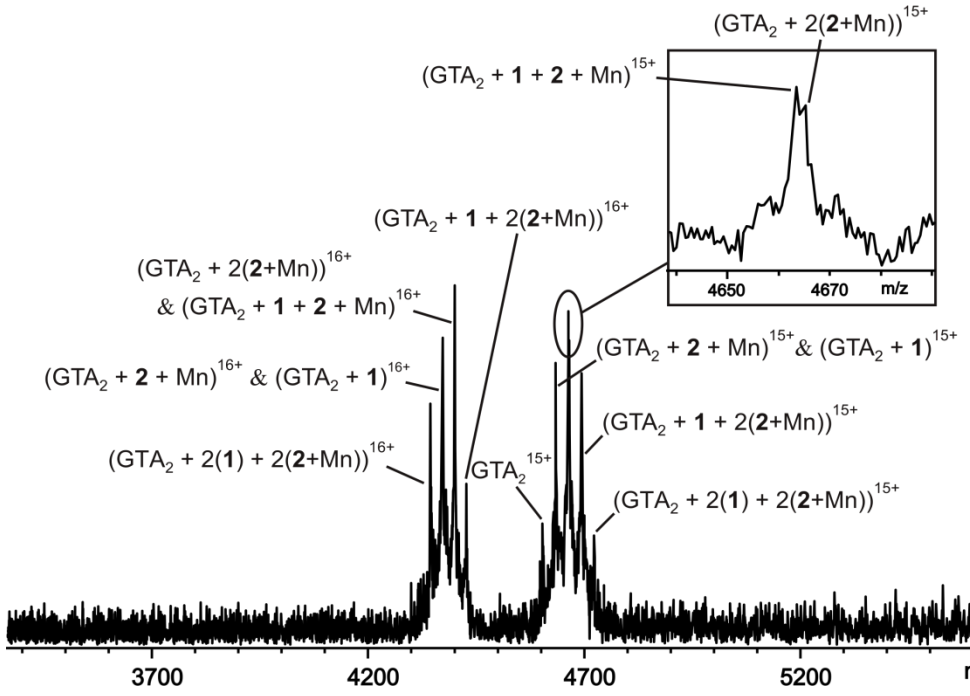

Figure 2.4 Illustrative nanoES mass spectrum acquired for an aqueous solution of 14 μM GTA_2, 100 μM Mn^{2+}, 20 μM **1**, 50 μM **2** and 10 mM ammonium acetate (pH 7) at 37 °C.

the van't Hoff plots for **1** binding in the absence and presence of **2** intersect at ~37 °C, indicating that binding of **1** is independent of **2** at this temperature. These results may suggest that native donor-acceptor substrate binding to the GTs is also non-cooperative at physiological temperature. Taken on their own, the present findings imply that the GTs bind to their donor and acceptor substrates in a random fashion. This contrasts with an ordered mechanism of binding, which is supported by kinetic data.[6] The apparent discrepancy between the thermodynamic and kinetic data may be due to acceptor binding in a non-catalytically competent mode, or that **2**, and not the native nucleotide-sugar donor, was used in the binding measurements. Acceptor affinity measurements performed in the presence of inert nucleotide-sugar donor analogs are now needed to assess whether the galactose moiety of the native donor substrate influences acceptor binding.

Analysis of crystallographic data obtained for free GTB and GTB bound with **1** and **2**,[3,25] suggests a possible explanation for the changes in ΔH_a and ΔS_a values measured for **1** binding with the GTs in the absence or presence of **2**. Specifically, in the presence of **2**, the binding of **1** results in a conformational change in GTB in which two disordered loop regions (residues 175–195 and 345–354) adopt a more ordered structure. A similar conformational change is not observed upon binding of **1** or **2** separately. In the absence of other effects, loop ordering in GTB upon binding of **1** is consistent with the reduction in the ΔS_a value. Similarly, the more favorable ΔH_a term is consistent with the formation of new or stronger intramolecular interactions which accompanies the

conformational change in GTB. Furthermore, Alfaro et al. reported in their crystallographic study that structures containing Arg^{176} show significantly more order than the corresponding structures containing Gly^{176}, supporting the lower ΔS_a values of GTA than those of GTB.[25] While other structural studies on glycosyltransferases have shown substrate-induced conformational changes upon donor substrate binding,[26] this is the first example where the effects of the conformational changes on the association thermodynamic parameters are quantified.

2.4 Conclusions

Using the temperature-controlled ES-MS assay, we have measured the affinities of GTA and GTB for the disaccharide acceptor **1** in the presence and absence of **2**. At physiological temperature, binding of **1** is independent of **2**, while positive cooperativity is observed at lower temperatures. It is proposed that structural changes in the GTs, upon binding of **1** in the presence of **2**, are primarily responsible for the decrease in ΔH_a and ΔS_a values. The present study, the first demonstration of the *direct* ES-MS binding assay for quantifying the association thermodynamic parameters for a multi-substrate enzyme system, highlights the benefits of having direct insight into binding stoichiometry when determining the thermodynamic parameters for multi-component complexes.

2.5 Literature Cited

(1) Breton, C.; Snajdrova, L.; Jeanneau, C.; Koca, J.; Imberty, A. *Glycobiology* **2006**, *16*, 29R–37R.

(2) Schuman, B.; Alfaro, J. A.; Evans, S. V. *Top. Curr. Chem*. **2007**, *272*, 217–257.

(3) Patenaude, S. I.; Seto, N. O. L.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol*. **2002**, *9*, 685–690.

(4) Seto, N. O. L.; Palcic, M. M.; Hindsgaul, O.; Bundle, D. R.; Narang, S. A. *Eur. J. Biochem*. **1995**, *234*, 323–328.

(5) Angulo, J.; Langpap, B.; Blume, A.; Biet, T.; Meyer, B.; Krishna, N. R.; Peters, H.; Palcic, M. M.; Peters, T. *J. Am. Chem. Soc*. **2006** *128*, 13529–13538.

(6) Kamath, V. P.; Seto, N. O. L.; Compston, C. A.; Hindsgaul, O.; Palcic, M. M. *Glycoconjugate J*. **1999**, *16*, 599–606.

(7) Bundle, D. R.; Sigurskjold, B. W. *Methods Enzymol.* **1994**, *247*, 288–305.

(8) Lundquist, J. J.; Toone, E. J. *Chem. Rev*. **2002**, *102*, 555–578.

(9) Schriemer, D. C.; Hindsgaul, O. *Comb. Chem. High Throughput Screen*. **1998**, *1*, 155–170.

(10) Zhang, B. Y.; Chan, N.; Schriemer, D. C.; Palcic, M. M.; Hindsgaul, O. *Glycobiology* **1999**, *9*, 1103–1103.

(11) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom*. **2002**, *216*, 1–27.

(12) Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev*. **2004**, *23*, 368–389.

(13) Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605–611.

(14) Loo, J. A.; Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y. M.; Kaufman, S. L.; Chernushevich, I. V. *J. Am. Soc. Mass. Spectrom*. **2005**, *16*, 998–1008.

(15) Wang, W. J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2003**, *75*, 4945–4955.

(16) Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc*. **2004**, *126*, 4786–4787.

(17) Marcus, S. L.; Polakowski, R.; Seto, N. O. L.; Leinala, E.; Borisova, S.; Blancher, A.; Roubinet, F.; Evans, S. V.; Palcic, M. M. *J. Biol. Chem*. **2003**, *278*, 12403–12405.

(18) Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; Mackenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Nat. Acad. Sci. USA*. **1994**, *91*, 6423–6427.

(19) Lowary, T. L.; Hindsgaul, O. *Carbohydr. Res*. **1993**, *249*, 163–195.

(20) Shoemaker, G. K.; Kitova, E. N.; Palcic, M. M.; Klassen, J. S. *J. Am. Chem. Soc*. **2007**, *129*, 8674–8675.

(21) Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S. *Anal. Chem*. **2006**, *78*, 3010–3018.

(22) Soya, N.; Shoemaker, G. K.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2009**, *19*, 1224–1234.

(23) Zhang, Y. N.; Wang, P. G.; Brew, K. *J. Biol. Chem*. **2001**, *276*, 11567−11574.

(24) Boix, E.; Zhang, Y. N.; Swaminathan, G. J.; Brew, K.; Acharya, K. R. *J. Biol. Chem*. **2002**, *277*, 28310−28318.

(25) Alfaro, J. A.; Ruixiang, Z. B.; Persson, M.; Letts, J. A.; Polakowski, R.; Bai, Y.; Borisova, S. N.; Seto, N. O. L.; Lowary, T.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2008**, *283*, 10097−10108.

(26) Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E. *Trends. Biochem. Sci*. **2005**, *30*, 53−62.

Chapter 3

Comparative Study of Substrate and Product Binding to the Human ABO(H) Blood Group Glycosyltransferases[†]

3.1 Introduction

Glycosyltransferases (GTs) represent a large and diverse group of enzymes of which there is over 200 known sequences in humans.[1] These enzymes catalyze the biosynthesis of the wide array of glycoconjugates found in nature through the stereo- and regiospecific transfer of monosaccharides from sugar-nucleotide donors to specific acceptor substrates. A remarkable feature of GTs is the specificity by which they carry out their function, with each GT typically responsible for the formation of one specific glycosidic linkage. The resulting glycoconjugates play a significant role in many life processes such as cell growth, cell-cell and cell-matrix interactions,[2] bacterial, viral, and fungal infections,[3] inflammation,[4] tumorigenesis[5–10] as well as many other disorders.[11–13] Understanding the mechanism of GTs and their interactions with their natural substrates could lead to the development of new therapeutics against various infections and disorders.[14,15]

[†] A version of this chapter has been published: Soya, N.; Shoemaker, G. K.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2009**, *19*, 1224–1234.

The human blood group GTs, α-(1→3)-N-acetylgalactosaminyltransferase (GTA; EC 2.4.1.40) and α-(1→3)-galactosyltransferase (GTB; EC 2.4.1.37), are two highly homologous GTs that catalyze the transfer of a monosaccharide from an activated sugar nucleotide donor to the H-antigen, α-L-Fuc*p*-(1→2)-β-D-Gal*p*-OR (**1**, the structures of the ligands are given in Figure 3.1), terminated oligosaccharides (where R is a glycoprotein or glycolipid, Figure 1.1). A remarkable feature of these two enzymes is that, despite only differing in four of 354 amino acids (Arg/Gly 176, Gly/Ser 235, Leu/Met 266, and Gly/Ala 268),[16] they possess different donor specificities. GTA utilizes a uridine 5′-diphospho-N-acetylgalactosamine (UDP-GalNAc, **9**) donor to transfer a GalNAc moiety to **1** forming the A trisaccharide (**7**) whereas GTB transfers a Gal residue from a uridine 5′-diphosphogalactose (UDP-Gal, **10**) donor forming the B trisaccharide (**8**). Both of these reactions proceed with a net retention of configuration of the anomeric centre of the transferred sugar and are thus referred to as retaining glycosyltransferases. These enzymes have been studied using a wide range of techniques such as X-ray crystallography,[17–21] kinetic analysis,[22,23] and NMR.[24,25] These studies have led to important insights into the molecular basis for substrate recognition and donor specificity and have resulted in the emergence of GTA and GTB as model systems for this important class of enzymes.

Despite this wealth of experimental data, several issues regarding these enzymes remain unresolved. For example, the native oligomeric forms of the recombinant soluble fragments of the GTs are unclear. Dimers are observed in X-ray structures while both monomer and dimer species have been observed in non-

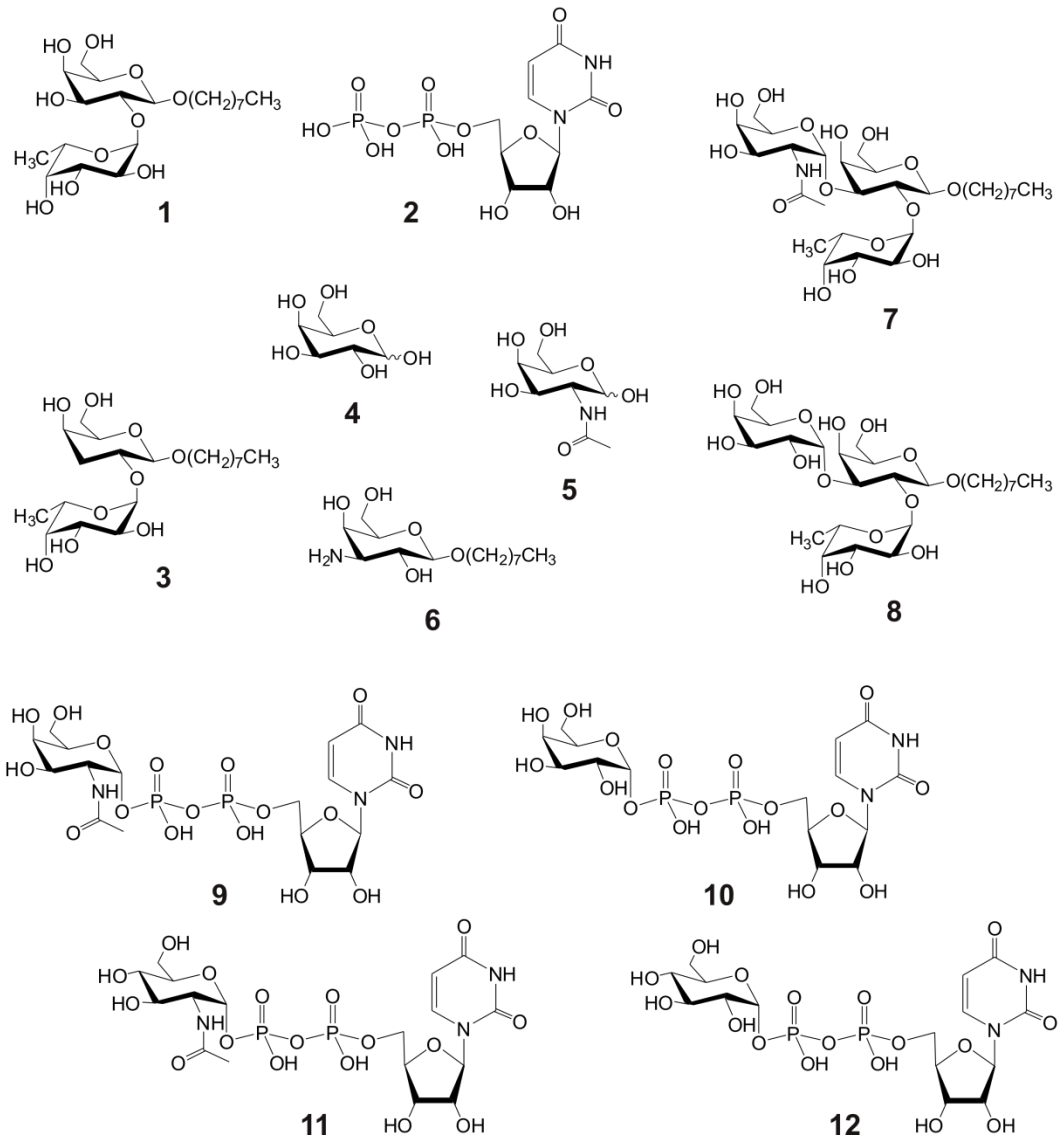

Figure 3.1 Structures of ligands used in the ES-MS binding measurements: α-L-Fuc*p*-(1→2)-β-D-Gal*p*-O-$(CH_2)_7CH_3$ (**1**), uridine 5′-diphosphate (**2**), α-L-Fuc*p*-(1→2)-β-D-3-deoxy-Gal*p*-O-$(CH_2)_7CH_3$ (**3**), D-galactose (**4**) N-acetyl-D-galactosamine (**5**), 3-amino-3-deoxy-β-D-Gal*p*-O-$(CH_2)_7CH_3$ (**6**), α-D-GalNAc*p*-(1→3)[α-L-Fuc*p*-(1→2)]-β-D-Gal*p*-O-$(CH_2)_7CH_3$ (**7**), α-D-Gal*p*-(1→3)[α-L-Fuc*p*-(1→2)]-β-D-Gal*p*-O-$(CH_2)_7CH_3$ (**8**), uridine 5′-diphospho-N-acetylgalactosamine (**9**), uridine 5′-diphosphogalactose (**10**), uridine 5′-diphospho-N-acetylglucosamine (**11**), uridine 5′-diphosphoglucose (**12**).

denaturing gel electrophoresis.[19] There also remain outstanding questions pertaining to the enzyme's mechanism, such as the nature and precise role of the divalent metal cation that is required for optimal enzymatic activity of the enzymes. Additionally, information regarding the binding of native donors to GTA and GTB is limited since attempts to soak or co-crystallize these enzymes with native donor molecules have proven difficult.

Mass spectrometry (MS), combined with electrospray ionization (ES), has become an important tool for quantifying the stoichiometry and affinity of protein-ligand complexes, including protein-carbohydrate interactions, *in vitro*.[26–28] Recently, our laboratories investigated the binding of **1** with recombinant soluble forms of GTA and GTB using the *direct* ES-MS binding assay.[29] Association constants (K_a) were measured for **1** binding to GTA and GTB homodimers, in the absence and presence of uridine 5′-diphosphate (UDP, **2**) and a metal ion cofactor, Mn^{2+}. From a van't Hoff analysis of the temperature dependence of the K_a values, the corresponding enthalpies and entropies of association (ΔH_a, ΔS_a) were determined. The presence of bound **2** and Mn^{2+} in the donor binding sites have a marked influence on the thermodynamic parameters, with both ΔH_a and ΔS_a decreasing significantly. However, for each GT, K_a is unchanged at physiological temperature. The differences in the ΔH_a and ΔS_a values determined in the presence and absence of bound **2** were attributed to structural changes in the GTs induced by the simultaneous binding of **1** and **2**.

In the present work, we seek to develop a more complete understanding of the behaviour of GTA and GTB in aqueous solution, their interactions with donor

and acceptor substrates and trisaccharide products and to more precisely define the role of a metal cofactor therein. To that end we have utilized the *direct* ES-MS binding assay to carry out the first comprehensive and quantitative study of the association of GTA and GTB with their native donor and acceptor substrates, a series of substrate analogues and fragments and the trisaccharide products, in the absence and presence of metal cofactor.

3.2 Experimental

3.2.1 Protein Preparation and General Reagents

Recombinant soluble fragments of GTA (monomer MW 34,519 Da) and GTB (monomer MW 34,483 Da) composed of a full C-terminal and catalytic domain as well as a truncated N-terminal domain were over-expressed in *E. coli* BL21 cells and purified using procedures described previously.[22,23,30] To facilitate ES-MS analysis, enzyme solutions were exchanged into a 50 mM ammonium acetate buffer (pH 7) using an Amicon ultracentrifugation filter (Millipore, Billerica, MA, USA) with a MWCO of 10 kDa. The purity of the dialyzed GT solutions was confirmed by performing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in 10% polyacrylamide gel under denaturing conditions (Figure 3.2). Protein concentrations were determined by lyophilizing an aliquot of the stock solution and weighing the mass of the corresponding protein pellet.

The carbohydrate ligands (**1, 3** and **6**) used in this study were obtained from Dr. Ole Hindsgaul (Carlsberg Laboratory) while compounds **7** and **8** were

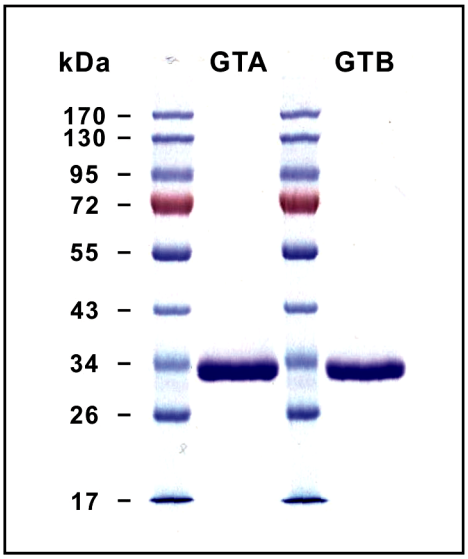

Figure 3.2 SDS-PAGE profiles of GTs in 10% acrylamide gel under denaturing conditions (reducing with DTT and heating at 95 °C for 5 min). 5 µg of GTs was analyzed by SDS-PAGE after purification and dialysis. The gel was stained with coomassie brilliant blue. The single bands identified for GTA and GTB correspond to the monomers of GTA and GTB (MW 34,519 Da and 34,483), respectively.

enzymatically synthesized from compound **1** using GTA and GTB.[22,30] The monosaccharides (**4** and **5**), the native donor substrates (**9** and **10**), UDP (**2**) and donor analogues (uridine 5′-diphospho-N-acetylglucosamine, **11** and uridine 5′-diphosphoglucose, **12**) were purchased from Sigma-Aldrich Canada (Oakville, ON, Canada).

A single chain variable fragment, scFv (MW 26,539 Da), of the carbohydrate-binding IgG antibody of Se155-4 was produced using recombinant technology[31] and was used as a reference protein in the ES-MS binding measurements.

3.2.2 Mass Spectrometry

ES-MS measurements were performed using an Apex II 9.4 tesla Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (Bruker, Billerica, MA) equipped with a custom built temperature-controlled nanoflow ES (nanoES) device.[32] The ES solutions were prepared from aqueous stock solutions of proteins and ligands with known concentration. Unless otherwise indicated, aqueous ammonium acetate was added to the ES solution to yield a final buffer concentration of 10 mM. ES was performed using borosilicate tubing (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in positive ion mode was established by applying a voltage of ~1000 V to a platinum wire inserted inside the glass tip. The solution flow rate was typically ~20 nL/min. The droplets and gaseous ions produced by ES were introduced into

the mass spectrometer through a stainless steel capillary (i.d. 0.43 mm) maintained at an external temperature of 66 °C. The ions sampled by the capillary (+50 V) was transmitted through a skimmer (–2 V) and stored electrodynamically in an rf hexapole for 2 s. Ions were ejected from the hexapole and accelerated to –2700 V into the superconducting magnet, decelerated, and introduced into the ion cell. The trapping plates of the cell were maintained at 1.6 V throughout the experiment. The typical base pressure for the instrument was ~5 × 10^{-10} mbar. Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. The time domain signals, consisting of the sum of 50–100 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier transformation.

3.2.3 Determination of Enzyme-Substrate Binding Constants by ES-MS

The GTA_2 and GTB_2 homodimers (≡ GT_2) possess two equivalent active sites where donor and acceptor substrate molecules can bind. The relevant reactions and equilibrium expressions for the binding of ligand (L), which may be the native or modified donor or acceptor substrate, the enzymatic product, or an inhibitor, to GT_2 are given below:

$$GT_2 + L \rightleftharpoons (GT_2 + L) \qquad K_{a,1} = \frac{[GT_2 + L]}{[GT_2][L]} \tag{3.1a}$$

$$(GT_2 + L) + L \rightleftharpoons (GT_2 + 2L) \qquad K_{a,2} = \frac{[GT_2 + 2L]}{[GT_2 + L][L]} \tag{3.1b}$$

The equilibrium concentrations [GT_2], [GT_2 + L] and [GT_2 + 2L] were determined from the relative abundance of the corresponding ions observed in the ES mass spectrum and the equations of mass balance given below:

$$[GT_2]_o = [GT_2] + [GT_2 + L] + [GT_2 + 2L] \quad (3.2a)$$

$$[L]_o = [L] + [GT_2 + L] + 2[GT_2 + 2L] \quad (3.2b)$$

where $[GT_2]_o$ is an initial concentration of GT_2, and $[L]_o$ is an initial concentration of ligand. After substituting the equilibrium concentrations with the corresponding concentration ratios (R_i), the K_a values were calculated from *eqs* 3:

$$K_{a,1} = \frac{R_1}{[L]_o - \frac{(R_1 + 2R_2)[GT_2]_o}{1 + R_1 + R_2}} \quad (3.3a)$$

$$K_{a,2} = \frac{R_2}{R_1([L]_o - \frac{(R_1 + 2R_2)[GT_2]_o}{1 + R_1 + R_2})} \quad (3.3b)$$

The concentration ratios R_1 (= [GT_2 + L]/[GT_2]) and R_2 (= [GT_2 + 2L]/[GT_2]) were calculated from the ion abundance (*Ab*) ratios, determined from the ES mass spectrum, using *eqs* 4:

$$R_1 = \frac{\sum_n Ab(GT_2 + L)^{n+}}{\sum_n Ab(GT_2)^{n+}} \quad (3.4a)$$

$$R_2 = \frac{\sum_n Ab(GT_2 + 2(L))^{n+}}{\sum_n Ab(GT_2)^{n+}} \quad (3.4b)$$

Because the GTA and GTB homodimers possess two equivalent substrate binding sites, the values of $K_{a,1}$ and $K_{a,2}$ are related by statistical factors that reflect the number of equivalent free and occupied binding sites in the reactant

and product, respectively. The intrinsic K_a (the affinity of each binding site) can be determined from either $K_{a,1}$ or $K_{a,2}$:

$$K_a = \frac{K_{a,1}}{2} \tag{3.5a}$$

$$K_a = 2 \times K_{a,2} \tag{3.5b}$$

The intrinsic K_a values reported in the present study correspond to an average of the values determined from $K_{a,1}$ and $K_{a,2}$.

The determination of K_a values for donor binding (in the presence of Mn^{2+}) is complicated by the occurrence of substrate hydrolysis, which leads to formation of **2**, *vide infra*. Prior to complete hydrolysis, the solution will contain both donor and **2**, each of which can bind to GT_2. As a result, species corresponding to GT_2 bound to one or two molecules of donor or to **2**, or to one molecule of donor and **2** may exist in solution. The corresponding equations of mass balance are given below, where donor is represented as L_1 and **2** as L_2:

$$[GT_2]_o = [GT_2] + [GT_2 + L_1] + [GT_2 + L_2] + [GT_2 + L_1 + L_2] + [GT_2 + 2L_1] + [GT_2 + 2L_2] \tag{3.6a}$$

$$[L_1]_o = [L_1] + [L_2] + [GT_2 + L_1] + [GT_2 + L_2] + 2[GT_2 + L_1 + L_2] + 2[GT_2 + 2L_1] + 2[GT_2 + 2L] \tag{3.6b}$$

The corresponding K_a expressions for donor binding are given below:

$$K_{a,1} = \frac{R_{1,0}}{[L]_o - [L_2] - \frac{(R_{1,0} + R_{0,1} + 2R_{1,1} + 2R_{2,0} + 2R_{0,2})[GT_2]_o}{1 + R_{1,0} + R_{0,1} + R_{1,1} + R_{2,0} + R_{0,2}}} \tag{3.7a}$$

$$K_{a,2} = \frac{R_{2,0}}{R_{1,0}([\mathrm{L}]_o - [\mathrm{L}_2] - \frac{(R_{1,0} + R_{0,1} + 2R_{1,1} + 2R_{2,0} + 2R_{0,2})[\mathrm{GT}_2]_o}{1 + R_{1,0} + R_{1,0} + R_{1,1} + R_{2,0} + R_{0,2}})} \quad (3.7b)$$

where the concentration ratios ($R_{i,j}$) are $R_{1,0}$ (= [GT_2 + L_1]/[GT_2]), $R_{2,0}$ (= [GT_2 + $2L_1$]/[GT_2]), $R_{0,1}$ (= [GT_2 + L_2]/[GT_2]), $R_{0,2}$ (= [GT_2 + $2L_2$]/[GT_2]) and $R_{1,1}$ (= [GT_2 + L_1 + L_2]/[GT_2]). Assuming that the ES-MS measurements are fast compared to the changes in solution concentrations arising from donor hydrolysis, the $R_{i,j}$ values can be calculated from the ion abundance ratios, as determined from the mass spectrum. The general expression for $R_{i,j}$ is given below:

$$R_{i,j} = \frac{\sum_n Ab(\mathrm{GT}_2 + i\mathrm{L}_1 + j\mathrm{L}_2)^{n+}}{\sum_n Ab(\mathrm{GT}_2)^{n+}} \quad (3.8)$$

In order to solve eq 7, [L_2], the concentration of free **2** in solution, must be known. This concentration, at a given time, was determined from the measured ratio, $R_{0,1}$ and $K_{a,1}(L_2)$, the binding constant for the association of L_2 (**2**) to GT_2, using *eq* 9:

$$[\mathrm{L}_2] = \frac{R_{0,1}}{K_{a,1}(L_2)} \quad (3.9)$$

The $K_{a,1}(L_2)$ value was determined in a separate measurement performed on solutions of GT_2 with **2** alone.

The reference protein method was used to correct the ES mass spectra for the occurrence of nonspecific protein-ligand binding.[33] Briefly, this method involves the addition of a reference protein (P_{ref}), which exhibits no specific binding with the ligands of interest, to the ES solution. It was shown previously that the extent of nonspecific binding of neutral and acidic carbohydrates, as well

as **2**, is independent of the size and structure of the protein.[33–35] As a result, the distribution of ligands bound nonspecifically to P_{ref} can be used to correct the measured intensities of the free and ligand-bound GT_2 ions for the occurrence nonspecific ligand binding, leading to more reliable values of K_a.

3.3 Results and Discussion

3.3.1 Native Oligomeric Form of the Recombinant GTA and GTB at Neutral pH

As highlighted in a recent review,[36] various GTs are known to undergo dimerization or oligomerization *in vitro* and *in vivo*. Some GTs form homo-complexes, while others associate with different GTs or non-GT proteins to form hetero-complexes. Evidence for the homodimerization of wild type GTA and GTB, extracted from human plasma, was first reported by Nagai et al.[37,38] The molecular weights of GTA and GTB estimated from Sephadex gel filtration measurements were found to be nearly twice that of the corresponding monomer. Crystallographic data reported by Lee et al. suggested that the recombinant soluble forms of GTA and GTB, which were used in the present study, can also form homodimers in solution.[19] While acknowledging that the dimers identified by X-ray analysis of the crystal structures could be influenced by crystal packing, their analysis suggested that dimerization occurs though the N-terminal residues of stem region. Further support for the formation of dimers was obtained from polyacrylamide gel electrophoresis performed on GTA - the estimated molecular weight determined under non-denaturing conditions corresponded to that of GTA

dimer.[19] While these studies provided evidence that the recombinant soluble forms of GTA and GTB, which are commonly used for biophysical and biochemical studies, can form stable homodimers in aqueous solution, it remains unclear whether the GTs exists exclusively as homodimer at physiological pH and temperature or whether a distribution of oligomeric forms is present.

ES-MS has emerged as a powerful and direct tool for studying protein assembly in solution.[39–41] To further investigate the oligomeric state of the recombinant forms of GTA and GTB, ES mass spectra were acquired for solutions of enzymes at concentrations ranging from 1 to 230 μM. Shown in Figures 3.3a and 3.3b are illustrative ES mass spectra measured for aqueous solutions containing GTA (10 μM) and GTB (10 μM), respectively, at pH 7 in 10 mM ammonium acetate buffer. All at concentrations investigated, the only protein ions detected by ES-MS correspond exclusively to the multiply protonated homodimer of GTA and GTB, $(GTA_2)^{n+}$ and $(GTB_2)^{n+}$ at n = 15 – 17 and n = 14 – 16, respectively (Figures 3.3 and 3.4). The measured molecular weights of GTA_2 (69,039 ± 3 Da) and GTB_2 (68,963 ± 4 Da) agree with the theoretical values for GTA (69,040 Da) and GTB homodimer (68,966 Da), calculated from their respective amino acid sequences. The complete absence of ions corresponding to monomeric forms of GTA or GTB, as well as multimers larger than dimer, suggests that these enzymes exist exclusively as homodimers in aqueous solution at concentrations >1 μM at pH 7 and 24 °C.

Using a temperature-controlled nanoES device, the thermal stability of the homodimers was investigated by ES-MS. Notably, at temperatures up to 60 ºC,

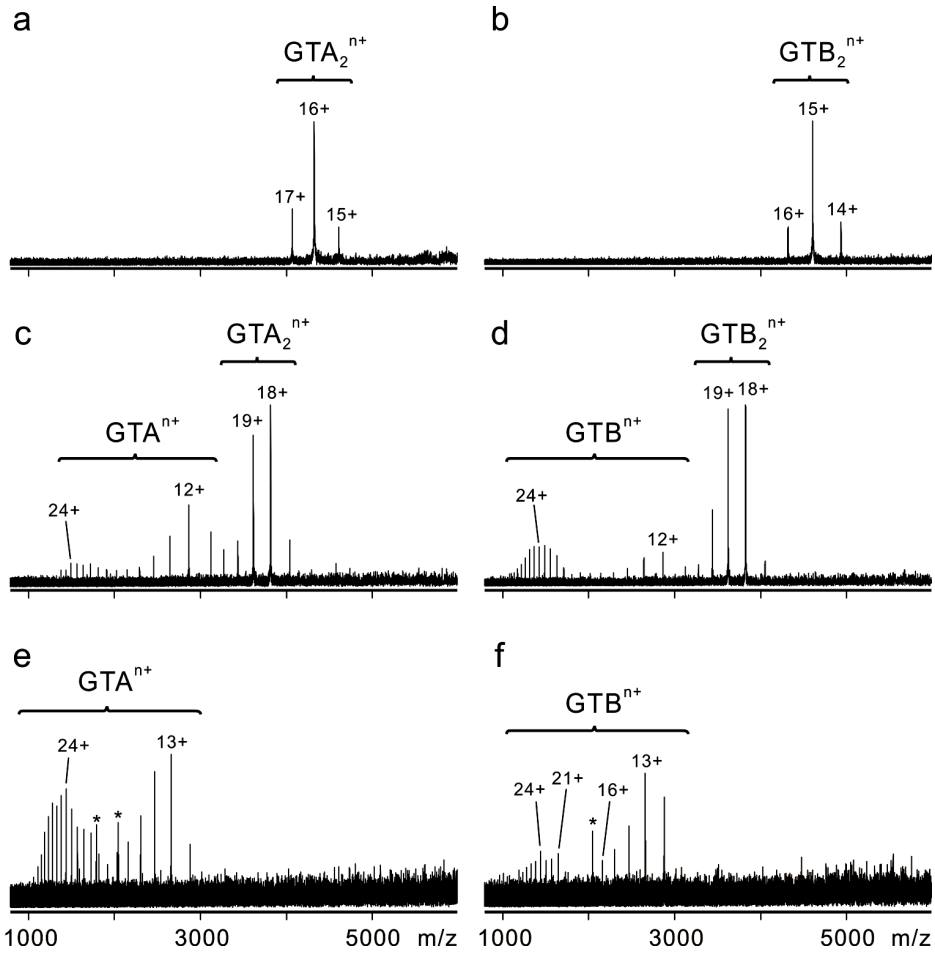

Figure 3.3 Illustrative nanoES mass spectra of aqueous solutions consisting of (a) GTA_2 (5 μM) or (b) GTB_2 (5 μM) in 10 mM ammonium acetate (pH 7); (c) GTA_2 (5 μM) or (d) GTB_2 (5 μM) in 1% acetic acid (pH 3.7) and (e) GTA_2 (5 μM) or (f) GTB_2 (5 μM) in 0.5% formic acid (pH 2.9). Peaks labeled with * correspond to impurities.

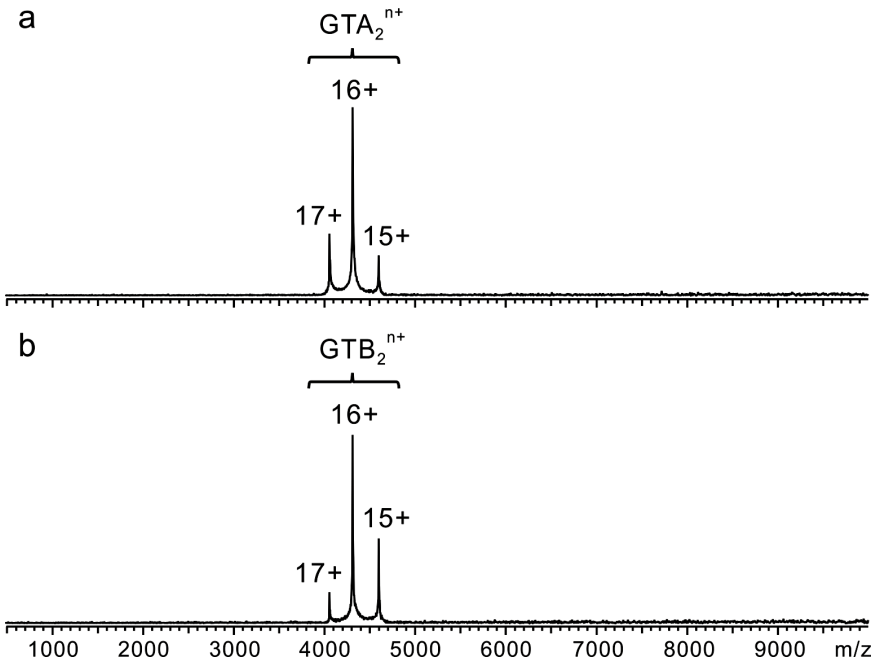

Figure 3.4 Illustrative nanoES mass spectra of aqueous solutions consisting of (a) GTA (230 μM) or (b) GTB (230 μM) in 10 mM ammonium acetate (pH 7).

only ions corresponding to the GTA_2 and GTB_2 were detected (data not shown). The absence of monomer ions at the elevated temperatures suggests that the homodimers are stabilized by relatively strong, possibly ionic, intermolecular interactions. Acidification of the GT solutions resulted in the appearance of ions corresponding to the GT monomers, i.e., GTA^{n+} and GTB^{n+} ions. At pH 3.7 (1% acetic acid), ions corresponding to both monomer and dimer were detected (Figures 3.3c and 3.3d). Notably, acidification results in a shift towards higher charge states for the $(GTA_2)^{n+}$ and $(GTB_2)^{n+}$ ions, consistent with partial acid-induced unfolding of the monomers within the dimer. Furthermore, two distinct charge state distributions are evident for the GTA^{n+} and GTB^{n+} ions, one centered at +12 and one at +24, suggesting the presence of two different solution conformations. Presumably, the distribution centered at +12 corresponds to partially unfolded monomer, while the broader distribution centered at higher charge state corresponds to the fully-denatured monomer.[42] A further reduction in pH to 2.9 (by the addition of 0.5% formic acid) resulted in the complete disappearance of the $(GTA_2)^{n+}$ and $(GTB_2)^{n+}$ ions, leaving only GTA^{n+} and GTB^{n+} ions. Again, two distinct charge state distributions are evident (Figure 3.3e and 3.3f). Taken together, the ES-MS results represent the first direct experimental evidence that these recombinant forms of GTA and GTB exist as homodimers in aqueous solution at neutral pH. Furthermore, dimerization occurs via the formation of non-covalent intermolecular interactions, rather than covalent bonding.

3.3.2 Monosaccharide and Acceptor Analogue Binding

Recent measurements of the binding constants for the acceptor disaccharide **1** with GTA_2 and with GTB_2 at pH 7 and 24 ºC revealed that each homodimer possesses two thermodynamically equivalent acceptor binding sites and that the two enzymes exhibit similar intrinsic affinities for **1**, $(3.2 \pm 0.3) \times 10^4$ M^{-1} (GTA), $(1.7 \pm 0.3) \times 10^4$ M^{-1} (GTB).[29] The similarity in the acceptor affinities is suggestive of similar intermolecular interactions operating in solution. This conclusion is supported by crystal structure data obtained for the (GTA + **1**) and (GTB + **1**) complexes, which reveal that the H-bond interactions between both GTs and **1** are identical (Figure 3.5), with the exception of a putative van der Waals between **1** and Met 266 in GTB.[17] There is no such contact between **1** and Leu 266 of GTA. As noted above, the presence of **2** and Mn^{2+} in the active sites has a significant effect on the thermodynamic parameters for association of **1** and leads to an approximately 5-fold and 10-fold enhancement in affinity for GTA and GTB, respectively, at 24 °C.[29] The enhanced in the binding was attributed to structural changes in the GTs induced by the simultaneous binding of **1** and **2**.

To further investigate the nature of acceptor substrate recognition by GTA and GTB, affinities for a disaccharide acceptor analogue (**3**) and several monosaccharides (**4**–**6**), in the absence and presence of **2** and Mn^{2+}, were measured at neutral pH and 24 ºC using the ES-MS assay. Illustrative ES mass spectra acquired for solutions of GTA and GTB with **3**–**6** in the absence and presence of **2** and Mn^{2+} are shown in Figures 3.6 and 3.7 and in Figures 3.8,

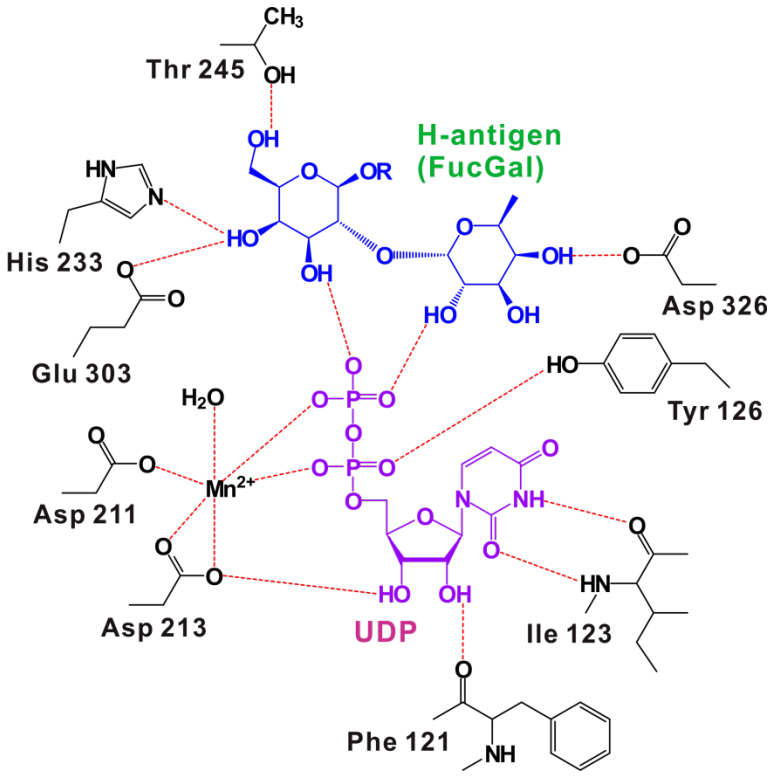

Figure 3.5 The H-bond interactions between GT, the disaccharide H-antigen (**1**) and UDP (**2**) in the presence of Mn^{2+} observed in their crystal structures.[17]

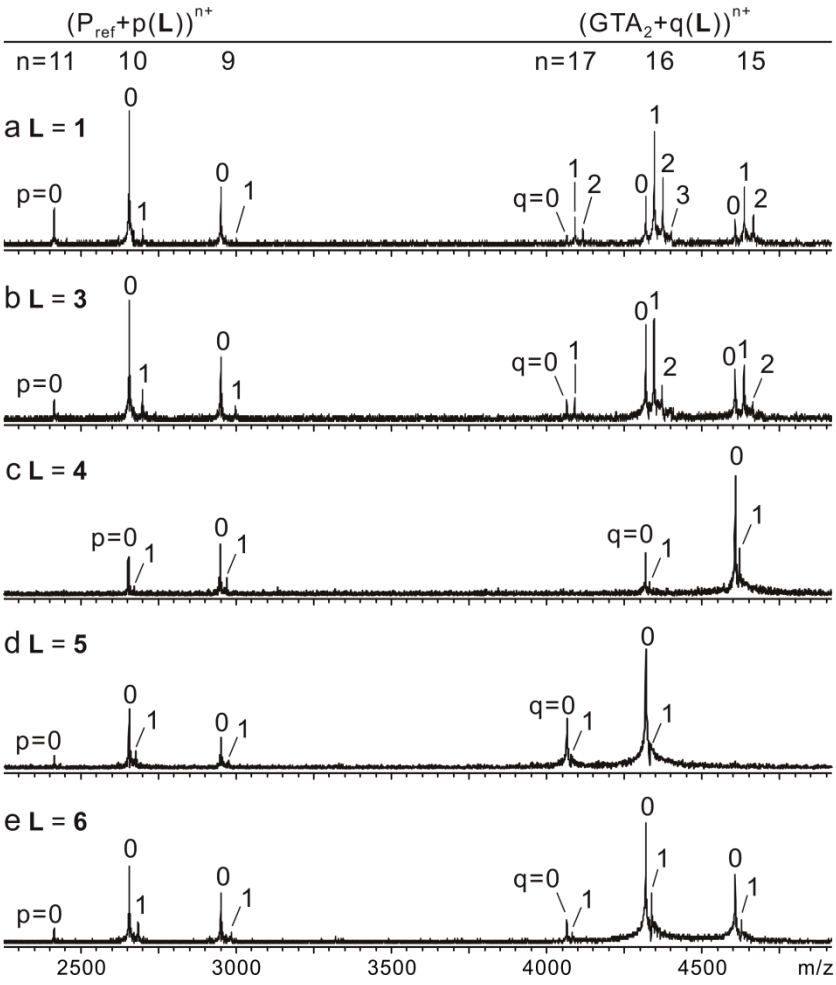

Figure 3.6 Illustrative nanoES mass spectra of solutions consisting of GTA_2 (8 μM) with (a) acceptor **1** (50 μM), (b) acceptor analogue **3** (50 μM), monosaccharides (c) **4** (100 μM), (d) **5** (100 μM) and (e) **6** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES process. All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of ligands bound (nonspecifically) to P_{ref} is indicated by *p*, and the number of ligands bound to GTA_2 is indicated by *q*.

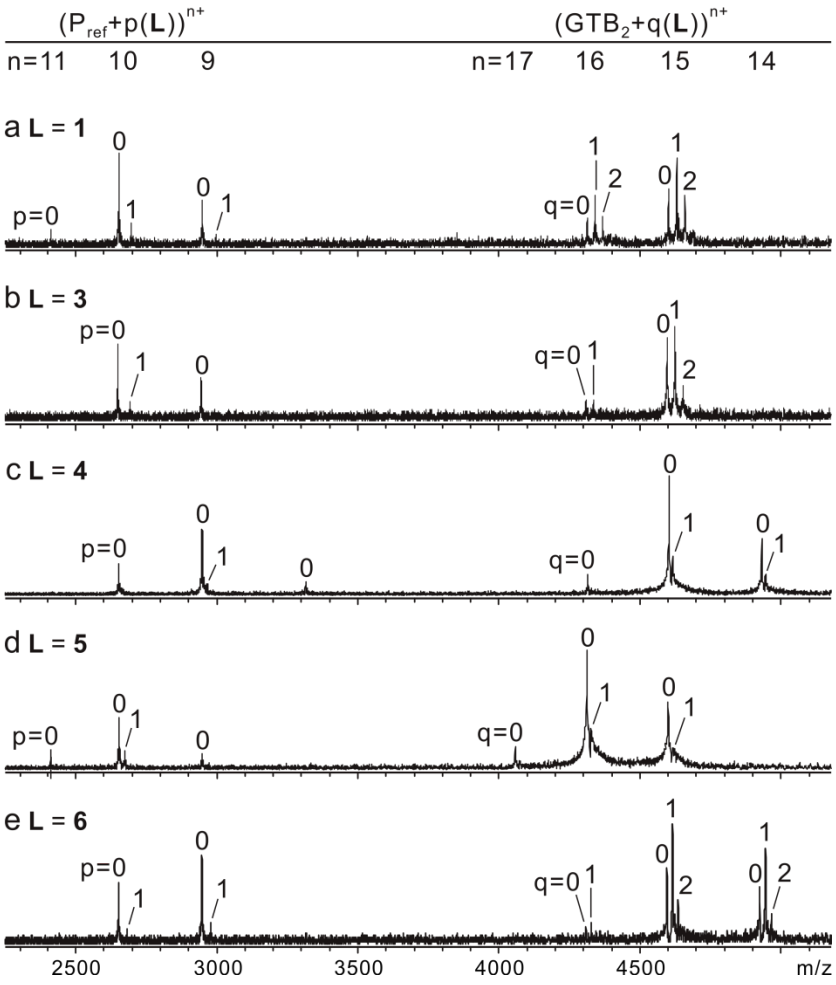

Figure 3.7 Illustrative nanoES mass spectra of solutions consisting of GTB_2 (8 μM) with (a) acceptor **1** (50 μM), (b) acceptor analogue **3** (50 μM), monosaccharides (c) **4** (100 μM), (d) **5** (100 μM) and (e) **6** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES process. All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of ligands bound (nonspecifically) to P_{ref} is indicated by *p*, and the number of ligands bound to GTB_2 is indicated by *q*.

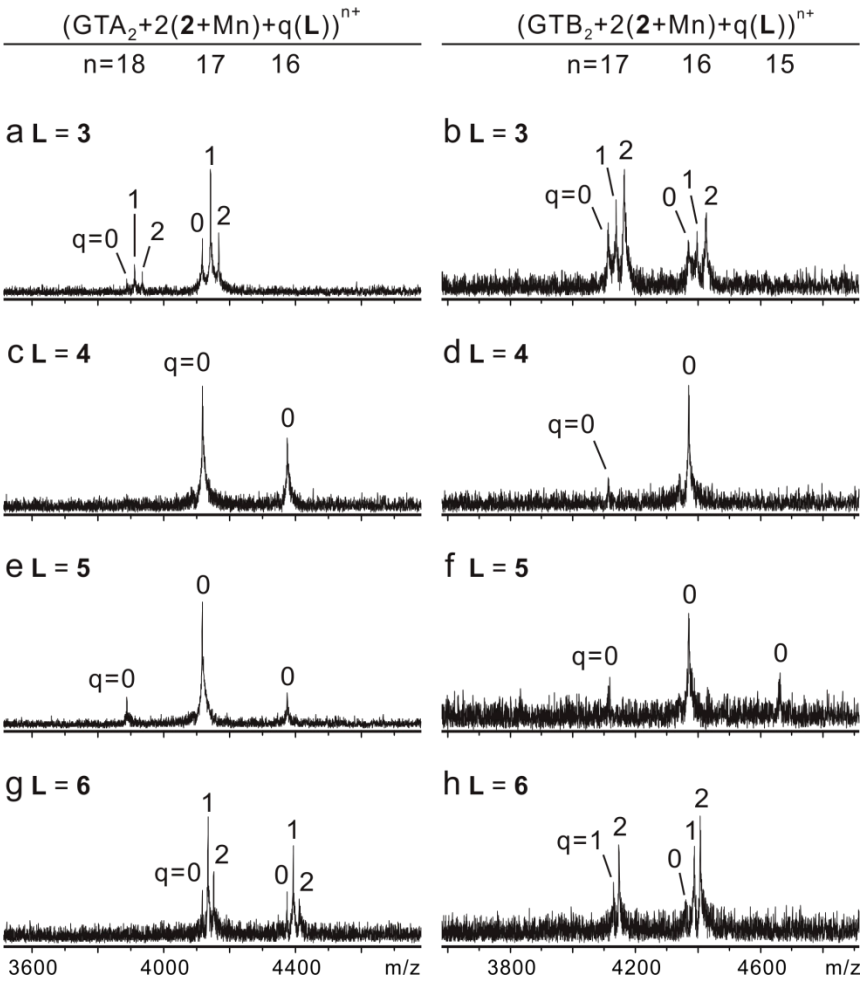

Figure 3.8 Illustrative nanoES mass spectra of aqueous solutions consisting of **2** (50 μM) and 100 μM Mn^{2+} and (a) GTA_2 with **3** (5 μM), (b) GTB_2 with **3** (5 μM), (c) GTA_2 with **4** (100 μM), (d) GTB_2 with **4** (100 μM), (e) GTA_2 with **5** (100 μM), (f) GTB_2 with **5** (100 μM), (g) GTA_2 with **6** (20 μM) or (h) GTB_2 with **6** (5 μM). The concentrations of GTA and GTB were 8 μM for all experiments. All solutions were prepared in 10 mM ammonium acetate buffer (pH 7). The number of ligands bound to the GT ions is indicated by *q*.

respectively. The intrinsic binding constants determined from the ES-MS measurements are listed in Table 3.1.

The disaccharide **3** is a monodeoxy analogue of **1** lacking the C3 OH group of Gal, which is the site of attachment for the monosaccharide that is transferred from the donor to the acceptor substrate. This disaccharide was previously shown to be an effective competitive inhibitor of GTA and GTB, with K_i values of 68 μM and 14 μM, respectively.[43] According to the crystal structure of (GT + **1**) complexes,[17] there are no direct H-bond interactions between the O-3 hydroxyl group of **1** and either GT. Moreover, Nguyen et al. reported that the monodeoxy analogue **3** engages in the same binding interactions with the GTs as does **1**.[18] Not surprisingly, then, the intrinsic binding constants measured for **3** with GTA ($(1.0 \pm 0.2) \times 10^4$ M^{-1}) and GTB ($(9.1 \pm 0.9) \times 10^3$ M^{-1}) are similar to those measured for **1** (Figures 3.6b and 3.7b). These results support the hypothesis that **1** and **3** bind to the GTs with similar intermolecular interactions, at least in the absence of native donor or a donor analogue in the active site. Notably, a dramatic 200–250 fold enhancement in affinity of GTA ($(2.4 \pm 0.3) \times 10^6$ M^{-1}) and GTB ($(1.9 \pm 0.3) \times 10^6$ M^{-1}) for **3** was measured upon introduction of **2** and Mn^{2+} into the GT active sites (Figures 3.8a and 3.8b). These higher affinities are consistent with the potent inhibitory effect of the disaccharide established from kinetic measurements.[43] Our laboratory is currently exploring the origin of the high affinities determined for **3** using molecular modeling tools.

According to a structural analysis carried out by Letts et al, Gal (**4**) can bind to the acceptor sites of GTA and GTB, and the intermolecular interactions

Table 3.1 Intrinsic association constants (K_a) for acceptor substrate (**1**), acceptor analogues (**3**–**6**) and A-, B-trisaccharides (**7**, **8**) binding with GTA_2 and GTB_2 in the absence and presence of 50 μM UDP (**2**) and 100 μM Mn^{2+} determined using the ES-MS assay. All values were determined in aqueous ammonium acetate (10 mM) solutions at pH 7 and 24 °C.[a,b]

Subt.	**GTA** K_a **(M^{-1})**	**GTA + 2 + Mn^{2+}** K_a **(M^{-1})**	**GTB** K_a **(M^{-1})**	**GTB + 2 + Mn^{2+}** K_a **(M^{-1})**
1	$3.2 \pm 0.3 \times 10^4$ [d]	$1.6 \pm 0.3 \times 10^5$ [d]	$1.7 \pm 0.3 \times 10^4$ [d]	$1.5 \pm 0.3 \times 10^5$ [d]
3	$1.0 \pm 0.2 \times 10^4$	$2.4 \pm 0.3 \times 10^6$	$9.1 \pm 0.9 \times 10^3$	$1.9 \pm 0.9 \times 10^6$
4	$5.2 \pm 0.8 \times 10^2$	NB [c]	$4 \pm 2 \times 10^2$	NB [c]
5	NB [c]	NB [c]	NB [c]	NB [c]
6	$1.0 \pm 0.4 \times 10^3$	$1.4 \pm 0.5 \times 10^5$	$1.7 \pm 0.7 \times 10^4$	$4.0 \pm 0.9 \times 10^6$
7	$1.4 \pm 0.4 \times 10^3$	$1.2 \pm 0.2 \times 10^3$	NB [c]	NB [c]
8	$1.6 \pm 0.1 \times 10^4$	NB [c]	$2.7 \pm 0.3 \times 10^3$	NB [c]

a. The reported K_a values are the average value of five measurements.

b. The reported errors are one standard deviation.

c. NB = no specific binding detected.

d. Values taken from ref. 29.

are similar to those involving the Gal residue in **1**.[20] The exception being that Glu 303 of GTB interacts with O-3 and O-4 hydroxyl group of **4** via water bridging, which is not observed in the Gal residue in **1**. In GTA, there are no such bridging interactions, nor direct interactions between the O-3 and O-4 hydroxyl groups of **4** and Glu 303. According to the ES-MS measurements, GTA and GTB exhibit only a weak affinity for **4**, $(5.2 \pm 0.8) \times 10^2\ M^{-1}$ and $(4 \pm 2) \times 10^2\ M^{-1}$, respectively (Figures 3.6c and 3.7c). These results suggest that Gal contributes only modestly to the overall affinities of the GTs for **1** in the absence of donor. In the case of GalNAc (**5**), no specific binding with GTA or GTB was detected (Figures 3.6d and 3.7d). One possible explanation for this observation is that the N-acetyl group is too bulky for the monosaccharide to be properly accommodated by the acceptor binding site. In contrast, both GTA and GTB bind to **6**, a derivative of Gal in which the 3′ hydroxyl group has been replaced with an amino group. GTB exhibits an affinity ($(1.7 \pm 0.7) \times 10^4\ M^{-1}$) (Figure 3.7e) that is identical, within error, to that for **1**; while GTA exhibits a markedly lower affinity ($(1.0 \pm 0.4) \times 10^3\ M^{-1}$) (Figure 3.6e). These results may reflect the ability of the amino group of **6** to form a direct interaction with GTB, but not GTA. However, this hypothesis remains to be verified experimentally or computationally.

Interestingly, the presence of **2** and Mn^{2+} in the donor sites has non-uniform effects on the binding of **4**–**6**. In the case of **4** and **5**, no specific binding to the GTs was detected (Figures 3.8c –f), while for **6**, the presence of **2** and Mn^{2+} resulted in a profound enhancement in binding (Figures 3.8g and h). In the case of GTA, a 150-fold increase in affinity ($(1.4 \pm 0.5) \times 10^5\ M^{-1}$) was measured; for

GTB the effect is even more pronounced with a 200-fold increase in affinity ((4.0 ± 0.9) × 10^6 M^{-1}). These results suggest that **6** could serve as an effective inhibitor of the GTs, particularly for GTB.

3.3.3 Donor Binding

Despite much effort, attempts to soak or co-crystallize wild-type GTA or GTB with their respective donor substrate (**9** and **10**) have been unsuccessful to date. As a result, direct insights into the nature of the GT-donor substrate interactions are lacking and indirect experimental methods or molecular modeling have been used. According to enzymatic activities measured for a series of GTA/GTB mutants, residues 266 and 268 (Leu/Met 266 and Gly/Ala 268) confer donor specificity.[44–46] The results of a molecular modeling study of the complexes of GTA and GTB with their native donors suggest that residue 266 (Leu/Met 266) is responsible for the only differences in the direct interactions between the GTs and their donors – Leu 266 can interact with acetamido group of **9** and Met 266 can interact with the hydroxyl group of **10**.[17] Recently, Blume et al. utilized a fragment based approach combined with affinity measurements performed at 15 °C using NMR spectroscopy to identify which functionalities are important for donor recognition by GTB in the presence of Mg^{2+}.[25] Their NMR experiments revealed that uracil is the minimum structural element of **10** that binds to GTB and that ribose and β-phosphate moieties increased the binding constant, whereas the Gal residue reduces binding. Among the donor and donor analogues investigated (**2**, **9**–**12**), **2** was found to be the best binder with a dissociation

constant (K_d) of 5 – 15 μM (the corresponding range of K_a values is $0.7 - 2 \times 10^5$ M^{-1}). Notably, these binding data correspond to relative K_d values, based on an estimated value determined for **10** (17 μM) from enzyme kinetic measurements in the presence of acceptor at 37 °C.[46]

To our knowledge, equilibrium constants for GTA and GTB interacting with their donor substrate have not been directly measured. To address this deficiency and to obtain new insights into donor recognition by GTA, as well as GTB, the *direct* ES-MS binding assay was used to quantify the interactions of **2** and **9–12** with the GTA and GTB, in the absence and presence of Mn^{2+} (Table 3.2).

Shown in Figure 3.9 are illustrative ES mass spectra obtained for solutions of GTA and GTB, respectively, with **2**. In the absence of Mn^{2+}, neither of the GTs exhibits a measurable affinity for **2** (Figure 3.9a and b). However, upon addition of excess Mn^{2+} (100 μM) both GTs readily bind to **2** and ions corresponding to $(GT_2 + \mathbf{2} + Mn)^{n+}$ and $(GT_2 + 2(\mathbf{2} + Mn))^{n+}$ complexes were detected (Figure 3.9c and d). The intrinsic affinities of GTA and GTB for **2** (in the presence of Mn^{2+}) are $(4.1 \pm 0.7) \times 10^5$ M^{-1} and $(1.9 \pm 0.3) \times 10^5$ M^{-1}, respectively. Notably, the affinity determined for GTB falls within the range of values reported by Blume et al.[25] The dramatic enhancement observed in the presence of Mn^{2+} can be explained by the added stabilization resulting from the recognition of the phosphate moiety of **2** by the GTs through a DXD motif (Figure 3.5).[17] Considering the donor specificity of GTA and GTB, it is interesting that the GTs would exhibit moderately strong binding affinity for **2** (complexed with Mn^{2+}).

Table 3.2 Association constants (K_a) for GTA_2 and GTB_2 binding with UDP (**2**) and native donors (**9**, **10**) and donor analogues (**11**, **12**) binding with GTA_2 and GTB_2 in the absence and presence of 100 μM Mn^{2+} determined using the ES-MS assay. All values were determined in aqueous ammonium acetate (10 mM) solutions at pH 7 and 24 °C.[a,b]

Subt.	**GTA** K_a **(M^{-1})**	**GTA + Mn^{2+}** K_a **(M^{-1})**	**GTB** K_a **(M^{-1})**	**GTB + Mn^{2+}** K_a **(M^{-1})**
2	NB [c]	$4.1 \pm 0.7 \times 10^5$	NB [c]	$1.9 \pm 0.4 \times 10^5$
				$(0.7 - 2) \times 10^5$ [f]
9	$2.9 \pm 0.9 \times 10^3$	NA [d]	NB [c]	$2.8 \pm 0.2 \times 10^3$
		3.3×10^5 [h]		2×10^4 [f]
				1.4×10^4 [h]
10	$5.6 \pm 0.8 \times 10^2$	$5.5 \pm 0.4 \times 10^4$	$1.1 \pm 0.2 \times 10^3$	$2.0 \pm 0.7 \times 10^4$
		2.7×10^5 [h]		6×10^4 [g]
				7.7×10^4 [h]
11	– [e]	$1.3 \pm 0.3 \times 10^4$	– [e]	$1.8 \pm 0.2 \times 10^3$
				2×10^4 [f]
12	– [e]	$2.0 \pm 0.7 \times 10^4$	– [e]	$1.2 \pm 0.3 \times 10^4$
				$(5 - 10) \times 10^4$ [f]

a. The reported K_a values are the average value of five measurements.

b. The reported error corresponds to one standard deviation.

c. NB = no specific binding detected.

d. NA = unable to determine binding constant due to rapid hydrolysis of ligand.

e. Binding measurements not performed.

f. Values taken from ref. 25.

g. Value taken from reference from ref. 46.

h. Values taken from reference ref. 23.

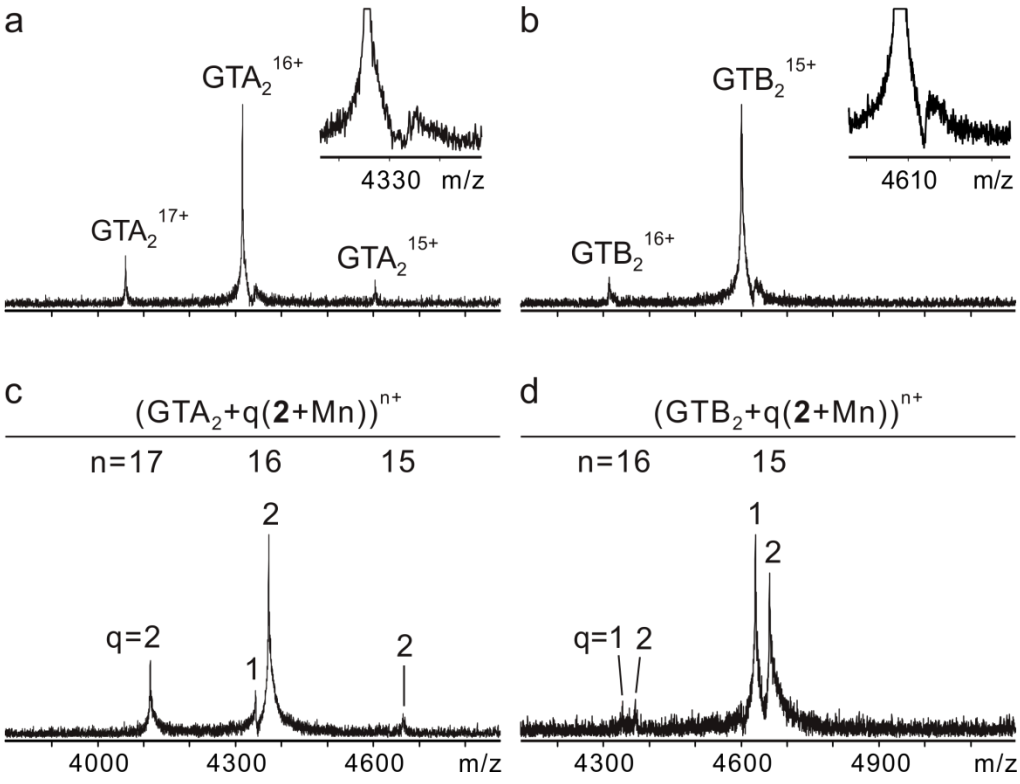

Figure 3.9 Illustrative nanoES mass spectra of solutions consisting of (a) GTA_2 with **2** (40 μM) in the presence of EDTA (40 μM), (b) GTB_2 with **2** (40 μM) in the presence of EDTA (40 μM), (c) GTA_2 with **2** (40 μM) in the presence of Mn^{2+} (100 μM) or (d) GTB_2 with **2** (40 μM) in the presence of Mn^{2+} (100 μM). The concentrations of GTA and GTB were 9 μM for all experiments. All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of **2** bound to the GT ions is indicated by *q*.

This is a structural element of native sugar-nucleotide donors and the product of enzymatic reaction of GTs and, in principle, **2** could significantly inhibit the enzymatic reaction. However, as described below, the presence of **2** in the donor binding site of GTB effectively destabilizes the trisaccharide product, promoting its release from the active site.

Although neither GTA nor GTB exhibit any measurable affinity for **2** in the absence of Mn^{2+}, they do possess a weak affinity for their native donors. Shown in Figures 3.10 and 3.11 are illustrative ES mass spectra acquired for solutions of GTA with **9** and GTB with **10**, respectively. In both cases, ions corresponding to GT_2 bound to one or two donor molecules were observed. The measured affinities are $(2.9 \pm 0.9) \times 10^3$ M^{-1} (GTA) and $(1.1 \pm 0.2) \times 10^3$ M^{-1} (GTB). These results indicate that the sugar moiety of the native donors contribute favorably to binding, at least in the absence of a metal cofactor. Cross donor binding (i.e., GTA with **10** and GTB with **9**) was also investigated and it was found that GTA exhibits a weak but measurable affinity for **10** ($5.6 \pm 0.8 \times 10^2$ M^{-1}); while GTB does not bind to **9** (Figures 3.10b and 3.11a). As mentioned above two (Leu/Met 266 and Gly/Ala 268) out of the four critical amino acids that differentiate GTA from GTB cause the altered donor specificity in the enzymes. It was shown by molecular modeling that the bulkier Met 266 group in GTB hinders the binding of **9**,[17] which is in qualitative agreement with the lack of binding between **9** and GTB. In the case of GTA, however, the less bulky Leu 266 can accommodate the acetamido group of **9**, allowing GTA to bind to both donor substrates, although preferably to its native donor.

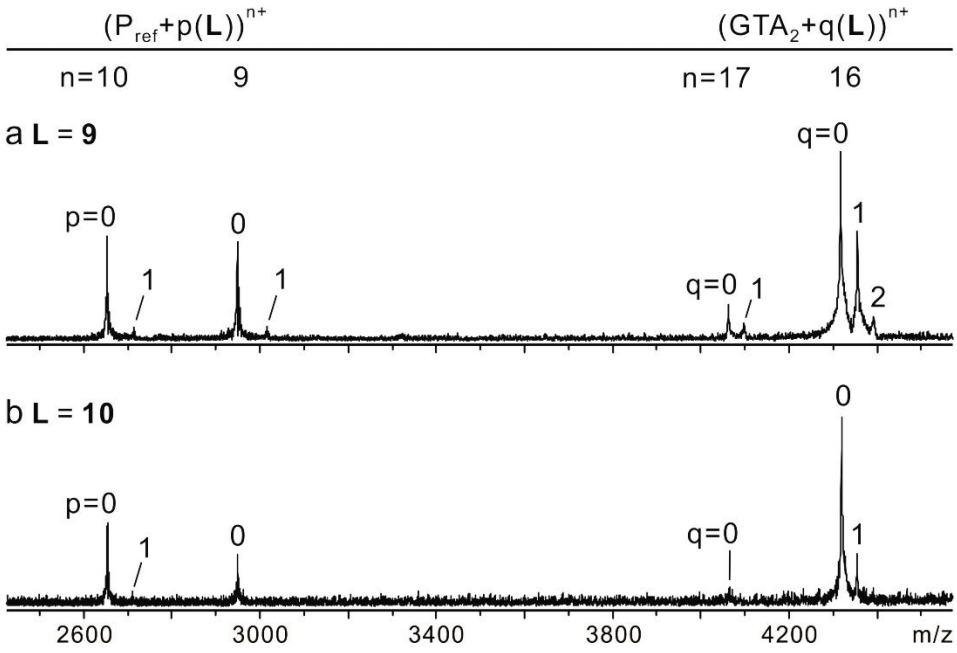

Figure 3.10 Illustrative nanoES mass spectra of solutions consisting of GTA_2 (9 μM) with (a) native donor UDP-GalNAc **9** (80 μM) and (b) the alternate donor UDP-Gal **10** (80 μM). A reference protein, P_{ref} (scFv), was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES process. All solutions were prepared in 20 μM EDTA and 10 mM ammonium acetate buffer (pH 7). The number of donor substrate bound (nonspecifically) to the P_{ref} ions is indicated by *p*, and the number of donor substrates bound to the GTA_2 ions is indicated by *q*.

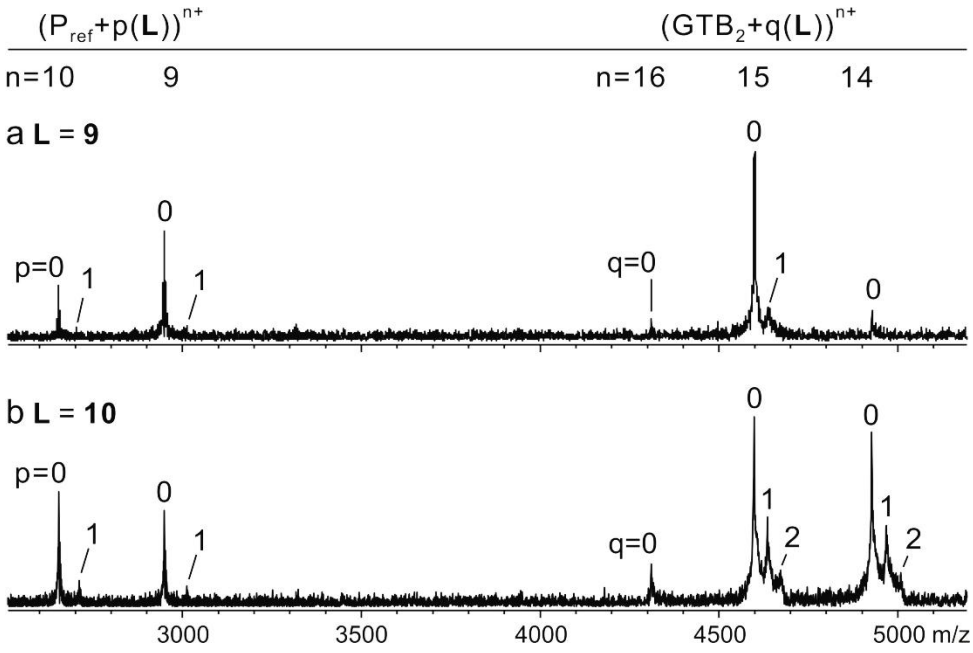

Figure 3.11 Illustrative nanoES mass spectra of solutions consisting of GTB_2 (9 μM) with (a) the alternate donor UDP-GalNAc **9** (80 μM) and (b) the native donor UDP-Gal **10** (80 μM). A reference protein, P_{ref} (scFv), was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES process. All solutions were prepared in 20 μM EDTA and 10 mM ammonium acetate buffer (pH 7). The number of donor substrate bound (nonspecifically) to the P_{ref} ions is indicated by *p*, and the number of donor substrates bound to the GTB_2 ions is indicated by *q*.

The determination of binding constants for the GTs with their native donor substrate in the presence of Mn^{2+} is complicated by enzyme-catalyzed substrate hydrolysis, which leads to time-dependent changes in substrate concentration and the formation of **2**, which also binds to the GTs. In the presence of GTA, hydrolysis of **9** is very rapid under the experimental conditions used, making it impossible to measure a K_a value for this interaction. Shown in Figure 3.12a is a mass spectrum acquired for a solution of 9 μM GTA_2, 40 μM **9** and 100 μM Mn^{2+} in 10 mM ammonium acetate. The mass spectrum, which was measured immediately after preparing the ES solution (<3 min reaction time), revealed only the presence of $(GTA_2 + \mathbf{2} + Mn)^{n+}$ and $(GTA_2 + 2(\mathbf{2} + Mn))^{n+}$ ions. From the fraction of bound GTA_2 and the known binding constant for **2**, the total concentration of **9** in solution was determined to be 2.4 μM. In other words, >90% of **9** was hydrolyzed at the time of ES-MS analysis.

In the presence of GTB, hydrolysis of **10** also occurs. However, the rate of hydrolysis of **10** was sufficiently slow that a binding constant could be determined by ES-MS. Shown in Figures 3.13a and b are ES mass spectra acquired for a solution of 9 μM GTB_2 and 40 μM **10** in the presence of 100 μM Mn^{2+} at ~5 and ~25 min after mixing, respectively. At the shorter reaction times, ions corresponding to GTB_2 with intact donor were detected e.g., $(GTB_2 + \mathbf{10} + Mn)^{n+}$ (Figure 3.13a). However, after ~25 minutes, only ions corresponding to bound **2** were detected, indicating that **10** was completely hydrolyzed (Figure 3.13b). The reason for the difference in the hydrolysis rates of **9** and **10** (in the presence of GTA and GTB, respectively) is not known. However, it is interesting to note that

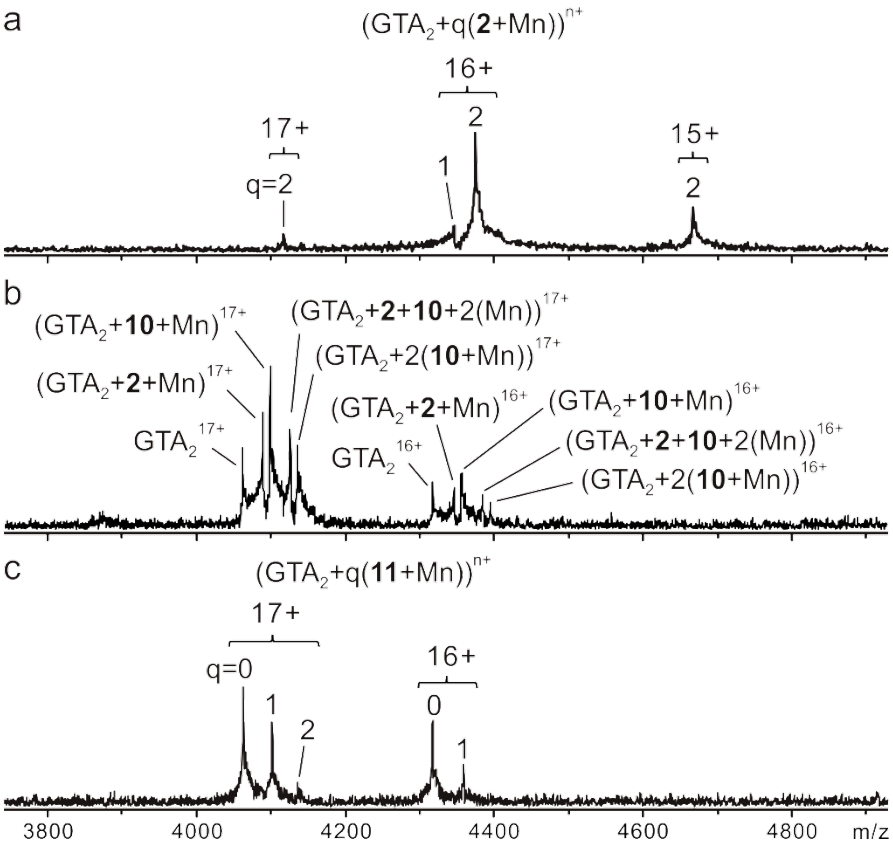

Figure 3.12 Illustrative nanoES mass spectra of solutions consisting of GTA_2 (9 μM) with (a) **9** (40 μM) in the presence of Mn^{2+} (100 μM), the mass spectrum was measured immediately after preparing the solution, (b) **10** (40 μM) in the presence of Mn^{2+} (100 μM) after 5 minutes preparing the solution and (c) **11** (40 μM) in the presence of Mn^{2+} (100 μM). All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of ligands bound to GTA_2 is indicated by *q*.

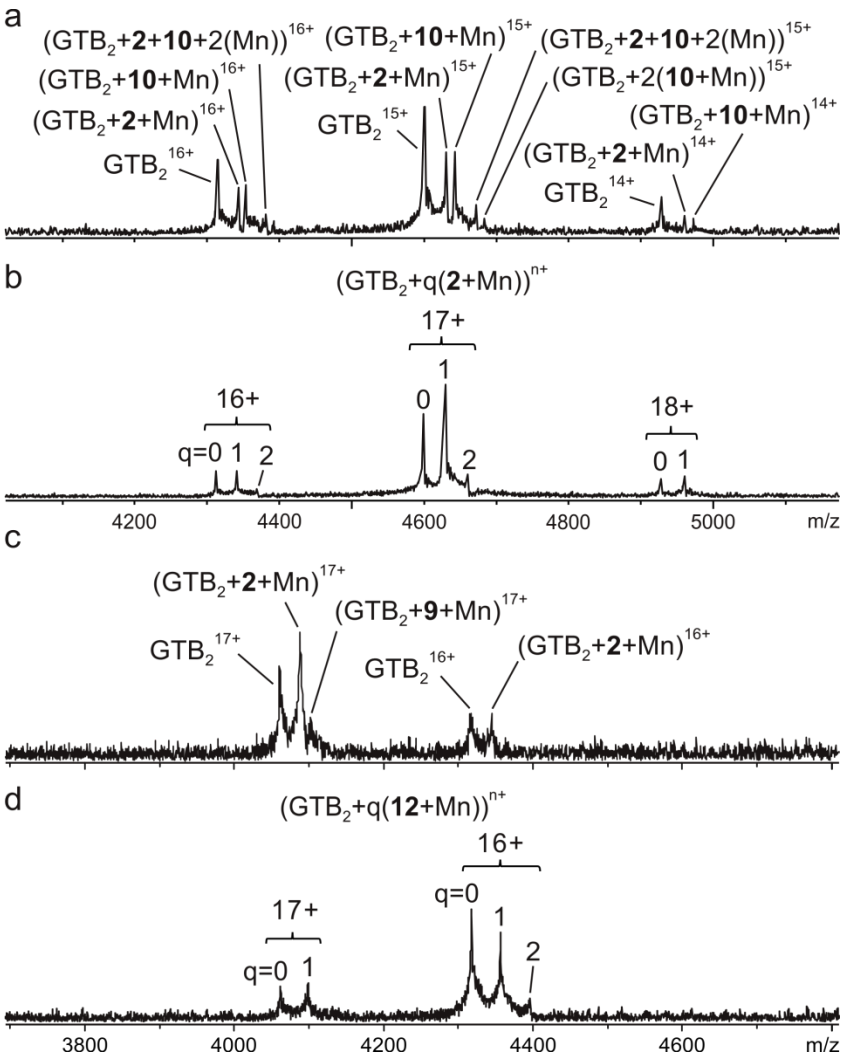

Figure 3.13 Illustrative nanoES mass spectrum of a solution consisting of GTB_2 (9 μM) with (a) **10** (40 μM) in the presence of Mn^{2+} (100 μM), the mass spectrum was measured 5 minutes after preparing the solution, (b) **10** (40 μM) in the presence of Mn^{2+} (100 μM), the mass spectrum was measured 25 minutes after preparing the solution, (c) **9** (40 μM) in the presence of Mn^{2+} (100 μM) 30 minutes after preparing the solution and (d) **12** (40 μM) in the presence of Mn^{2+} (100 μM). All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of ligands bound to the GTB_2 ions is indicated by *q*.

the relative rates of donor hydrolysis mirror the corresponding enzymatic activities – at 37 °C, k_{cat} for the enzymatic reaction involving **1** and **9** (GTA) is 17.5 s^{-1} and 5.1 s^{-1} for the enzymatic reaction involving **1** and **10** (GTB).[23]

Given the slower hydrolysis kinetics for **10** (in the presence of GTB), it was possible to estimate the donor binding constant using the procedure outlined in the Experimental section. Briefly, the concentration of **10** in solution at a given time was established from the relative abundance of bound **2** and the $K_{a,1}$ and $K_{a,2}$ values measured for **2** binding to GTB. Using this procedure, the affinity for GTB with **10** was determined to be $(2.0 \pm 0.7) \times 10^4$ M^{-1}. This corresponds to a 20-fold increase in affinity upon introduction of Mn^{2+} but a 10-fold reduction in binding relative to **2** (in the presence of Mn^{2+}). These results indicate that the monosaccharide moiety of the donor contributes unfavorably to binding. In the absence of Mn^{2+}, GTA binds only very weakly to **10**, while GTB exhibits no measurable affinity of **9**. However, both GTs display moderate affinities in the presence of Mn^{2+}. Interestingly, GTA and GTB were also found to cause the hydrolysis of the cross donor **10** and **9**, respectively, although the reactions were not as rapid as was observed for native donor (Figures 3.12b and 3.13c).

To further examine the specificity of GTA and GTB donor recognition, affinities were measured for the nucleotide donor analogues UDP-GlcNAc (**11)** and UDP-Glc (**12**) in the presence of **2** and Mn^{2+}. It was shown previously that **11** and **12** are poor donors for GTA and GTB, with enzymatic activities <1% of those measured for the native donors.[47] Unlike the native donors, the donor analogues do not undergo hydrolysis in the presence of GTs (Figure 3.12c and 3.13d). The

measured affinities of GTA and GTB for **11** are (1.3 ± 0.3) × 10^4 M^{-1} and (1.8 ± 0.2) × 10^3 M^{-1}, respectively, while the measured affinities for **12** are (2.0 ± 0.7) × 10^4 M^{-1} and (1.2 ± 0.3) × 10^4 M^{-1}, respectively. Interestingly, the affinities determined for **11** and **12** are similar to the values measured for **9** and **10**, a result that suggests that the native donors and donor analogues bind to the GTs through similar intermolecular interactions. It follows that the affinity of GTA for **9** is expected to be similar to that of GTA for **11**, or ~1 × 10^4 M^{-1}.

In contrast to the agreement in the binding constants determined for **2**, the ES-MS derived affinities of GTB for **9**–**12** are consistently smaller (by as much as a factor of 10) than the values reported by Blume et al.[25] Part, but not all, of the discrepancy can be explained by the lower temperature used in the previous study. For example, reducing the solution temperature to 15 °C results in a 2-fold increase in K_a for the affinities GTB for **9** (5 × 10^3 M^{-1}) and **10** (4 × 10^4 M^{-1}). It is possible that the remaining differences are due to partial donor hydrolysis. Because **2** binds more strongly to GTB than the other ligands (**9**–**12**), partial hydrolysis of the ligands will lead to artificially high binding constants in the NMR-based approach. Although it was reported that no hydrolysis was detected for any of the ligands investigated (in the presence of Mg^{2+}), conversion of as little as 5–10% of the donor, resulting perhaps from the presence of trace amounts of Mn^{2+}, which was used in the GT purification, would account for the differences in K_a values.

3.3.4 Trisaccharide Binding

The A (**7**) and B (**8**) trisaccharides are the products of their respective enzymatic reactions. Shown in Figures 3.14a and b are illustrative ES mass spectra obtained for solutions of GTA with **9**, and GTB with **10**, respectively, in the presence of **1** and Mn^{2+}. The ions corresponding to **7** and **8** were clearly detected demonstrating the enzymes maintain their activity under mass spectrometry conditions used. The affinities of GTA and GTB for the **7** and **8** were also determined by the ES-MS assay and the values are listed in Table 3.1. Illustrative ES mass spectra acquired for solutions of GTA and GTB with **7** and **8** are shown in Figures 3.15a,b and 3.16a,b, respectively. In the absence of **2** and Mn^{2+}, GTA binds only weakly to **7** ($(1.4 \pm 0.4) \times 10^3$ M^{-1}) (Figure 3.14a), while GTB exhibits no measurable affinity for the trisaccharide (Figure 3.16a). In contrast, both GTA and GTB exhibit an affinity for **8** (Figures 3.15b and 3.16b). Curiously though, the affinity of GTA ($(1.7 \pm 0.2) \times 10^4$ M^{-1}) for **8** is four-fold higher than that of GTB ($(2.7 \pm 0.3) \times 10^3$ M^{-1}) and is similar to its affinity for native disaccharide **1**. It is known that the critical amino acid changes of Leu/Met 266 and Gly/Ala 268 lead to a somewhat smaller binding site cleft in GTB compared to GTA,[17] which could explain why the GTA binding pocket can accommodate both trisaccharide products but GTB can't. These findings, on their own, suggest that trisaccharide B (**8**) could serve as an inhibitor for GTA. However, as described below, this potential inhibitory effect is mitigated by the presence of **2** in the active site.

As interesting and previously unexplored question is whether the two enzymatic products, A or B trisaccharide and **2** (and metal cofactor), can bind

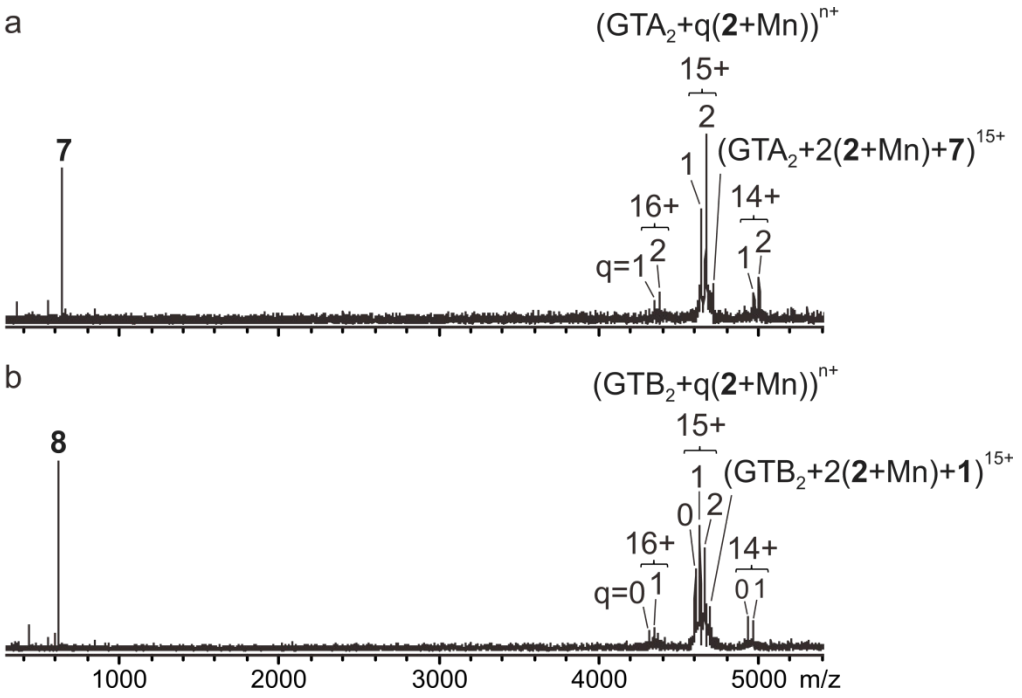

Figure 3.14 Illustrative nanoES mass spectrum of a solution consisting of (a) GTA_2 (8 μM), **1** (20 μM), **9** (20 μM), Mn^{2+} (100 μM) and (b) GTB_2 (8 μM), **1** (20 μM), **10** (20 μM), Mn^{2+} (100 μM) in 10 mM ammonium acetate buffer (pH 7). The number of **2** bound to the GT ions is indicated by *q*.

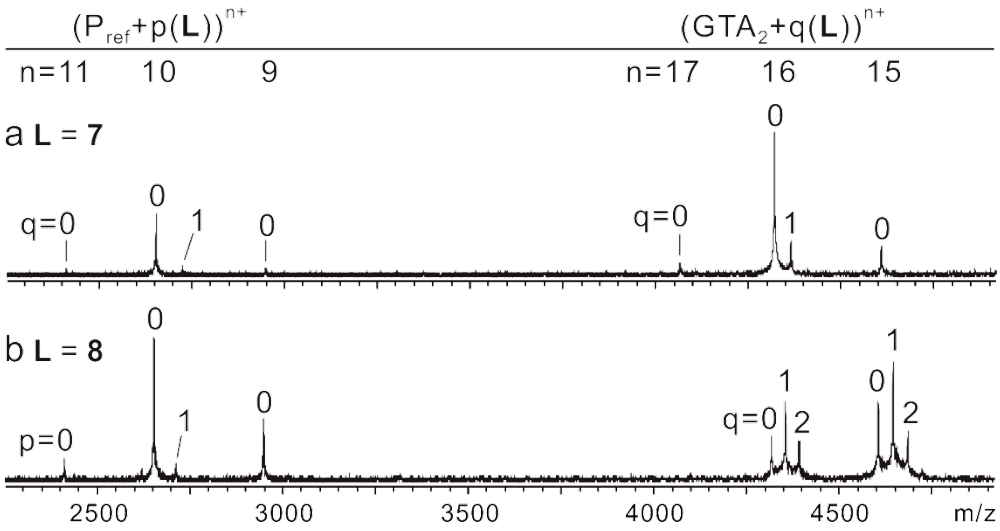

Figure 3.15 Illustrative nanoES mass spectra of solutions consisting of GTA_2 (8 μM) with trisaccharide (a) **7** (50 μM) and (b) **8** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES process. All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of trisaccharides bound (nonspecifically) to P_{ref} is indicated by *p*, and the number of trisaccharides bound to GTA_2 is indicated by *q*.

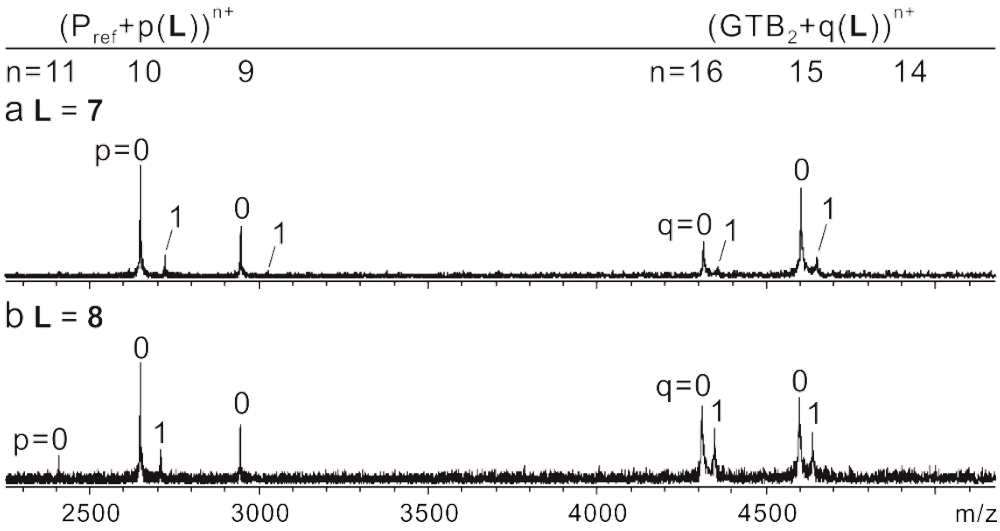

Figure 3.16 Illustrative nanoES mass spectra of solutions consisting of GTB_2 (8 μM) with trisaccharide (a) **7** (50 μM) and (b) **8** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES process. All solutions were prepared in a 10 mM ammonium acetate buffer (pH 7). The number of trisaccharides bound (nonspecifically) to P_{ref} is indicated by *p*, and the number of trisaccharides bound to GTB_2 is indicated by *q*.

simultaneously in the GT active sites. To answer this question, ES-MS was performed on solutions containing GTA or GTB with their respective trisaccharide product, **2** and Mn^{2+}. Illustrative mass spectra of solutions of GTA and GTB with **7** or **8** in the absence or presence of **2** (and Mn^{2+}) are shown in Figure 3.17 and 3.18, respectively. Importantly, the ES-MS data reveal that the presence of **2** in the active sites prevents appreciable association (K_a <10^3 M^{-1}) of either trisaccharide to GTB (Figures 3.18a and b). Similarly, GTA was found not to bind significantly to **8** in the presence of **2** and Mn^{2+} (Figure 3.17d). However, the ES-MS data clearly show that GTA is able to bind simultaneously to **2**, **7** and Mn^{2+} (Figure 3.17c), though the binding is considerably weaker ($(1.2 \pm 0.2) \times 10^3$ M^{-1}) than that of the acceptor substrate (**1**) ($(1.6 \pm 0.3) \times 10^5$ M^{-1}). In fact, the affinities measured for **7** in the presence and absence of **2** are identical, within error. In order to better understand the differential effects of **2** (and metal cofactor) on GTA and GTB recognition of the trisaccharide products, a molecular modeling study has been undertaken.

3.4 Conclusions

The present work represents the first quantitative, comparative study of the interactions between the human blood group glycosyltransferases, GTA and GTB, and their donor substrate, donor and acceptor analogues, and trisaccharide products. The binding measurements, undertaken with the *direct* ES-MS binding assay, yield a number of new insights into these models GTs and their interactions with substrate.

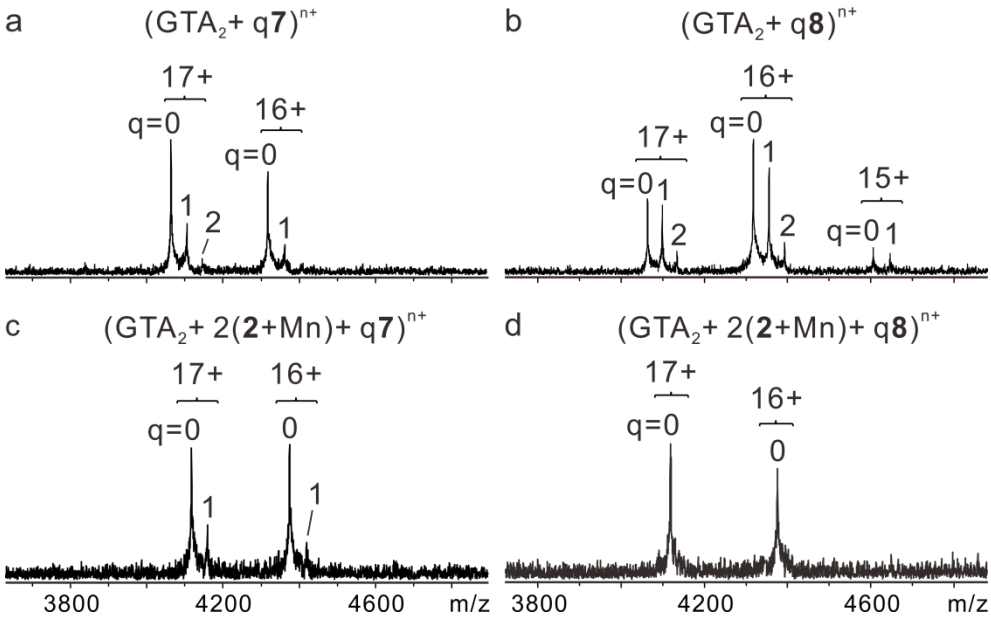

Figure 3.17 Illustrative nanoES mass spectrum of a solution consisting of GTA_2 (9 μM) with (a) **7** (50 μM) and (b) **8** (30 μM) or GTA_2 (9 μM), **2** (50 μM) and Mn^{2+} (100 μM) with (c) **7** (50 μM) and (d) **8** (30 μM). All solutions were prepared in 10 mM ammonium acetate buffer (pH 7). The number of trisaccharide ligands bound to the GTA_2 ions is indicated by *q*.

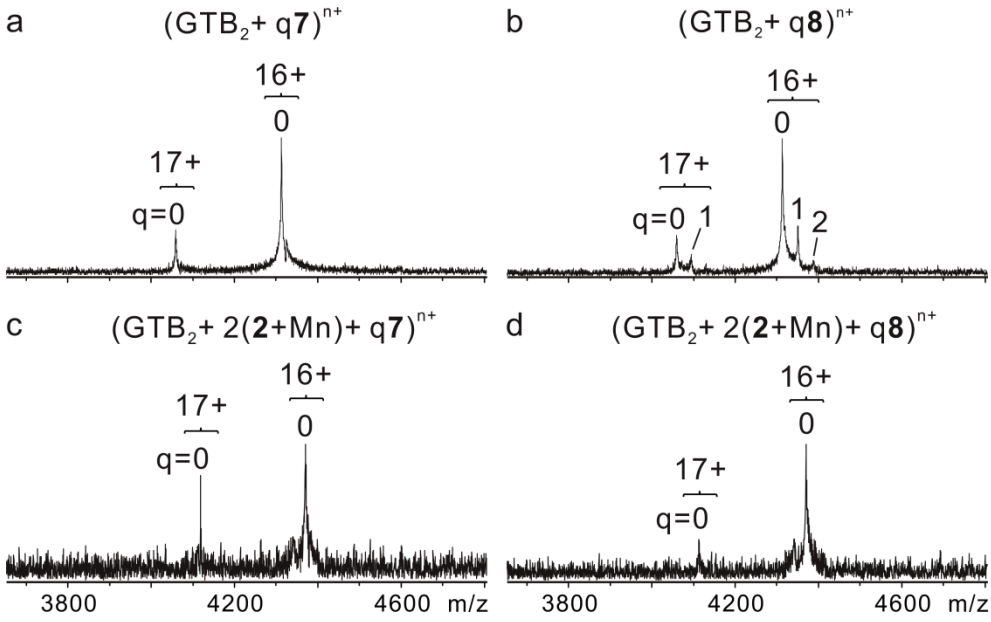

Figure 3.18 Illustrative nanoES mass spectrum of a solution consisting of GTB_2 (9 μM) with (a) **7** (30 μM) and (b) **8** (30 μM), or GTB_2 (9 μM), **2** (50 μM) and Mn^{2+} (100 μM) with (c) **7** (30 μM) and (d) **8** (30 μM). All solutions were prepared in 10 mM ammonium acetate buffer (pH 7). The number of trisaccharide ligands bound to the GTB_2 ions is indicated by *q*.

1) First, it was shown that the recombinant forms of GTs used in this study, as well as by a number of other laboratories, exist preferentially as homodimers in aqueous solution at neutral pH and temperatures ranging from 15 to 60 °C. Furthermore, it was conclusively demonstrated that dimerization occurs via non-covalent interactions as opposed to covalent bonds.

2) Neither of the GTs exhibit any appreciable affinity for Gal (**4**) or GalNAc (**5**), but bind very strongly to a derivative of Gal (**6**) in which the 3′ hydroxyl group has been replaced with an amino group. In fact, when UDP (**2**) and Mn^{2+} are present in the active site, the measured affinities are comparable (in the case of GTA) or larger than those determined for the natural disaccharide acceptor.

3) In the absence of a divalent metal ion cofactor (Mn^{2+}), neither GTA nor GTB bind to **2**. However, both exhibit relatively strong binding in the presence of Mn^{2+}. Similarly, the GTs exhibit weak or no binding to their native donors (**9**, **10**) in the absence of Mn^{2+}. Notably, both **9** and **10** were found to undergo hydrolysis (in the presence of *either* GTA or GTB) upon addition of Mn^{2+}. In the case of **9**, the hydrolysis reaction was too rapid (>90% of donor reacted within 3 minutes) to allow for the determination of the GTA binding constant. The hydrolysis of **10** (in the presence of GTB) was significantly slower and it was possible to extract a binding constant using a modified data analysis approach. Interestingly, cross donor binding was observed for both GTs (in the presence of Mn^{2+}). In

fact, the interaction between **10** and GTA is stronger than the interaction with GTB. Both GTs also bind to the donor analogues (**11**, **12**) with affinities that are similar to those measured for **9** and **10**. These results suggest that the native donors and donor analogues bind to the GTs through similar interactions.

4) Finally, binding of the A and B trisaccharides (**7**, **8**) to the GTs was investigated for the first time. In the absence of **2** and Mn^{2+}, both GTA and GTB recognize their respective trisaccharide products and GTA exhibits relatively strong binding to **7**. However, only GTA retains weak affinity for its trisaccharide product in the presence of **2** and Mn^{2+}.

3.5 Literature Cited

(1) Breton, C.; Mucha, J.; Jeanneau, C. *Biochimie* **2001**, *83*, 713–718.

(2) Crocker, P. R.; Feize, T. *Curr. Opin. Struct. Biol*. **1996**, *6*, 679–691.

(3) Sacks, D.; Kamhawi, S. *Annu. Rev. Microbiol*. **2001**, 55, 453–483.

(4) Lowe, J. B. *Curr. Opin. Cell Biol*. **2003**, *15*, 531–538.

(5) Hakomori, S.; Kannagi, R. *J. Natl. Cancer Inst*. **1983**, *71*, 231–251.

(6) Laferte, S.; Dennis, J. W. *Cancer Res*. **1988**, 48, 4743–4748.

(7) Saitoh, O.; Wang, W. C.; Lotan, R.; Fukuda, M. *J. Biol. Chem*. **1992**, *267*, 5700–5711.

(8) Kim, Y. J.; Varki, A. *Glycoconj. J*. **1997**, *14*, 569–576.

(9) Couldrey, C.; Green, J. E. *Breast. Cancer Res*. **2000**, *2*, 321–323.

(10) Hakomori, S. *Proc. Natl. Acad. Sci. USA*. **2002**, *99*, 10231–10233.

(11) Hounsell, E. F.; Feizi, T. *Med. Biol.* **1982**, *60*, 227–236.

(12) Furukawa, K.; Okajima, T. *Biochim. Biophys. Acta*. **2002**, *1573*, 377–381.

(13) Hanßke, B.; Thiel, C.; Lübke, T.; Hasilik, M.; Höning, S.; Peters, V.; Heidemann, P. H.; Hoffmann, G. F.; Berger, E. G.; Figura, K. V.; Körner, C. *J. Clin. Invest*. **2002**, *109*, 725–733.

(14) Kleene, R.; Berger, E. G. *Biochim. Biophys. Acta*. **1993**, *1154*, 283–325.

(15) Lazar, K.; Walker, S. *Curr. Opin. Chem. Biol*. **2002**, *6*, 786–793.

(16) Yamamoto, F.; Clausen, H.; White, T.; Marken, J.; Hakomori, S. *Nature* **1990**, *345*, 229–233.

(17) Patenaude, S. I.; Seto, N. O. L.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol*. **2002**, *9*, 685–690.

(18) Nguyen, H. P.; Seto, N. O. L.; Cai, Y.; Leinala, E. K.; Borisova, S. N.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2003**, *278*, 49191–49195.

(19) Lee, H. J.; Barry, C. H.; Borisova, S. N.; Seto, N. O. L.; Zheng, R. X. B.; Blancher, A.; Evans, S. V.; Palcic, M. M. *J. Biol. Chem*. **2005**, *280*, 525–529.

(20) Letts, J. A.; Rose, N. L.; Fang, Y. R.; Barry, C. H.; Borisova, S. N.; Seto, N. O. L.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2006**, *281*, 3625–3632.

(21) Alfaro, J. A.; Zheng, R. B.; Persson, M.; Letts, J. A.; Polakowski, R.; Bai, Y.; Borisova, S. N.; Seto, N. O. L.; Lowary, T.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2008**, *283*, 10097–10108.

(22) Seto, N. O. L.; Palcic, M. M.; Compston, C. A.; Li, H.; Bundle, D. R.; Narang, S. A. *J. Biol. Chem*. **1997**, *272*, 14133–14138.

(23) Marcus, S. L.; Polakowski, R.; Seto, N. O. L.; Leinala, E.; Borisova, S.; Blancher, A.; Roubinet, F.; Evans, S. V.; Palcic, M. M. *J. Biol. Chem*. **2003**, *278*, 12403–12405.

(24) Angulo, J.; Langpap, B.; Blume, A.; Biet, T.; Meyer, B.; Krishna, N. R.; Peters, H.; Palcic, M. M.; Peters, T. *J. Am. Chem. Soc*. **2006**, *128*, 13529–13538.

(25) Blume, A.; Angulo, J.; Biet, T.; Peters, H.; Benie, A. J.; Palcic, M. M.; Peters, T. *J. Biol. Chem*. **2006**, *281*, 32728–32740.

(26) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom*. **2002**, *216*, 1–27.

(27) Wang, W. J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2003**, *75*, 4945–4955.

(28) Yu, Y.; Sweeney, M. D.; Saad, O. M.; Leary, J. A. *J. Am. Soc. Mass Spectrom*. **2006**, *17*, 524–535.

(29) Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**, *18*, 587–592.

(30) Seto, N. O. L.; Palcic, M. M.; Hindsgaul, O.; Bundle, D. R.; Narang, S. A. *Eur. J. Biochem*. **1995**, *234*, 323–328.

(31) 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. USA*. **1994**, *91*, 6423–6427.

(32) Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc*. **2004**, *126*, 4786–4787.

(33) Wang, W. J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2005**, *77*, 3060−3071.

(34) Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S. *Anal. Chem*. **2006**, *78*, 3010−3018.

(35) Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472−481.

(36) Seko, A. *Trends Glycosci. Glycotechnol*. **2006**, *18*, 209−230.

(37) Nagai, M.; Davè, V.; Kaplan, B. E.; Yoshida, A. *J. Biol. Chem*. **1978**, *253*, 377−379.

(38) Nagai, M.; Davè, V.; Muensch, H.; Yoshida, A. *J. Biol. Chem*. **1978**, *253*, 380−381.

(39) Loo, J. A. *Mass Spectrom. Rev*. **1997**, *16*, 1−23.

(40) Sobott, F.; Hernandez, H.; McCammon, M. G.; Tito, M. A.; Robinson, C. V. *Anal. Chem*. **2002**, *74*, 1402−1407.

(41) Heck, A. J. R.; Van Den Heuvel, R. H. H. *Mass Spectrom. Rev*. **2004**, *23*, 368−389.

(42) Lafitte, D.; Heck, A. J. R.; Hill, T. J.; Jumel, K.; Harding, S. E.; Derrick, P. J. *Eur. J. Biochem*. **1999**, *261*, 337−344.

(43) Lowary, T. L.; Hindsgaul, O. *Carbohydr. Res*. **1993**, *249*, 163−195.

(44) Yamamoto, F.; Hakomori, S. *J. Biol. Chem*. **1990**, *265*, 19257−19262.

(45) Yamamoto, F.; McNeill, P. D. *J. Biol. Chem*. **1996**, *271*, 10515−10520.

(46) Seto, N. O. L.; Compston, C. A.; Evans, S. V.; Bundle, D. R.; Narang, S. A.; Palcic, M. M. *Eur. J. Biochem*. **1999**, *259*, 770−775.

(47) Seto. N. O. L.; Compston, C. A.; Szpacenko, A.; Palcic, M. M. *Carbohydr. Res*. **2000**, *324*, 161−169.

Chapter 4

Trapping and Characterization of Covalent Intermediates of Mutant Retaining Glycosyltransferases[†]

4.1 Introduction

Glycosyltransferases (GTs) are a large family of enzymes that are encoded by about 1–2% of the coding regions of all genomes sequenced to date. These enzymes catalyze the biosynthesis of glycosidic linkages by the stereo- and regiospecific transfer of monosaccharides from donors, typically nucleotide sugars, to diverse acceptors such as saccharides, lipids, proteins, DNA and natural products.[1] GTs biosynthesize highly abundant biopolymers such as cellulose,[2] amylose[3] and chitin,[4] as well as therapeutically important glycosylated natural product antibiotics[5,6] and anti-cancer agents.[7,8] They play important roles in a variety of normal biological processes,[9] and are implicated in many diseases and infections.[10,11] Therefore, GTs are attractive drug targets, and a detailed understanding of their mechanisms aids in the design of new therapeutic inhibitors.[12]

GTs are classified as either inverting or retaining, depending on whether the anomeric configuration of the transferred monosaccharide in the product is

[†]A version of this chapter has been published: Soya, N.; Fang, Y.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2011**, *21*, 547–552.

inverted or remains the same as that in the donor substrate. It is generally accepted that inverting GTs transfer monosaccharides from donor to acceptor, in a single step S_N2-like reaction, aided by an enzymatic base catalyst.[13,14] In contrast, the mechanism by which retaining GTs transfer monosaccharides with net retention of configuration has resisted elucidation, despite extensive experimental and computational studies. A double displacement mechanism, in which the monosaccharide is first transferred to the enzyme forming a glycosyl-enzyme intermediate and subsequently transferred to the acceptor (Figure 4.1a), has been suggested by analogy to the well characterized retaining glycosidases.[13–15] This mechanism is supported by the chemical rescue of an inactive α-(1→3)-galactosyltransferase mutant, in which the putative enzymatic nucleophile Glu^{317} was replaced by Ala.[16] However, conclusive evidence for the formation of covalent glycosyl-enzyme intermediates has been elusive.[13–17] An alternative mechanism, referred to as S_Ni, proceeds via a front side single displacement forming an oxocarbenium ion-like transition state (Figure 4.1b).[13–15,18–20] While such a mechanism can explain the retention of stereochemistry at the anomeric centre without invoking a glycosyl-enzyme intermediate, the formation of this transition state would be sterically and entropically unfavorable. Recently, a new mechanism, referred to as S_Ni-like, involving the formation of a short-lived ion pair intermediate has been suggested.[14,21] The intermediate undergoes a small shift in position in the enzyme active site allowing the acceptor to attack the donor on the same face that the nucleotide departs from (Figure 4.1c).

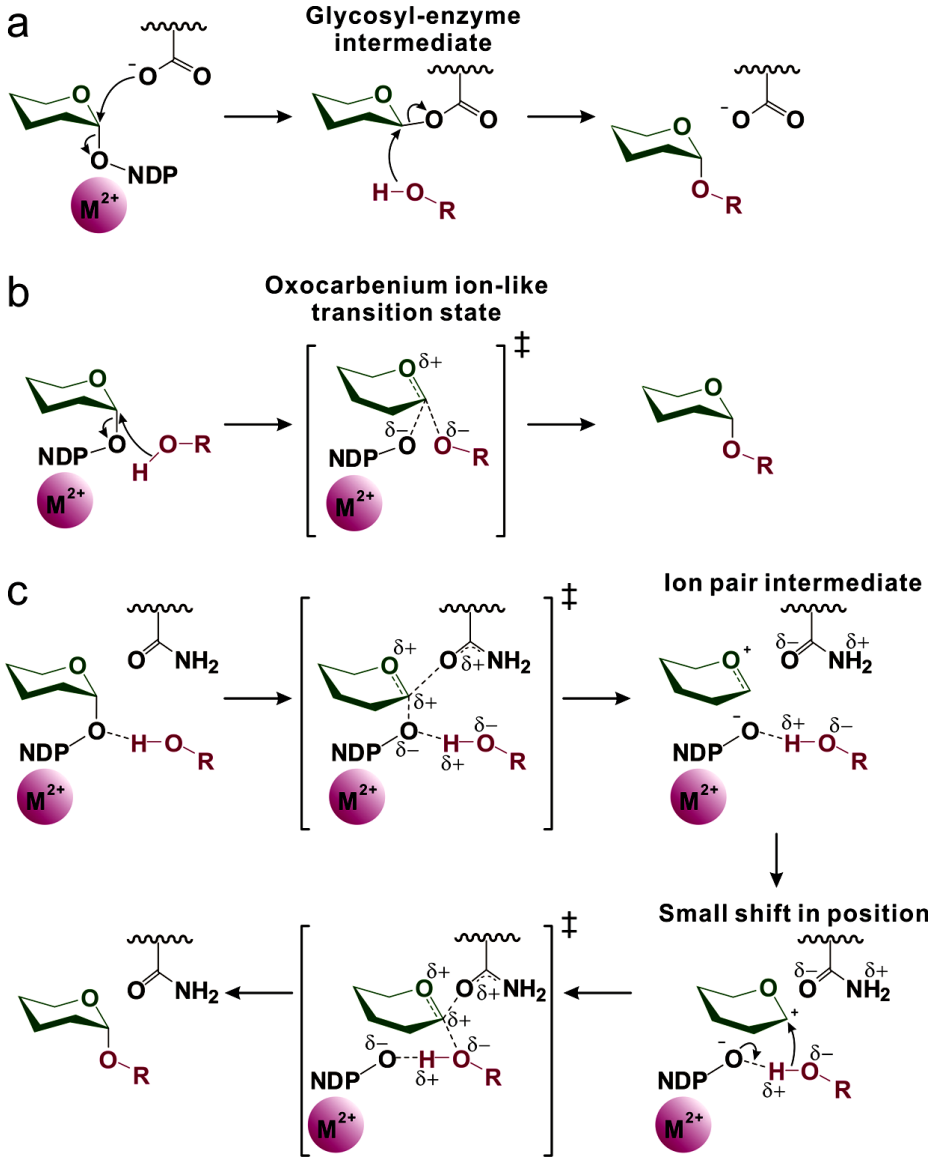

Figure 4.1 Previously proposed enzymatic mechanisms for retaining glycosyltransferases. (a) Double displacement mechanism which occurs via a covalent glycosyl-enzyme intermediate.[15] (b) S_Ni mechanism involving an oxocarbenium ion-like transition state.[15] (c) S_Ni-like mechanism proceeding with the formation of short-lived ion pair intermediate.[14]

ABO is the most important blood group system in transplantation and transfusion medicine. The A and B antigenic determinants are cell surface oligosaccharides biosynthesized by the glycosyltransferases, α-(1→3)-N-acetylgalactosaminyltransferase (GTA; EC 2.4.1.40) and α-(1→3)-galactosyltransferase (GTB; EC 2.4.1.37). GTA and GTB are highly homologous, differing in only four of 354 amino acids (Arg/Gly 176, Gly/Ser 235, Leu/Met 266, and Gly/Ala 268), but they possess distinct donor specificities.[22] GTA catalyzes the transfer of N-acetylgalactosamine (GalNAc) from uridine 5′-diphospho-N-acetylgalactosamine (UDP-GalNAc, **9**, Figure 3.1) to α-L-Fuc*p*-(1→2)-β-D-Gal*p*-OR (H-disaccharide antigen, where R may be a protein or lipid) acceptor resulting in A-trisaccharide, and GTB catalyzes the transfer of galactose (Gal) from uridine 5′-diphosphogalactose (UDP-Gal, **10**, Figure 3.1) to α-L-Fuc*p*-(1→2)-β-D-Gal*p*-OR resulting in the B-trisaccharide antigen (Figure 1.1).

The structures, activities and binding properties of recombinant soluble forms of these enzymes have been extensively studied.[23–28] GTA and GTB are believed to undergo extensive structural rearrangement during a catalytic cycle.[26,28] Moreover, UDP has been suggested to have a role in deprotonation of the 3-OH of the acceptor activating it for nucleophilic attack.[14] When conformational changes and/or substrate contacts are impaired the catalytic turnover drops dramatically, although these reactions occur with retention of stereochemistry.[26,29,30] In the absence of acceptor, GTA and GTB also catalyze the slow hydrolysis of donor, which can be considered as transfer to water rather than to acceptor, with retention of configuration.[31] The slow rate of hydrolysis is

attributed to the inability of the enzyme to form all contacts in the closed conformation.

Based on an analysis of X-ray crystal structures and the results of molecular modeling, Patenaude et al. proposed that, in both GTA and GTB, Glu^{303} could serve as the catalytic nucleophile for a double displacement mechanism.[23] A 30,000-fold decrease in the enzymatic activity of GTB upon replacement of Glu^{303} with Ala provides additional support for the proposal that this position plays a critical role in enzymatic activity.[23] Here, to test the hypothesis that residue Glu^{303} serves as the catalytic nucleophile in both GTA and GTB, the enzymatic properties of single point mutants, in which Glu^{303} is replaced with Cys (i.e., GTA_{E303C} and GTB_{E303C}) to trap the covalent glycosyl-enzyme intermediates, were investigated using by electrospray ionization mass spectrometry (ES-MS) measurements. Direct experimental evidence for the formation of a covalent glycosyl-enzyme intermediate by GTA_{E303C} and GTB_{E303C} is reported. Using tandem MS, the site of glycosylation was identified in both mutants. Evidence that exposure of the glycosyl-enzyme intermediates to a disaccharide acceptor results in the formation of the corresponding enzymatic trisaccharide products was also obtained. These findings suggest that GTA_{E303C} and GTB_{E303C} may operate by a double displacement mechanism.

4.2 Experimental

4.2.1 Chemicals and Reagents

The donor substrates UDP-GalNAc (**9**), UDP-Gal (**10**), iodoacetamide and ammonium bicarbonate were purchased from Sigma-Aldrich (Oakville, Canada). Ammonium acetate was purchased from Caledon Laboratory Chemicals (Georgetown, Canada). Dithiothreitol (DTT), urea, formic acid, HPLC grade water, HPLC grade methanol (MeOH) and HPLC grade acetonitrile (ACN) were purchased from Fisher Scientific (Ottawa, Canada). Sequencing grade modified trypsin was purchased from Promega (Madison, WI). The H-disaccharide acceptor (α-L-Fuc*p*-(1→2)-β-D-Gal*p*-O-$(CH_2)_7CH_3$) used in this study was donated by Dr. Ole Hindsgaul (Carlsberg Laboratory).

4.2.2 Generation of Mutants

Site-directed mutagenesis was carried out by the two fragment method as previously described[24] using wild-type GTA and GTB (aa 63 – 354) plasmid DNA as a template in the plasmid vector pCWΔlac. The primers used for mutagenesis, with codon substitutions in bold, were as follows; forward primer 5′-TGG CAC GAC **TGC** TCC CAC CTG AAC AAA TAC CTG-3′, reverse primer 5′-CAG GTG GGA **GCA** GTC GTG CCA AAC AGC TTC GAT AC-3′. *E.coli* BL21 Gold Cells were used for expression.

4.2.3 Enzyme Preparation and Characterization

Recombinant soluble fragments of GTA_{E303C} and GTB_{E303C} mutants (MW 34,494 Da and 34,457 Da, respectively) composed of the catalytic domain, as well as a part of the stem region were over-expressed in *E. coli* BL21 cells and purified

using the same procedures as described previously for wild-type GTs.[24] Kinetic studies were also performed as described.[24] The enzymatic synthesis of trisaccharide products by the GT_{E303C} mutants and the characterization of products were carried out using procedures described previously.[24] To facilitate ES-MS analysis, the enzyme solutions were exchanged into a 50 mM ammonium acetate (pH 7) using an Amicon ultracentrifugation filter (Millipore, Billerica, MA) with a MWCO of 10 kDa. Protein concentrations were determined by lyophilizing an aliquot of the stock solution and weighing the mass of the corresponding protein pellet. Dialyzed GTA_{E303C} and GTB_{E303C} were stored −80 °C until used.

To trap the covalent GTA_{E303C}→GalNAc intermediate, GTA_{E303C} (80 μM) was incubated with **9** (100 μM) in the presence of excess Mn^{2+} (10 mM) in 50 mM MOPS buffer (pH 7) at 37 °C for 2 h to produce GTA_{E303C} covalently bound to GalNAc (i.e. GTA_{E303C}→GalNAc). The extent of conversion of GTA_{E303C} to GTA_{E303C}→GalNAc was confirmed by ES-MS. Before ES-MS analysis, the solutions were exchanged 3 times with 50 mM ammonium acetate (pH 7) in the presence of 5 mM EDTA, to remove excess Mn^{2+}, and then 5 times with 50 mM ammonium acetate (pH 7) using Amicon Ultra-0.5 centrifugal filter (Millipore, Billerica, MA) with MWCO of 10 kDa. The dialyzed solution was acidified using 10% acetic acid to adjust the pH to 3.3 prior to ES-MS analysis.

Once GTA_{E303C} was fully converted to GTA_{E303C}→GalNAc, the solution was dialyzed with 50 mM MOPS buffer (pH 7) using Amicon Ultra-0.5 centrifugal filter with MWCO of 10 kDa to remove unreacted **9** and UDP (**2**). The H-antigen acceptor (20 mM) was added to the GTA_{E303C}→GalNAc solution, and

the mixture was incubated at 37 °C for varying lengths of time prior to ES-MS analysis. ES-MS analysis was carried out as described above.

The formation of A-trisaccharide was confirmed in a separate experiment. GTA_{E303C} (180 μM) was incubated with **9** (200 μM) in the presence of excess Mn^{2+} (10 mM) in 50 mM MOPS buffer (pH 7) at 37 °C for 2.5 h. Following the conversion of GTA_{E303C} to GTA_{E303C}→GalNAc, the solution was dialyzed with 50 mM MOPS buffer (pH 7) using Amicon Ultra-0.5 centrifugal filter with MWCO of 10 kDa to remove unreacted **2** and **9**. The H-antigen acceptor (500 μM) was added to the purified GTA_{E303C}→GalNAc solution, and the mixture was incubated at 37 °C for 48 h. The A-trisaccharide product (α-D-GalNAc*p*-(1→3)-[α-L-Fuc*p*-(1→2)]-β-D-Gal*p*-O-$(CH_2)_7CH_3$) was isolated using C18 Zip-tips (Millipore, Billerica, MA). The reversed-phase bed was washed 5 times with 10 μL of 75% MeOH, equilibrated 10 times with 10 μL of water prior to applying the reaction mixture. The bed was washed 10 times with 10 μL of water, prior to eluting the bound trisaccharide with 10 μL of MeOH. The MeOH was removed by evaporation using a vacuum centrifuge, and dried product was stored in –20 °C until used. The dried product was dissolved in 5 mM ammonium acetate (pH 7) before ES-MS analysis.

4.2.4 In-solution Protein Digestion

Glycosylated GT_{E303C} mutants were digested using trypsin with or without alkylation (carbamidomethylation) of cysteine residues. For alkylation, GT_{E303C} mutants were dissolved in 100 μL of 6 M urea, 100 mM Tris-HCl buffer (pH 8),

reduced with 10 mM DTT for 1 h at room temperature, alkylated with 50 mM iodoacetamide for 1 h in the dark at room temperature. The reaction mixture was dialyzed against 100 mM NH_4HCO_3 buffer (pH 8) using an Amicon ultracentrifugation filter (Millipore, Billerica, MA) with a 5 kDa MWCO to remove excess reagents and was subsequently dried in a vacuum centrifuge.

The GT_{E303C} samples were dissolved in 100 mM NH_4HCO_3 buffer (pH 8) (the GT_{E303C} final concentration was between 1–10 μM) and digested for 4 h at 37 °C using modified trypsin at 1:50 enzyme-to-protein molar ratio. The reaction was quenched by adding formic acid (pH 3).

4.2.5 Fourier-transform Ion Cyclotron Resonance Mass Spectrometry

ES-MS measurements were performed using an Apex II 9.4 tesla Fourier-transform ion cyclotron resonance mass spectrometer (FTICR-MS) (Bruker, Billerica, MA). The ES solutions were prepared from aqueous stock solutions of proteins and substrates with known concentration. Unless otherwise indicated, aqueous ammonium acetate was added to the ES solution to yield a final buffer concentration of 10 mM. ES was performed using borosilicate tubing (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in positive ion mode was established by applying a voltage of ~1000 V to a platinum wire inserted inside the glass tip. The solution flow rate was typically ~20 nL/min. The droplets and gaseous ions produced by ES were introduced into the mass spectrometer through a stainless steel capillary (i.d. 0.43 mm)

maintained at an external temperature of 66 °C. The ions sampled by the capillary (+50 V) was transmitted through a skimmer (–2 V) and stored electrodynamically in an rf hexapole for 4–8 s. Ions were ejected from the hexapole and accelerated to –2700 V into the superconducting magnet, decelerated, and introduced into the ion cell. The trapping plates of the cell were maintained at 1.6 V throughout the experiment. The typical base pressure for the instrument was ~5 x 10^{-10} mbar. Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. The time domain signal, consisting of the sum of 50–100 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier transformation.

4.2.6 Ultra-performance Liquid Chromatography Tandem Mass Spectrometry

The tryptic digests were analyzed using a hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer, Q-TOF Premier (Waters, Milford, MA), equipped with nanoACQUITY Ultra-performance liquid chromatography (UPLC) system (Waters, Milford, MA). Briefly, 2 µL of the peptide solution was injected into a microprecolumn C_{18} cartridge that was connected to an Atlantis dC18 column, 75 µm × 150 mm, 3 µm particle size (Waters, Milford, MA). Solvent A consisted of 0.1% formic acid in water, and solvent B consisted of 0.1% formic acid in ACN. Tryptic peptides were separated using a 90 min gradient (1–5% solvent B for 5 min, 5–95% solvent B for 60 min, 95–2% solvent

B for 25 min) and electrosprayed into the mass spectrometer at a flow rate 350 nL/min. Real-time mass correction in MS and MS/MS mode was carried out using NanoLockSpray^{TM} (Waters, Milford, MA). Argon gas was used for the collision-induced dissociation (CID) experiments. The data analysis was carried out using MassLynx version 4.1 (Waters, Milford, MA).

4.3 Results and Discussion

As was found by ES-MS for wild-type GTA and GTB (i.e., GTA_{WT} and GTB_{WT}),[27] the GTA_{E303C} and GTB_{E303C} mutants exist preferentially as homodimers in aqueous solution at neutral pH and 37 °C (Figure 4.2). Incubation of GTA_{WT} or GTB_{WT} in the presence of their respective donor substrates (**9** or **10**) and excess Mn^{2+}, which serves as the metal ion cofactor, at 37 °C and pH 7 for 30 min results in the enzyme-catalyzed hydrolysis of the donor giving uridine 5′-diphosphate (UDP, **2**) and the corresponding monosaccharide (GalNAc or Gal) (Figure 4.3a and 4.4a).[27] For GTB_{WT}, the k_{cat} for UDP-Gal donor hydrolysis at 25 °C is 2 h^{-1} and the reaction occurs with retention of configuration of the released α-D-galactose.[31] In the case of GT_{E303C}, ions corresponding to the free GT_{E303C} homodimer and the GT_{E303C} homodimer bound with donors (**9** or **10** + Mn), or UDP, as (**2** + Mn), were identified under the same conditions as above (Figure 4.3b and 4.4b). More importantly, however, ions with mass-to-charge ratios consistent with those expected for the GT_{E303C} homodimers covalently bound to a monosaccharide (i.e., $(GTA_{E303C})_2$→GalNAc or $(GTB_{E303C})_2$→Gal,

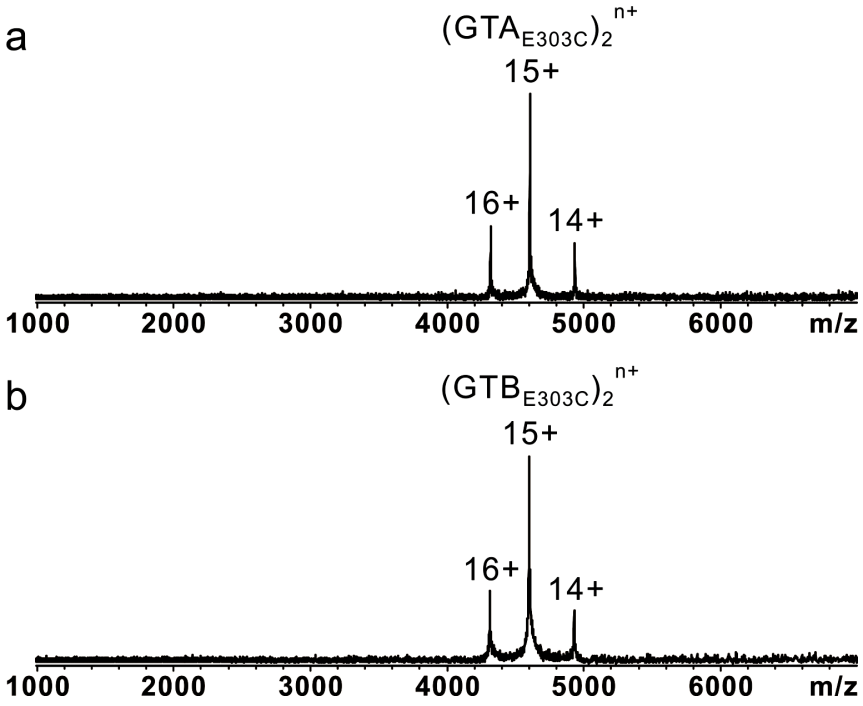

Figure 4.2 ES mass spectra of aqueous solutions consisting of (a) GTA_{E303C} (8 μM) or (b) GTB_{E303C} (8 μM) in 10 mM ammonium acetate (pH 7).

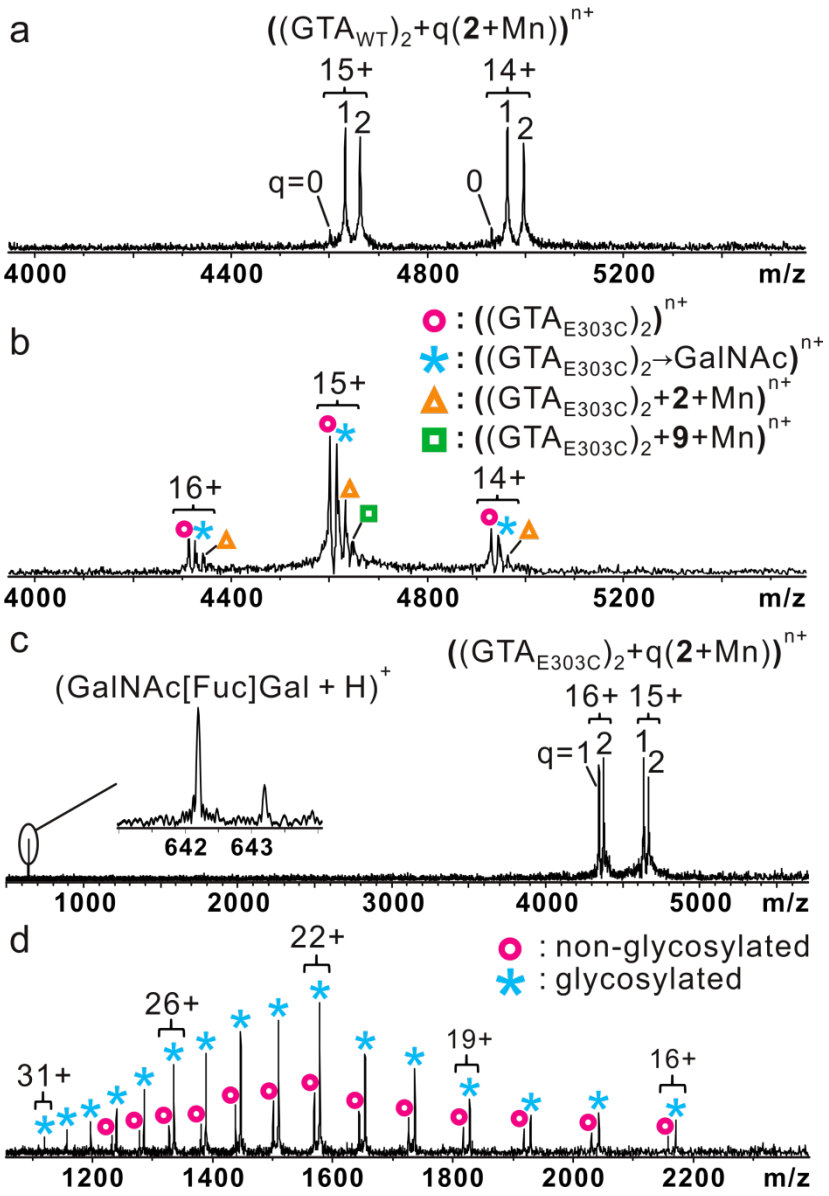

Figure 4.3 ES mass spectra. (a) GTA_{WT} (15 μM) or (b) GTA_{E303C} (15 μM), with **9** (25 μM) and Mn^{2+} (100 μM). Both solutions were incubated at 37 °C and pH 7 for 30 min. (c) GTA_{E303C} (15 μM), **9** (25 μM), acceptor (25 μM) and Mn^{2+} (100 μM), incubated at 37 °C and pH 7 for 4 h. (d) GTA_{E303C} (15 μM), **9** (200 μM) and Mn^{2+} (200 μM), acidified by 0.5% formic acid (pH 3) after the incubation at 37 °C and pH 7 for 4 h.

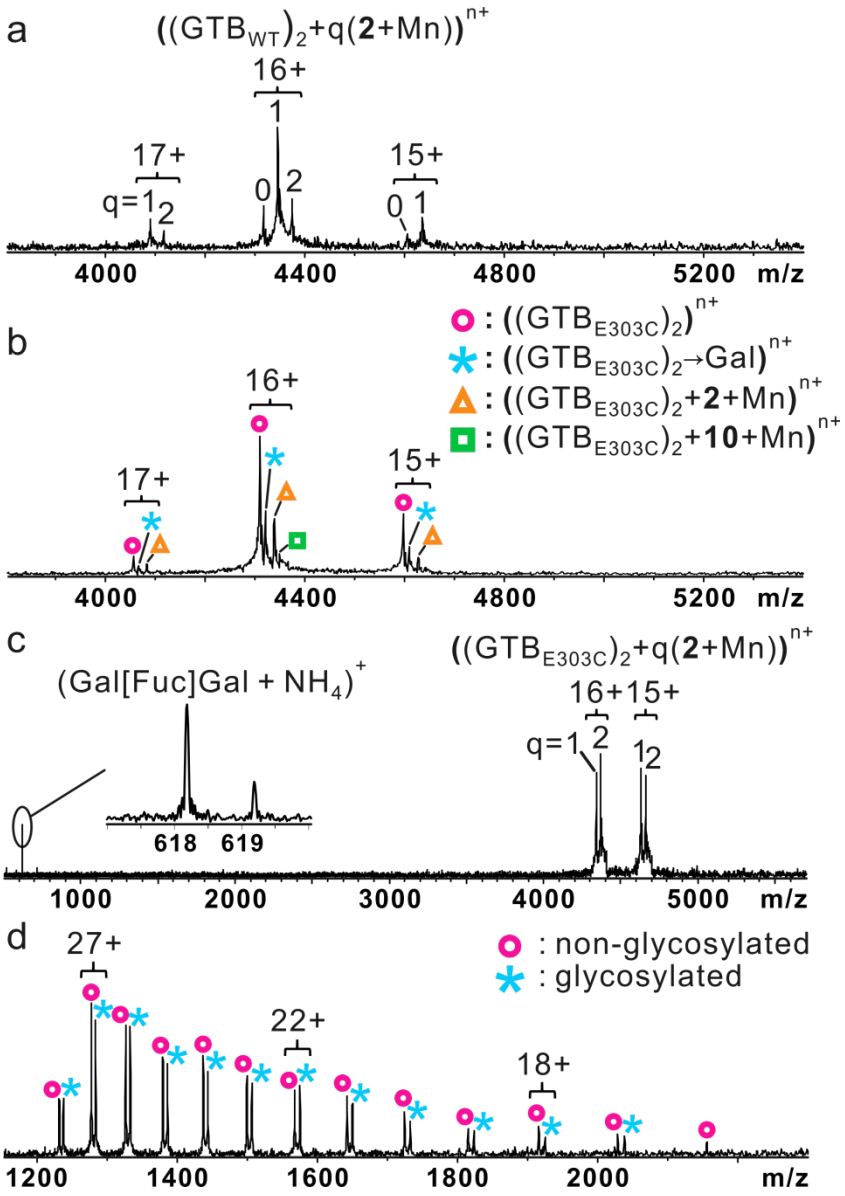

Figure 4.4 ES mass spectra. (a) GTB_{WT} (15 μM) or (b) GTB_{E303C} (15 μM), with **10** (25 μM) and Mn^{2+} (100 μM). Both solutions were incubated at 37 °C and pH 7 for 30 min. (c) GTB_{E303C} (15 μM), **10** (25 μM), acceptor (25 μM) and Mn^{2+} (100 μM), incubated at 37 ºC and pH 7 for 4 h. (d) GTB_{E303C} (15 μM), **10** (200 μM) and Mn^{2+} (200 μM) acidified by 0.5 % formic acid (pH 3) after the incubation at 37 ºC and pH 7 for 4 h.

respectively) were detected (Table 4.1). Similar results were obtained upon incubation of GTA_{E303C} with **10** and GTB_{E303C} with **9**, although significantly longer incubation times (~24 h) were required to produce detectable levels of the corresponding GTA_{E303C}→Gal or GTB_{E303C}→GalNAc (Figures 4.5). Furthermore, when both GT_{E303C} mutants were incubated in the presence of their respective donor, acceptor (α-L-Fuc*p*-(1→2)-β-D-Gal*p*-O$(CH_2)CH_3$) and excess Mn^{2+}, ion signal for the glycosyl-enzyme intermediates was no longer detected by ES-MS (Figure 4.3c and 4.4c). Instead, ions corresponding to the enzymatic products, the A- and B-trisaccharides and **2**, which binds as (**2** + Mn) to the GT_{E303C} dimers, were detected. The k_{cat} values for the formation of products for GTA_{E303C} and GTB_{E303C} are 0.9 s^{-1} and 1.3 s^{-1}, respectively, compared to 17 s^{-1} and 5 s^{-1} for the corresponding wild-type enzymes.[24]

Using a higher concentration of Mn^{2+} (10 mM), the GTA_{E303C} dimer was shown to be fully converted to intermediate, $(GTA_{E303C})_2$→2GalNAc, within 2 h, comparable to the rate of donor hydrolysis by GTB_{WT} (Figure 4.6). UDP is not present and the conformation of this intermediate is unknown. Importantly, exposure of the purified intermediate to acceptor resulted in the slow regeneration of free enzyme (Figures 4.6b–d) and the formation of the trisaccharide product (Figure 4.6f).

Taken together, these results demonstrate that monosaccharides are indeed transferred from the covalent GTA_{E303C}→GalNAc species to disaccharide acceptors, producing trisaccharide products. Moreover, this result suggests that the covalent GTA_{E303C}→GalNAc and GTB_{E303C}→Gal species is catalytically

Table 4.1 The theoretical and measured (by ES-MS) molecular masses of the GT_{WT} and GT_{E303C} species.

Enzyme	Theoretical mass (Da)	Measured mass (Da)
$(GTA_{WT})_2$	69,040	69,039
$(GTA_{WT})_2$ + **2** + Mn	69,497	69,500
$(GTA_{WT})_2$ + 2(**2** + Mn)	69,954	69,959
GTA_{E303C}	34,494	34,493
GTA_{E303C}→GalNAc	34,697	34,694
GTA_{E303C}→Gal	34,656	34,652
$(GTA_{E303C})_2$	68,988	68,987
$(GTA_{E303C})_2$→GalNAc	69,191	69,195
$(GTA_{E303C})_2$ + **2** + Mn	69,445	69,447
$(GTA_{E303C})_2$ + 2(**2** + Mn)	69,902	69,906
$(GTA_{E303C})_2$ + **9** + Mn	69,648	69,647
$(GTB_{WT})_2$	68,966	68,963
$(GTB_{WT})_2$ + **2** + Mn	69,423	69,427
$(GTB_{WT})_2$ + 2(**2** + Mn)	69,880	69,876
GTB_{E303C}	34,457	34,453
GTB_{E303C}→GalNAc	34,660	34,658
GTB_{E303C}→Gal	34,619	34,614
$(GTB_{E303C})_2$	68,914	68,917
$(GTB_{E303C})_2$→Gal	69,076	69,079
$(GTB_{E303C})_2$ + **2** + Mn	69,371	69,375
$(GTB_{E303C})_2$ + 2(**2** + Mn)	69,828	69,828
$(GTB_{E303C})_2$ + **10** + Mn	69,533	69,535

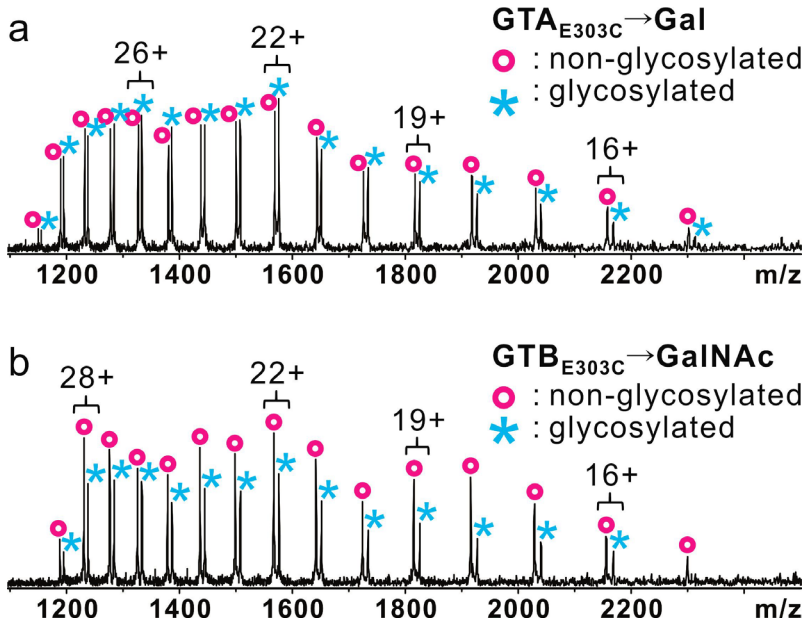

Figure 4.5 ES mass spectra of the monomer forms of glycosyl-GT_{E303C} intermediate. (a) GTA_{E303C} (15 μM), **10** (200 μM) and Mn^{2+} (200 μM), (b) GTB_{E303C} (15 μM), **9** (200 μM) and Mn^{2+} (200 μM). All solutions were acidified with 0.5% formic acid (pH 3) after the incubation at 37 ºC and pH 7 for 24 h.

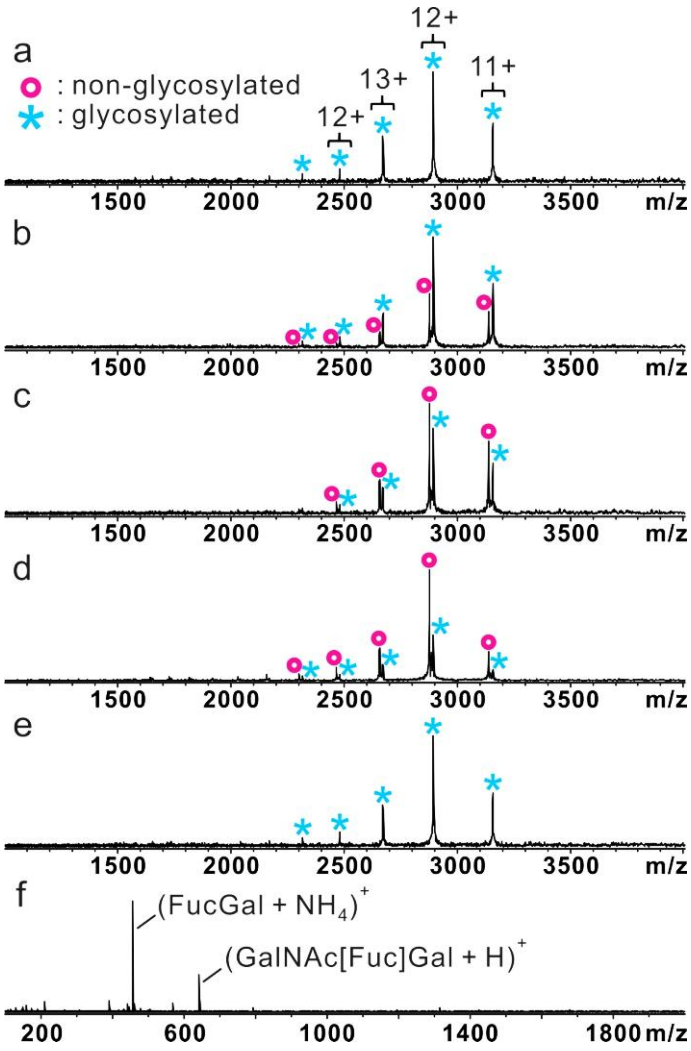

Figure 4.6 ES mass spectra. (a) GTA_{E303C} (80 μM) incubated with **9** (100 μM) and Mn^{2+} (10 mM) at 37 °C and pH 7 for 2 h. (b) Purified intermediate $(GTA_{E303C})_2 \rightarrow$2GalNAc incubated with disaccharide acceptor (20 mM) at 37 °C and pH 7 for 12 h, (c) for 24 h and (d) for 48 h. (e) Purified intermediate $(GTA_{E303C})_2 \rightarrow$2GalNAc incubated at 37 °C and pH 7 for 48 h in the absence of acceptor. All solutions were acidified to pH 3.3 prior to ES-MS analysis. (f) Enzymatically synthesized A-trisaccharide isolated from solution of purified $(GTA_{E303C})_2 \rightarrow$2GalNAc following incubation with acceptor (500 μM) at 37 °C and pH 7 for 48 h.

competent and could represent intermediates in the enzymatic reactions. The slow product formation is not surprising given the lack of UDP, which is believed to be necessary for deprotonation of acceptor. Moreover, in the absence of donor or UDP, the enzyme is unable to make critical contacts with acceptor and undergo the conformational changes required in a normal catalytic cycle.[26,28] Thus, while the timescales for product formation are lower than those of a normal catalytic cycle, and while therefore the relevance to normal catalysis might be questioned, they are not unexpected given the complexity of glycosyltransferase reactions.

To confirm that the glycosyl-enzyme intermediates were in fact covalent in nature, solutions of GT_{E303C} and their respective donor were acidified (to pH 3) with the addition of 0.5% formic acid following 4 h of incubation. Notably, ions corresponding to free monomeric GT_{E303C}, as well as the monomeric form of the glycosyl-enzyme intermediate were detected by ES-MS (Figure 4.3d and 4.4d). The absence of dimeric GT_{E303C} ions and the broad charge state distribution observed for the monomeric GT_{E303C} ions are consistent with the acid denaturation of the enzymes. These findings confirm that both GTA_{E303C} and GTB_{E303C} are indeed glycosylated in the presence of the respective donor.

To identify the site of glycosylation in GTA_{E303C}→GalNAc and GTB_{E303C}→Gal, the intermediates, along with the non-glycosylated mutants, were digested in solution with trypsin. The resulting glycopeptides were separated from the tryptic peptides by reversed-phase liquid chromatography and analyzed by tandem MS (MS/MS) performed using CID. Shown in Figure 4.7a is the deconvoluted CID mass spectrum measured for the triply protonated glycopeptide

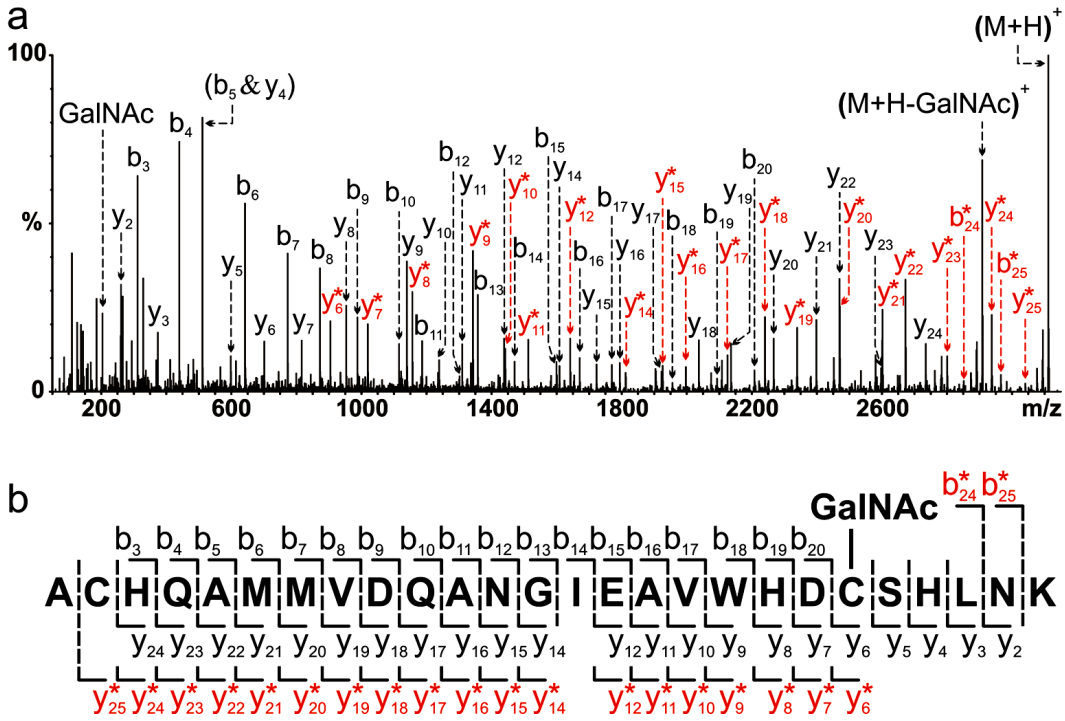

Figure 4.7 (a) Deconvoluted MS/MS spectrum of the glycosylated tryptic peptide ^{283}ACHQAMMVDQANGIEAVWHDCSHLNK^{308} from the GTA_{E303C}→GalNAc intermediate. Only b- and y-ions are assigned. Peaks labeled with * correspond to ions containing GalNAc. (b) Product map indicating the cleavage sites identified from the CID MS/MS analysis.

^{283}ACHQAMMVDQANGIEAVWHDCSHLNK^{308} (where Cys^{303} is underlined) obtained from the GTA_{E303C}→GalNAc intermediate. The mass spectrum reveals an extensive series of y-ions resulting from the fragmentation of the glycosylated peptide, as well as the deglycosylated peptide, wherein cleavage of the glycosidic bond accompanies peptide bond cleavage (Figure 4.7b and Table 4.2). Notably, the smallest y-ion identified by MS/MS for the glycosylated peptide is y_6, while for the non-glycosylated peptides, y_2–y_5 ions were also detected. Furthermore, MS/MS analysis of the glycopeptide following treatment with iodoacetamide revealed that carbamidomethylation occurred at Cys^{284} but not Cys^{303} (Figure 4.8), indicating that only Cys^{284} possesses a free thiol group in the glycopeptides. Together, these findings confirm Cys^{303} as the site of glycosylation in GTA_{E303C}. In a similar manner, Cys^{303} was established as the site of glycosylation in GTB_{E303C} (Figure 4.9, 4.10 and Table 4.2).

4.4 Conclusions

In summary, this study demonstrates that direct ES-MS measurements, combined with trypsin digestion followed by MS/MS analysis of the fragments, provide a straightforward and effective strategy for identifying and characterizing covalent enzyme intermediates. Using this approach, direct evidence for the formation of covalent glycosyl-enzyme intermediates in the enzymatic reactions of mutants of GTA and GTB, GTA_{E303C} and GTB_{E303C}, was obtained. Furthermore, the sites of glycosylation for the two enzymes were conclusively identified. The existence of covalent intermediates raises the possibility that these

Table 4.2 Summary of theoretical and measured m/z values for the b- and y-ions identified from CID of tryptic glycosylated peptides.

^{283}ACHQAMMVDQANGIEAVWHDC(GalNAc)SHLNK^{308}					
Assignment	**Theoretical m/z**	**Measured m/z**	**Assignment**	**Theoretical m/z**	**Measured m/z**
b_3	312.11	312.12	y_2	261.16	261.17
b_4	440.17	440.19	y_3	374.24	374.26
b_5	511.21	511.22	y_4	511.30	511.22
b_6	642.25	642.26	y_5	598.33	598.35
b_7	773.29	773.30	y_6	701.34	701.37
b_8	872.36	872.39	y_7	816.37	816.38
b_9	987.39	987.42	y_8	953.43	953.47
b_{10}	1115.44	1115.50	y_9	1139.51	1139.54
b_{11}	1186.48	1186.51	y_{10}	1238.57	1238.59
b_{12}	1300.52	1300.53	y_{11}	1309.61	1309.65
b_{13}	1357.55	1357.56	y_{12}	1438.65	1438.71
b_{14}	1470.63	1470.64	y_{14}	1608.76	1608.80
b_{15}	1599.67	1599.73	y_{15}	1722.80	1722.79
b_{16}	1670.71	1670.74	y_{16}	1793.84	1793.89
b_{17}	1769.78	1769.73	y_{17}	1921.90	1921.90
b_{18}	1955.86	1955.84	y_{18}	2036.92	2036.91
b_{19}	2092.92	2092.98	y_{19}	2135.99	2136.09
b_{20}	2207.94	2208.03	y_{20}	2267.03	2267.11
b_{24}^*	2851.21	2851.19	y_{21}	2398.07	2398.08

b_{25}^{*}	2965.25	2965.32	y_{22}	2469.11	2469.16
GalNAc	204.09	204.09	y_{23}	2597.17	2597.21
M+H–GalNAc	2908.27	2908.31	y_{24}	2734.23	2734.33
M+H	3111.35	3111.45	y_{6}^{*}	904.42	904.45
			y_{7}^{*}	1019.45	1019.48
			y_{8}^{*}	1156.51	1156.55
			y_{9}^{*}	1342.59	1342.61
			y_{10}^{*}	1441.65	1441.71
			y_{11}^{*}	1512.69	1512.75
			y_{12}^{*}	1641.73	1641.78
			y_{14}^{*}	1811.84	1811.83
			y_{15}^{*}	1925.88	1925.82
			y_{16}^{*}	1996.92	1997.02
			y_{17}^{*}	2124.98	2125.05
			y_{18}^{*}	2240.00	2240.09
			y_{19}^{*}	2339.07	2339.15
			y_{20}^{*}	2470.11	2470.27
			y_{21}^{*}	2601.15	2601.17
			y_{22}^{*}	2672.19	2672.21
			y_{23}^{*}	2800.25	2800.28
			y_{24}^{*}	2937.31	2937.36
			y_{25}^{*}	3040.32	3040.46

^{283}AC'HQAMMVDQANGIEAVWHDC(GalNAc)SHLNK^{308}					
Assignment	**Theoretical m/z**	**Measured m/z**	**Assignment**	**Theoretical m/z**	**Measured m/z**
b_2'	232.08	232.07	y_2	261.16	261.14
b_3'	369.13	369.11	y_3	374.24	374.22
b_4'	497.19	497.17	y_4	511.30	511.27
b_5'	568.23	568.21	y_5	598.33	598.30
b_6'	699.27	699.25	y_6	701.34	701.31
b_7'	830.31	830.27	y_7	816.37	816.32
b_8'	929.38	929.33	y_8	953.43	953.38
b_9'	1044.41	1044.35	y_9	1139.51	1139.44
b_{10}'	1172.46	1172.40	y_{10}	1238.57	1238.49
b_{11}'	1243.50	1243.46	y_{11}	1309.61	1309.52
b_{12}'	1357.54	1357.46	y_{12}	1438.65	1438.52
b_{13}'	1414.57	1414.50	y_{13}	1551.74	1551.62
b_{14}'	1527.65	1527.58	y_{14}	1608.76	1608.67
b_{15}'	1656.69	1656.58	y_{15}	1722.80	1722.70
b_{16}'	1727.73	1727.67	y_{17}	1921.90	1921.77
b_{19}'	2149.94	2149.80	y_{18}	2036.92	2036.76
b_{24}^{*},	2908.23	2908.11	y_{19}	2135.99	2135.86
b_{25}^{*},	3022.27	3022.08	y_{20}	2267.03	2266.86
GalNAc	204.09	204.08	y_{21}	2398.07	2397.98
M+H–GalNAc	2965.30	2965.17	y_{22}	2469.11	2468.96

M+H	3168.38	3168.24	y_{23}	2597.17	2597.02
			y_{24}	2734.23	2734.06
			y_6^*	904.42	904.38
			y_7^*	1019.45	1019.41
			y_8^*	1156.51	1156.45
			y_9^*	1342.59	1342.50
			y_{10}^*	1441.65	1441.55
			y_{11}^*	1512.69	1512.58
			y_{12}^*	1641.73	1641.60
			y_{14}^*	1811.84	1811.67
			y_{15}^*	1925.88	1925.70
			y_{16}^*	1996.92	1996.81
			y_{17}^*	2124.98	2124.86
			y_{18}^*	2240.00	2239.85
			y_{19}^*	2339.07	2338.91
			y_{20}^*	2470.11	2470.07
			y_{21}^*	2601.15	2601.05
			y_{22}^*	2672.19	2672.00
			y_{23}^*	2800.25	2800.06
			y_{24}^*	2937.31	2937.15
			$y_{25}^{*,}$	3097.34	3097.21

^{283}ACHQAMMVDQANGIEAVWHDC(Gal)SHLNK^{308}					
Assignment	**Theoretical m/z**	**Measured m/z**	**Assignment**	**Theoretical m/z**	**Measured m/z**
b_3	312.11	312.12	y_1	147.11	147.12
b_4	440.17	440.19	y_2	261.16	261.17
b_5	511.21	511.23	y_3	374.24	374.25
b_6	642.25	642.27	y_4	511.30	511.23
b_7	773.29	773.32	y_5	598.33	598.35
b_8	872.36	872.41	y_6	701.34	701.39
b_9	987.39	987.42	y_7	816.37	816.42
b_{10}	1115.44	1115.51	y_8	953.43	953.47
b_{11}	1186.48	1186.54	y_9	1139.51	1139.57
b_{12}	1300.52	1300.54	y_{10}	1238.57	1238.59
b_{13}	1357.55	1357.57	y_{11}	1309.61	1309.65
b_{14}	1470.63	1470.73	y_{12}	1438.65	1438.70
b_{15}	1599.67	1599.82	y_{14}	1608.76	1608.82
b_{16}	1670.71	1670.73	y_{15}	1722.80	1722.81
b_{17}	1769.78	1769.81	y_{16}	1793.84	1793.82
b_{18}	1955.86	1956.00	y_{17}	1921.90	1921.97
b_{19}	2092.92	2093.02	y_{18}	2036.92	2037.01
b_{20}	2207.94	2208.01	y_{19}	2135.99	2135.99
b_{24}^*	2810.18	2810.37	y_{20}	2267.03	2267.05
b_{25}^*	2924.22	2924.32	y_{21}	2398.07	2398.09

M+H–Gal	2908.27	2908.36	y_{22}	2469.11	2469.17
M+H	3070.32	3070.44	y_{23}	2597.17	2597.38
			y_{24}	2734.23	2734.22
			y_6^*	863.39	863.45
			y_7^*	978.42	978.44
			y_8^*	1115.48	1115.51
			y_9^*	1301.56	1301.62
			y_{10}^*	1400.63	1400.69
			y_{11}^*	1471.66	1471.68
			y_{12}^*	1600.71	1600.74
			y_{14}^*	1770.81	1770.72
			y_{15}^*	1884.86	1884.83
			y_{16}^*	1955.89	1956.00
			y_{17}^*	2083.95	2083.99
			y_{18}^*	2198.98	2199.12
			y_{19}^*	2298.05	2298.18
			y_{20}^*	2429.09	2429.20
			y_{21}^*	2560.13	2560.16
			y_{22}^*	2631.16	2631.27
			y_{23}^*	2759.22	2759.32
			y_{24}^*	2896.28	2896.34
			y_{25}^*	2999.29	2999.51

^{283}AC'HQAMMVDQANGIEAVWHDC(Gal)SHLNK^{308}					
Assignment	**Theoretical m/z**	**Measured m/z**	**Assignment**	**Theoretical m/z**	**Measured m/z**
b_2'	232.08	232.07	y_1	147.11	147.11
b_3'	369.13	369.12	y_2	261.16	261.14
b_4'	497.19	497.17	y_3	374.24	374.23
b_5'	568.23	568.20	y_4	511.30	511.29
b_6'	699.27	699.23	y_5	598.33	598.31
b_7'	830.31	830.28	y_6	701.34	701.33
b_8'	929.38	929.35	y_7	816.37	816.34
b_9'	1044.41	1044.35	y_8	953.43	953.38
b_{10}'	1172.46	1172.43	y_9	1139.51	1139.49
b_{11}'	1243.50	1243.44	y_{10}	1238.57	1238.53
b_{12}'	1357.54	1357.48	y_{11}	1309.61	1309.52
b_{13}'	1414.57	1414.50	y_{12}	1438.65	1438.52
b_{14}'	1527.65	1527.62	y_{14}	1608.76	1608.67
b_{15}'	1656.69	1656.60	y_{15}	1722.80	1722.73
b_{16}'	1727.73	1727.58	y_{17}	1921.90	1921.85
b_{17}'	1826.80	1826.68	y_{18}	2036.92	2036.87
b_{19}'	2149.94	2149.81	y_{19}	2135.99	2135.91
b_{24}^{*}'	2867.20	2867.10	y_{20}	2267.03	2266.92
b_{25}^{*}'	2981.24	2981.08	y_{21}	2398.07	2397.97
M+H–Gal	2965.30	2965.22	y_{22}	2469.11	2469.06

M+H	3127.35	3127.23	y_{23}	2597.17	2597.11
			y_{24}	2734.23	2734.08
			y_6^*	863.39	863.36
			y_7^*	978.42	978.39
			y_8^*	1115.48	1115.43
			y_9^*	1301.56	1301.51
			y_{10}^*	1400.63	1400.56
			y_{11}^*	1471.66	1471.59
			y_{12}^*	1600.71	1600.62
			y_{14}^*	1770.81	1770.70
			y_{15}^*	1884.86	1884.67
			y_{16}^*	1955.89	1955.87
			y_{17}^*	2083.95	2083.88
			y_{18}^*	2198.98	2198.88
			y_{19}^*	2298.05	2297.90
			y_{20}^*	2429.09	2429.00
			y_{21}^*	2560.13	2560.09
			y_{22}^*	2631.16	2631.12
			y_{23}^*	2759.22	2759.09
			y_{24}^*	2896.28	2896.14
			$y_{25}^{*,}$	3056.31	3056.28

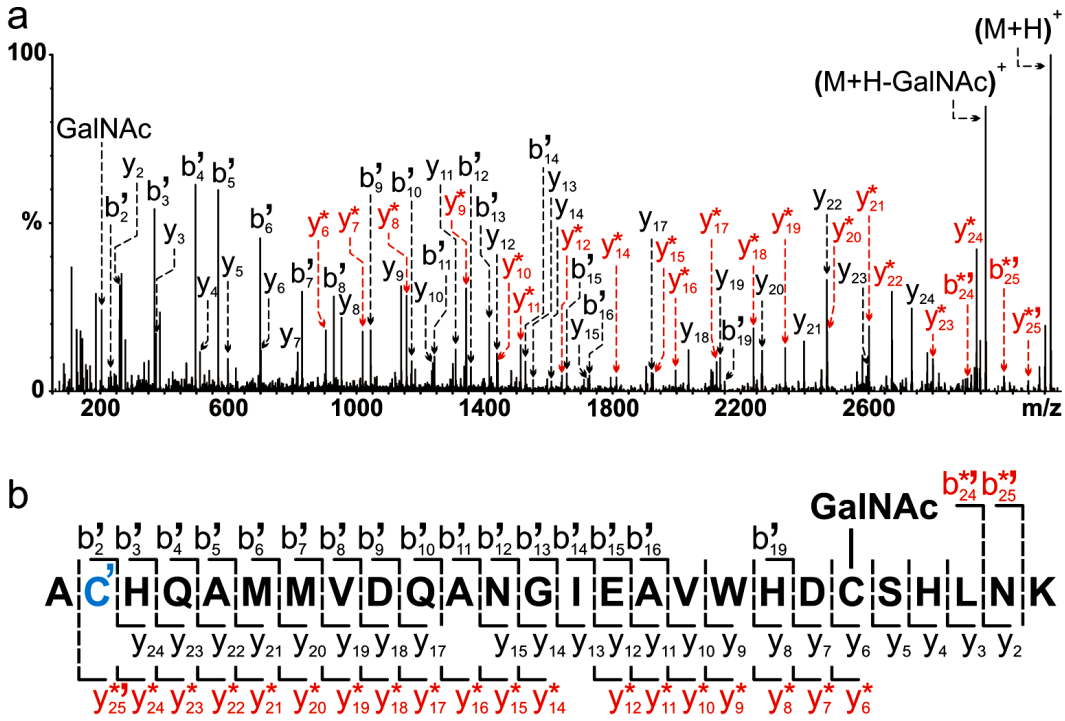

Figure 4.8 (a) Deconvoluted MS/MS spectrum of a glycosylated tryptic peptide ^{283}ACHQAMMVDQANGIEAVWHDCSHLNK^{308} obtained from the GTA_{E303C}→GalNAc intermediate under alkylation of cysteine residues. Only b- and y-ions were assigned. Peaks labeled with * and ’ correspond to ions containing GalNAc attachment and alkylated cysteine residue, respectively. (b) Product map indicating the cleavage sites identified from the CID MS/MS analysis.

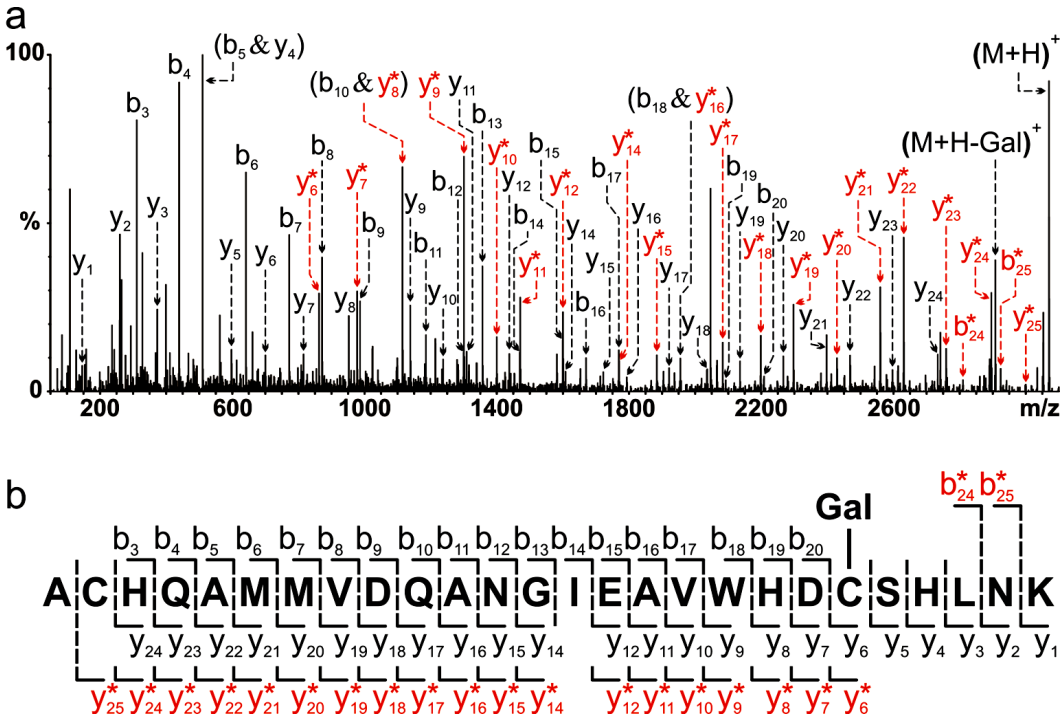

Figure 4.9 (a) Deconvoluted MS/MS spectrum of a glycosylated tryptic peptide ^{283}ACHQAMMVDQANGIEAVWHDCSHLNK^{308} obtained from the GTB_{E303C}→Gal intermediate. Only b- and y-ions were assigned. Peaks labeled with ***** correspond to ions containing Gal attachment. (b) Product map indicating the cleavage sites identified from the CID MS/MS analysis.

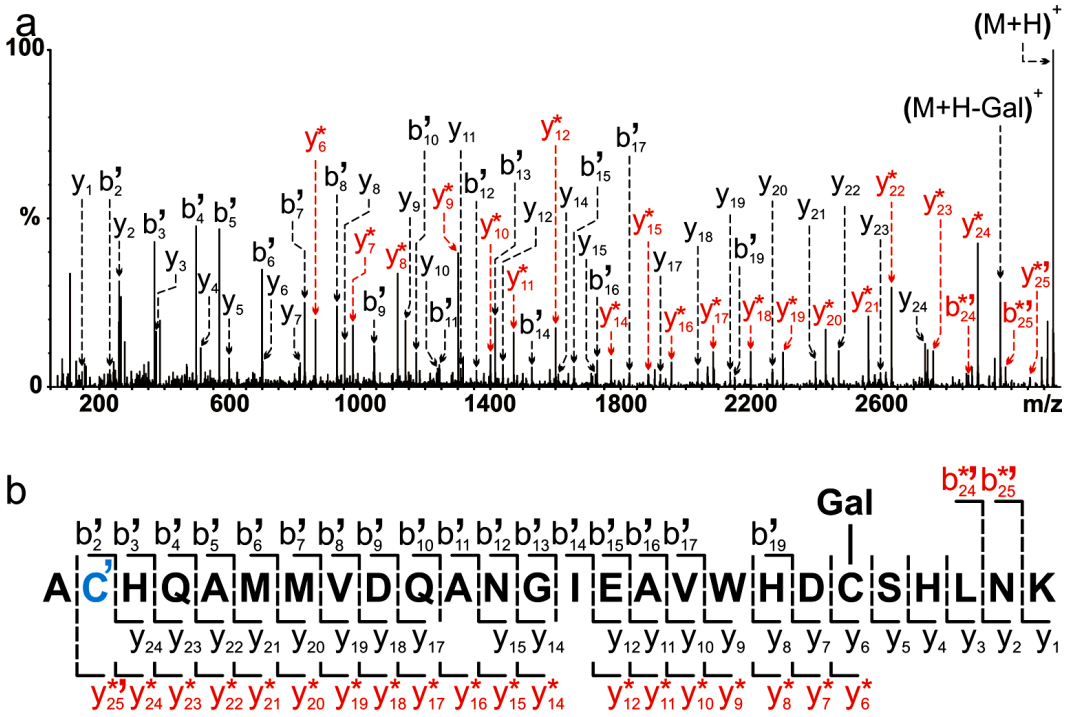

Figure 4.10 (a) Deconvoluted MS/MS spectrum of a glycosylated tryptic peptide ^{283}ACHQAMMVDQANGIEAVWHDCSHLNK^{308} obtained from the GTB_{E303C}→Gal intermediate under alkylation of cysteine residues. Only b- and y-ions were assigned. Peaks labeled with * and ' correspond to ions containing Gal attachment and alkylated cysteine residue, respectively. (b) Product map indicating the cleavage sites identified from the CID MS/MS analysis.

mutant retaining GTs operate by a double displacement enzymatic mechanism. As shown here, GTA and GTB, as well as other GT6 family enzymes, have a carboxylate properly oriented to act as a catalytic nucleophile. Other retaining GTs may lack such a residue or have Gln or Asn at an equivalent position.[14] In these cases, either conformational changes occur to correctly position a catalytic nucleophile relative to substrates for a double displacement mechanism or, possibly, alternate mechanisms, e.g., the S_{Ni}-like mechanism, are operational.

4.5 Literature Cited

(1) Taniguchi, N.; Honke, K.; Fukuda, M. *Handbook of Glycosyltransferases and Related Genes.* **2002**, Tokyo: Springer.

(2) Mutwil, M.; Debolt, S.; Persson, S. *Curr. Opin. Plant. Biol*. **2008**, *11,* 252−257.

(3) Ball, S. G.; Morell, M. K. *Annu. Rev. Plant Biol*. **2003**, *54*, 207−233.

(4) Merzendorfer, H. *J. Comp. Physiol. B*. **2006**, *176*, 1−15.

(5) Walsh, C.; Freel Meyers C. L.; Losey H. C. *J. Med. Chem*. **2003**, *46*, 3425−3436.

(6) Křen, V.; Řezanka, T. *FEMS Microbiol. Rev*. **2008**, *32*, 858−889.

(7) Griffith, B. R.; Langenhan, J. M.; Thorson, J. S. *Curr. Opin. Biotechnol*. **2005**, *16*, 622−630.

(8) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I. K.; Li, L.; Thorson, J. S. *Science* **2006**, *313*, 1291−1294.

(9) Ohtsubo, K.; Marth, J. D. *Cell* **2006**, *126*, 855−867.

(10) Freeze, H. H.; Aebi, M. *Curr. Opin. Struct. Biol*. **2005**, *15*, 490–498.

(11) Vigerust, D. J.; Shepherd, V. L. *Trends Microbiol*. **2007**, *15*, 211–218.

(12) Hinou, H.; Nishimura, S. *Curr. Top Med. Chem*. **2009**, *9*, 106–116.

(13) Breton, C.; Šnajdrová, L.; Jeanneau, C.; Koča, J.; Imberty, A. *Glycobiology* **2006**, *16*, 29R–37R.

(14) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem*. **2008**, *77*, 521–555.

(15) Davies, G. J. *Nat. Struct. Biol*. **2001**, *8*, 98–100.

(16) Monegal, A.; Planas, A. *J. Am. Chem. Soc*. **2006**, *128*, 16030–16031.

(17) Lairson, L. L.; Chiu, C. P. C.; Ly, H. D.; He, S.; Wakarchuk, W. W.; Strynadka, N. C. J.; Withers, S. G. *J. Biol. Chem*. **2004**, *279*, 28339–28344.

(18) Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C. J. *Nat. Struct. Biol*. **2001**, *8*, 166–175.

(19) Gibson, R. P.; Turkenburg, J. P.; Charnock, S. J.; Lloyd, R.; Davies, G. J. *Chem. Biol*. **2002**, *9*, 1337–1346.

(20) Flint, J.; Taylor, E.; Yang, M.; Bolam, D. N.; Tailford, L. E.; Martinez-Fleites, C.; Dodson, E. J.; Davis, B. G.; Gilbert, H. J.; Davies, G. J. *Nat. Struct. Mol. Biol*. **2005**, *12*, 608–614.

(21) Goedl, C.; Nidetzky, B. *ChemBioChem* **2009**, *10*, 2333–2337.

(22) Yamamoto, F.; Clausen, H.; White, T.; Marken, J.; Hakomori, S. *Nature* **1990**, *345*, 229–233.

(23) Patenaude, S. I.; Seto, N. O. L.; Borisova, S. N.; Szpacenko, A.; Marcus, S. L.; Palcic, M. M.; Evans, S. V. *Nat. Struct. Biol*. **2002**, *9*, 685–690.

(24) Marcus, S. L.; Polakowski, R.; Seto, N. O. L.; Leinala, E.; Borisova, S.; Blancher, A.; Roubinet, F.; Evans, S. V.; Palcic, M. M. *J. Biol. Chem*. **2003**, *278*, 12403–12405.

(25) Angulo, J.; Langpap, B.; Blume, A.; Biet, T.; Meyer, B.; Krishna, N. R.; Peters, H.; Palcic, M. M.; Peters, T. *J. Am. Chem. Soc*. **2006**, *128*, 13529–13538.

(26) Alfaro, J. A.; Zheng, R. B.; Persson, M.; Letts, J. A.; Polakowski, R.; Bai, Y.; Borisova, S. N.; Seto, N. O. L.; Lowary, T. L.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2008**, *283*, 10097–10108.

(27) Soya, N.; Shoemaker, G. K.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2009**, *19*, 1224–1234.

(28) Pesnot, T.; Jørgensen, R.; Palcic, M. M.; Wagner, G. K. *Nat. Chem. Biol*. **2010**, *6*, 321–323.

(29) Yazer, M. H.; Palcic, M. M. *Transfus. Med. Rev.* **2005**, *19*, 210–216.

(30) Letts, J. A.; Rose, N. L.; Fang, Y. R.; Barry, C. H.; Borisova, S. N.; Seto, N. O. L.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2006**, *281*, 3625–3632.

(31) Sindhuwinata, N.; Munoz, E.; Munoz, F. J.; Palcic, M. M.; Peters, H.; Peters, T. *Glycobiology* **2010**, *20*, 718–723.

Chapter 5

Enzyme Kinetics of the Human Neuraminidase 3 (NEU3) using Electrospray Mass Spectrometry: Substrate Recognition of NEU3 Requires Hydrophobic Aglycone[†]

5.1 Introduction

Sialic acid is a monosaccharide composed of nine carbon backbone and exists widely in nature, usually found at the non-reducing position in glycoproteins, glycolipids or oligosaccharide in mammalian cells.[1–3] Terminal sialic acid residues attached to glycans or glycoconjugates are known to play important roles in various biological and pathological processes such as cell-cell interactions,[4,5] viral and bacterial infections[6–8], inflammation[9,10] and cancers.[11–14] The removal of sialic acid residues is catalyzed by *exo*-neuraminidases (or sialidases: EC 3.2.1.18), and it is believed that the biological and pathological processes are controlled by the enzymes through the regulation of cellular sialic acid contents.[15,16]

The human neuraminidases (NEU) consist of a family of four isoforms (NEU1 – NEU4), which differ in their subcellular localization, roles and

[†] Parts of this chapter have been published: Sandbhor, M.S.; Soya, N.; Albohy, A.; Zheng, R.B.; Cartmell, J.; Bundle, D.R; Klassen, J.S.; Cairo, C.W. *Biochemistry* **2011**, *50*, 6753–6762.

enzymatic properties including the substrate specificity.[16] Among these isoforms, Neu3 is known as a plasma membrane-associated sialidase, which gains an increasing interest as participating in diseases and cell surface functions. Recent research reported that up-regulation of NEU3 has been observed in various carcinomas.[17–21] The increased NEU3 expression results in the suppression of apoptosis in cancer cells, which may imply that NEU3 involve in the malignant process.[17–21] Moreover, overexpression of human NEU3 in transgenic mice developed diabetic phenotype, suggesting that NEU3 involves in insulin signaling by modulation of gangliosides.[22,23] Therefore, NEU3 is an potential therapeutical target, and recent research efforts has been devoted to the design and synthesis of potential NEU3 inhibitors[24–26] as well as to the development and application of NEU3 siRNAs as novel therapeutical tools.[27]

NEU3 is known to exhibit a substrate preference particularly for gangliosides. A number of groups have observed that the enzyme hydrolyzes α(2→3)-sialosides with a β-3- or β-4-linkage to galactose (e.g. GM3, GD1a, GT1b).[19,28–33] Additionally, NEU3 has been found to hydrolyze α(2→8)-sialosides (e.g. GD3, GD1a, GD1b, GT1b),[28,31] and may even preferentially cleave this linkage.[34] Although some reports contradict this finding,[29] glycolipids which are α(2→3)-sialylated, but contain a branch point at the adjacent galactose residue (β-D-GalNAc*p*-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc-(1→1)-Cer), as found in GM1 and GM2, are poor substrates for NEU3.[28,30,31] However, despite its apparent importance in disease and membrane structure, to date the substrate specificity of NEU3 has not been well defined at molecular

level. In fact, recent works have primarily examined the enzymatic activity of NEU3 using natural ganglioside substrates. A detailed study of the NEU3 activity using synthetic glycolipid substrates analogues could aid in an understanding of the substrate recognition by NEU3, and thus the design of new therapeutical inhibitors.

Widely used techniques to investigate the sialidase activity are a fluorimetric assay and a radiochemical assay.[29,32,35–37] 2′-(4-methylumbelliferyl-α-D-*N*-acetyl-neuraminic acid (4MU-NA) is a common substrate to elucidate the activity of NEU3 in a fluorimetric assay (Figure 5.1). Neu5Ac is attached to a chromophore in 4MU-NA, and the product formation (the release of 4-methylumbelliferone) catalyzed by NEU3 is monitored by measuring fluorescence emission with excitation at 340 nm and emission at 455 nm.[37] In radiochemical assay, [^{3}H]-labeled glycolipids are utilized, and the release of radioactive label from the substrates results in the change in radiation intensities, which is monitored as product formation. Although those assays are well established and widely employed, there are several limitations. The critical drawback in both fluorimetric and radiochemical assays is that these assays depend on chromogenic and radioactive substrates. When the availability of chromogenic and radioactive-labeled substrates is limited, those must be prepared by labor-intensive, multistep synthesis. Moreover, the detection of substrates and/or products without chromophore is impossible using a spectroscopic detector, and the modification by chromogenic agent often alters enzyme kinetics in a fluorimetric assay. A radioactive-labeled substrate is more preferred due to the

similar properties to an original substrate, and they usually exhibit identical enzymatic kinetics. However, the use of radioisotopes requires careful operation and generates hazardous radioactive waste.

MS combined with electrospray ionization (ES) has emerged as a powerful tool particularly to study nonvolatile biological molecules.[38–40] ES is a "soft" ionization technique that enables to transfer intact biological molecules from solution to gas-phase, and gaseous ions generated by ES are separated depending on their mass-to-charge ratio (m/z) by an MS analyzer and detected. Moreover, the emergence of low flow variant nanoflow ES (nanoES) has significantly enhanced the performance of ES-MS.[41,42] The remarkable features of nanoES are that lower sample consumption (typically 2 ~ 3 μL) and smaller amount (pico-gram level) of analytes are sufficient for analysis.[42] ES-MS has been applied to steady-state kinetics of numerous enzymatic systems[43-45] and can be an excellent alternative technique for sialidase kinetics due to no requirement of modified or labeled substrates, and analysis can be carried out with smaller amount of enzymes and substrates owing to nanoES.

In the present work, the application of an electrospray ionization mass spectrometry (ES-MS)-based assay to enzyme kinetics of NEU3 with native and synthetic glycolipid substrates is described. Recently, a recombinant expression system for NEU3 was developed by Albohy et al.[37], and this system was used to explore the substrate tolerance of the enzyme in this study. The investigation into the influence of modification in aglycone and Neu5Ac residue to the enzymatic activity of NEU3 was aimed. In particular, the role of the lipid chain in substrate

recognition was explored using substrates with a modified aglycone, as previous reports have suggested that the enzyme has specificity for glycolipid substrates.[28] Moreover, it was tested if the active site of the protein could tolerate modifications of the Neu5Ac residue that would enable labeling studies.[46] Using a series of synthetic glycolipid analogs, this study revealed that the lipid chain hydrophobicity, but not its identity, was critical for NEU3 recognition. Moreover, this study showed that the modification of Neu5Ac at the *C*9 position can be tolerated by the enzyme.

5.2 Materials and Methods

5.2.1 Chemicals and Reagents

Ammonium acetate was purchased from Caledon Laboratory Chemicals (Georgetown, Canada). GM3 was purchased from Avanti polar lipids and stored under argon at –20 °C. Ammonium bicarbonate, sodium cholate and 4MU-NA was purchased from Sigma Aldrich (St Louis, MO). Purified human NEU3 as a maltose-binding protein (MBP)-fusion protein (MBP-NEU3)[37] and synthetic glycolipid substrates (**13a–f**, **14a–e**) used in this study were provided by Dr. Christopher Cairo (University of Alberta). The structures of the substrates are given in Figure 5.1.

5.2.2 Mass Spectrometry

ES-MS measurements were carried out using an Apex-Qe 9.4 Tesla Fourier transform ion cyclotron resonance mass spectrometer (FTICR-MS)

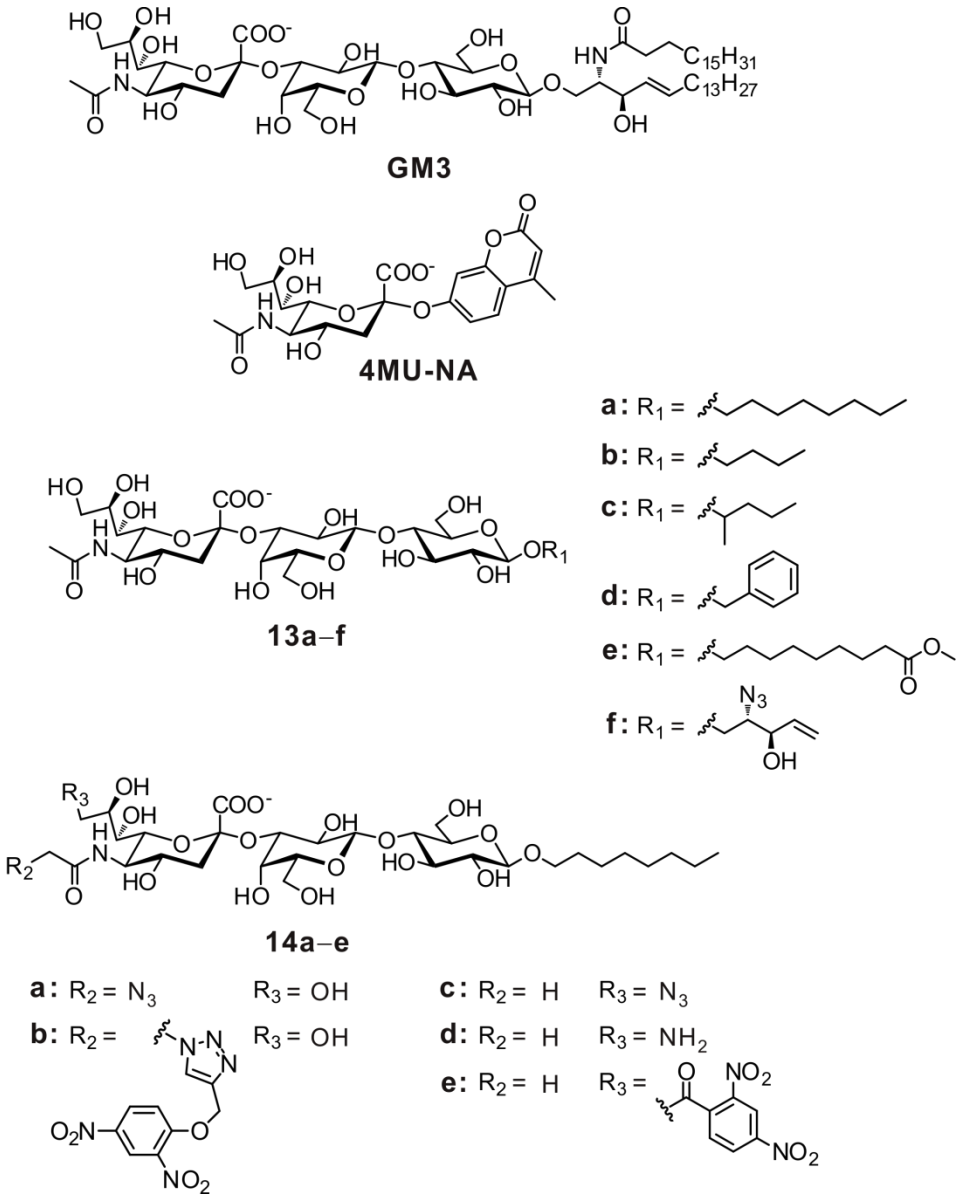

Figure 5.1 Structures of substrates used in NEU3 kinetics: α-D-Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→1)-Cer (**GM3**), 2′-(4-methylumbelliferyl-α-D-*N*-acetyl-neuraminic acid (**4MU-NA**), 3′-sialyllactoside with modified aglycone (**13a–f**), 3′-sialyllactoside with modified Neu5Ac (**14a–e**).

(Bruker, Billerica, MA) equipped with an external nanoES source. ES was performed using a borosilicate tube (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-97 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in negative ion mode was established by applying a voltage of ~ –900 V to a platinum wire inserted inside the ES tip. The solution flow rate was typically ~20 nL/min. The droplets and gaseous ions produced by ES were introduced into the mass spectrometer through a glass sampling capillary (i.d. 0.5 mm). Nitrogen gas at a flow rate of 2.0 L/min and a temperature of 90 °C was used as drying gases to facilitate the production of gas-phase ions. The capillary entrance voltage was held at 0 V, and the capillary exit was at –340 V. A deflector voltage of –250 V was used. Gaseous ions were transmitted through the first funnel and skimmer held at –190 V and –20 V, respectively, then through the second funnel and skimmer held at –7.6 and –4.5 V, respectively. The ions were stored electrodynamically in an rf hexapole for 0.5 s followed by further accumulation in a hexapole collision cell for 0.4 s. Following accumulation, the ions were transferred from the hexapole collision cell, and through a series of ion optics, introduced into the ICR cell. The front and back trapping plates of the cell were maintained at –0.65 and –0.95 V, respectively, throughout experiment. The typical base pressure for the instrument was ~1 × 10^{-10} mbar. Data acquisition and analysis was performed using the ApexControl software, version 4.0 (Bruker Daltonics). A minimum of 30 transients with 512 K data points per transient were used for each acquisition.

5.2.3 Enzyme Kinetics

Steady-state kinetic studies were performed for 4MU-NA hydrolysis by MBP-NEU3 using ES-MS. The reactions were prepared by mixing 4MU-NA (25, 50, 75, 100, 150, 200, 250, 300 μM) and purified MBP-NEU3 (0.83 μM) in 0.1 M ammonium acetate buffer (pH = 5.0), with incubation at 37 °C. The reactions were stopped at predetermined periods (16, 20, 24, 28, 32, 36, and 40 min) by adding an aliquot (5 μL) of the reaction mixture to 35 μL of 0.1 M ammonium bicarbonate buffer (pH = 10) containing internal standard (5 μM). The quenched solutions were stored at −20 °C until analysis by ES-MS. All solutions were diluted (1:8) using Milli-Q water prior to ES-MS analysis.

For k_{rel} determinations, the reactions were prepared by mixing the indicated substrate (100 μM) with purified MBP-NEU3 (0.83 μM) in 0.1 M ammonium acetate buffer (pH = 5.0), or buffer containing 3.5 mM sodium cholate, with incubation at 37 °C. The reactions were stopped at predetermined periods (4 to 120 min) and measured as described above. Rates were determined by fitting the initial linear portion of the curves, with a minimum of four points over 20 min

5.2.4 Calibration Curve

Calibration curves were used for the quantification of enzymatic products.[47–50] The signal intensity of the product and an internal standard (I_P and I_{IS}, respectively) were used to determine an intensity ratio (I_P/I_{IS}). Calibration curves were generated by plotting I_P/I_{IS} against a known concentration ratio of product and internal standard ([P]/[IS]) to provide a slope (*m*) and intercept (*int*).

The experimental concentration of the product, [P], was then estimated using the following equation:[49]

$$[\mathrm{P}] = \left(\frac{I_{\mathrm{P}}}{I_{\mathrm{IS}}} - Int \right) \times \frac{[\mathrm{IS}]}{m} \tag{5.1}$$

In reactions of 4MU-NA, GM3, and **13a**–**f**, 5-*N*-acetyl neuraminic acid (Neu5Ac, Figure 5.2a) was monitored as the product ion (m/z 308.1) with 9-azido-9-deoxy-Neu5Ac (Figure 5.2c) as the internal standard (m/z 333.1). In reactions of **14a–e**, the formation of octyl-β-D-lactoside (Figure 5.2b) was monitored as the product ion (m/z 453.2) with benzyl-β-D-lactoside (Figure 5.2d) as an internal standard (m/z 431.2). Calibration curves for Neu5Ac and octyl-β-lactoside are provided in Figure 5.3.

5.3 Results and Discussion

5.3.1 Analysis of NEU3 Kinetics using ES-MS

A number of groups have employed ES-MS to study enzyme kinetics including sulfotransferases,[50,51] glutathione *S*-transferase,[52] hexokinase,[53] phosphoglucomutase[54] and others.[48,55,56] The considerable advantages of ES-MS are the low consumption of enzyme and synthetic substrates (often they are precious) as well as the capability to detect virtually any products and/or substrates that are ionizable. However, one of the major issues for the application of ES-MS to enzyme kinetics is that the correlation between signal intensities in mass spectra and concentration in electrosprayed solution should be considered to

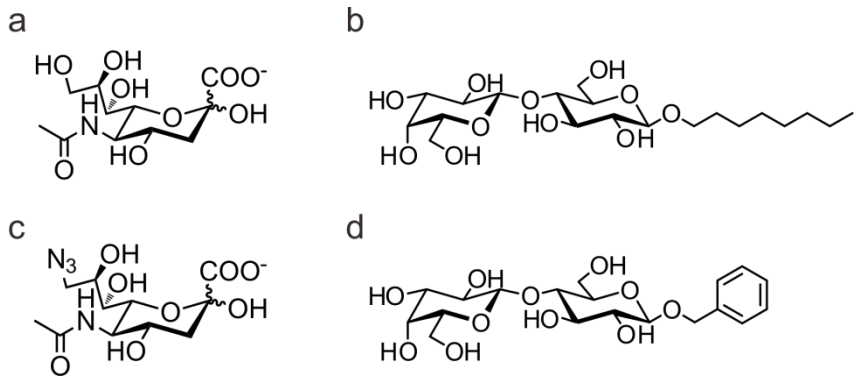

Figure 5.2 Structures of enzymatic products (a) Neu5Ac for **13a–f**, (b) octyl-β-D-lactoside for **14a–e** and internal standards (c) 9-azido-9-deoxy-Neu5Ac, (d) benzyl-β-D-lactoside used for calibration curves and quantification of product formation.

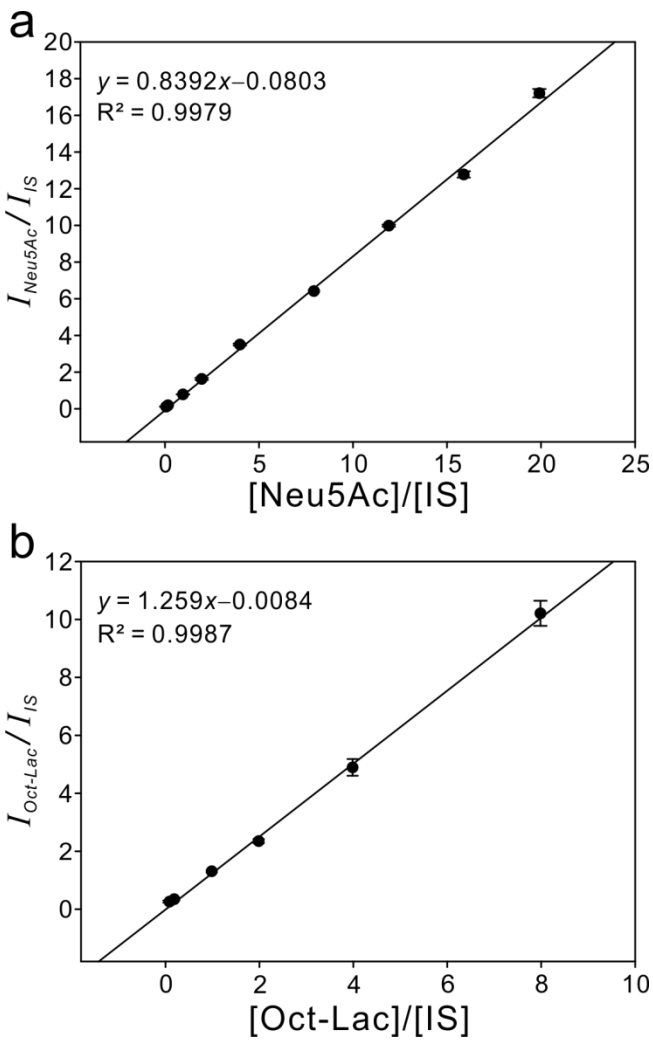

Figure 5.3 Calibration curves for (a) Neu5Ac and (b) octyl-β-D-lactoside. All points represent five replicates and error is shown as the standard deviation.

quantify the concentration of enzymatic products. The correlation between signal intensities and concentration is given by the following equation:[57]

$$[\mathrm{M}] = R_M \times I_M \tag{5.2}$$

where [M] is concentration of analyte, I_M is signal intensity and R_M is a response factor of analyte. The difference in the response factors originates from the difference in ES ionization efficiency, transmission efficiency in mass analyzer and detector efficiency.[57–59] In cases where response factors differ, the relative abundance of the gas phase ions measured by ES-MS will not reflect the concentrations of the individual species in solution.[57–60] Moreover, the response factors can differ in each measurement, which results in varied signal intensities for even same concentration of analytes. Therefore, the response factors influence the accuracy and precision in quantification of analytes.

To implement a quantitative kinetic assay with ES-MS, an internal standard was used to cancel out any effects relating to response factors. The product formation was quantified using calibration curves, which was generated using an internal standard with similar properties and functionality, therefore similar response factors, to the expected enzymatic products (Figure 5.2a and 5.2b). In the present study, 9-azido-9-deoxy-Neu5Ac and benzyl-β-D-lactoside were selected as the internal standards (Figure 5.2c and 5.2d). To validate the ES-MS assay, the steady-state kinetics of the fluorogenic substrate, 4MU-NA, was determined by acquiring initial rates of hydrolysis (v_0) at a range of substrate concentrations (Figure 5.4). The Michaelis-Menten constant K_m = 39 ± 7 μM and the maximum reaction rate V_{max} = 0.24 ± 0.01 μM/min were obtained using ES-

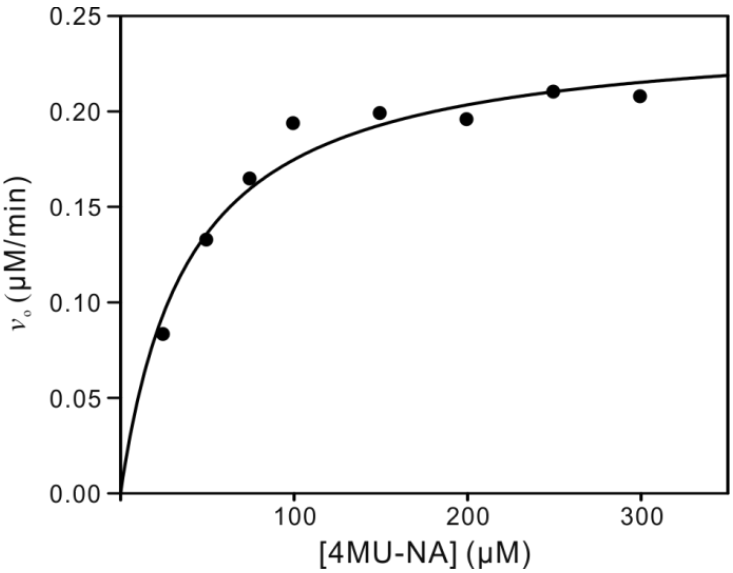

Figure 5.4 Steady-state kinetics of 4MU-NA hydrolysis by NEU3 observed by ES-MS. The kinetic parameters determined were K_m = 39 ± 7 μM and V_{max} = 0.243 ± 0.011 μM/min. These values were in good agreement with those using a fluorescence-based assay (K_m = 45 ± 3 μM and V_{max} = 5.6 ± 0.1 fluorescence units/s or 0.314 ± 0.006 μM/min).[37]

MS assay, which were in close agreement with those determined using a fluorescence-based assay (K_m = 45 ± 3 μM and V_{max} = 0.314 ± 0.006 μM/min).[37] After confirming that the ES-MS assay could provide accurate kinetic measurements, the activity of the remaining substrates was analyzed through the use of k_{rel} based on the initial velocity (v_0) of product formation. The use of a k_{rel} analysis has been used in a variety of enzymatic systems,[61,62] including NEUs.[28,63] This approach relies on pseudo-first order conditions where all parameters are kept constant between substrates, in particular the enzyme and substrate concentrations. Although 4MU-NA is a substrate that is generally used for a fluorescence-based assay, a native substrate of NEU3, the ganglioside GM3, was also examined (Figure 5.5). As expected, the reaction rate of GM3 catalyzed by NEU3 was faster than 4MU-NA, with an approximately 4.5-fold increase in relative rate (Table 5.1).[28,31] For comparison, all rates to that of GM3 (k_{rel} = 1.00 ± 0.04) was normalized.

Gangliosides typically have remarkably low critical micelle concentrations,[64] and a previous report has suggested that NEU3 activity is sensitive to the presence of surfactants.[29] In contrast, Albohy et al. have previously seen only minor effects of surfactants on the enzymatic activity of purified MBP-NEU3 with 4MU-NA as a substrate.[37] Using the ES-MS enzyme assay, it was confirmed that, indeed, the presence of surfactants (0.5 and 1.5 mM sodium cholate) increased the relative rates of reaction for GM3 but had only a minor effect on 4MU-NA (Figure 5.5b and Table 5.1). These results are typical

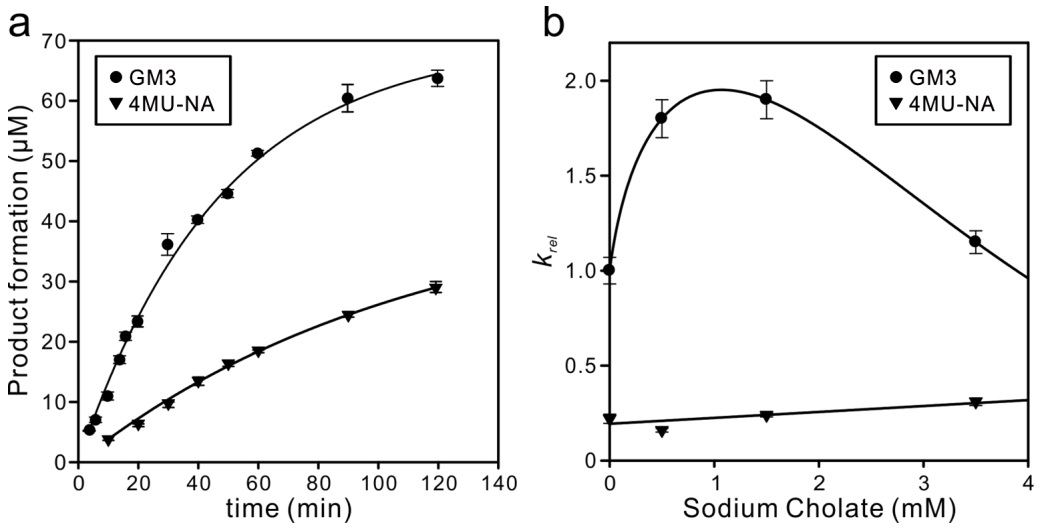

Figure 5.5 (a) Hydrolysis of 4MU-NA and GM3 by NEU3 were monitored using the ES-MS. Curves were fit to an exponential for clarity. (b) The relative rates of 4MU-NA and GM3 hydrolysis in the presence of sodium cholate (0.5–3.5 mM). Initial rates were used to determine k_{rel}. All points represent five replicates and error is shown as the standard deviation.

Table 5.1 Hydrolysis rates of substrates catalyzed by NEU3 determined by the ES-MS kinetic assay and the calculated octanol-water partition coefficients (log P).

substrates	v_0 [μM/min]	k_{rel}	*Log P* [a]
4MU-NA	0.25 ± 0.02	0.21 ± 0.02	2.42
GM3	1.20 ± 0.03	1.00 ± 0.04	14.76
13a	1.11 ± 0.03	0.93 ± 0.03	5.01
13b	0.19 ± 0.01	0.16 ± 0.01	2.88
13c	0.37 ± 0.02	0.31 ± 0.02	3.41
13d	0.26 ± 0.01	0.22 ± 0.01	2.68
13e	0.61 ± 0.02	0.51 ± 0.02	3.90
13f	0.37 ± 0.02	0.31 ± 0.02	0.81
14a	1.28 ± 0.03	1.07 ± 0.04	5.01
14b	0.003 ± 0.001	0.003 ± 0.001	5.01
14c	0.35 ± 0.01	0.29 ± 0.01	5.01
14d	0.08 ± 0.01	0.06 ± 0.01	5.01
14e	0.14 ± 0.01	0.12 ± 0.01	5.01

a. Log P values were calculated only for the aglycone portion of the substrate.

b. The reported v_0 values are the average value of five measurements.

c. The reported error corresponds to one standard deviation.

for the cleavage of ganglioside substrates by NEUs, which often have enhance activity with mixed micelles.[65–67]

5.3.2 Cleavage of α(2→3)-sialosides with Modified Aglycone

To examine the role of the lipid aglycone, the rates of hydrolysis for a series of six sialyllactosides (**13a–f**) relative to GM3 (Figure 5.6) were determined. Interestingly, the activity of an octyl-sialyl-β-lactoside (**13a**, k_{rel} 0.93 ± 0.03), was similar to that of GM3. However, truncation of the octyl chain in **13a** to a butyl chain (**13b**, k_{rel} 0.16 ± 0.01) resulted in a large reduction in the rate of hydrolysis. Increasing the bulk of the lipid with a branching methyl group doubled the activity (**13c**, k_{rel} 0.31 ± 0.02), while attaching a benzyl group in place of the butyl chain showed less of an increase (**13d**, k_{rel} 0.22 ± 0.01). A long hydrophobic chain with a methyl ester group had a relative rate approximately half of GM3 (**13e**, k_{rel} 0.51 ± 0.02). Finally, it was tested if the presentation of the amino and hydroxyl groups of the core sphingolipid were important to NEU3 recognition using a truncated sphingosine analog as the aglycone of sialyllactose, **13f**. Although the truncated sphingosine derivative (**13f**, k_{rel} 0.31 ± 0.02) was more active than the butyl derivative, **13b**, it had activity comparable to that of the 2-pentyl derivative, **13c**. Based on these results, enzyme activity appeared to track with the hydrophobicity of the linker, rather than specific functional groups. When the calculated octanol-water partition coefficients (log P) of each of the aglycone alkyl groups were compared, a clear relation between increased hydrophobicity of the alkyl chain and improved enzyme activity was observed

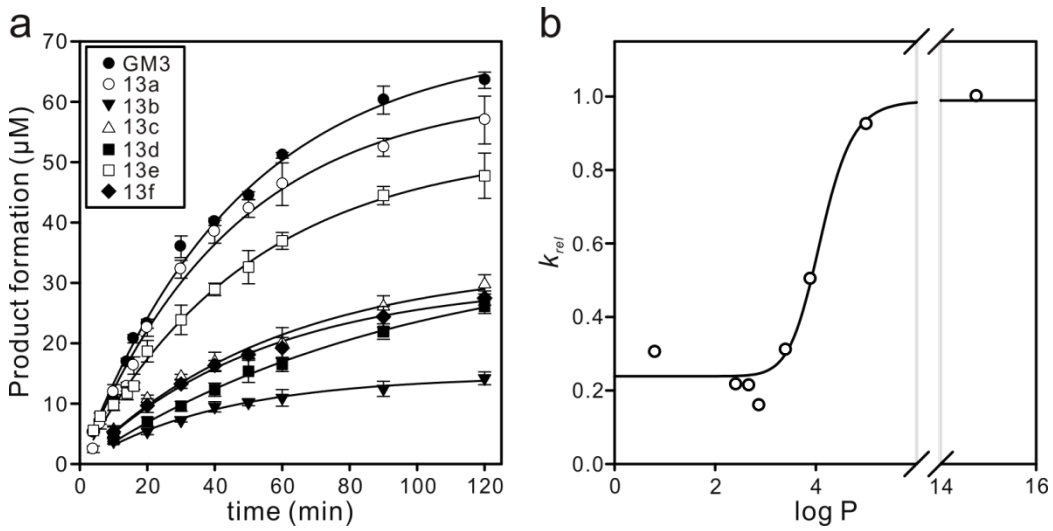

Figure 5.6 NEU3 cleavage of α(2→3)-sialosides with a modified aglycone. (a) The time course of synthetic glycolipid hydrolysis by NEU3 was monitored by ES-MS. Initial rates were used to determine k_{rel} and values are shown in Table 1. All points represent five replicates and error is shown as the standard deviation. (b) Relative rates of NEU3 hydrolysis are dependent on the hydrophobicity of the acyl chain. Calculated log P values (octanol-water) of **GM3**, **4MU-NA**, and **13a-f** were plotted versus the measured k_{rel} values.

(Figure 5.6b).[68] Glycolipids with an alkyl chain of log P > 4 (**GM3**, **13a**) were excellent substrates, while those with log P < 3 were poor substrates (Table 5.1). It is not expected that the observed dependency on the lipid chain is due to aggregation of the substrate, as the concentration used (100 μM) is well below the critical micelle concentration of octylglycoside.[69]

Although the very earliest studies of NEU3 identified the clear preference of the enzyme for glycolipids,[28] this is the first study to examine the role of the lipid in substrate recognition. This study showed a direct dependence upon the hydrophobicity of the aglycone, where increased hydrophobicity resulted in increased NEU3 activity. Although we cannot rule out that this dependence is due to the formation of micelles by the substrate, assays were conducted at concentrations of substrate that are expected to be below the CMC for octylglycosides, which we expect to be the lower limit for the compounds in the lipid series 1a–f. Additionally, substrates which included features of the sphingosine headgroup (**13f**) did not show improvement over similar length alkyl chains (**13c**); and replacement of the ceramide moiety (**GM3**) with an octyl chain (**13a**) showed only minor loss in activity. Together, these data are consistent with a two-site model of NEU3 recognition. In addition to the glycone recognition site that has been proposed by homology,[37] a second hydrophobic subsite must recognize general hydrophobic properties of the substrate.

5.3.3 **Cleavage of α(2→3)-sialosides with Modified Neu5Ac**

Fluorogenic substrates present an attractive strategy for probing the activity of glycosidase enzymes.[70] However, the design of substrates with appended chromophores requires some foreknowledge of the enzyme tolerance for substrate modifications. There are no reports of NEU3 hydrolysis with substrates containing modified Neu5Ac residues; therefore, two series of compounds were designed to explore the tolerance of the enzyme for these types of modifications (Figure 5.7). Modification of the *N*5-Ac group to azidoacetate resulted in a moderate improvement in hydrolysis by NEU3 relative to GM3 (**14a**, k_{rel} 1.07 ± 0.04). This observation is consistent with the previous study of NEU3 inhibitors by Zou et al.,[26] where the NEU3 inhibitory potency was improved by introducing an *N*5-azido modification to 2,3-didehydro-N-acetyl-neuraminidic acid. They suggested that the azide may enhance the binding to NEU3, and the increase in hydrolysis rate of **14a** might attribute to the enhancement of binding. Based on the result of ES-MS assay, NEU3 did not tolerate the presence of larger groups, such as a triazole, at the *N*5-Ac position; this modification leaded to almost undetectable substrate activity (**14b**, k_{rel} 0.003± 0.001). This is consistent with the homology model of NEU3, which predicts a small hydrophobic pocket for the *N*-acyl sidechain.[37]

Secondly, the modifications of the *C*9 position in the Neu5Ac residue were examined (Figure 5.7b). This study revealed that the *C*9-azido was a moderate substrate (**14c**, k_{rel} 0.29 ± 0.01); however, reduction to the *C*9-amino analog led to a dramatic decrease in the reaction rate (**14d**, k_{rel} 0.06 ± 0.01). The positive charge by the protonation state of the amine group under assay conditions

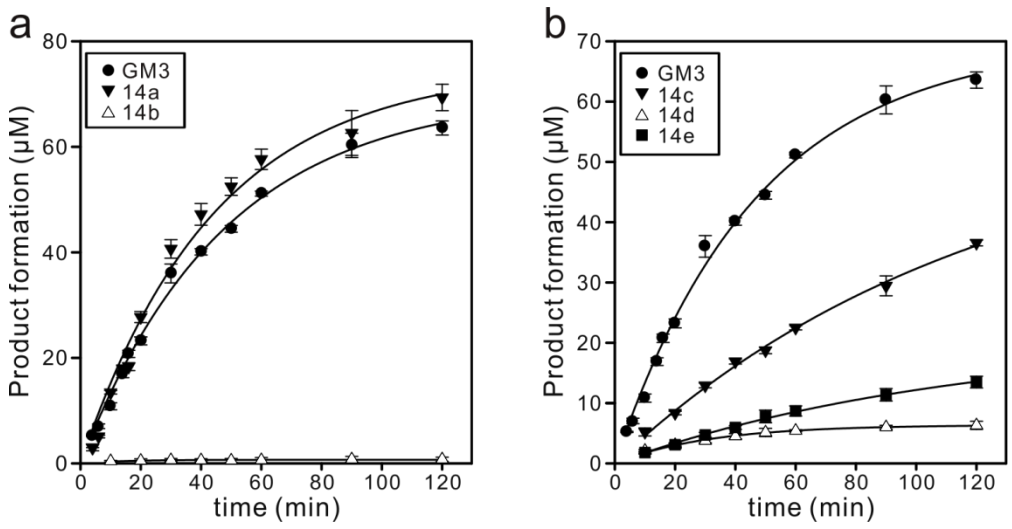

Figure 5.7 NEU3 cleavage of α(2→3)-sialosides with a modified Neu5Ac. Initial rates of hydrolysis for (a) *N*5-Ac and (b) *C*9 modified sialosides were used to determine k_{rel} and values are shown in Table 5.1. All points represent five replicates and error is shown as the standard deviation.

might result in this drop in activity. Interestingly, our result showed that coupling of the *C*9-amido derivative, **14d**, to 2,4-dinitrobenzoic acid (**14e**), resulted in increased activity (**14e**, k_{rel} 0.12 ± 0.01). Zou et al. suggested that the binding pocket for the glycerol sidechain is relatively large in NEU3 and can accommodate larger groups.[26] Confirmation of this finding with NEU3 substrates finding suggests that incorporation of fluorophores, quenchers, or other labels into glycolipid substrates for NEU3 could be possible at *C*9. However, modification of the *N*5-Ac position would likely obscure enzyme activity.

5.4 Conclusion

Membrane-associated sialidase NEU3 is an important regulator of membrane composition, and recent studies revealed that NEU3 appears to involve in cancer progression and development of diabetes. However, little is known about its substrate recognition to date. Using a series of synthetic and natural substrates, it was revealed that NEU3 requires a hydrophobic aglycone for substrate recognition. There was no evidence that features of the ceramide headgroup are recognized by the enzyme, or that these observations were confounded by micelle formation. Finally, the modifications of α(2→3)-Neu5Ac residue with an azide at *C*9 or *N*5-Ac remained substrates. However, larger modifications at *N*5-Ac resulted in inactive substrates. In at least one trisaccharide (**14e**) the substrate activity was maintained, suggesting that this position could be exploited in the design of labelled substrates.

In the present study, ES-MS was applied to explore the enzyme kinetics of NEU3 with synthetic glycolipid substrates. The use of ES-MS for monitoring the course of enzyme kinetics allowed us to carry out the analysis with smaller amounts of the substrate, and to quantitatively observe the cleavage of even weak substrates. Moreover, ES-MS enables us to use a variety of natural and synthetic substrates. In conclusion, this study demonstrated that ES-MS can be an alternative methodology for sialidase kinetics.

5.5 Literature Cited

(1) Varki, A. *Glycobiology.* **1992**, *2*, 25–40.

(2) Schauer, R. *Glycoconj. J.* **2000**, *17*, 485–499.

(3) Vimr, E. R.; Kalivoda, K. A.; Deszo, E. L.; Steenbergen, S. M. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 132–153.

(4) Schauer, R. *Curr. Opin. Struct. Biol.* **2009**, *19*, 507–514.

(5) Lopez, P. H.; Schnaar R. L. *Curr. Opin. Struct. Biol.* **2009**, *19*, 549–557.

(6) Sakarya, S.; Öncü, S. *Med. Sci. Monit.* **2003**, *9*, RA76–82.

(7) Suzuki, Y.; Ito, T.; Suzuki, T.; Holland Jr., R. E.; Chambers, T. M.; Kiso, M.; Ishida, H.; Kawaoka, Y. *J. Virol.* **2000**, *74*, 11825–11831.

(8) Viswanathan, K.; Chandrasekaran, A.; Srinivasan, A.; Raman, R.; Sasisekharan, V. Sasisekharan, R. *Glycoconj. J.* **2010**, *27*, 561–570.

(9) Lasky, L. *Science* **1992**, *258*, 964–969.

(10) Lowe, J. B. *Curr. Opin. Cell. Biol.* **2003**, *15*, 531–538.

(11) Mabry, E. W.; Carubell, R. *Experientia* **1972**, *28*, 182–183.

(12) Hakomori, S. *Proc. Natl. Acad. Sci. USA.* **2002**, *99*, 10231−10233.

(13) Ohyama, C. *Int. J. Clin. Oncol*. **2008**, *13*, 308−313.

(14) Yin, J.; Miyazaki, K.; Shaner, R. L.; Merrill Jr., A. H.; Kannagi, R. *FEBS Lett.* **2010**, *584*, 1872−1878.

(15) Monti, E.; Preti, A.; Venerando, B.; Borsani, G. *Neurochem. Res.* **2002**, *27*, 649−663.

(16) Miyagi, T. *Trends. Glycosci. Glycotech.* **2010**, *22*, 162−172.

(17) Kakugawa, Y.; Wada, T.; Yamaguchi, K.; Yamanami, H.; Ouchi, K.; Sato, I.; Miyagi, T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10718−10723.

(18) Miyagi, T.; Wada, T.; Yamaguchi, K.; Hata, K. *Glycoconj. J.* **2004**, *20*, 189−198.

(19) Kato, K.; Shiga, K.; Yamaguchi, K.; Hata, K.; Kobayashi, T.; Miyazaki, K.; Saijo, S.; Miyagi, T. *Biolchem. J.* **2006**, *394*, 647−656.

(20) Miyagi, T. *Proc. Jpn. Acad. Ser. B.* **2008**, *84*, 407−418.

(21) Shiozaki, K.; Yamaguchi, K.; Sato, I.; Miyagi, T. *Cancer. Sci.* **2009**, *100*, 588−594.

(22) Sasaki, A.; Hata, K.; Suzuki, S.; Sawada, M.; Wada, T.; Yamaguchi, K.; Obinata, M.; Tateno, H.; Suzuki, H.; Miyagi, T. *J. Biol. Chem*. **2003**, *278*, 27896−27902.

(23) Miyagi, T.; Wada, T.; Yamaguchi, K.; Hata, K.; Shiozaki, K. *J. Biochem.* **2008**, *144*, 279−285.

(26) Magesh, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. *J. Mol. Graphics Modell.* **2006**, *25*, 196−207.

(25) Magesh, S.; Savita, V.; Moriya, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. *Bioorg. Med. Chem.* **2009**, *17*, 4595–4603.

(26) Zou, Y.; Albohy, A.; Sandbhor, M.; Cairo, C. W. *Bioorg. Med. Chem. Lett*. **2010**, *20*, 7529–7533.

(27) Wada, T.; Hata, K.; Yamaguchi, K.; Shiozaki, K.; Koseki, K.; Moriya, S.; Miyagi, T. *Oncogene* **2007**, *26*, 2483–2490.

(28) Miyagi, T.; Wada, T.; Iwamatsu, A.; Hata, K.; Yoshikawa, Y.; Tokuyama, S.; Sawada, M. *J. Biol. Chem.* **1999**, *274*, 5004–5011.

(29) Monti, E.; Bassi, M. T.; Papini, N.; Riboni, M.; Manzoni, M.; Venerando, B.; Croci, G.; Preti, A.; Ballabio, A.; Tettamanti, G.; Borsani, G. *Biochem. J.* **2000**, *349*, 343–351.

(30) Wang, Y.; Yamaguchi, K.; Shimada, Y.; Zhao, X. J.; Miyagi, T. *Eur. J. Biochem.* **2001**, *268*, 2201–2208.

(31) Ha, K. T.; Lee, Y. C.; Cho, S. H.; Kim, J. K.; Kim, C. H. *Mol. Cells* **2004**, *17*, 267–273.

(32) Papini, N.; Anastasia, L.; Tringali, C.; Croci, G.; Bresciani, R.; Yamaguchi, K.; Miyagi, T.; Preti, A.; Prinetti, A.; Prioni, S.; Sonnino, S.; Tettamanti, G.; Venerando, B.; Monti, E. *J. Biol. Chem.* **2004**, *279*, 16989–16995.

(33) Wang, J.; Wu, G.; Miyagi, T.; Lu, Z. H.; Ledeen, R. W. *J. Neurochem.* **2009**, *111*, 547–554.

(34) Azuma, Y.; Sato, H.; Higai, K.; Matsumoto, K. *Biol. Pharm. Bull.* **2007**, *30*, 1680–1684.

(35) Venerando, B.; Fiorilli, A.; Croci, G. L.; Tettamanti, G. *Blood* **1997**, *90*, 2047–2056.

(36) Zanchetti, G.; Colombi, P.; Manzoni, M.; Anastasia, L.; Caimi, L.; Borsani, G.; Venerando, B.; Tettamanti, G.; Preti, A.; Monti, E.; Bresciani, R. *Biochem. J.* **2007**, *408*, 211–219.

(37) Albohy, A.; Li, M. D.; Zheng, R. B.; Zou, C.; Cairo, C. W. *Glycobiology* **2010**, *20*, 1127–1138.

(38) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64–71.

(39) Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, *65*, 972A–986A.

(40) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2001**, *20*, 362–387.

(41) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1–68.

(42) Karas, M.; Bahr, U.; Dülcks, T. *Fresenius. J. Anal. Chem.* **2000**, *366*, 669–676.

(43) Liesener, A.; Karst, U. *Anal. Bioanal. Chem.* **2005**, *382*, 1451–1464.

(44) Shipovskov, S.; Reimann, C. T. *Analyst* **2007**, *132*, 397–402.

(45) Greis, K. D. *Mass Spectrom. Rev.* **1996**, *26*, 324–339.

(46) Cairo, C. W.; Key, J. A.; Sadek, C. M. *Curr. Opin. Chem. Biol.* **2010**, *14*, 57–63.

(47) Hsieh, F. Y. L.; Tong, X.; Wachs, T.; Ganem, B.; Henion, J. *Anal. Biochem.* **1995**, *229*, 20–25.

(48) Pi, N.; Meyers, C. L. F.; Pacholec, M.; Walsh, C. T.; Leary, J. A. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 10036–10041.

(49) Pi, N.; Yu, Y.; Mougous, J. D.; Leary, J. A. *Protein Sci.* **2004** *13*, 903–912.

(50) Danan, L. M.; Yu, Z. H.; Ludden, P. J.; Jia, W. T.; Moore, K. L.; Leary, J. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1633–1642.

(51) Pi, N.; Hoang, M. B.; Gao, H.; Mougous, J. D.; Bertozzi, C. R.; Leary, J. A. *Anal. Biochem.* **2005**, *341*, 94–104.

(52) Ge, X.; Sirich, T. L.; Beyer, M. K.; Desaire, H.; Leary, J. A. *Anal. Chem.* **2001**, *73*, 5078–5082.

(53) Gao, H.; Petzold, C. J.; Leavell, M. D.; Leary, J. A. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 916–924.

(54) Gao, H.; Leary, J. A. *Anal. Biochem.* **2004**, *329*, 269–275.

(55) Zechel, D. L.; Konermann, L.; Withers, S. G.; Douglas, D. J. *Biochemistry* **1998**, *37*, 7664–7669.

(56) Bothner, B.; Chavez, R.; Wei, J.; Strupp, C.; Phung, Q.; Schneemann, A.; Siuzdak, G. *J. Biol. Chem.* **2000**, *275*, 13455–13459.

(57) Gabelica, V.; Rosu, F.; Pauw, E. D. *Anal. Chem.* **2009**, *81*, 6708–6715.

(58) Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, *406*, 11–35.

(59) Gabelica, V.; Galic, N.; Rosu, F.; Houssier, C.; Pauw, E. D. *J. Mass Spectrom.* **2003**, *38*, 491–501.

(60) Sterner, J. L.; Johnston, M. V.; Nicol, G. R.; Ridge, D. P. *J. Mass Spectrom.* **2000**, *35*, 385–391.

(61) Mehrotra, K. N.; Dauterman, W. C. *J. Neurochem.* **1963**, *10*, 119–123.

(62) Roberts, M. F.; Adamich, M.; Robson, R. J.; Dennis, E. A. *Biochemistry* **1979**, *18*, 3301–3308.

(63) Sato, K.; Hanagata, G.; Kiso, M.; Hasegawa, A.; Suzuki, Y. *Glycobiology* **1998**, *8*, 527–532.

(64) Sonnino, S.; Cantù, L.; Corti, M.; Acquotti, D.; Venerando, B. *Chem. Phys. Lipids* **1994**, *71*, 21–45.

(65) Tringali, C.; Papini, N.; Fusi, P.; Croci, G.; Borsani, G.; Preti, A.; Tortora, P.; Tettamanti, G.; Venerando, B.; Monti, E. *J. Chem. Biol.* **2004**, *279*, 3169–3179.

(66) Barton, N. W.; Lipovac, V.; Rosenberg, A. *J. Biol. Chem.* **1975**, *250*, 8462–8466.

(67) Gatt, S.; Gazit, B.; Barenholz, Y. *Biochem. J.* **1981**, *193*, 267–273.

(68) Dunn, W. J.; Hansch, C. *Chem. Biol. Interact* **1974**, *9*, 75–95.

(69) Sonnino, S.; Cantù, L.; Corti, M.; Acquotti, D.; Venerando, B. *Chem. Phys. Lipids* **1994**, *71*, 21–45.

(70) Zhou, X. T.; Forestier, C.; Goff, R. D.; Li, C. H.; Teyton, L.; Bendelac, A.; Savage, P. B. *Org. Lett.* **2002**, *4*, 1267–1270.

Chapter 6

Conclusion and Future Work

This thesis has discussed the application of ES-MS based methodologies to examine GT-substrate and GT-product complexes, the catalytic mechanism of retaining GTs and the enzyme kinetics of NEU3. To close the entire story, this chapter highlights the results of each chapter and addresses potential research directions.

6.1 Summaries of Chapters

Chapter 1 discussed the background and motivation of this thesis. As an introduction, the biological significance in carbohydrates and carbohydrate-modifying enzymes was described. Moreover, fundamental aspects of ES, FTICR as well as the *direct* ES-MS binding assay developed by the Klassen group and ES-MS kinetic assay utilized in this work were presented.

In Chapter 2, the temperature-controlled ES-MS binding assay[1] was utilized in order to investigate the cooperativity in donor-acceptor binding. The presence of UDP and a metal cofactor Mn^{2+} in binding sites significantly enhanced the affinities of H-antigen acceptor to GTs at lower temperatures, suggesting positive cooperativity in donor-acceptor binding. Although the direct

determination of K_a at physiological temperature was not possible under the experimental conditions in this work, estimated K_a values from van't Hoff plots suggested non-cooperativity in donor-acceptor binding at physiological temperature. Further analysis revealed that both enthalpy (ΔH_a) and entropy (ΔS_a) of the association of H-antigen to GTs decreased significantly as binding sites are occupied by UDP and Mn^{2+}. This decrease in ΔH_a and ΔS_a were attributed to the conformational change of two disordered loops in GTs induced by the binding of both acceptor and donor.[2]

In Chapter 3, the first comparative thermodynamic study of GTA and GTB interacting with donor substrates, donor and acceptor analogues, and trisaccharide products *in vitro* was reported. The binding constants, measured at 24 °C with the *direct* ES-MS binding assay, provided new insights into these model GTs and their interactions with substrate and product. Notably, the recombinant soluble forms of GTA and GTB used in this study were shown to exist as homodimers, stabilized by non-covalent interactions at neutral pH. In the absence of Mn^{2+}, neither GTA nor GTB exhibited any appreciable affinity for their native donors, UDP-GalNAc and UDP-Gal, respectively. Upon introduction of Mn^{2+}, both donors underwent enzyme-catalyzed hydrolysis in the presence of either GTA or GTB. Hydrolysis of UDP-GalNAc in the presence of GTA proceeded very rapidly under the solution conditions investigated and a binding constant could not be directly measured. In contrast, the rate of hydrolysis of UDP-Gal in the presence of GTB was significantly slower and, utilizing a modified approach to analyze the ES-MS data, a binding constant of 2×10^4 M^{-1}

was established. GTA and GTB bound the donor analogues, UDP-GlcNAc and UDP-Glc, with affinities comparable to those of UDP-GalNAc and UDP-Gal (GTB only), suggesting that the native donors and donor analogues bind to the GTA and GTB through similar interactions. The binding constant determined for GTA and UDP-GlcNAc (~1×10^4 M^{-1}), therefore, provided an estimate for the binding constant for GTA and UDP-GalNAc. Binding of GTA and GTB with the A and B trisaccharide products was also investigated for the first time. In the absence of UDP and Mn^{2+}, both GTA and GTB recognized their respective trisaccharide products but with a low affinity ~10^3 M^{-1}; the presence of UDP and Mn^{2+} had no effect on A-trisaccharide binding but precludes B-trisaccharide binding.

In Chapter 4, direct detection of covalent glycosyl-enzyme intermediates for mutants of GTA and GTB by ES-MS was reported. Incubation of mutants of GTA or GTB, in which the putative catalytic nucleophile Glu^{303} was replaced with Cys (i.e. GTA_{E303C} and GTB_{E303C}), with their respective donor substrate resulted in a covalent intermediate. Tandem MS analysis using collision-induced dissociation confirmed Cys^{303} as the site of glycosylation. Exposure of the glycosyl-enzyme intermediates to a disaccharide acceptor resulted in the formation of the corresponding enzymatic trisaccharide products. These findings suggested that the GTA_{E303C} and GTB_{E303C} mutants may operate by a double displacement mechanism.

In Chapter 5, the time course of substrate cleavage by NEU3 was measured using an ES-MS enzyme assay to obtain relative rates (k_{rel}). NEU3

substrate activity was directly dependent upon the hydrophobicity of the aglycone, but had no apparent requirement for features of the ceramide head-group. In addition, the substrates with incorporated azide groups in the Neu5Ac residue at either *C*9 or the *N*5-Ac position were substrates, and in the case of the *N*5-azidoacetyl derivative the activity was superior to GM3. However, the incorporation of larger aryl groups was tolerated only at *C*9, but not at *N*5-Ac. In conclusion, a two-site model for enzyme recognition, requiring interaction at both the Neu5Ac residue and the hydrophobic aglycone was proposed.

6.2 Proposed Research Directions

6.2.1 Trapping Glycosyl-enzyme Intermediate by Wild-type GTs

In chapter 4, trapping and characterization of covalent glycosyl-enzyme intermediates of mutant GTA and GTB were described. Although this work supports the possibility of the double displacement mechanism in retaining glycosyltransferases, the direct detection of covalent intermediates of wild-type GTA and GTB will be stronger evidence for the mechanism. In our ES-MS analysis of the solutions containing wild-type GT, donor substrate and Mn^{2+}, ions corresponding to the glycosyl-enzyme intermediates of wild-type GTs have not been observed. This might suggest that the rate of deglycosylation steps is significantly greater than the rate of glycosylation step. If this is true, the accumulation of the covalent intermediates is required by reducing the rate of deglycosylation in order to detect them by ES-MS.

As mentioned in Chapter 4, retaining glycosidases are believed to operate

the double displacement mechanism involving a covalently bound glycosyl-enzyme intermediate (Figure 6.1). This is supported by the direct detection of their covalent intermediates by reducing the rate of deglycosylation step.[3,4] One of successful strategies to slow down the deglycosylation step of glycosyl-enzyme intermediates in retaining glycosidases is the application of fluorinated substrates.[3,4] Since the deglycosylation steps of retaining glycosidases are believed to proceed through oxocarbenium ion-like transition states, the substitution of an electronegative fluorine for a hydroxyl group at *C*2 should inductively destabilize both transition states, resulting in a decrease in the rates of deglycosylation steps.[3,4] The utility of fluorinated sugar-nucleotide donors[5] could be a possible approach to trap glycosyl-enzyme intermediates by wild-type GTA and GTB. Moreover, decreasing temperature of ES solutions using a temperature-controlled nanoES device[1] could be another approach to reduce the rate of deglycosylation step.

6.2.2 Development of a Novel Kinetic Assay using nanoES-MS

In Chapter 5, the application of nanoES-MS to enzyme kinetics of NEU3 was discussed. The determined K_m and V_{max} for 4MU-NA hydrolysis by the nanoES-MS kinetic assay were in good agreement with those using a fluorescence-based assay. Although the nanoES-MS kinetic assay has advantages over conventional spectroscopic and radiochemical assays, including that MS-based techniques do not require chromophoric or radioactive labeling, and substrates as well as products can be detected simultaneously,[6,7] the improvement

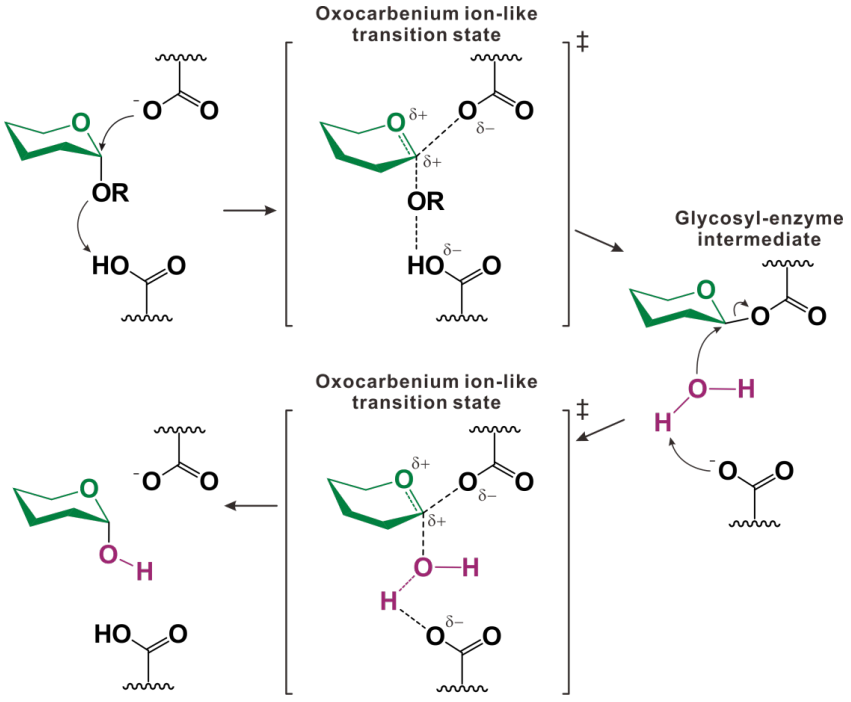

Figure 6.1 The double displacement mechanism of retaining glycosidases, which involves oxocarbenium ion-like transition state in glycosylation and deglycosylation steps, and occurs via a covalent glycosyl-enzyme intermediate.

of the assay is desirable. In particular, the nanoES-MS kinetic assay used in the present work requires the lengthy procedures for the sample preparations prior to analysis and cleaning of Pt-wire as well as a syringe at each measurement. These procedures are time-consuming.

Research efforts in our group would be directed towards the development of a novel on-line nanoES device, which is combined with nano-flow pump (or nanoliter syringe pump), minimized mixing chambers to mix enzyme and substrates and fused-silica capillary. The nano-flow pump will be used to deliver substrate and enzyme solutions to the reaction chamber, the enzymatic reaction will be initiated in the chamber. The mixture will be directly delivered to nanoES tip via fused-silica capillary, or the reaction will be quenched in a second mixing chamber by delivering the mixture and quenching solution to the second chamber before nanoES tip. The reacting time will be governed by the length of capillary and solution flow rate. We expect that the novel on-line nanoES device will facilitate kinetic studies with smaller amount of enzymes and substrates.

6.3 Literature Cited

(1) Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *J. Am. Chem. Soc*. **2004**, *126*, 4786–4787.

(2) Alfaro, J. A.; Ruixiang, Z. B.; Persson, M.; Letts, J. A.; Polakowski, R.; Bai, Y.; Borisova, S. N.; Seto, N. O. L.; Lowary, T.; Palcic, M. M.; Evans, S. V. *J. Biol. Chem*. **2008**, *283*, 10097−10108.

(3) Mosi, R. M.; Withers, S. G. *Methods Enzymol.* **2002**, *354*, 64–84.

(4) Wicki, J.; Rose, D. R.; Withers, S. G. *Methods Enzymol.* **2002**, *354*, 84−105.

(5) Burkart, M. D.; Vincent, S. P.; Düffels, A.; Murray, B. W.; Ley, S. V.; Wong, C. *Bioorg. Med. Chem.* **2000**, *8*, 1937−1946.

(6) Liesener, A.; Karst, U. *Anal. Bioanal. Chem.* **2005**, *382*, 1451−1464.

(7) Shipovskov, S.; Reimann, C. *Analyst* **2007**, *132*, 397−402.