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

Naoto Soya,[2] Glen K. Shoemaker,[2] Monica M. Palcic[3] and John S. Klassen[1,2]

[1] Author to whom correspondence should be addressed

[2] Alberta Ingenuity Centre for Carbohydrate Science and Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

[3] Carlsberg Laboratory, Gamle Carlsberg Vej 10, 2500 Valby, Denmark

Keywords: protein-oligosaccharide complexes/association constants/protein dimerization/human ABO(H) blood group glycosyltransferases/electrospray ionization mass spectrometry

Abstract

The first comparative thermodynamic study of the human blood group glycosyltransferases, α-(1→3)-N-acetylgalactosaminyltransferase (GTA) and α-(1→3)-galactosyltransferase (GTB), interacting with donor substrates, donor and acceptor analogues, and trisaccharide products *in vitro* is reported. The binding constants, measured at 24 °C with the direct electrospray ionization mass spectrometry (ES-MS) assay, provide new insights into these model GTs and their interactions with substrate and product. Notably, the recombinant forms of GTA and GTB used in this study are shown to exist as homodimers, stabilized by non-covalent interactions at neutral pH. In the absence of divalent metal ion, neither GTA nor GTB exhibit any appreciable affinity for their native donors (UDP-GalNAc, UDP-Gal). Upon introduction of Mn^{2+}, both donors undergo enzyme-catalyzed hydrolysis in the presence of either GTA or GTB. Hydrolysis of UDP-GalNAc in the presence of GTA proceeds 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 is significantly slower and, utilizing a modified approach to analyze the ES-MS data, a binding constant of 2 x 10^4 M^{-1} was established. 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. The binding constant determined for GTA and UDP-GlcNAc (~1 x 10^4 M^{-1}), therefore, provides 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

recognize their respective trisaccharide products but with a low affinity ~10^3 M^{-1}; the presence of UDP and Mn^{2+} has no effect on A trisaccharide binding but precludes B-trisaccharide binding.

Introduction

Glycosyltransferases (GTs) represent a large and diverse group of enzymes of which there is over 200 known sequences in humans (Breton, *et al*. 2001). 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 (Crocker and Feize 1996), bacterial, viral, and fungal infections (Sacks and Kamhawi 2001), inflammation (Lowe 2003), tumorigenesis (Couldrey and Green 2000, Hakomori 2002, Hakomori and Kannagi 1983, Kim and Varki 1997, Laferte and Dennis 1988, Saitoh, *et al*. 1992) as well as many other disorders (Furukawa and Okajima 2002, Hanßke, *et al*. 2002, Hounsell and Feizi 1982). 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 (Kleene and Berger 1993, Lazar and Walker 2002).

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 α-L-Fuc*p*-(1→2)-β-D-Gal*p*-OR (H-disaccharide, **1**) terminated oligosaccharides (where R is a glycoprotein or glycolipid). 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) (Yamamoto, *et al*. 1990), 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 (**6**) whereas GTB transfers a Gal residue from a uridine 5′-diphosphogalactose (UDP-Gal, **10**) donor forming the B trisaccharide (**7**). 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, (Alfaro, *et al*. 2008, Lee, *et al*. 2005, Letts, *et al*. 2006, Nguyen, *et al*. 2003, Patenaude, *et al*. 2002) kinetic analysis, (Marcus, *et al*. 2003, Seto, *et al*. 1997) and NMR (Angulo, *et al*. 2006, Blume, *et a*l. 2006). 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-denaturing gel electrophoresis (Lee, *et al*. 2005). 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* (Daniel, *et al.* 2002, Wang *et al*. 2003, Yu, *et al*. 2006). Recently, our laboratories investigated the binding of **1** with recombinant soluble forms of GTA and GTB using the direct ES-MS assay (Shoemaker, *et al*. 2008). Association constants (K_a) were measured for **1** binding to GTA and GTB homodimers, in the absence and presence of uridine 5′-diphosphate (UDP, **8**) 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 **8** 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 **8** were attributed to structural changes in the GTs induced by the simultaneous binding of **1** and **8**.

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 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.

Materials and methods

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 (Marcus, *et al*. 2003, Seto, *et al*. 1995, Seto, *et al*. 1997). 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 molecular weight cut-off 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 S1, Supplementary Materials). 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, 2** and **5**) used in this study were obtained from Dr. Ole Hindsgaul (Carlsberg Laboratory) while compounds **6** and **7** were enzymatically synthesized from compound **1** using GTA and GTB (Seto, *et al*. 1995, Seto, *et al*. 1997). The monosaccharides (**3** and **4**), the native donor substrates (**9** and **10)**, the donor fragment (**8**) and donor analogues (uridine 5′-diphospho-N-acetylglucosamine, **11** and uridine 5′-diphosphoglucose, **12**) were purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). The structures of the ligands are given in Figure 1. A single chain variable fragment, scFv (26 539 Da), of the carbohydrate-binding IgG antibody of Se155-4 was produced using recombinant technology (Zdanov, *et al*. 1994) and was used as a reference protein in the ES-MS binding measurements.

Mass Spectrometry

ES-MS measurements were performed using an Apex II 9.4 tesla Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker, Billerica, MA) equipped with a custom built temperature-controlled nanoflow ES (nanoES) device (Daneshfar, *et al*. 2004). 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 ~$5x10^{-10}$ 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.

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:

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

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

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:

$$[GT_2]_o = [GT_2] + [GT_2 + L] + [GT_2 + 2L] \tag{2a}$$

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

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:

$$K_{a,1} = \frac{R_1}{[L]_o - \frac{(R_1 + 2R_2)[GT_2]_o}{1 + R_1 + R_2}} \tag{3a}$$

$$K_{a,2} = \frac{R_2}{R_1([L]_o - \frac{(R_1 + 2R_2)[GT_2]_o}{1 + R_1 + R_2})} \tag{3b}$$

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:

$$R_1 = \frac{\sum_{n} Ab(GT_2 + L)^{n+}}{\sum_{n} Ab(GT_2)^{n+}} \quad (4a)$$

$$R_2 = \frac{\sum_{n} Ab(GT_2 + 2(L))^{n+}}{\sum_{n} Ab(GT_2)^{n+}} \quad (4b)$$

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 $K_{a,2}$:

$$K_a = \frac{K_{a,1}}{2} \quad (5a)$$

$$K_a = 2 \times K_{a,2} \quad (5b)$$

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 **8**, *vide infra*. Prior to complete hydrolysis, the solution will contain both donor and **8**, 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 **8**, or to one molecule of donor and **8** may exist in solution.

The corresponding equations of mass balance are given below, where donor is represented as L_1 and **8** as L_2:

$$[GT_2]_o = [GT_2] + [GT_2 + L_1] + [GT_2 + L_2] + [GT_2 + L_1 + L_2] + [GT_2 + 2L_1] + [GT_2 + 2L_2] \quad (6a)$$

$$[L_1]_o = [L_1] + [L_2] + [GT_2 + L_1] + [GT_2 + L_2] + 2[GT_2 + L_1 + L_2] + 2[GT_2 + 2L_1] + 2[GT_2 + 2L_2] \quad (6b)$$

The corresponding K_a expressions for donor binding are given below:

$$K_{a,1} = \frac{R_{1,0}}{[L]_o - [L_2] - \frac{(R_{1,0} + R_{0,1} + 2R_{1,1} + 2R_{2,0} + 2R_{0,2})[GT_2]_o}{1 + R_{1,0} + R_{0,1} + R_{1,1} + R_{2,0} + R_{0,2}}} \quad (7a)$$

$$K_{a,2} = \frac{R_{2,0}}{R_{1,0}([L]_o - [L_2] - \frac{(R_{1,0} + R_{0,1} + 2R_{1,1} + 2R_{2,0} + 2R_{0,2})[GT_2]_o}{1 + R_{1,0} + R_{1,0} + R_{1,1} + R_{2,0} + R_{0,2}})} \quad (7b)$$

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:

$$R_{i,j} = \frac{\sum_n Ab(GT_2 + iL_1 + jL_2)^{n+}}{\sum_n Ab(GT_2)^{n+}} \quad (8)$$

In order to solve eq 7, $[L_2]$, the concentration of free **8** 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 (**8**) to GT_2, using eq 9:

$$[L_2] = \frac{R_{0,1}}{K_{a,1}(L_2)} \quad (9)$$

The $K_{a,1}(L_2)$ value was determined in a separate measurement performed on solutions of GT_2 with **8** alone.

The reference protein method was used to correct the ES mass spectra for the occurrence of nonspecific protein-ligand binding (Sun, *et al*. 2006). 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 **8**, is independent of the size and structure of the protein (Wang, *et al*. 2005; Sun, *et al*. 2006; Sun, *et al.* manuscript in preparation). 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.

Results and Discussion

Native oligomeric form of the recombinant GTA and GTB at neutral pH

As highlighted in a recent review (Seko 2006), various GTs are known to undergo dimerizaton 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.* (Nagai, *et al*. 1978, Nagai, *et al*. 1978). 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 (Lee, *et al*. 2005). 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 (Lee, *et al*. 2005). 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 (Heck and van den Heuvel 2004, Loo 1997, Sobott, *et al*. 2002). 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 2a and 2b 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 2 and S2). 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, 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 2c and 2d). 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 centred at higher charge state corresponds to the fully-denatured monomer (Lafitte, *et al*. 1999). 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 2e and 2f). 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.

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) (Shoemaker, *et al*. 2008). 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 intermolecular interactions between both GTs and **1** are identical, with the exception of a putative van der Waals between **1** and Met 266 in GTB (Patenaude, *et al*. 2002). There is no such contact between **1** and Leu 266 of GTA. As noted above, the presence of **8** 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 (Shoemaker, *et al*. 2008). The enhanced in the binding was attributed to structural changes in the GTs induced by the simultaneous binding of **1** and **8**.

To further investigate the nature of acceptor substrate recognition by GTA and GTB, affinities for a disaccharide acceptor analogue (**2**) and several monosaccharides (**3-5**), in the absence and presence of **8** 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 **2 - 5** in the absence and presence of **8** and Mn^{2+} are shown in Figures S3 and

S4 (Supplementary materials) and in Figures 3, respectively. The intrinsic binding constants determined from the ES-MS measurements are listed in Table 1.

The disaccharide **2** 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 (Lowary and Hindsgaul 1993). According to the crystal structure of (GT + **1**) complexes (Patenaude, *et al*. 2002), 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 **2** engages in the same binding interactions with the GTs as does **1** (Nguyen, *et al*. 2003). Not surprisingly, then, the intrinsic binding constants measured for **2** 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 S1b and S2b). These results support the hypothesis that **1** and **2** 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 **2** was measured upon introduction of **8** and Mn^{2+} into the GT active sites (Figures 3a and 3b). These higher affinities are consistent with the potent inhibitory effect of the disaccharide established from kinetic measurements (Lowary and Hindsgaul, 1993). Our laboratory is currently exploring the origin of the high affinities determined for **2** using molecular modeling tools.

According to a structural analysis carried out by Letts *et al*, Gal (**3**) can bind to the acceptor sites of GTA and GTB, and the intermolecular interactions are similar to

those involving the Gal residue in **1** (Letts, *et al*. 2006). The exception being that Glu 303 of GTB interacts with O-3 and O-4 hydroxyl group of **3** 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 **3** and Glu 303. According to the ES-MS measurements, GTA and GTB exhibit only a weak affinity for **3**, $(5.2 \pm 0.8) \times 10^2\ M^{-1}$ and $(4 \pm 2) \times 10^2\ M^{-1}$, respectively (Figures S3c and S4c). 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 (**4**), no specific binding with GTA or GTB was detected (Figures S3d and S4d). 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 **5**, 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 S4e) 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 S3e). These results may reflect the ability of the amino group of **5** to form a direct interaction with GTB, but not GTA. However, this hypothesis remains to be verified experimentally or computationally.

Interestingly, the presence of **8** and Mn^{2+} in the donor sites has non-uniform effects on the binding of **3** – **5**. In the case of **3** and **4**, no specific binding to the GTs was detected (Figures 3c – 3f), while for **5**, the presence of **8** and Mn^{2+} resulted in a profound enhancement in binding (Figures 3g and 3h). 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) x 10^6 M^{-1}). These results suggest that **5** could serve as an effective inhibitor of the GTs, particularly for GTB.

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 (Yamamoto and Hakomori 1990, Yamatomo and McNeill 1996, Seto, *et al*. 1999). 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** (Patenaude *et al*. 2002). 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+} (Blume, *et al*. 2006). 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 (**8** - **12**), **8** 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 x 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 (Seto, *et al*. 1999).

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 assay was used to quantify the interactions of **8 - 12** with the GTA and GTB, in the absence and presence of Mn^{2+} (Table 2).

Shown in Figure 4 are illustrative ES mass spectra obtained for solutions of GTA and GTB, respectively, with **8**. In the absence of Mn^{2+}, neither of the GTs exhibits a measurable affinity for **8** (Figure 4a and 4b). However, upon addition of excess Mn^{2+} (100 μM) both GTs readily bind to **8** and ions corresponding to $(GT_2 + \mathbf{8} + Mn)^{n+}$ and $(GT_2 + 2(\mathbf{8} + Mn))^{n+}$ complexes were detected (Figure 4c and 4d). The intrinsic affinities of GTA and GTB for **8** (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*. (Blume, *et al*. 2006). 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 **8** by the GTs through a DXD motif (Patenaude *et al*. 2002). Considering the donor specificity of GTA and GTB, it is interesting that the GTs would exhibit moderately strong binding affinity for **8** (complexed with Mn^{2+}). This is a structural element of native sugar-nucleotide donors and the product of enzymatic reaction of GTs and, in principle, **8** could significantly inhibit the enzymatic reaction. However, as described below, the presence of **8** 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 **8** in the absence of Mn^{2+}, they do possess a weak affinity for their native donors. Shown in Figures S5a and S6b 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 S5b and S6a). 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** (Patenaude, *et al*. 2002), 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.

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 **8**, 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 5a 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 $(GTA_2 + \mathbf{8} + Mn)^{n+}$ and $(GTA_2 + 2(\mathbf{8} + Mn))^{n+}$ ions. From the fraction of bound GTA_2 and the known binding constant for **8**, 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 6a and 6b 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 6a). However, after ~25 minutes, only ions corresponding to bound **8** were detected, indicating that **10** was completely hydrolyzed (Figure 6b). 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 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) (Marcus, *et al*. 2003).

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 **8** and the $K_{a,1}$ and $K_{a,2}$ values measured for **8** binding to GTB. Using this procedure, the affinity for GTB with **10** was determined to be (2.0 ± 0.7) x 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 **8** (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 5b and 6c).

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 **8** 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 (Seto, *et al*. 2000). Unlike the native donors, the donor analogues do not undergo hydrolysis in the presence of GTs (Figure 5c and 6d). The measured affinities of GTA and GTB for **11** are (1.3 ± 0.3) x 10^4 M^{-1} and (1.8 ± 0.2) x 10^3 M^{-1}, respectively, while the measured affinities for **12** are (2.0 ± 0.7) x 10^4 M^{-1} and (1.2 ± 0.3) x 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 x 10^4 M^{-1}.

In contrast to the agreement in the binding constants determined for **8**, 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*. (Blume, *et al*. 2006). 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** (4×10^4 M^{-1}) and **10** (5×10^3 M^{-1}). It is possible that the remaining differences are due to partial donor hydrolysis. Because **8** 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.

Trisaccharide binding

The A (**6**) and B (**7**) trisaccharides are the products of their respective enzymatic reactions. Shown in Figures S7a and S7b 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 **6** and **7** were clearly detected demonstrating the enzymes maintain their activity under mass spectrometry conditions used. The affinities of GTA and GTB for the **6** and **7** were also determined by the ES-MS assay and the values are listed in Table 1. Illustrative ES mass spectra acquired for solutions of GTA and GTB with **6** and **7** are shown in Figures S8a,b and S9a,b), respectively. In the absence of **8** and Mn^{2+}, GTA binds only weakly to **6** ($(1.4 \pm 0.4) \times 10^3$ M^{-1}) (Figure S8a), while GTB

exhibits no measurable affinity for the trisaccharide (Figure S9a). In contrast, both GTA and GTB exhibit an affinity for **7** (Figures S8b and S9b). Curiously though, the affinity of GTA ((1.7 ± 0.2) x 10^4 M^{-1}) for **7** is four-fold higher than that of GTB ((2.7 ± 0.3) x 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（Patenaude, *et al*. 2002), which could explain why the GTA binding pocket can accommodate both trissacharide products but GTB can't. These findings, on their own, suggest that trisaccharide B (**7**) could serve as an inhibitor for GTA. However, as described below, this potential inhibitory effect is mitigated by the presence of **8** in the active site.

As interesting and previously unexplored question is whether the two enzymatic products, A or B trisaccharide and **8** (and metal cofactor), can bind 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, **8** and Mn^{2+}. Illustrative mass spectra of solutions of GTA and GTB with **6** or **7** in the absence or presence of **8** (and Mn^{2+}) are shown in Figure 7 and 8, respectively. Importantly, the ES-MS data reveal that the presence of **8** in the active sites prevents appreciable association (K_a <10^3 M^{-1}) of either trisaccharide to GTB (Figures 8c and 8d). Similarly, GTA was found not to bind significantly to **7** in the presence of **8** and Mn^{2+} (Figure 7d). However, the ES-MS data clearly show that GTA is able to bind simultaneously to **6**, **8** and Mn^{2+} (Figure 7c), though the binding is considerably weaker ((1.2 ± 0.2) x 10^3 M^{-1}) than that of the acceptor substrate (**1**) ((1.6 ± 0.3) x 10^5 M^{-1}). In fact, the affinities measured for **6** in the presence and absence of **8** are identical, within error. In order to better understand the

differential effects of **8** (and metal cofactor) on GTA and GTB recognition of the trisaccharide products, a molecular modeling study has been undertaken.

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 assay, yield a number of new insights into these models GTs and their interactions with substrate.

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 (**3**) or GalNAc (**4**), but bind very strongly to a derivative of Gal (**5**) in which the 3′ hydroxyl group has been replaced with an amino group. In fact, when UDP (**8**) 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 **8**. 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 (**6**, **7**) to the GTs was investigated for the first time. In the absence of **8** 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 **8** and Mn^{2+}.

Acknowledgements

The authors acknowledge the Natural Sciences and Engineering Research Council of Canada and the Alberta Ingenuity Centre for Carbohydrate Science for generous funding.

Abbreviations

ES-MS, electrospray ionization mass spectrometry; FT-ICR, Fourier-transform ion cyclotron resonance; GTA, α-(1→3)-N-acetylgalactosaminyltransferase; GTB, α-(1→3)-galactosyltransferase; nanoES, nanoflow electrospray ionization; scFv, single chain variable fragment; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; 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.

References

Alfaro, JA., Zheng, RB., Persson, M., Letts, JA., Polakowski, R., Bai, Y., Borisova, SN., Seto, NOL., Lowary, T., Palcic, MM. and Evans, SV. 2008. ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes. *J Biol Chem*. 283:10097-10108.

Angulo, J., Langpap, B., Blume, A., Biet, T., Meyer, B., Krishna, NR., Peters, H., Palcic, MM. and Peters, T. 2006. Blood group B galactosyltransferase: Insights into substrate binding from NMR experiments. *J Am Chem Soc*. 128:13529-13538.

Breton, C., Mucha, J., Jeanneau, C. 2001. Structural and functional features of glycosyltransferases. *Biochimie*. 83:713-718.

Blume, A., Angulo, J., Biet, T., Peters, H., Benie, AJ., Palcic, MM., Peters, T. 2006. Fragment-based screening of the donor substrate specificity of human blood group B galactosyltransferase using saturation transfer difference NMR. *J Biol Chem*. 281:32728-32740.

Couldrey, C., Green, JE. 2000. Metastases: the glycan connection. *Breast Cancer Res*. 2:321-323.

Crocker, PR., Feize, T. 1996. Carbohydrate recognition systems: functional triads in cell-cell interactions. *Curr Opin Struct Biol*. 6:679-691.

Daneshfar, R., Kitova, EN., Klassen, JS. 2004. Determination of protein-ligand association thermochemistry using variable-temperature nanoelectrospray mass spectrometry. *J Am Chem Soc*. 126:4786-4787.

Daniel, JM., Friess, SD., Rajagopalan, S., Wendt, S., Zenobi, R. 2002. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *Int J Mass Spectrom*. 216:1-27

Furukawa, K., Okajima, T. 2002. Galactosyltransferase I is a gene responsible for progeroid variant of Ehlers-Danlos syndrome: molecular cloning and identification of mutations. *Biochim Biophys Acta*. 1573:377-381.

Hakomori, S. 2002. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc Natl Acad Sci USA*. 99:10231-10233.

Hakomori, S., Kannagi, R. 1983. Glycosphingolipids as tumor-associated and differentiation markers. *J Natl Cancer Inst*. 71:231-251.

Hanßke, B., Thiel, C., Lübke, T., Hasilik, M., Höning, S.; Peters, V., Heidemann, PH., Hoffmann, GF., Berger, EG., Figura, KV., Körner, C. 2002. Deficiency of UDP-galactose:N-acetylglucosamine β-1,4-galactosyltransferase I causes the congenital disorder of glycosylation type IId. *J Clin Invest*. 109:725-733.

Heck, AJR., Van Den Heuvel, RHH. 2004. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom Rev*. 23:368-389.

Hounsell, EF., Feizi, T. 1982. Gastrointestinal mucins. Structures and antigenicities of their carbohydrate chains in health and disease. *Med Biol* 60:227-236.

Kim, YJ., Varki, A. 1997. Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glycoconj J*. 14:569-576.

Kleene, R., Berger, EG. 1993. The molecular and cell biology of glycosyltransferases. *Biochim Biophys Acta*. 1154:283-325.

Laferte, S., Dennis, JW. 1988. Glycosylation-dependent collagen-binding activities of two membrane glycoproteins in MDAY-D2 tumor cells. *Cancer Res*. 48:4743-4748.

Lafitte, D., Heck, AJR., Hill, TJ. Jumel, K., Harding, SE., Derrick, PJ. 1999. Evidence of noncovalent dimerization of calmodulin. *Eur J Biochem*. 261:337-344.

Lazar, K., Walker, S. 2002. Substrate analogues to study cell-wall biosynthesis and its inhibition. *Curr Opin Chem Biol*. 6:786-793.

Lee, HJ., Barry, CH., Borisova, SN., Seto, NOL., Zheng, RXB., Blancher, A., Evans, SV., Palcic, MM. 2005. Structural basis for the inactivity of human blood group O^2 glycosyltransferase. *J Biol Chem*. 280:525-529.

Letts, JA., Rose, NL., Fang, YR., Barry, CH., Borisova, SN., Seto, NOL., Palcic, MM., Evans, SV. 2006. Differential recognition of the type I and II H antigen acceptors by the human ABO(H) blood group A and B glycosyltransferases. *J Biol Chem*. 281:3625-3632.

Loo, JA. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom Rev*. 16:1-23.

Lowary, TL., Hindsgaul, O. 1993. Recognition of synthetic deoxy and deoxyfluoro analogs of the acceptor α-L-Fuc*p*-(1→2)-β-D-gal*p*-OR by the blood-group A and B gene-specified glycosyltransferases. *Carbohydr Res*. 249:163-195.

Lowe, JB. 2003. Glycan-dependent leukocyte adhesion and recruitment in inflammation. *Curr Opin Cell Biol*. 15:531-538.

Marcus, SL., Polakowski, R., Seto, NOL., Leinala, E., Borisova, S., Blancher, A., Roubinet, F., Evans, SV., Palcic, MM. 2003. A single point mutation reverse the

donor specificity of human blood group B-synthesizing galactosyltransferase. *J Biol Chem*. 278:12403-12405.

Nagai, M., Davè, V., Kaplan, BE., Yoshida, A. 1978. Human blood group glycosyltransferases: I. Purification of *N*-acetylgalactosaminyltransferase. *J Biol Chem*. 253:377-379.

Nagai, M., Davè, V., Muensch, H., Yoshida, A. 1978. Human blood group glycosyltransferases: II. Purification of galactosyltransferase. *J Biol Chem*. 253:380-381.

Nguyen, HP., Seto, NOL., Cai, Y., Leinala, EK., Borisova, SN., Palcic, MM., Evans, SV. 2003. The influence of an intramolecular hydrogen bond in differential recognition of inhibitory acceptor analogs by human ABO(H) blood group A and B glycosyltransferases. *J Biol Chem*. 278:49191-49195.

Patenaude, SI., Seto, NOL., Borisova, SN., Szpacenko, A., Marcus, SL., Palcic, MM., Evans, SV. 2002. The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat Struct Biol*. 9:685-690.

Sacks, D., Kamhawi, S. 2001. Molecular aspects of parasite-vector and vector-host interactions in Leishmaniasis. *Annu Rev Microbiol*. 55:453-483.

Saitoh, O., Wang, WC., Lotan, R., Fukuda, M. 1992. Differential glycosylation and cell surface expression of lysosomal membrane glycoproteins in sublines of a human colon cancer exhibiting distinct metastatic potentials. *J Biol Chem*. 267:5700-5711.

Seko, A. 2006. Complex formation of glycosyltransferases and their biological significance. *Trends Glycosci Glycotechnol*. 18:209-230.

Seto, NOL., Compston, CA., Evans, SV., Bundle, DR., Narang, S. A., Palcic, MM. 1999. Donor substrate specificity of recombinant human blood group A, B and hybrid A/B glycosyltransferases expressed in *Escherichia coli*. *Eur J Biochem*. 259:770-775.

Seto. NOL., Compston, CA., Szpacenko, A., Palcic, MM. 2000. Enzymatic synthesis of blood group A and B trisaccharide analogues. *Carbohydr Res*. 324:161-169.

Seto, NOL., Palcic, MM., Compston, CA., Li, H., Bundle, DR., Narang, SA. 1997. Sequential interchange of four amino acids from blood group B to blood group A glycosyltransferase boosts catalytic activity and progressively modifies substrate recognition in human recombinant enzymes. *J Biol Chem*. 272:14133-14138.

Seto, NOL., Palcic, MM., Hindsgaul, O., Bundle, DR., Narang, SA. 1995. Expression of a recombinant human glycosyltransferase from a synthetic gene and its utilization for synthesis of the human blood group B trisaccharide. *Eur J Biochem*. 234:323-328.

Shoemaker, GK., Soya, N., Palcic, MM., Klassen, JS. 2008. Temperature-dependent cooperativity in donor-acceptor substrate binding to the human blood group glycosyltransferases. *Glycobiology*. 18:587-592.

Sobott, F., Hernandez, H., McCammon, MG., Tito, MA., Robinson, CV. 2002. A tandem mass spectrometer for improved tansmission and analysis of large macromolecular assemblies. *Anal Chem*. 74:1402-1407.

Sun, JX., Kitova, EN., Wang, WJ., Klassen, JS. 2006. Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal Chem*. 78:3010-3018.

Sun, JX., Kitova, EN., Sun, N., Klassen, JS. 2007. Method for identifying nonspecific protein-protein interactions in nanoelectrospray ionization mass spectrometry. *Anal Chem*. 79:8301-8311.

Wang, WJ., Kitova, EN., Klassen, JS. 2003. Influence of solution and gas phase processes on protein-carbohydrate binding affinities determined by nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem*. 75:4945-4955.

Wang, WJ., Kitova, EN., Klassen, JS. 2005. Nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization. Factors influencing their formation and stability. *Anal Chem*. 77:3060-3071.

Yamamoto, F., Clausen, H., White, T., Marken, J., Hakomori, S. 1990. Molecular genetics basis of the histo-blood group ABO system. *Nature*. 345:229-233.

Yamamoto, F., Hakomori, S. 1990. Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. *J Biol Chem*. 265:19257-19262.

Yamamoto, F., McNeill, PD. 1996. Amino acid residue at codon 268 determines both activity and nucleotide-sugar donor substrate specificity of human histo-blood group A and B transferases. *J Biol Chem*. 271:10515-10520.

Yu, Y., Sweeney, MD., Saad, OM., Leary, JA. 2006. Potential inhibitors of chemokine function: analysis of noncovalent complexes of CC chemokine and small polyanionic molecules by ESI FT-ICR mass spectrometry. *J Am Soc Mass Spectrom*. 17:524-535.

Zdanov, A., Li, Y., Bundle, DR., Deng, SJ., Mackenzie, CR., Narang, SA., Young, NM., Cygler, M. 1994. Structure of a single-chain antibody variable domain (Fv) fragment complexed with a carbohydrate antigen at 1.7-Å resolution. *Proc Natl Acad Sci USA*. 91:6423-6427.

Table 1. Intrinsic association constants (K_a) for acceptor substrate (**1**), acceptor analogues (**2** - **5**) and A-, B-trisaccharides (**6**, **7**) binding with GTA_2 and GTB_2 in the absence and presence of 50 μM UDP (**8**) 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]

Substrate	**GTA** K_a **(M^{-1})**	**GTA + 8 + Mn^{2+}** K_a **(M^{-1})**	**GTB** K_a **(M^{-1})**	**GTB + 8 + 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]
2	$1.0 \pm 0.2 \times 10^4$	$2.4 \pm 0.3 \times 10^6$	$9.1 \pm 0.9 \times 10^3$	$1.9 \pm 0.9 \times 10^6$
3	$5.2 \pm 0.8 \times 10^2$	NB [c]	$4 \pm 2 \times 10^2$	NB [c]
4	NB [c]	NB [c]	NB [c]	NB [c]
5	$1.0 \pm 0.4 \times 10^3$	$1.4 \pm 0.5 \times 10^5$	$1.7 \pm 0.7 \times 10^4$	$4.0 \pm 0.9 \times 10^6$
6	$1.4 \pm 0.4 \times 10^3$	$1.2 \pm 0.2 \times 10^3$	NB [c]	NB [c]
7	$1.6 \pm 0.1 \times 10^4$	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 Shoemaker, *et al*. 2008.

Table 2. Association constants (K_a) for GTA_2 and GTB_2 binding with UDP (**8**) 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]

Substrate	GTA K_a (M^{-1})	GTA + Mn^{2+} K_a (M^{-1})	GTB K_a (M^{-1})	GTB + Mn^{2+} K_a (M^{-1})
8	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 Blume, *et al*. 2006.

g. Value taken from reference from Seto, *et al*. 1999.

h. Values taken from reference from Marcus, *et al*. 2003.

Figure captions

Figure 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**), α-L-Fuc*p*-(1→2)-β-D-3-deoxy-Gal*p*-O-$(CH_2)_7CH_3$ (**2**), D-galactose (**3**) N-acetyl-D-galactosamine (**4**), 3-amino-3-deoxy-β-D-Gal*p*-O-R (**5**), α-D-GalNAc*p*-(1→3)[α-L-Fuc*p*-(1→2)]-β-D-Gal*p*-O-$(CH_2)_7CH_3$ (**6**), α-D-Gal*p*-(1→3)[α-L-Fuc*p*-(1→2)]-β-D-Gal*p*-O-$(CH_2)_7CH_3$ (**7**), uridine 5′-diphosphate (**8**), uridine 5′-diphospho-N-acetylgalactosamine (**9**), uridine 5′-diphosphogalactose (**10**), uridine 5′-diphospho-N-acetylglucosamine (**11**), uridine 5′-diphosphoglucose (**12**).

Figure 2. 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. Illustrative nanoES mass spectra of aqueous solutions consisting of **8** (50 μM) and 100 μM Mn^{2+} and (a) GTA_2 with **2** (5 μM), (b) GTB_2 with **2** (5 μM), (c) GTA_2 with **3** (100 μM), (d) GTB_2 with **3** (100 μM), (e) GTA_2 with **4** (100 μM), (f) GTB_2 with **4** (100 μM), (g) GTA_2 with **5** (20 μM) or (h) GTB_2 with **5** (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*.

Figure 4. Illustrative nanoES mass spectra of solutions consisting of (a) GTA_2 with **8** (40 μM) in the presence of EDTA (40 μM), (b) GTB_2 with **8** (40 μM) in the

presence of EDTA (40 μM), (c) GTA_2 with **8** (40 μM) in the presence of Mn^{2+} (100 μM) or (d) GTB_2 with **8** (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 **8** bound to the GT ions is indicated by *q*.

Figure 5. 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 6. 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*.

Figure 7. Illustrative nanoES mass spectrum of a solution consisting of GTA_2 (9 μM) with (a) **6** (50 μM) and (b) **7** (30 μM) or GTA_2 (9 μM), **8** (50 μM) and Mn^{2+}

(100 μM) with (c) **6** (50 μM) and (d) **7** (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 8. Illustrative nanoES mass spectrum of a solution consisting of GTB_2 (9 μM) with (a) **6** (30 μM) and b) **7** (30 μM), or GTB_2 (9 μM), **8** (50 μM) and Mn^{2+} (100 μM) with (c) **6** (30 μM) and (d) **7** (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*.

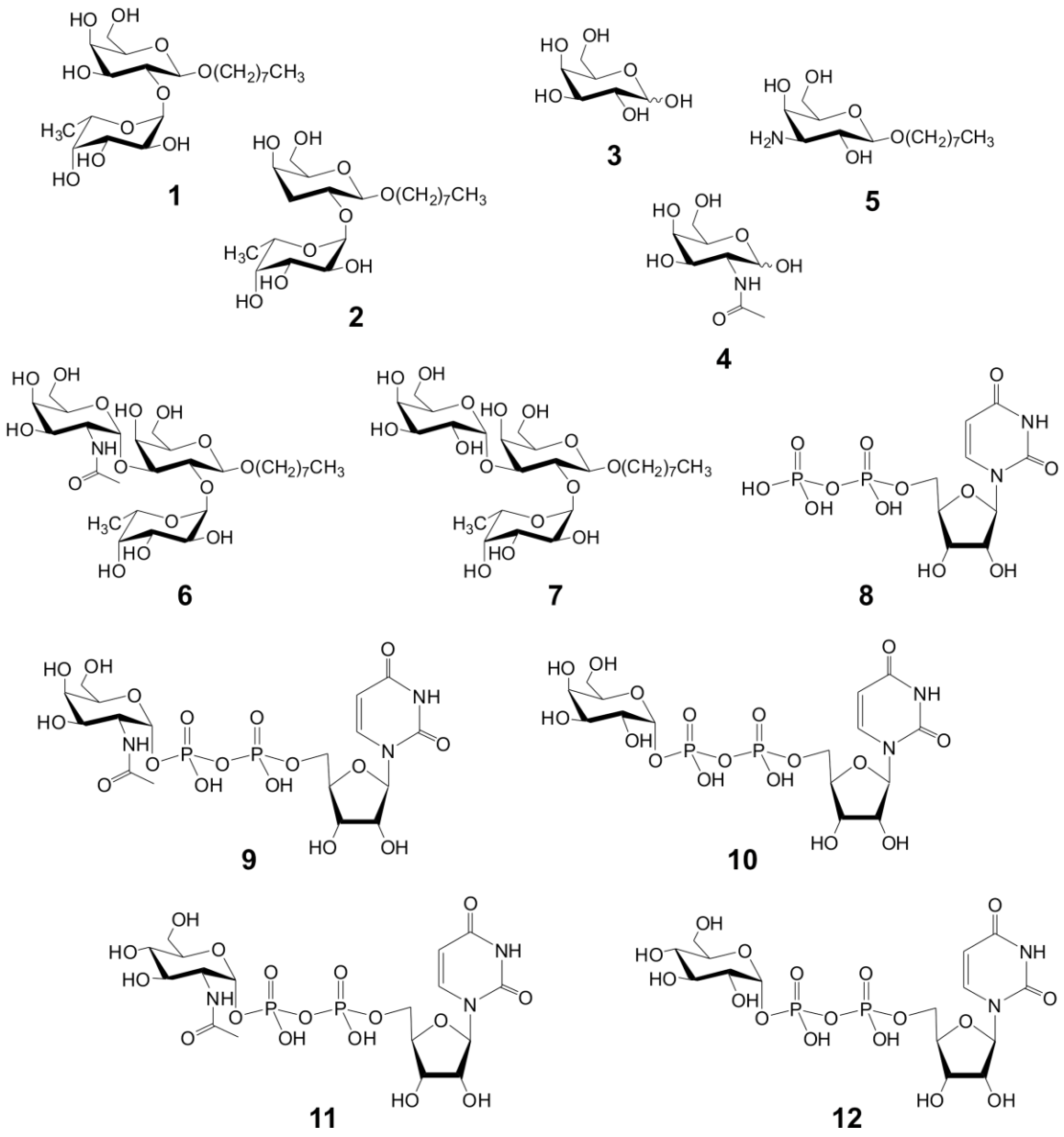

Figure 1

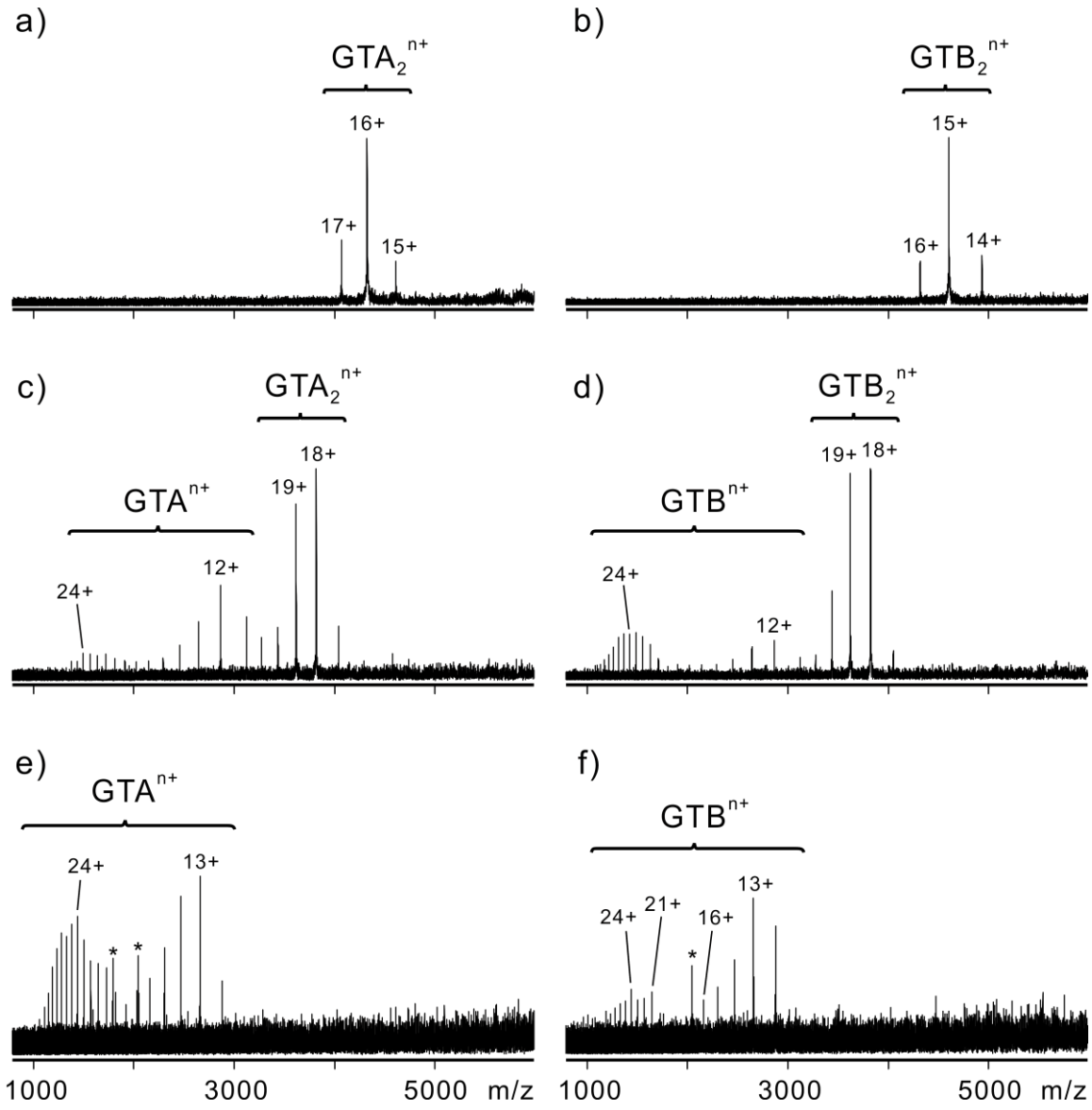
a)
GTA2 n+
16+
17+
15+
b)
GTB2 n+
15+
16+
14+
c)
GTA2 n+
18+
19+
GTA n+
12+
24+
d)
GTB2 n+
19+
18+
GTB n+
24+
12+
e)
GTA n+
24+
13+
*
*
f)
GTB n+
24+
21+
16+
13+
*
1000
3000
5000 m/z
1000
3000
5000 m/z

Figure 2

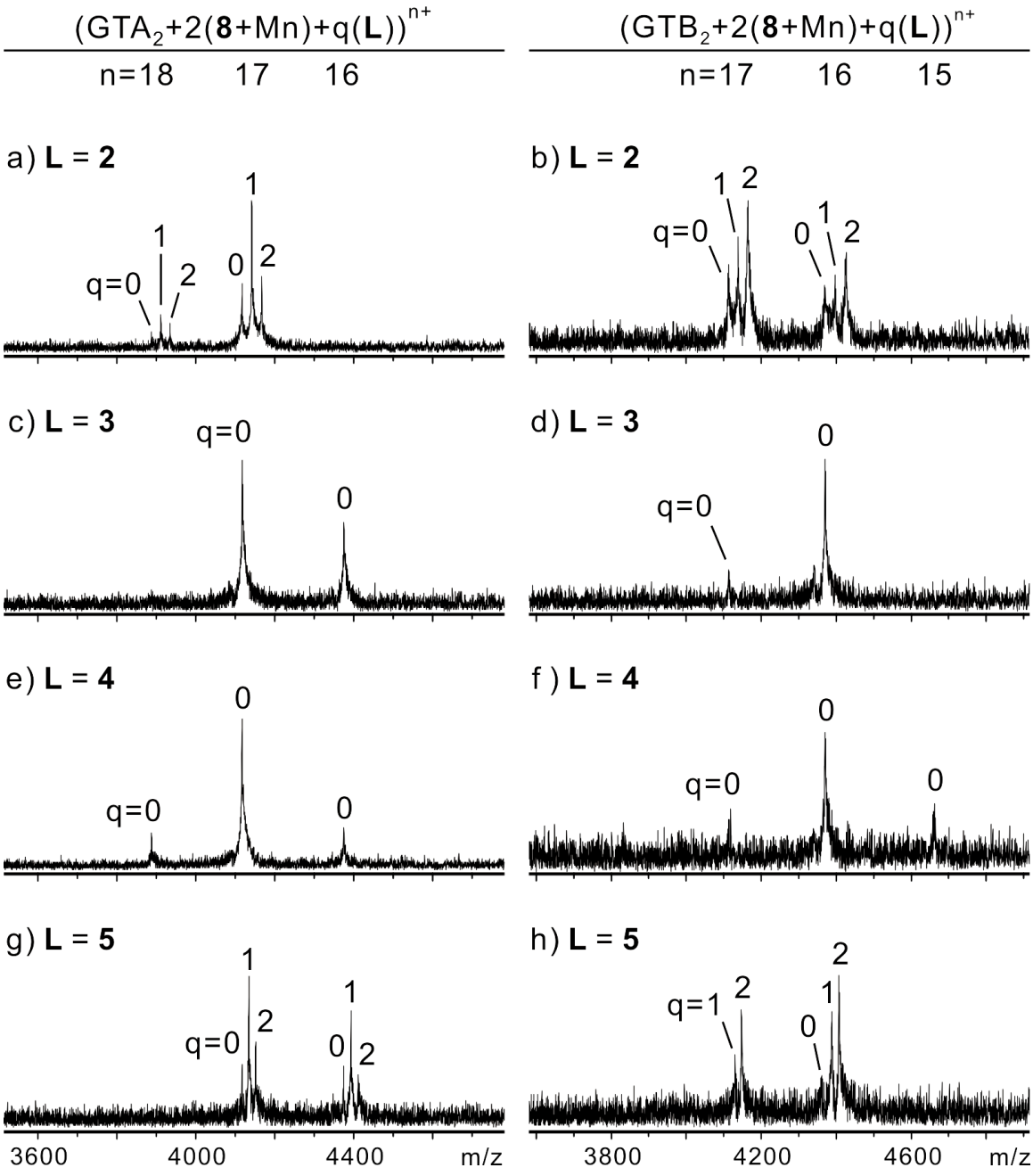
(GTA2+2(8+Mn)+q(L))n+
n=18 17 16
(GTB2+2(8+Mn)+q(L))n+
n=17 16 15
a) L = 2
b) L = 2
c) L = 3
d) L = 3
e) L = 4
f) L = 4
g) L = 5
h) L = 5
q=0
q=1
0
1
2
3600
4000
4400
3800
4200
4600
m/z

Figure 3

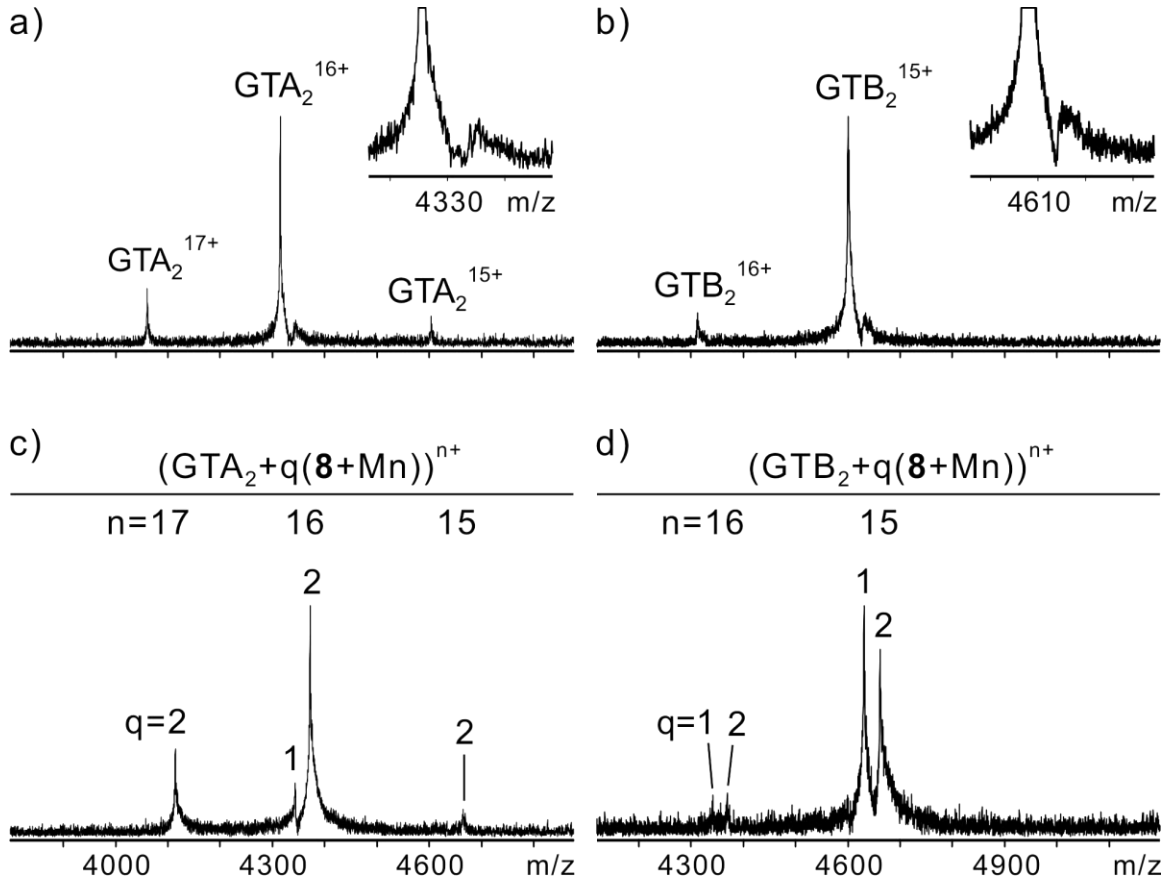
a)
GTA2 16+
GTA2 17+
GTA2 15+
4330 m/z
b)
GTB2 15+
GTB2 16+
4610 m/z
c)
(GTA2+q(8+Mn))n+
n=17 16 15
q=2
1
2
2
4000 4300 4600 m/z
d)
(GTB2+q(8+Mn))n+
n=16 15
q=1 2
1
2
4300 4600 4900 m/z

Figure 4

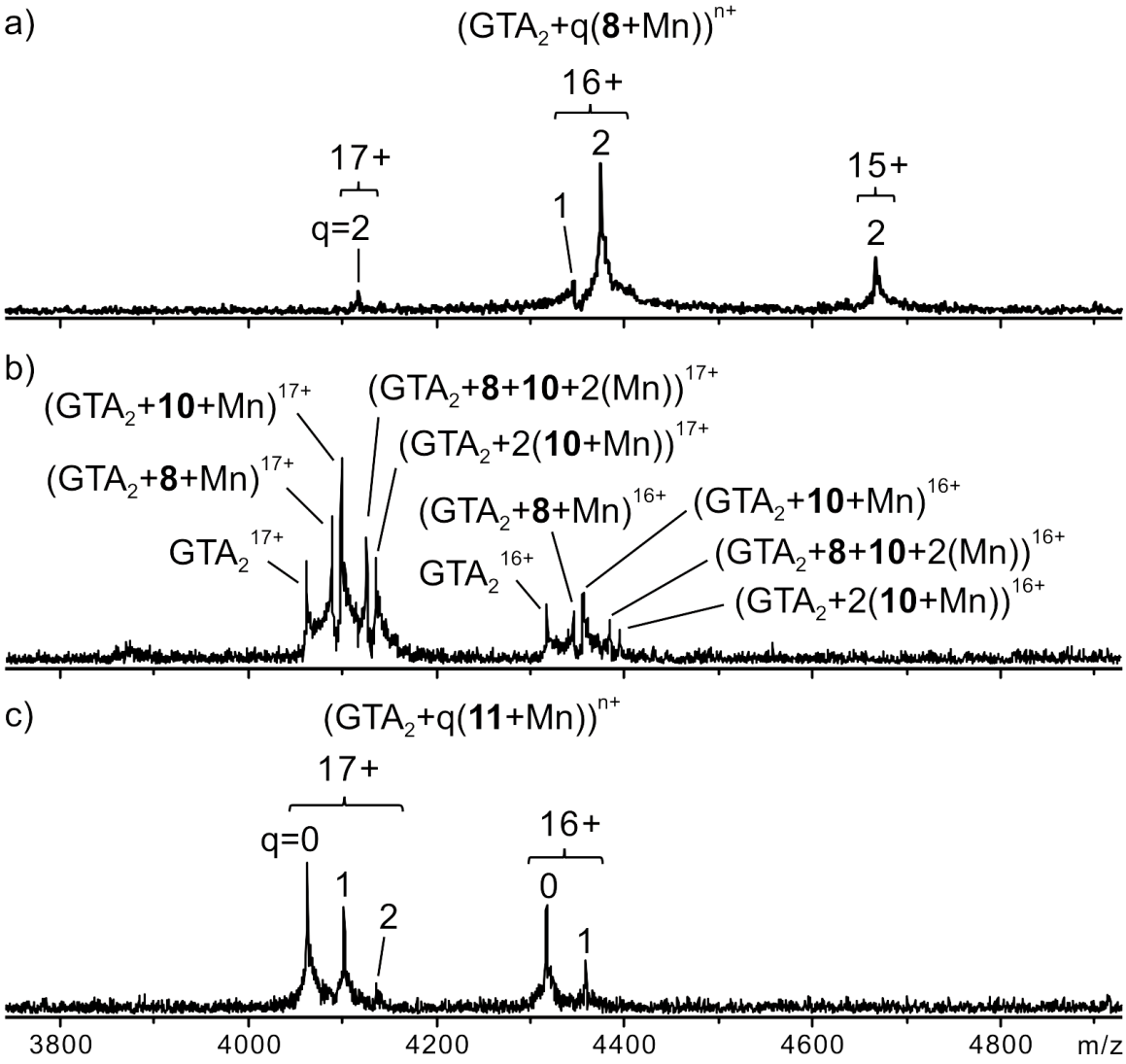
a)
(GTA2+q(8+Mn))n+
16+
2
17+
q=2
1
15+
2
b)
(GTA2+10+Mn)17+
(GTA2+8+10+2(Mn))17+
(GTA2+2(10+Mn))17+
(GTA2+8+Mn)17+
(GTA2+8+Mn)16+
(GTA2+10+Mn)16+
GTA2 17+
GTA2 16+
(GTA2+8+10+2(Mn))16+
(GTA2+2(10+Mn))16+
c)
(GTA2+q(11+Mn))n+
17+
q=0
1
2
16+
0
1
3800
4000
4200
4400
4600
4800
m/z

Figure 5

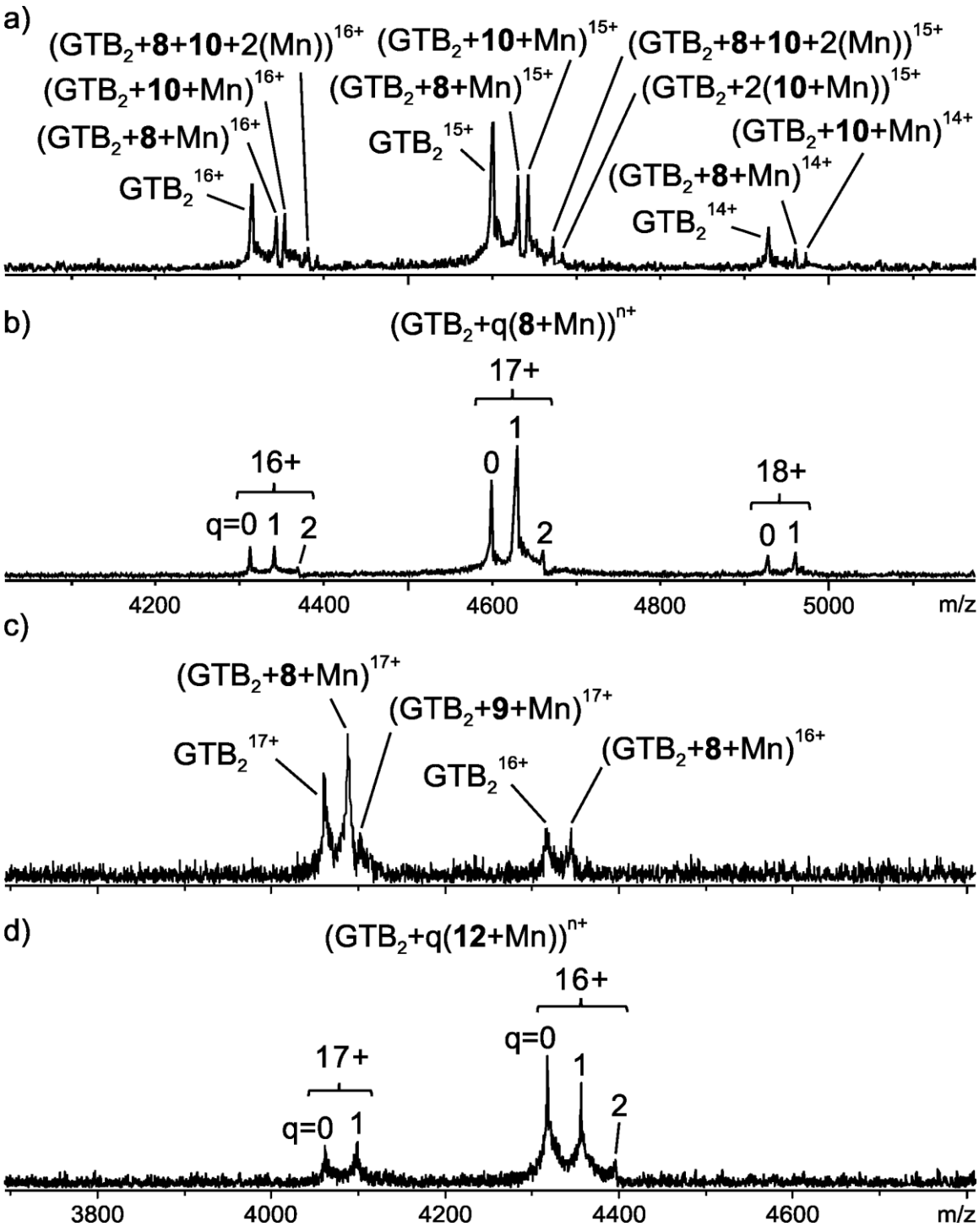
a)
(GTB2+8+10+2(Mn))16+
(GTB2+10+Mn)15+
(GTB2+8+10+2(Mn))15+
(GTB2+10+Mn)16+
(GTB2+8+Mn)15+
(GTB2+2(10+Mn))15+
(GTB2+8+Mn)16+
GTB2 15+
(GTB2+10+Mn)14+
GTB2 16+
(GTB2+8+Mn)14+
GTB2 14+
b)
(GTB2+q(8+Mn))n+
17+
1
16+
0
18+
q=0 1 2
2
0 1
4200
4400
4600
4800
5000
m/z
c)
(GTB2+8+Mn)17+
(GTB2+9+Mn)17+
GTB2 17+
(GTB2+8+Mn)16+
GTB2 16+
d)
(GTB2+q(12+Mn))n+
16+
q=0
17+
1
q=0 1
2
3800
4000
4200
4400
4600
m/z

Figure 6

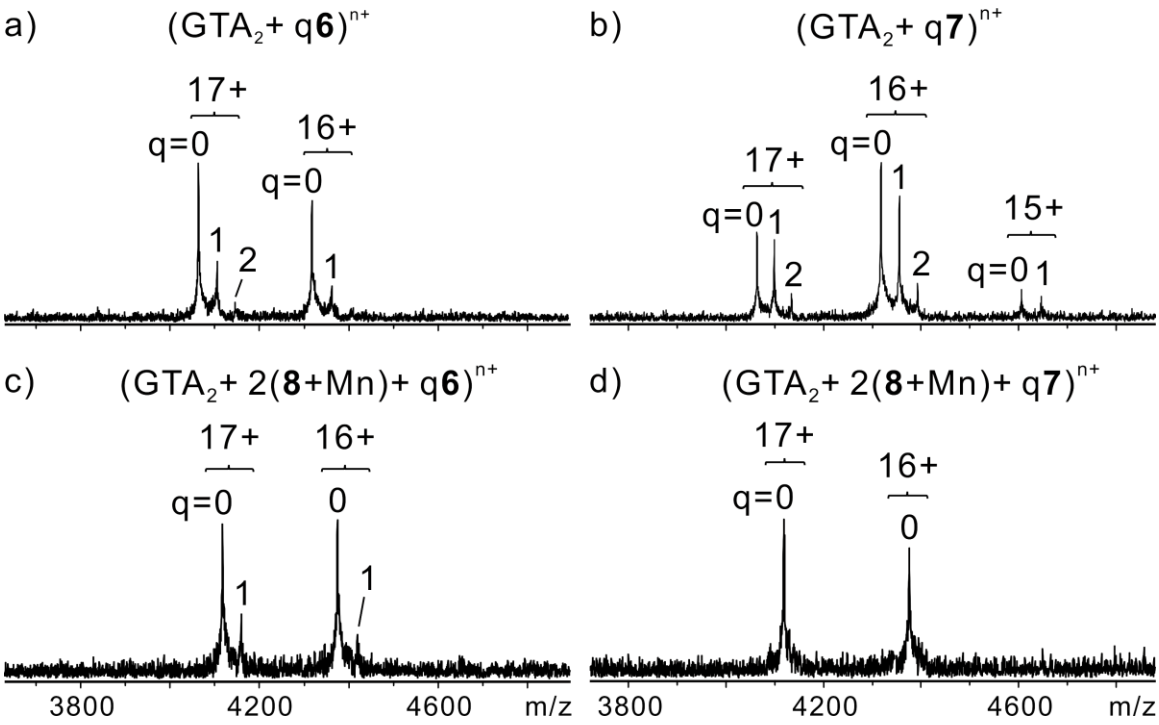
a)
(GTA2+ q6)n+
17+
16+
q=0
q=0
1
2
1
b)
(GTA2+ q7)n+
16+
17+
q=0
q=0
1
15+
1
2
2
q=0
1
c)
(GTA2+ 2(8+Mn)+ q6)n+
17+
16+
q=0
0
1
1
3800
4200
4600
m/z
d)
(GTA2+ 2(8+Mn)+ q7)n+
17+
16+
q=0
0
3800
4200
4600
m/z

Figure 7

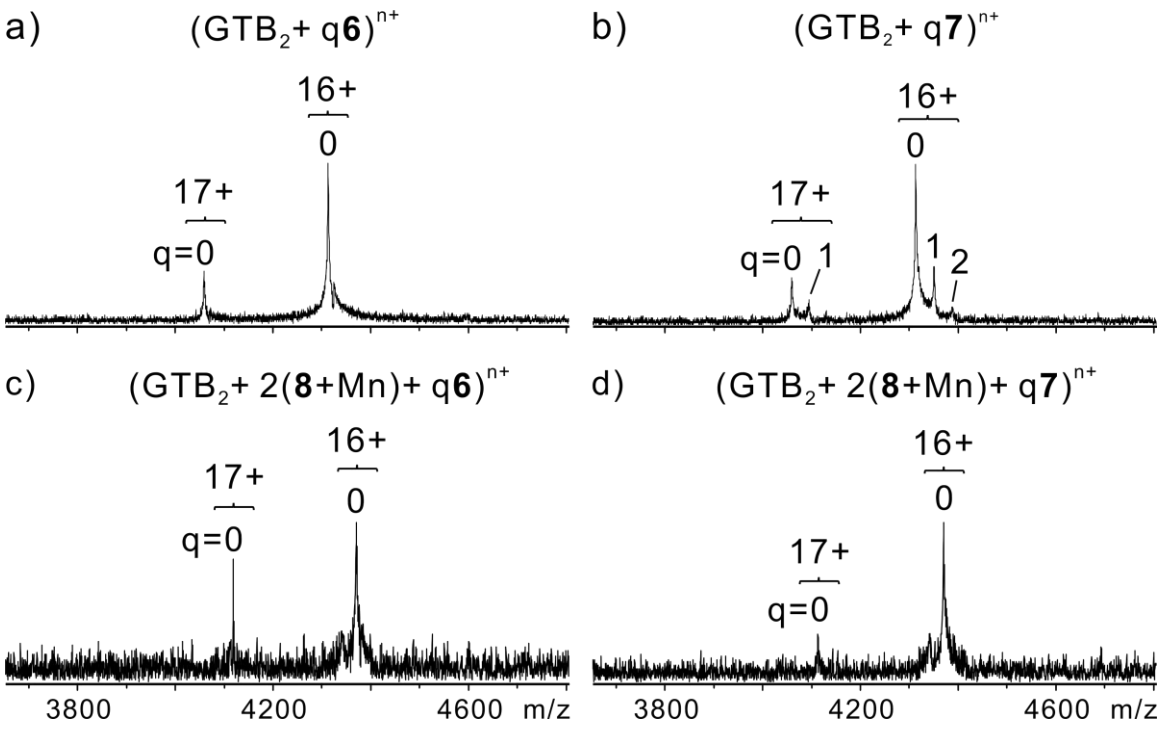
a)
(GTB2+ q6)n+
16+
0
17+
q=0
b)
(GTB2+ q7)n+
16+
0
17+
q=0
1
1
2
c)
(GTB2+ 2(8+Mn)+ q6)n+
16+
0
17+
q=0
3800
4200
4600
m/z
d)
(GTB2+ 2(8+Mn)+ q7)n+
16+
0
17+
q=0
3800
4200
4600
m/z

Figure 8

Supplementary Materials

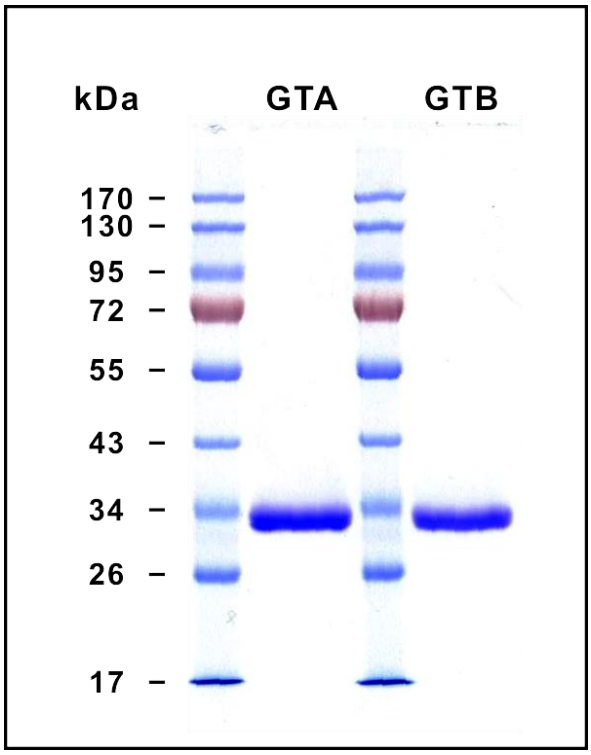

Figure S1. SDS-PAGE profiles of GTs in 10% acrylamide gel under denaturing conditions (reducing with DTT and heating at 95 °C for 5 min). 5 μg of GTs was analyzed by SDS-PAGE after purification and dialysis. The single bands identified for GTA and GTB correspond to the monomers of GTA and GTB (MW 34 519 Da and 34 483), respectively.

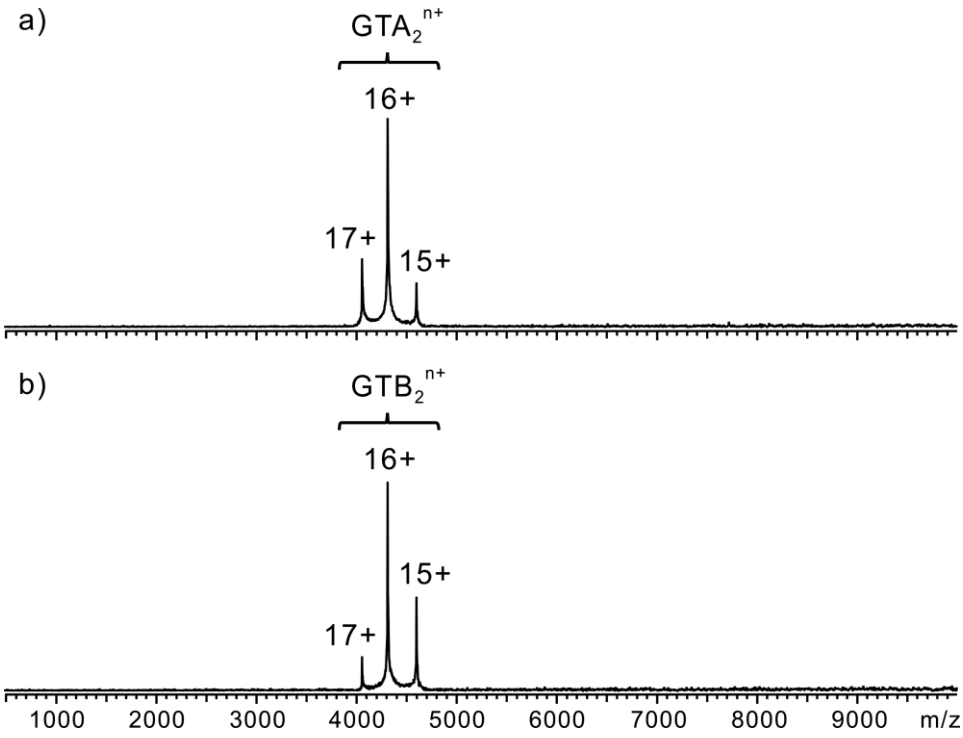

Figure S2. Illustrative nanoES mass spectra of aqueous solutions consisting of (a) GTA (230 μM) or (b) GTB (230 μM) in 10 mM ammonium acetate (pH 7).

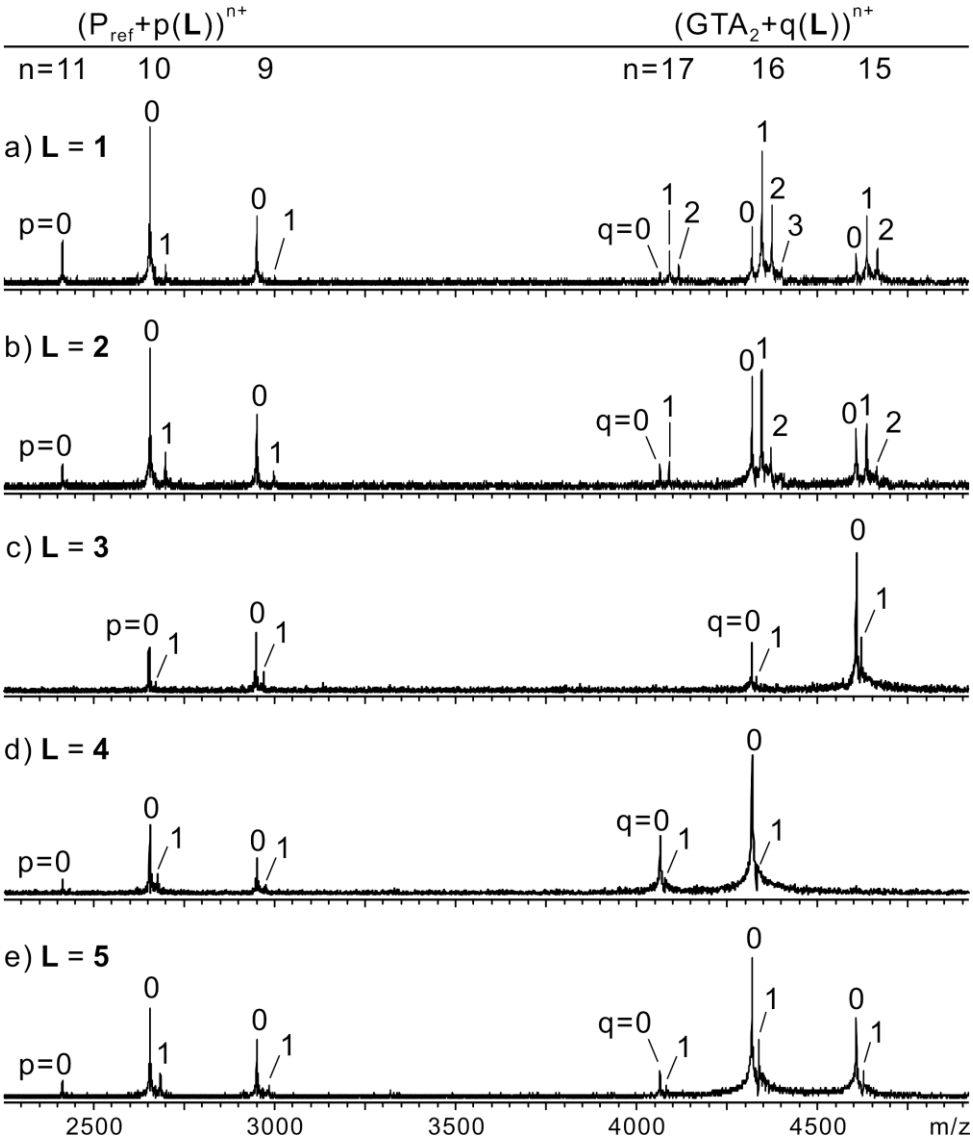

Figure S3. Illustrative nanoES mass spectra of solutions consisting of GTA_2 (8 μM) with (a) acceptor **1** (50 μM), (b) acceptor analogue **2** (50 μM), monosaccharides (c) **3** (100 μM), (d) **4** (100 μM) and (e) **5** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration 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*.

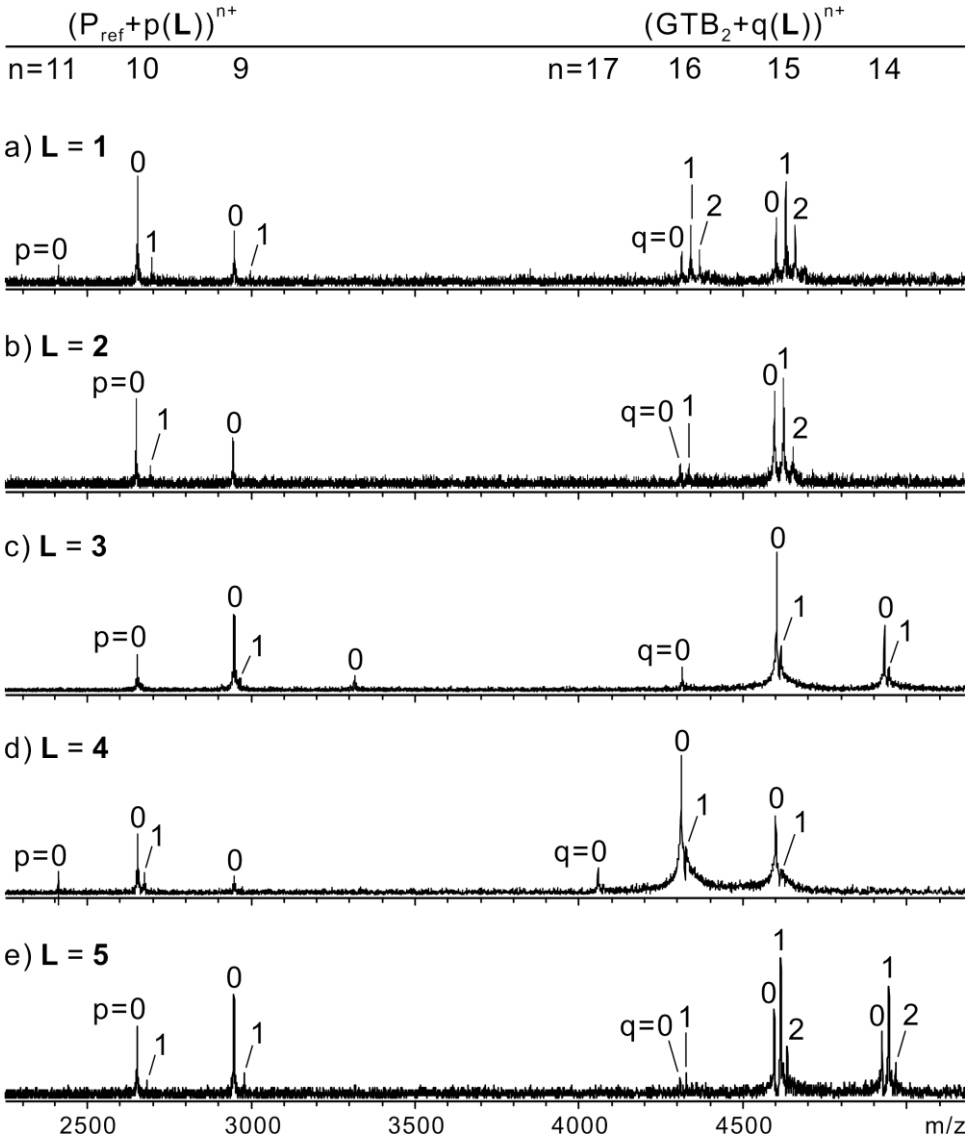

Figure S4. Illustrative nanoES mass spectra of solutions consisting of GTB_2 (8 μM) with (a) acceptor **1** (50 μM), (b) acceptor analogue **2** (50 μM), monosaccharides (c) **3** (100 μM), (d) **4** (100 μM) and (e) **5** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration 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*.

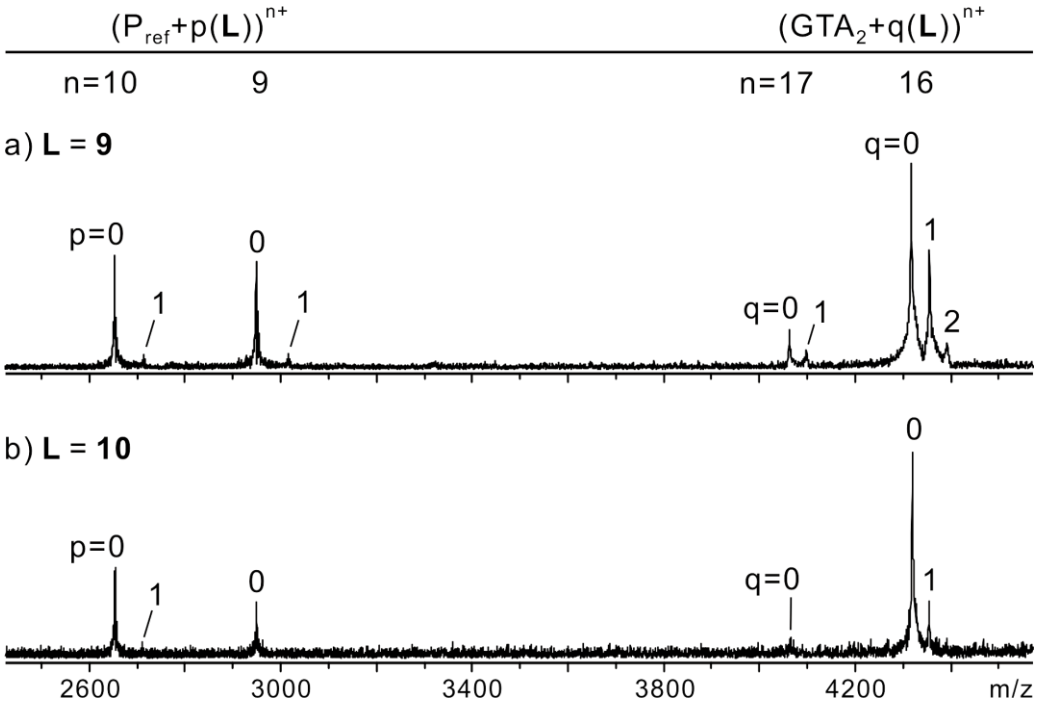

Figure S5. Illustrative nanoES mass spectra of solutions consisting of GTA_2 (9 μM) with (a) native donor UDP-GalNAc **9** (80 μM) and (b) the alternate donor UDP-Gal **10** (80 μM). A reference protein, P_{ref} (scFv), was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific 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*.

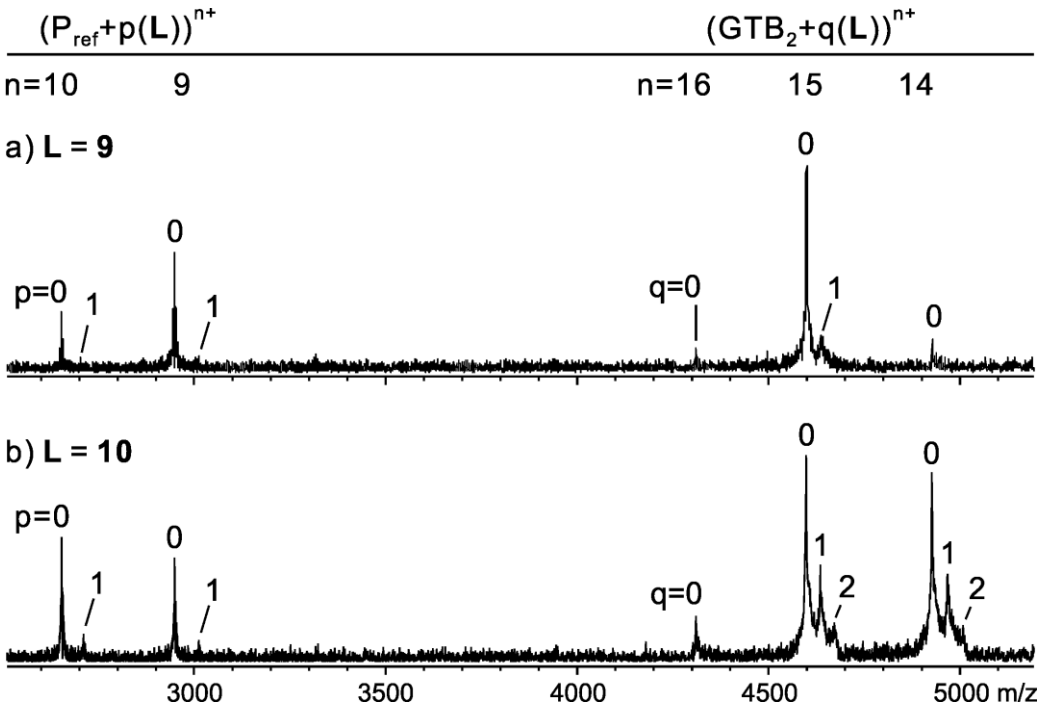

Figure S6. Illustrative nanoES mass spectra of solutions consisting of GTB_2 (9 μM) with (a) the alternate donor UDP-GalNAc **9** (80 μM) and (b) the native donor UDP-Gal **10** (80 μM). A reference protein, P_{ref} (scFv), was added to each solution at a concentration of 5 μM to quantify the extent 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*.

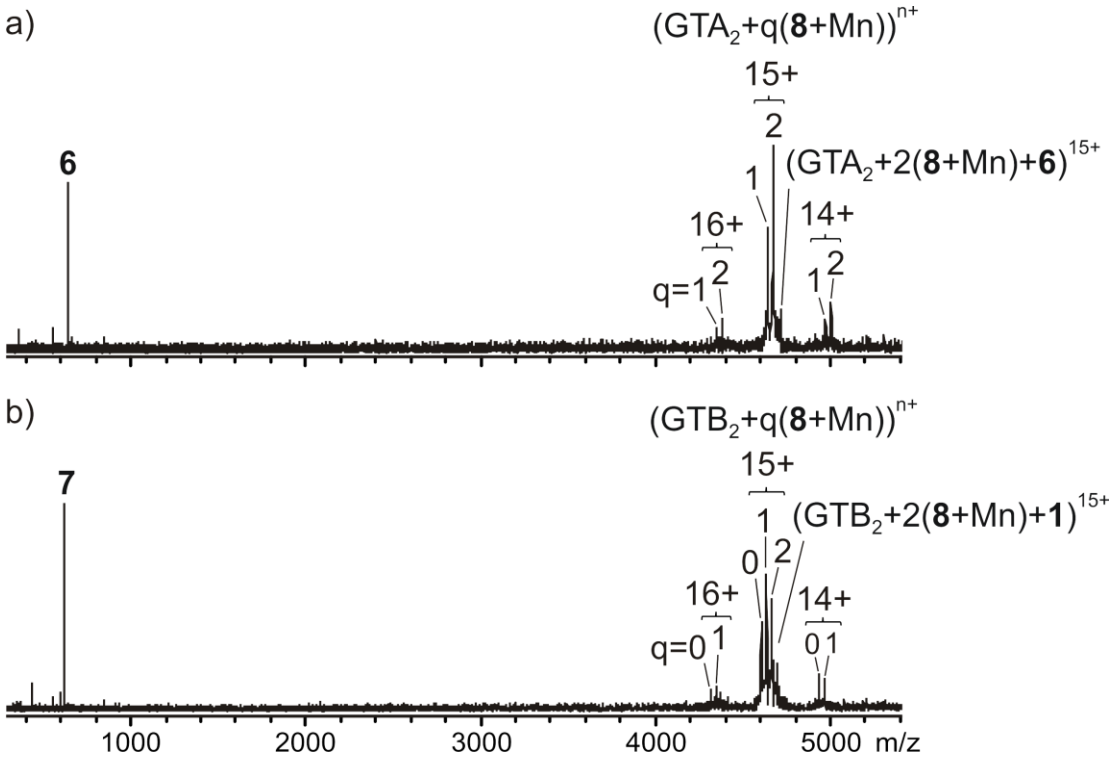

Figure S7. Illustrative nanoES mass spectrum of a solution consisting of (a) GTA_2 (8 µM), **1** (20 µM), **9** (20 µM), Mn^{2+} (100 µM) and (b) GTB_2 (8 µM), **1** (20 µM), **10** (20 µM), Mn^{2+} (100 µM) in 10 mM ammonium acetate buffer (pH 7). The number of **8** bound to the GT ions is indicated by *q*.

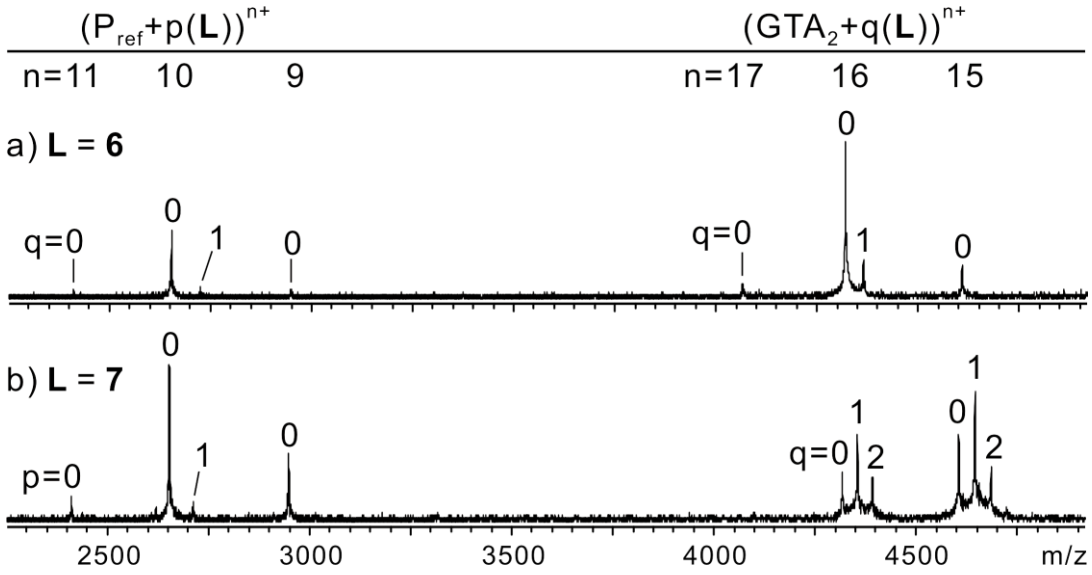

Figure S8. Illustrative nanoES mass spectra of solutions consisting of GTA_2 (8 μM) with trisaccharide (a) **6** (50 μM) and (b) **7** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES 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*.

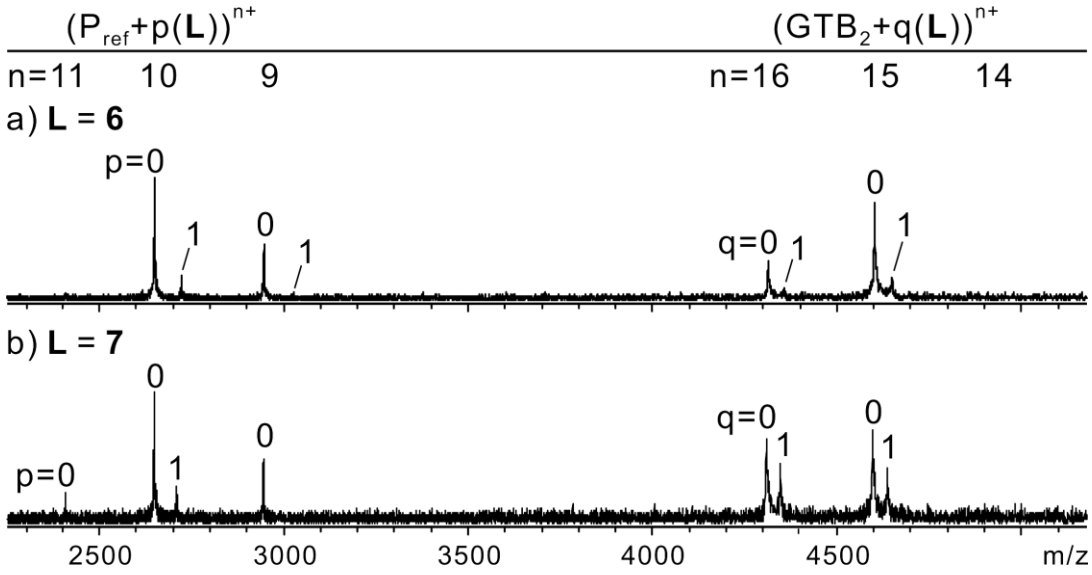

Figure S9. Illustrative nanoES mass spectra of solutions consisting of GTB_2 (8 μM) with trisaccharide (a) **6** (50 μM) and (b) **7** (50 μM). A reference protein, P_{ref}, was added to each solution at a concentration of 5 μM to quantify the extent of nonspecific protein-ligand binding during the ES 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*.