

Recognition of Human Milk Oligosaccharides by Bacterial Exotoxins

Amr El-Hawiet,^{1,2} Elena N. Kitova,¹ and John S. Klassen^{1*}

*¹Alberta Glycomics Centre and Department of Chemistry, University of
Alberta, Edmonton, Alberta, Canada T6G 2G2*

²Current address: Faculty of Pharmacy, University of Alexandria, Alexandria, Egypt

*Corresponding Author's address:

Department of Chemistry

University of Alberta

Edmonton, AB CANADA T6G 2G2

Email: john.klassen@ualberta.ca

Telephone: (780) 492-3501

Fax: (780) 492 8231

Keywords: toxin / human milk oligosaccharide / affinity / electrospray ionization mass spectrometry

Abstract

The affinities of the most abundant oligosaccharides found in human milk for four bacterial exotoxins (from *Vibrio cholerae* and pathogenic *Escherichia coli*) were quantified for the first time. Association constants (K_a) for a library of twenty human milk oligosaccharides (HMOs) binding to Shiga toxin type 2 holotoxin (Stx2) and the B subunit homopentamers of cholera toxin, heat labile toxin and Shiga toxin type 1 (CTB₅, HLTB₅ and Stx1B₅) were measured at 25 °C and pH 7 using the direct electrospray ionization mass spectrometry assay. Notably, all four bacterial toxins bind to a majority of the HMOs tested and five of the HMOs (2'-fucosyllactose, lacto-N-tetraose, lacto-N-fucopentaose I, lacto-N-fucopentaose II and lacto-N-fucopentaose III) are ligands for all four toxins. These five HMOs are also reported to bind to other bacterial toxins (e.g. toxin A and toxin B of *Clostridium difficile*). In all cases, the HMO affinities (apparent K_a) are relatively modest ($\leq 15\,000\text{ M}^{-1}$). However, at the high concentrations of HMOs typically ingested by infants, a significant fraction of these toxins, if present, are expected to be bound to HMOs. Binding measurements carried out with 2'-fucosyllactose or lacto-N-fucopentaose I, together with a high affinity ligand based on the native carbohydrate receptor, revealed that all four toxins possess HMO binding sites that are distinct from those of the native receptors, although evidence of competitive binding was found for lacto-N-fucopentaose I with Stx2 and 2'-fucosyllactose and lacto-N-fucopentaose I with HLTB₅. Taken together, the results of this study suggest that, while HMOs are expected to bind extensively to these bacterial toxins, it is unlikely that HMO binding will effectively inhibit their interactions with their cellular receptors.

Introduction

Breast milk, which contains a variety of bioactive components, including proteins, glycoproteins, free oligosaccharides (referred to as human milk oligosaccharides, HMOs) and fat globules (Lara-Villoslada et al. 2007; Martin et al. 2003), is not only an essential source of nutrition for infants but also serves to protect breast-fed infants against a variety of infectious diseases, e.g. enteric bacterial and viral infections (Newburg et al. 2005). Several mechanisms have been proposed to explain the protective effects of human milk, such as agglutination of pathogen cells (Howie et al. 1990), suppression of host inflammation (Ruiz-Palacios et al. 2003), modulation of host cell growth (Morrow et al. 2004) and the selective promotion of bifidobacteria growth in the gastrointestinal tract of infants (Zivcovic et al. 2010). It has also been suggested that the competitive binding of HMOs to the carbohydrate recognition domains of pathogen-generated proteins (e.g. surface lectins and toxins) plays a role. HMOs, which are structurally similar to some intestinal mucosal cell surface glycans (Kobata 2010), could act as decoys and disrupt the binding of microbial lectins to host cell receptors, thus preventing infection of the host by these organisms (Idota et al. 1995; Newburg et al. 2009; Bode et al. 2012; Liu et al. 2013; Smilowitz et al. 2014; Yu et al. 2014). Although the protective effects of HMOs against enteric infections are widely recognized, there have been few detailed studies of the nature of the HMO interactions (Dingle et al. 2008; El-Hawiet et al. 2011; Mandal et al. 2012; Vasile et al. 2014).

Here, we report the results of HMO binding measurements performed on exotoxins from pathogenic bacteria (cholera toxin (CT), *Escherichia coli* heat-labile enterotoxin (type 1) (HLT) and Shiga toxin type 1 (Stx1) and type 2 (Stx2)), against

which HMOs are known to have protective effects. These four toxins investigated belong to the family of AB₅ toxins, which consist of a catalytically active A subunit and five identical B subunits that assemble into a doughnut-shaped homopentameric B₅ structure that binds to the C-terminus of the A subunit and is responsible for cellular attachment (Fraser et al. 1994; Merrit et al. 1998; Fraser et al. 2004; Mudrak et al. 2010). Each B₅ complex possesses multiple binding sites (at least one per subunit) for their native carbohydrate receptors located on the surface of intestinal epithelial cells. The Shiga toxins recognize the globotriaose Gb₃, α -Gal(1-4) β -Gal(1-4) β -Glc-ceramide (Ling et al. 1998), while CT and HLT bind to the monosialylganglioside GM1, β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc-(1 \rightarrow 4)-[α -D-Neu5Ac-(2 \rightarrow 3)]- β -D-Gal-(1 \rightarrow 4)- β -D-Glc-ceramide (Merrit et al. 1994). The solved crystal structure of Stx1B₅ bound to a Gb₃ analogue (the methoxycarbonyloctyl glycoside of the P^k trisaccharide α -D-Gal-(1 \rightarrow 4)- β -D-Gal(1 \rightarrow 4)- β -D-Glc) revealed the presence of three binding sites per subunit (referred to as *site 1*, *site 2* and *site 3*), with *site 2* having the highest occupancy and, presumably, the highest affinity (St. Hilaire et al. 1994). Recently, the crystal structure of Stx2 bound to a disaccharide ligand α -D-GalNAc-(1 \rightarrow 4)- β -D-GalOMe (which corresponds to the terminal moiety of the O-polysaccharide of lipopolysaccharide expressed by *E. coli* strain O117) (Kale et al. 2008) was reported (Jacobson et al. 2014). The solved structure showed binding of the disaccharide at *site 1* and *site 2*. According to the crystal structures of CTB₅ and HLTB₅ bound to the GM1 pentasaccharide β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc-(1 \rightarrow 4)-[α -D-Neu5Ac-(2 \rightarrow 3)]- β -D-Gal-(1 \rightarrow 4)- β -D-Glc, there is one binding site per subunit for GM1 (Merrit et al. 1998; Holmner et al. 2011). However, HLTB₅ and CTB₅ are also known to recognize other carbohydrate ligands, such as H, A and B blood group antigens

(Mandal et al. 2012; Vasile et al. 2014; Holmner et al. 2007; Heggelund et al. 2012). Notably, the blood group antigen binding site is distinct from that of GM1, based on x-ray crystallography results obtained for HLTB₅. In the present work, the direct electrospray ionization mass spectrometry (ESI-MS) assay (El-Hawiet et al. 2012; Kitova et al. 2012) was used to quantify the binding of twenty commercially available HMOs to CTB₅, HLTB₅, Stx1B₅ and Stx2 in neutral aqueous solutions at 25 °C.

Results and discussion

ESI-MS binding measurements were performed on Stx2, CTB₅, HLTB₅ and Stx1B₅ and individual HMOs (**L1** – **L20**) in aqueous ammonium acetate solutions (100 mM, pH 7 and 25 °C). The structures of **L1** – **L20** are given in Table 1 and Figure 1. Representative ESI mass spectra acquired for solutions containing toxin and **L8** (α -L-Fuc-(1→2)- β -D-Gal-(1→3)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc) are shown in Figures S3 – S6 (Supplementary Data). Following correction of the mass spectra for the occurrence of nonspecific HMO-protein binding during the ESI process (Sun et al. 2006), $K_{a,app}$ values were calculated from the relative abundances of the free and ligand-bound toxin ions (Table II). To facilitate a comparison of the binding data, the $K_{a,app}$ values for the four toxins are also plotted in Figure 2. The $K_{a,app}$ values reported previously for recombinant subfragments of toxins A and B (TcdA2 and TcdB1, respectively) of *C. difficile* are included in Table I and Figure 2 for comparison purposes (El-Hawiet et al. 2011).

Inspection of Table II (and Figure 2) reveals that the bacterial toxins investigated in the present study bind to a majority of the HMOs. Although, in all cases, the measured HMO affinities are relatively low, with $K_{a,app}$ ranging from 600 M⁻¹ to 15 000 M⁻¹. Remarkably, CTB₅ and Stx2 bind to eighteen of the twenty HMOs tested, while Stx1B₅

and HLTB₅ bind to seventeen and sixteen HMOs, respectively. These results, taken together with those of an earlier study on TcdA2 and TcdB1, which showed that these two toxins bind to eight and eleven of these HMOs, respectively, (El-Hawiet et al. 2011), suggest that enteric bacterial toxins generally exhibit broad specificity for both neutral and acidic HMOs.

Each of the HMOs tested was found to bind to at least three of the toxins investigated and five neutral HMOs, **L1** (2'-fucosyllactose), **L7** (lactotetraose), **L8** (lacto-N-fucopentaose I), **L9** (lacto-N-fucopentaose II) and **L18** (lacto-N-fucopentaose III), bind to all six toxins. Notably, these highly “promiscuous” HMOs are among the most abundant HMOs found naturally among secretors (Hong et al. 2014; Urashima et al. 2012). For non-secretors, the concentrations of these oligosaccharides are smaller but still significant (Hong et al. 2014). Three of these HMOs (**L7**, **L8** and **L9**) are type I (containing lacto-N-biose motif Gal-(1→3)-β-D-GlcNAc) oligosaccharides, which are the dominant type found in humans (Urashima et al. 2012). The structural motif Gal-(1→4)-β-D-GlcNAc, which is found in **L18**, is of type II and is less abundant in human breast milk. It also interesting to note that the four monofucosylated HMOs possess a variety of fucose linkages (1→2 (**L1** and **L8**), 1-3 (**L18**) or 1-4 (**L9**)) and different fucosylation sites (Gal (**L1** and **L8**) or GlcNAc (**L9** and **L18**)). That **L7**, which is not fucosylated, binds to all six toxins further suggests that HMOs produced by Lewis-negative donors (which do not produce fucosylated HMOs) also contain structures that could, in principle, impart protection against bacterial pathogens. Taken together, these results suggest that, regardless of the secretor status and Lewis blood group of the mother (Smilowitz et al.

2014), many of the most abundant HMOs present in a given breast milk sample will bind to bacterial exotoxins and could, potentially, have an inhibitory role against infection.

As discussed above, one mechanism by which HMOs could inhibit bacterial toxins is by blocking their binding to cellular receptors. The measured $K_{a,app}$, while low compared to the affinity of ganglioside GM1 for CT and HLT ($\sim 10^9 \text{ M}^{-1}$) (MacKenzie et al. 1997), are similar in magnitude to many biologically-relevant protein-carbohydrate interactions, including the P^k trisaccharide interaction with Stx1 and Stx2 ($\sim 1500 \text{ M}^{-1}$) (Kitova et al. 2007). Therefore, it is realistic to expect that HMOs have the potential to effectively compete with low affinity, monovalent protein-carbohydrate receptor interactions. However, because AB_5 toxins, and most carbohydrate-binding exotoxins, have multiple ligand binding sites, they can engage in multivalent binding to their cellular receptors and, thus, achieve high affinity or avidity (Kitov et al. 2000; Dasgupta et al. 2014). For example, although the monovalent P^k trisaccharide interaction with Stx1 and Stx2 is weak, binding measurements suggest that toxins bind to surface-bound glycolipid Gb_3 with avidities in the 10^9 M^{-1} range (Fuchs et al. 1986). Therefore, if present at comparable concentrations, the HMOs would not effectively compete with the native receptors (assuming competitive binding) for the toxins. However, it is known that HMOs enter the intestinal tract at very high concentrations ($5 - 20 \text{ g L}^{-1}$ or $\sim 0.005 - 0.02 \text{ M}$) (Bode, 2012). Under these conditions, the possibility that even low affinity HMOs could inhibit even high affinity interactions cannot be dismissed outright.

As a first step towards estimating the effectiveness of HMOs at blocking high affinity interactions, theoretical calculations of the fraction of toxin bound to HMO were carried out using a single site (1:1) protein-ligand binding model with two competing

ligands - an HMO and a cellular receptor. Shown in Figure 3 are plots of the fraction of protein bound to HMO and to receptor calculated assuming apparent HMO affinities of 10^2 to 10^5 M^{-1} and apparent receptor affinities of 10^3 M^{-1} (low affinity), 10^6 M^{-1} (moderate affinity) and 10^9 M^{-1} (high affinity). The HMO, receptor and protein concentrations were taken to be 0.02 M, 0.1 μM and 0.1 μM , respectively. Inspection of Figure 3 shows that, in the presence of low or moderate affinity receptors, HMOs are expected to bind a significant fraction of the toxin; this is true even for the low affinity HMOs. It also shows that high affinity receptors will outcompete with the lowest affinity HMOs ($K_a \leq 10^3 \text{ M}^{-1}$). However, at these concentrations, the higher affinity HMOs are expected to preferentially bind to the toxin. The above analysis was carried out at a single set of concentrations and it is not possible to accurately predict the effective concentrations of pathogenic proteins and cellular receptors that would be present in the body. Consequently, the analysis was extended to a high affinity receptor (10^9 M^{-1}) over a range of concentrations (0.001 to 0.1 μM) (Figure S7). Inspection of the calculated surfaces reveals that, except at the highest receptor concentrations considered ($>0.05 \mu\text{M}$), the HMO ligands, regardless of their affinities, will effectively bind to toxin and, therefore, could inhibit binding to its cellular receptors and offer protection against infection.

It must be emphasized that the aforementioned analysis was carried out assuming that the HMO and receptor ligands bind competitively. However, the diversity of HMO structures recognized by the toxins investigated is somewhat surprising and raises the question of where the HMO binding sites are located and whether they differ from those of the native carbohydrate receptors. For AB₅ bacterial toxins there is little structural information available on HMO interactions. However, it is known that some histo-blood

group antigens, e.g. difucosyllactose (α -L-Fuc-(1→2)- β -D-Gal(1→4)[α -L-Fuc-(1→3)]- β -D-Glc) (**L3**) and α -D-GalNAc(1→3) α -L-Fuc-(1→2)- β -D-Gal(1→4)[α -L-Fuc-((1→3)]- β -D-Glc (α GalNAc-**L3**) and α -D-Gal(1→3) α -L-Fuc-(1→2)- β -D-Gal(1→4)[α -L-Fuc-(1→3)]- β -D-Glc (α Gal-**L3**), which are structurally similar to some HMOs, bind to CTB₅ at sites distinct from the GM1 binding site (Vasile et al. 2014; Heggelund et al. 2012). Additionally, the blood group A antigen analog (α GalNAc-**L3**) was shown to bind to a site that is distinct from the primary GM1 binding site of HLTB₅, one that involves residues Gly45, Thr47, Asn94 from one subunit and Gln3 from an adjacent subunit (Holmner et al. 2007). Based on these observations it is likely that both CTB₅ and HLTB₅ have HMO binding sites that are distinct from those of the native carbohydrate receptor, at least for HMOs containing the α -L-Fuc(1→2)- β -D-Gal(1→4)- β -D-Glc motif. To the best of our knowledge, there are no structural data available for Stx1 and Stx2 binding to HMOs or histo-blood group antigens.

To establish conclusively whether the four toxins investigated in the present study have HMO binding sites distinct from those of their native carbohydrate receptors, binding experiments were performed on solutions containing each toxin with an HMO ligand (**L1** or **L8**) and a high affinity ligand based on the native carbohydrate receptor, which was present at a concentration high enough to saturate the native receptor binding sites. For CTB₅ and HLTB₅, the GM1 pentasaccharide **L21** (Figure S1, Supplementary Data), which has affinities of $>10^6$ M⁻¹ for CTB₅ (Turnbull et al. 2004) and 10^8 M⁻¹ for HLTB₅ (Minke et al. 1999) was used. For Stx1B₅ and Stx2, the decavalent ligand **L22** (Figure S2, Supplementary Data), which forms a high affinity 1:1 complex with Stx1 (IC₅₀ of 10^{-6} M) and Stx2 (IC₅₀ of 10^{-6} M) (Prof DR Bundle, personal communication),

was employed. **L22** consists of a glycosyl scaffold with five arms ending with a dimer of α -D-Gal(1 \rightarrow 4)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (P^k) trisaccharide (Figure S2, Supplementary Data). The ten P^k trisaccharides can bind to *site 2* and *site 1* on each B subunit (Kitov et al. 2000); *site 2* is reported to be the highest affinity binding site for the P^k trisaccharide (St. Hilaire et al. 1994).

Representative ESI mass spectra acquired for the aqueous ammonium acetate solutions (100 mM, pH 7 and 25 °C) of each of the four toxins together with **L1** and either **L21** or **L22** are shown in Figure 4. Inspection of the ESI mass spectrum acquired for a solution of CTB₅ (5 μ M), **L21** (30 μ M) and **L1** (500 μ M) reveals the presence of ions corresponding to the (CTB₅ + 4**L21**), (CTB₅ + 5**L21**) complexes, at charge states +13 – +15, as well as ions corresponding to the mixed complexes (CTB₅ + 5**L21** + q **L1**), where $q = 1 - 3$ (Figure 4a). The appearance of nonspecific ($P_{\text{ref}} + q$ **L1**) ions, where $q = 1$ and 2, at charge states +8 – +11, indicates the occurrence of nonspecific binding of **L1** to the CTB₅ complexes during the ESI process. The distribution of **L1** bound specifically to (CTB₅ + 5**L21**) was obtained following correction of the mass spectrum using the P_{ref} method (Sun et al. 2006) (inset Figure 4a). It can be seen that one to three **L1** are bound specifically to (CTB₅ + 5**L21**) in solution. Similarly, the ESI mass spectrum acquired for a solution of HLTB₅ (17 μ M), **L21** (90 μ M) and **L1** (200 μ M) reveals the presence of ions corresponding to (HLTB₅ + 5**L21** + q **L1**), where $q = 1 - 3$ (Figure 4b). After correction for nonspecific binding, it is found that one or two **L1** bind specifically to HLTB₅ under these solution conditions (inset Figure 4b). Because the ESI-MS experiments were designed in such a way that the primary native receptor binding sites were saturated, or nearly so, by **L21**, the specific binding of **L1** must necessarily involve sites distinct from

those occupied by **L21**. Similar results were obtained with **L8** (Figures S8a and S8b, Supplementary Data).

Shown in Figures 4c and 4d are ESI mass spectra acquired for the Stx1B₅ (10 μM) or Stx2 (10 μM), respectively, with **L22** (15 μM) and **L1** (200 μM). Protonated ions corresponding to 1:1 toxin-**L22** complexes, i.e., (Stx1B₅ + **L22**), at charge states +13 – +15, and (Stx2 + **L22**), at charge states +16 – +19 were detected in both cases. Ions corresponding to the complexes (Stx1B₅ + **L22** + q **L1**), where $q = 1$ and 2, and (Stx2 + **L22** + q **L1**), where $q = 1 – 3$, were also detected. The distributions of specifically bound **L1**, obtained from the mass spectra following correction for nonspecific binding (inset, Figures 4c and 4d), confirms that **L1** can bind to sites distinct from *site 1* and *site 2* in both toxins. Similar results were obtained for **L8** (Figures S8c and S8d, Supplementary Data).

While the aforementioned binding data establish unambiguously that the four toxins possess HMO binding sites that are distinct from those of the native receptors, they do not reveal the location of these sites. Conclusively establishing the HMO binding sites will ultimately require the use of high resolution structural techniques, such as X-ray crystallography. The binding data also do not rule out the possibility of some HMOs interacting at the native receptor binding sites. Some insight into this possibility was gained by comparing $K_{a,app}$ values for **L1** and **L8** measured in the absence and presence of the high affinity ligands. From the direct ESI-MS measurements (Table I), the affinities of **L1** and **L8** for CTB₅, HLTB₅, Stx1B₅ and Stx2 are 2100 M⁻¹ and 1100 M⁻¹, 15000 M⁻¹ and 12000 M⁻¹, 1000 M⁻¹ and 1500 M⁻¹, and 2700 M⁻¹ and 3800 M⁻¹, respectively. In the presence of the high affinity ligand, the $K_{a,app}$ of **L1** for CTB₅ (2000 M⁻¹), Stx1B₅ (1100

M^{-1}) and Stx2 ($2700 M^{-1}$) are largely unchanged, while that for HLTB₅ ($2300 M^{-1}$) is significantly smaller. These results suggest that the native receptor binding sites for CTB₅, Stx1B₅ and Stx2 do not represent the major sites of interaction for **L1**. In contrast, the significant reduction in affinity observed with HLTB₅ reveals that the **L1** binds competitively with **L21**. However, given that the binding is not completely lost in the presence of **L21**, there must be additional binding sites for **L1**. The $K_{a,app}$ measured for **L8** in the presence of high affinity ligand lead to similar conclusions for CTB₅ ($1300 M^{-1}$) and Stx1B₅ ($800 M^{-1}$) binding. However, in contrast to what was observed for **L1**, the affinities of **L8** for both HLTB₅ ($500 M^{-1}$) and Stx2 ($1200 M^{-1}$) were significantly reduced in the presence of high affinity ligand. This is an intriguing finding, which suggests the possibility of **L8** binding at the native receptor sites of HLTB₅ and Stx2. The affinity data further point to significant differences in the ligand specificities of the native receptor binding sites of CTB₅ and HLTB₅ and also between the native receptor binding sites of Stx2 and Stx1B₅.

Taken together, the results of this binding study clearly establish that the toxins investigated exhibit a broad specificity for the most abundant oligosaccharides found in human milk. Moreover, given the high concentrations of HMOs typically ingested by infants, a significant fraction of these toxins, if present, would be bound to HMOs. However, the results of this study also show that the four toxins possess HMO binding sites that are distinct from those of the native receptors. Therefore, while some HMOs may bind competitively with native receptors, a significant and possibly dominant fraction of the bound HMOs will be located at other sites. Consequently, it seems unlikely that HMO binding to the toxins would effectively inhibit their interactions with their cellular

receptors. However, more extensive binding data, as well as inhibition studies, are needed to conclusively establish whether HMO binding to bacterial toxins confers protection against infection.

Conclusions

In summary, the affinities of twenty of the most abundant HMOs for CTB₅, HLTB₅, Stx1B₅ and Stx2 were quantified for the first time. Remarkably, it was found that each toxin binds to a majority of the HMOs tested, albeit with relatively low apparent affinities ($\leq 15\ 000\ \text{M}^{-1}$). Five of the HMOs (2'-fucosyllactose, lacto-N-tetraose, lacto-N-fucopentaose I, lacto-N-fucopentaose II and lacto-N-fucopentaose III) are ligands for all four toxins. These same five HMOs also bind to toxin A and toxin B of *C. difficile*. Based on these findings it can be reasonably proposed that many, if not most, enteric bacterial toxins exhibit broad specificity for both neutral and acidic HMOs. Moreover, significant fraction of bacterial toxins expected to be bound to HMOs given the high concentrations of the oligosaccharides typically ingested by infants during feeding. The results of binding measurements carried out using 2'-fucosyllactose or lacto-N-fucopentaose I, together with a high affinity ligand based on the native carbohydrate receptor, revealed some evidence of competitive binding. However, the binding measurements also established conclusively that the four toxins possess HMO binding sites that are distinct from those of the native receptors. Given these findings, it seems unlikely that HMO binding to bacterial toxins will effectively inhibit their interactions with their cellular receptors.

Materials and Methods

Carbohydrate ligands

The HMOs library (Figure 1 and Table I) consisted of 2'-fucosyllactose, α -L-Fuc-(1 \rightarrow 2)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L1**, MW 488.43 Da); 3-fucosyllactose, β -D-Gal-(1 \rightarrow 4)-[α -L-Fuc-(1 \rightarrow 3)]- β -D-Glc (**L2**, MW 488.43 Da); difucosyllactose, α -L-Fuc-(1 \rightarrow 2)- β -D-Gal-(1 \rightarrow 4)-[α -L-Fuc-(1 \rightarrow 3)]- β -D-Glc (**L3**, MW 634.57 Da); 3'-sialyllactose, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L4**, MW 633.53 Da); 6'-sialyllactose, α -D-Neu5Ac-(2 \rightarrow 6)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L5**, MW 633.53 Da); 3'-sialyl-3-fucosyllactose, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)-[α -L-Fuc-(1 \rightarrow 3)]- β -D-Glc (**L6**, MW 779.67 Da); lacto-N-tetraose, β -D-Gal-(1 \rightarrow 3)- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L7**, MW 708.62 Da); lacto-N-fucopentaose I, α -L-Fuc-(1 \rightarrow 2)- β -D-Gal-(1 \rightarrow 3)- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc, (**L8**, MW 853.76 Da); lacto-N-fucopentaose II, β -D-Gal-(1 \rightarrow 3)-[α -L-Fuc-(1 \rightarrow 4)]- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L9**, MW 853.76 Da); lacto-N-difucohexaose I, α -L-Fuc-(1 \rightarrow 2)- β -D-Gal-(1 \rightarrow 3)-[α -L-Fuc-(1 \rightarrow 4)]- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L10**, MW 999.90 Da); difucosyllacto-N-hexaose a, β -D-Gal-(1 \rightarrow 4)-[α -L-Fuc-(1 \rightarrow 3)]-GlcNAc-(1 \rightarrow 6)-[α -L-Fuc-(1 \rightarrow 2)- β -D-Gal-(1 \rightarrow 3)- β -D-GlcNAc-(1 \rightarrow 3)]- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L11**, MW 1365.67 Da); sialyllacto-N-tetraose a, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L12**, MW 998.74 Da); sialyllacto-N-tetraose b, α -D-Neu5Ac-(2 \rightarrow 6)-[β -D-Gal-(1 \rightarrow 3)]- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L13**, MW 998.74 Da); disialyllacto-N-tetraose, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[α -D-Neu5Ac-(2 \rightarrow 6)]- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L14**, MW 1290.46 Da); sialylfucosyllacto-N-tetraose, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[α -L-Fuc-(1 \rightarrow 4)]- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc (**L15**, MW 1145.38 Da); sialyllacto N-

fucopentaose V, α -L-Fuc-(1→2)- β -D-Gal-(1→3)-[α -D-Neu5Ac-(2→6)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc (**L16**, MW 1145.38 Da); lacto-N-neo-tetraose, β -D-Gal-(1→4)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc (**L17**, MW 708.62 Da); lacto-N-fucopentaose III, β -D-Gal-(1→4)-[α -L-Fuc(1→3)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc (**L18**, MW 853.76 Da); difucosyl para-lacto-N-hexaose, β -D-Gal-(1→3)-[α -L-Fuc-(1→4)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)-[α -L-Fuc-(1→3)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc (**L19**, MW 1365.67 Da); sialyl-lacto-N-tetraose c, α -D-Neu5Ac-(2→6)- β -D-Gal-(1→4)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc (**L20**, MW 998.74 Da). The HMOs were purchased from IsoSep AB (Sweden) (**L1-L20**). The vendor guaranteed the structures of the HMOs. The GM1 pentasaccharide β -D-Gal-(1→3)- β -D-GalNAc-(1→4)-[α -D-Neu5Ac-(2→3)]- β -D-Gal-(1→4)- β -D-Glc (**L21**, MW 998.34 Da) (Figure S1, Supplementary Data) was purchased from Elicityl (Crolles, France). Compound **L22** (MW 8418.51 Da, Figure S2) was a gift from Prof. D. Bundle (Univ. of Alberta). As shown in Figure S2, **L22** consists of a glycosyl scaffold with five arms ending with a dimer of α -D-Gal(1→4)- β -D-Gal-(1→4)- β -D-Glc (P^k) trisaccharide. Stock solutions were prepared by dissolving each ligand in ultrafiltered water (Milli-Q, Millipore), to give a final concentration of 1 mM, and stored at -20 °C until used.

Proteins

The B subunits of CT (monomer MW 11 607 Da, CTB₅ MW 58 035 Da) and HLT (monomer MW 11 790 Da, HLTB₅ MW 58 950 Da) were purchased from Sigma-Aldrich Canada (Oakville, ON). Shiga toxin 1 B subunit (monomer MW 7685, Stx1B₅ MW 38 450 Da) and Stx2 holotoxin (Stx2, 72 220 Da) were gifts from Prof. G. Armstrong (Univ. of Calgary). A single chain variable fragment (scFv, MW 26 539 Da) of the monoclonal

antibody Se155-4, which served as a reference protein (P_{ref}) for the ESI-MS binding assays, was produced as described previously (Zdanov et al. 1994). CTB₅ and HLTB₅ were stored as lyophilized solids at +4 °C, then dissolved in 100 mM aqueous ammonium acetate to give 200 μ M stock solutions. Stx1B₅ and Stx2 were stored in the Tris buffer (pH 7) at -80 °C and thawed and buffer exchanged into 100 mM aqueous ammonium acetate using Amicon Ultra-4 centrifugal filters with a molecular weight cut-off of 10,000 Da (Millipore Corp., Bedford, MA) just before ESI-MS experiments. Prior to ESI-MS analysis, the protein solutions were diluted with 100 mM ammonium acetate (pH 7.2) to the desired concentration.

Mass spectrometry

The ESI-MS binding measurements were carried out using a 9.4T ApexQe FTICR mass spectrometer (Bruker, Billerica, MA) and a Synapt G2 quadrupole-ion mobility separation-time-of-flight (Q-IMS-TOF) mass spectrometer (Waters UK Ltd., Manchester, UK). For both instruments, nanoflow ESI (nanoESI) was performed using borosilicate glass tips (1.0 mm o.d., 0.68 mm i.d.) pulled to \sim 5 μ m o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). A capillary voltage of \sim 1.0 kV was applied to a Pt wire in the nanoESI tip to carry out ESI. A brief description of the instrumental conditions used for the two mass spectrometers is given below.

ApexQe 9.4T FTICR mass spectrometer. The ESI source was equipped with a metal sampling capillary (0.5 mm i.d.) maintained at 280 V. Nitrogen gas at a flow rate of 2.0 L min^{-1} and 90 °C was used as a drying gas. Ions were transmitted through the first funnel and skimmer at 150 V and 20 V, respectively, and then through the second funnel and skimmer at 7.6 V and 5.3 V, respectively. The ions were stored electrostatically in an

rf hexapole for 0.6 s and then further accumulated in a hexapole collision cell for 0.5 s. Following accumulation, the ions were transferred into the ion cell for detection. The front and back trapping plates of the cell were maintained at 0.9 and 1.0 V, respectively, throughout the experiment. The typical base pressure for the instrument was $\sim 1 \times 10^{-10}$ mbar. Data acquisition and analysis were performed using ApexControl, version 4.0 (Bruker Daltonics). A minimum of 30 transients with 128K data points per transient were used for each acquisition.

Synapt G2 mass spectrometer. Mass spectra were obtained using a sampling cone voltage of 30 V and an extraction cone voltage of 2 V. The source wave velocity and wave height were 200 m/s and 0.2 V, respectively. Gas flow rates were 2 mL min⁻¹ (trap), 180 mL min⁻¹ (He cell), 90 mL min⁻¹ (ion mobility cell). A source backing pressure was 3.2 mbar and the source block temperature was 70 °C. For ion transmission, the Trap and Transfer ion guides were 5 V and 2 V, respectively. For every acquisition at least 200 scans were collected. Data acquisition and processing were carried out using MassLynx (v 4.1).

ESI-MS affinity measurements

The apparent association constants ($K_{a,app}$) for the binding of the four bacterial toxins to each of the twenty HMOs (**L1 – L20**) were measured by the direct ESI-MS assay (Wang et al. 2003; El-Hawiet et al. 2011; El-Hawiet et al. 2012; Kitova et al. 2012) using a minimum of two different protein and HMO concentrations and a minimum of six replicate measurements at each concentration were performed. The binding measurements were carried out on one HMO at a time. In all cases, a reference protein (P_{ref}) was added to the solution in order to correct the mass spectra for the occurrence of nonspecific carbohydrate-protein binding during the ESI process (Sun et al. 2006). Following

correction for nonspecific binding, $K_{a,app}$ was calculated from the ratios (R_q) of the total abundance (Ab) of HMO ligand-bound protein (PL_q) to free protein (P) ions (eq 1.1) measured by ESI-MS for solutions of known initial concentrations of protein ($[P]_0$) and ligand ($[L]_0$), eq 1.2:

$$R_q = \frac{Ab(PL_q)}{Ab(P)} = \frac{[PL_q]}{[P]} \quad (1.1)$$

$$K_{a,app} = \frac{[PL_1]}{[P][L]} = \frac{R_1}{[L]_0 - \frac{\sum_{1 \leq q \leq 5} q R_q}{1 + \sum_{1 \leq q \leq 5} R_q} [P]_0} \quad (1.2)$$

For the toxin binding measurements carried out on solutions containing an HMO (**L1** or **L8**) and a high affinity ligand that binds to the native receptor binding sites (**L21** for CTB₅ and HLTB₅, or **L22** for Stx1B₅ and Stx2), the solutions were prepared at concentrations that lead to saturation or near saturation of the native receptor binding sites (such (CTB₅ + 5**L21**), (HLT₅ + 5**L21**), (Stx1B₅ + **L22**) and (Stx2 + **L22**) were the major protein species detected) and produce detectable signal for the toxin-HMO complexes. The $K_{a,app}$ were calculated from the mass spectra, following correction for nonspecific binding, in the same manner as described above, with the exception that P corresponded to the (CTB₅ + 5**L21**), (HLT₅ + 5**L21**), (Stx1B₅ + **L22**) and (Stx2 + **L22**) species and PL_q to the mixed complexes (CTB₅ + 5**L21** + q **L1**), (HLT₅ + 5**L21** + q **L1**), (Stx1B₅ + **L22** + q **L1**) or (Stx2 + **L22** + q **L1**), and (CTB₅ + 5**L21** + q **L8**), (HLT₅ + 5**L21** + q **L8**), (Stx1B₅ + **L22** + q **L8**) or (Stx2 + **L22** + q **L8**).

Acknowledgements

The authors wish to thank Prof. D. Bundle (University of Alberta) for generously donating carbohydrates used in this study and Prof. G. Armstrong (University of Calgary) for donating proteins used in this study. The authors acknowledge the Alberta Glycomics Centre and the Natural Sciences and Engineering Research Council of Canada for funding.

References

- Angstrom J, Teneberg S, Karlsson KA. 1994. Delineation and comparison of ganglioside-binding epitopes for the toxins of *Vibrio cholera*, *Escherichia coli*, and *Clostridium tetani*: Evidence for overlapping epitopes. *Proc Natl Acad Sci USA* 91:11859-11863.
- Bode L. 2012. Human milk oligosaccharides: Every baby needs a sugar mama *Glycobiology* 22:1147–1162.
- Bode L, Jantscher-Krenn E. 2012. Structure-function relationships of human milk oligosaccharides. *Adv Nutr.* 3:383S-391S.
- Dasgupta S, Kitov PI, Sadowska JM, Bundle DR. 2014. Discovery of inhibitors of Shiga toxin type 2 by on-plate generation and screening of a focused compound library. *Angew Chem Int Ed Engl.* 53:1510-1515.
- Dingle T, Wee S, Mulvey GL, Greco A, Kitova EN, Sun J, Lin S, Klassen JS, Palcic MM, Ng KK, Armstrong GD. 2008. Functional properties of the carboxy-terminal host cell-binding domains of the two toxins, TcdA and TcdB, expressed by *Clostridium difficile*. *Glycobiology* 18:698-706.
- El-Hawiet A, Kitova EN, Kitov P, Eugenio L, Ng KKS, Mulvey GL, Dingle TC, Szpacenko A, Armstrong GD, Klassen JS. 2011. Binding of *Clostridium difficile* toxins to human milk oligosaccharides. *Glycobiology* 21:1217–1227.
- El-Hawiet A, Kitova EN, Klassen JS. 2012. Quantifying carbohydrate-protein interactions by electrospray ionization mass spectrometry analysis. *Biochemistry* 51:4244-4253.
- Kitova EN, El-Hawiet A, Schnier PD, Klassen JS. 2012. Reliable determinations of protein-ligand interactions by direct ESI-MS measurements. Are we there yet? *J Am Soc Mass Spectrom.* 23:431-441.
- Fraser ME, Chernaiia MM, Kozlov YV, James MN. 1994. Crystal-structure of the holotoxin from shigella-dysenteriae at 2.5-angstrom resolution. *Nat Struct Biol.* 1:59–64.
- Fraser ME, Fujinaga M, Cherney MM, Melton-Celsa AR, Twiddy EM, O'Brien AD, James MN. 2004. Structure of Shiga toxin type 2 (Stx2) from *Escherichia coli* O157:H7. *J Biol Chem.* 279:27511–27517.
- Fuchs G, Mobassaleh M, Donohue-Rolfe A, Montgomery RK, Grand RJ, Keusch GT. 1986. Pathogenesis of Shigella diarrhea - rabbit intestinal-cell microvillus membrane-binding site for Shigella toxin. *Infect Immun.* 53:372-377.
- Heggelund JE, Haugen E, Lygren B, Mackenzie A, Holmner Å, Vasile F, Reina JJ, Bernardi A, Kregel U. 2012. Both El Tor and classical cholera toxin bind blood group determinants. *Biochem Biophys Res Commun.* 418:431-435.

Holmner A, Askarieh G, Okvist M, Kregel U. 2007. Blood group antigen recognition by *Escherichia coli* heat-labile enterotoxin. *J Mol Biol.* 371:754–764.

Holmner A, Mackenzie A, Okvist M, Jansson L, Lebens M, Teneberg S, Kregel U. 2011. Crystal structures exploring the origins of the broader specificity of *Escherichia coli* heat-labile enterotoxin compared to cholera toxin. *J Mol Biol.* 406:387-402.

Hong Q, Ruhaak LR, Totten SM, Smilowitz JT, German JB, Lebrilla CB. 2014. Label-free absolute quantitation of oligosaccharides using multiple reaction monitoring. *Anal Chem.* 86:2640-2647.

Howie PW, Forsyth JS, Ogston SA, Clark A, Florey CD, 1990. Protective effect of breast feeding against infection. *Brit Med J.* 300:11–16.

Idota T, Kawakami H, Murakami Y, Sugawara M. 1995. Inhibition of cholera toxin by human milk fractions and sialyllactose. *Biosci Biotechnol Biochem.* 59:417-419.

Jacobson JM, Yin J, Kitov PI, Mulvey, G, Griener TP, James MNG, Armstrong G, Bundle DR. 2014. The crystal structure of Shiga toxin type 2 with bound disaccharide guides the design of a heterobifunctional toxin inhibitor. *J Biol Chem.* 289:885–894.

Kale RR, McGannon CM, Fuller-Schaefer C, Hatch DM, Flagler MJ, Gamage SD, Weiss AA, Iyer SS. 2008 Differentiation between structurally homologous Shiga 1 and Shiga 2 toxins by using synthetic glycoconjugates. *Angew Chem Int Ed Engl.* 47:1265–1268.

Kitov PI, Sadowska JM, Mulvey G, Armstrong GD, Ling H, Pannu NS, Read RJ, Bundle DR. 2000. Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands. *Nature* 403:669-672.

Kitova EN, Kitov PI, Paszkiewicz E, Kim J, Mulvey GL, Armstrong GD, Bundle DR, Klassen JS. 2007. Affinities of Shiga toxins 1 and 2 for univalent and oligovalent Pk-trisaccharide analogs measured by electrospray ionization mass spectrometry. *Glycobiology* 17:1127-1137.

Kitova EN, El-Hawiet A, Schnier PD, Klassen JS. 2012. Reliable determinations of protein-ligand interactions by direct ESI-MS measurements. Are we there yet? *J Am Soc Mass Spectrom.* 23:431-441.

Kobata A. Structures and application of oligosaccharides in human milk. 2010. *Proc Jpn Acad Ser B Phys Biol Sci.* 86:1–17.

Lara-Villoslada F, Olivares M, Sierra S, Rodriguez JM, Boza J, Xaus J. 2007. Beneficial effects of probiotic bacteria isolated from breast milk. *Br. J. Nutr.* 98 (Suppl 1):S96–S100.

Ling H, Boodhoo A, Hazes B, Cummings MD, Armstrong GD, Brunton JL, Read RJ. 1998. Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* 37:1777–1788.

- Liu B, Newburg DS. 2013. Human milk glycoproteins protect infants against human pathogens. *Breastfeed Med.* 8:354-362.
- MacKenzie CR, Hiramata T, Lee KK, Altman E, Young NM. 1994. Quantitative analysis of bacterial toxin affinity and specificity for glycolipid receptors by surface plasmon resonance. *J Biol Chem.* 272:5533-5538.
- Mandal PK, Branson TR, Hayes ED, Ross JF, Gavín JA, Daranas AH, Turnbull WB. 2012. Towards a structural basis for the relationship between blood group and the severity of El Tor cholera. *Angew Chem Int Ed.* 51:5143–5146.
- Martin R, Langa S, Reviriego C, Jimenez E, Marin ML, Xaus J, Fernandez L, Rodriguez JM. 2003. Human milk is a source of lactic acid bacteria for the infant gut. *J Pediatr.* 143:754–758.
- Merritt EA, Sixma TK, Kalk KH, van Zanten BA, Hol WG. 1994. Galactose-binding site in *Escherichia coli* heat-labile enterotoxin (LT) and Cholera toxin (CT). *Mol Microbiol.* 13:745–753.
- Merritt EA, Kuhn P, Sarfaty S, Erbe JL, Holmes RK, Hol WG. 1998. The 1.25 Å resolution refinement of the cholera toxin B-pentamer: evidence of peptide backbone strain at the receptor-binding site. *J Mol Biol.* 282:1043-1059.
- Minke WE, Roach C, Hol WG, Verlinde CLMJ. 1999. Structure-based exploration of the ganglioside GM1 binding sites of *Escherichia coli* heat-labile enterotoxin and cholera toxin for the discovery of receptor antagonists. *Biochemistry* 38:5684-5692.
- Mudrak B, Kuehn MJ. 2010. Heat-labile enterotoxin: beyond GM1 binding. *Toxins* 2:1445-1470.
- Newburg DS, Ruiz-Palacios GM, Morrow AL, 2005. Human milk glycans protect infants against enteric pathogens. *Annu Rev Nutr.* 25:37–58.
- Newburg DS. 2009. Neonatal protection by an innate immune system of human milk consisting of oligosaccharides and glycans. *J Anim Sci.* 87(13 Suppl):26-34.
- Ruiz-Palacios GM, Cervantes LE, Ramos P, Chavez-Munguia B, Newburg DS. 2003. *Campylobacter jejuni* binds intestinal H (O) antigen (Fuc α 1, 2Gal β 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection. *J Biol Chem.* 278:14112–14120.
- Smilowitz JT, Lebrilla CB, Mills DA, German JB, Freeman SL. 2014. Breast milk oligosaccharides: structure-function relationships in the neonate. *Annu Rev Nutr.* 34:143–169.
- St Hilaire PM, Boyd MK, Toone EJ. 1994. Interaction of the Shiga-like toxin type 1 B-subunit with its carbohydrate receptor. *Biochemistry* 33:14452-14463.

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

Turnbull WB, Precious BL, Homans SW. 2004. Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry. *J Am Chem Soc.* 126:1047-1054.

Urashima T, Asakuma S, Leo F, Fukuda K, Messer M, Oftedal OT. 2012. The predominance of type I oligosaccharides is a feature specific to human breast milk. *Adv Nutr.* 3: 473S-482S.

Vasile F, Reina JJ, Potenza D, Heggelund JE, Mackenzie A, Kregel U, Bernardi A. 2014. Comprehensive analysis of blood group antigen binding to classical and El Tor cholera toxin B-pentamers by NMR. *Glycobiology* 24:766-778.

Wang W, 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.

Yu Y, Lasanajak Y, Song X, Hu L, Ramani S, Mickum ML, Ashline DJ, Prasad BV, Estes MK, Reinhold VN, Cummings RD, Smith DF. 2014. Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses. *Mol Cell Proteomics.* 13: 2944-2960.

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 U S A.* 91: 6423-6427.

Zivkovic AM, German JB, Lebrilla CB, Mills DA. 2010. Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc Natl Acad Sci USA.* 108:4653-4658.

Table I. Structures of the HMOs (**L1** – **L20**).

HMO	Structure
L1	α -L-Fuc-(1→2)- β -D-Gal-(1→4)- β -D-Glc
L2	β -D-Gal-(1→4)-[α -L-Fuc-(1→3)]- β -D-Glc
L3	α -L-Fuc-(1→2)- β -D-Gal-(1→4)-[α -L-Fuc-(1→3)]- β -D-Glc
L4	α -D-Neu5Ac-(2→3)- β -D-Gal-(1→4)- β -D-Glc
L5	α -D-Neu5Ac-(2→6)- β -D-Gal-(1→4)- β -D-Glc
L6	α -D-Neu5Ac-(2→3)- β -D-Gal-(1→4)-[α -L-Fuc-(1→3)]- β -D-Glc
L7	β -D-Gal-(1→3)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
L8	α -L-Fuc-(1→2)- β -D-Gal-(1→3)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
L9	β -D-Gal-(1→3)-[α -L-Fuc-(1→4)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
L10	α -L-Fuc-(1→2)- β -D-Gal-(1→3)-[α -L-Fuc-(1→4)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
L11	β -D-Gal-(1→4)-[α -L-Fuc-(1→3)]-GlcNAc-(1→6)-[α -L-Fuc-(1→2)- β -D-Gal-(1→3)- β -D-GlcNAc-(1→3)]- β -D-Gal-(1→4)- β -D-Glc

- L12** α -D-Neu5Ac-(2→3)- β -D-Gal-(1→3)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L13** α -D-Neu5Ac-(2→6)-[β -D-Gal-(1→3)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L14** α -D-Neu5Ac-(2→3)- β -D-Gal-(1→3)-[α -D-Neu5Ac-(2→6)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L15** α -D-Neu5Ac-(2→3)- β -D-Gal-(1→3)-[α -L-Fuc(1→4)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L16** α -L-Fuc-(1→2)- β -D-Gal-(1→3)-[α -D-Neu5Ac-(2→6)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L17** β -D-Gal-(1→4)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L18** β -D-Gal-(1→4)-[α -L-Fuc(1→3)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L19** β -D-Gal-(1→3)-[α -L-Fuc-(1→4)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)-[α -L-Fuc-(1→3)]- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
- L20** α -D-Neu5Ac-(2→6)- β -D-Gal-(1→4)- β -D-GlcNAc-(1→3)- β -D-Gal-(1→4)- β -D-Glc
-

Table II. Apparent association constants, $K_{a,app}$ (units of 10^2 M^{-1}), for HMOs (**L1 – L20**) binding to CTB₅, HLTB₅, Stx1B₅, Stx2, TcdA2 and TcdB1 determined in 100 mM aqueous ammonium acetate at 25 °C and pH 7.2 by the direct ESI-MS assay.^{a,b}

HMO	$K_{a,app}$					
	CTB₅	HLTB₅	Stx1B₅	Stx2	TcdA2^c	TcdB1^c
L1	21 ± 6	150 ± 60	10 ± 1	27 ± 11	20 ± 8	12 ± 5
L2	9 ± 1	39 ± 12	12 ± 3	44 ± 2	NB	NB
L3	16 ± 4	47 ± 5	6 ± 2	70 ± 20	NB	10 ± 3
L4	15 ± 5	NB	16 ± 3	59 ± 10	NB	NB
L5	9 ± 3	55 ± 10	13 ± 5	41 ± 2	NB	NB
L6	NB	26 ± 5	NB	31 ± 17	7 ± 5	NB
L7	30 ± 2	13 ± 3	6 ± 1	22 ± 2	15 ± 5	10 ± 5
L8	11 ± 2	120 ± 40	15 ± 6	38 ± 5	8 ± 1	31 ± 2
L9	12 ± 2	110 ± 10	9 ± 3	35 ± 11	7 ± 2	8 ± 4
L10	11 ± 4	14 ± 4	110 ± 4	25 ± 8	NB	18 ± 9
L11	18 ± 6	NB	10 ± 3	NB	NB	21 ± 5
L12	NB	28 ± 2	8 ± 5	28 ± 2	NB	NB
L13	36 ± 12	26 ± 3	NB	27 ± 2	11 ± 2	NB
L14	19 ± 4	NB	NB	13 ± 2	7 ± 2	NB
L15	27 ± 7	15 ± 4	9 ± 2	35 ± 1	NB	11 ± 6
L16	27 ± 5	13 ± 2	7 ± 1	NB	NB	NB
L17	14 ± 4	12 ± 5	16 ± 4	32 ± 1	NB	15 ± 2

L18	30 ± 12	75 ± 8	13 ± 4	52 ± 12	17 ± 2	9 ± 4
L19	10 ± 4	10 ± 2	10 ± 4	29 ± 10	NB	NB
L20	15 ± 5	NB	8 ± 3	NB	NB	20 ± 5

a. Errors correspond to one standard deviation. b. NB \equiv No binding detected. c. Values taken from reference El-Hawiet et al. 2011.

Figure captions

Figure 1. Structures of HMOs (**L1-L20**); glucose:●, galactose:●, N-acetylglucosamine:■

sialic acid:◆, fucose:▲.

Figure 2. Plot of $K_{a,app}$ values measured for HMOs (**L1-L20**) binding to CTB₅, Stx1B₅, Stx2, HLTB₅, TcdA2 and TcdB1.

Figure 3. Plots showing the fraction of ligand-bound protein calculated using a 1:1 protein-ligand binding model with two competing ligands - an HMO (with affinities of 10^2 to 10^5 M⁻¹) and a cellular receptor with (a) low affinity (10^3 M⁻¹), (b) medium affinity (10^6 M⁻¹) or (b) high affinity (10^9 M⁻¹) for the protein. The concentrations of protein, cellular receptor and HMO were 0.1 μM, 0.1 μM and 0.02 M, respectively.

Figure 4. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (100 mM, pH 7.2 and 25 °C) of (a) CTB₅ (5 μM), **L21** (30 μM), **L1** (500 μM) and P_{ref} (4 μM); the normalized distribution of **L1** bound to (CTB₅ + 5**L21**) complex after correction for nonspecific ligand binding is shown in the inset, (b) HLTB₅ (17 μM), **L21** (90 μM), **L1** (200 μM) and P_{ref} (4 μM); the normalized distribution of **L1** bound to (HLTB₅ + 5**L21**) complex after correction for nonspecific ligand binding is shown in the inset, (c) Stx1B₅ (10 μM), **L22** (15 μM), **L1** (200 μM) and P_{ref} (0.5 μM); the normalized distribution of **L1** bound to (Stx1B₅ + **L22**) complex after correction for nonspecific ligand binding is shown in the inset, (d) Stx2 (10 μM), **L22** (15 μM), **L1** (200 μM) and P_{ref} (4 μM); the normalized distribution of **L1** bound to

(Stx2 + **L22**) complex after correction for nonspecific ligand binding is shown in the inset.

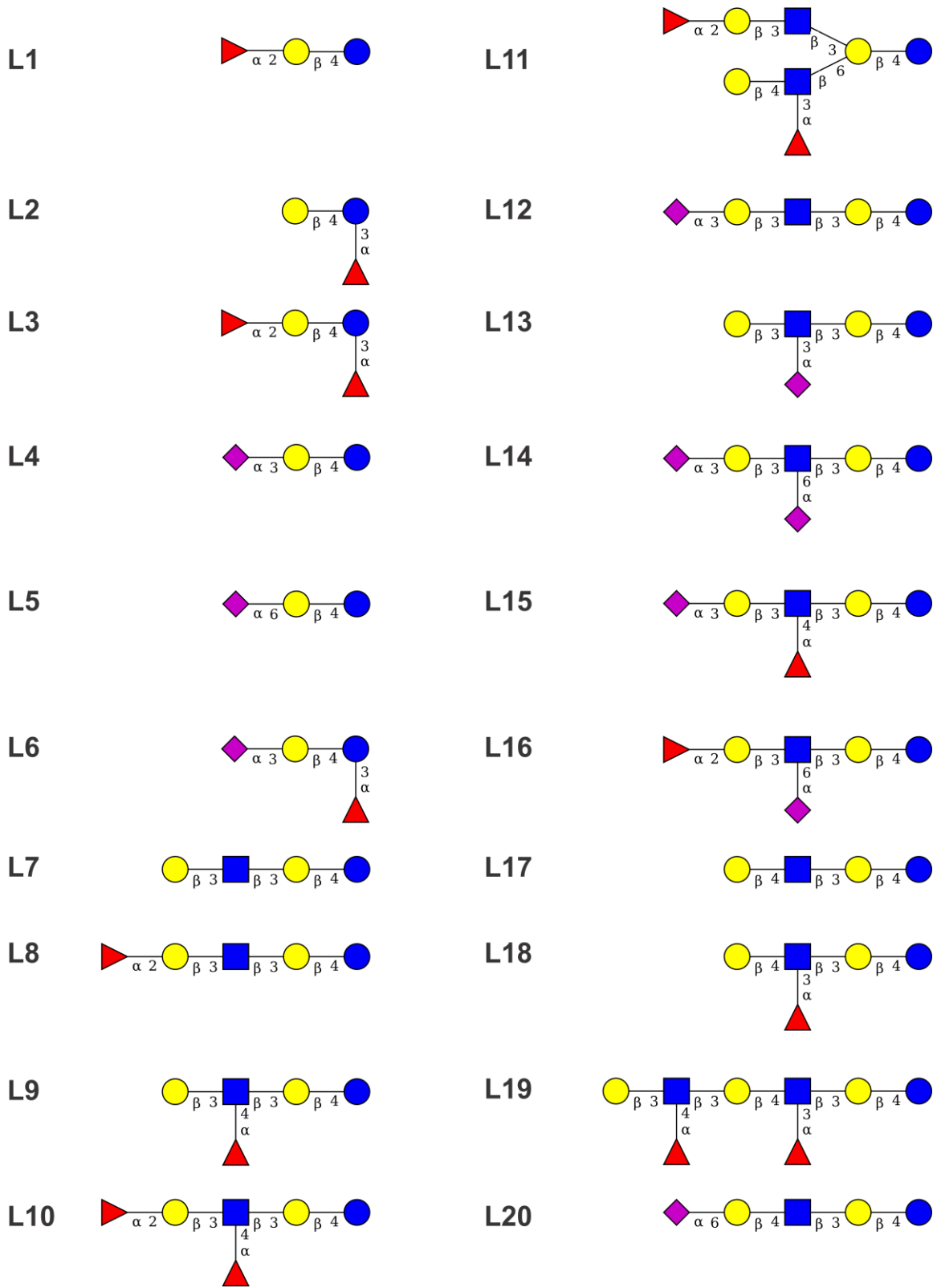


Figure 1

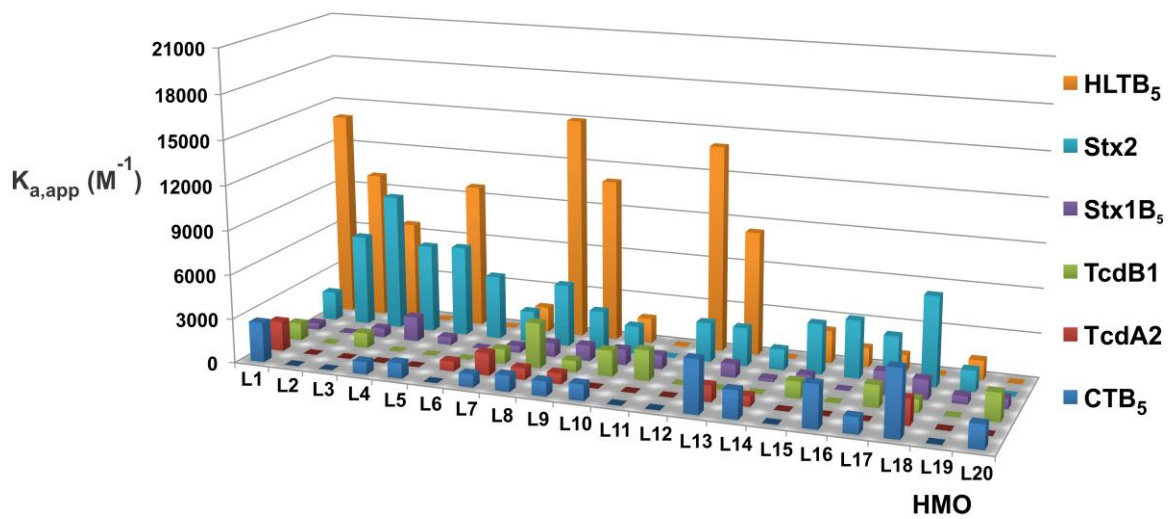


Figure 2

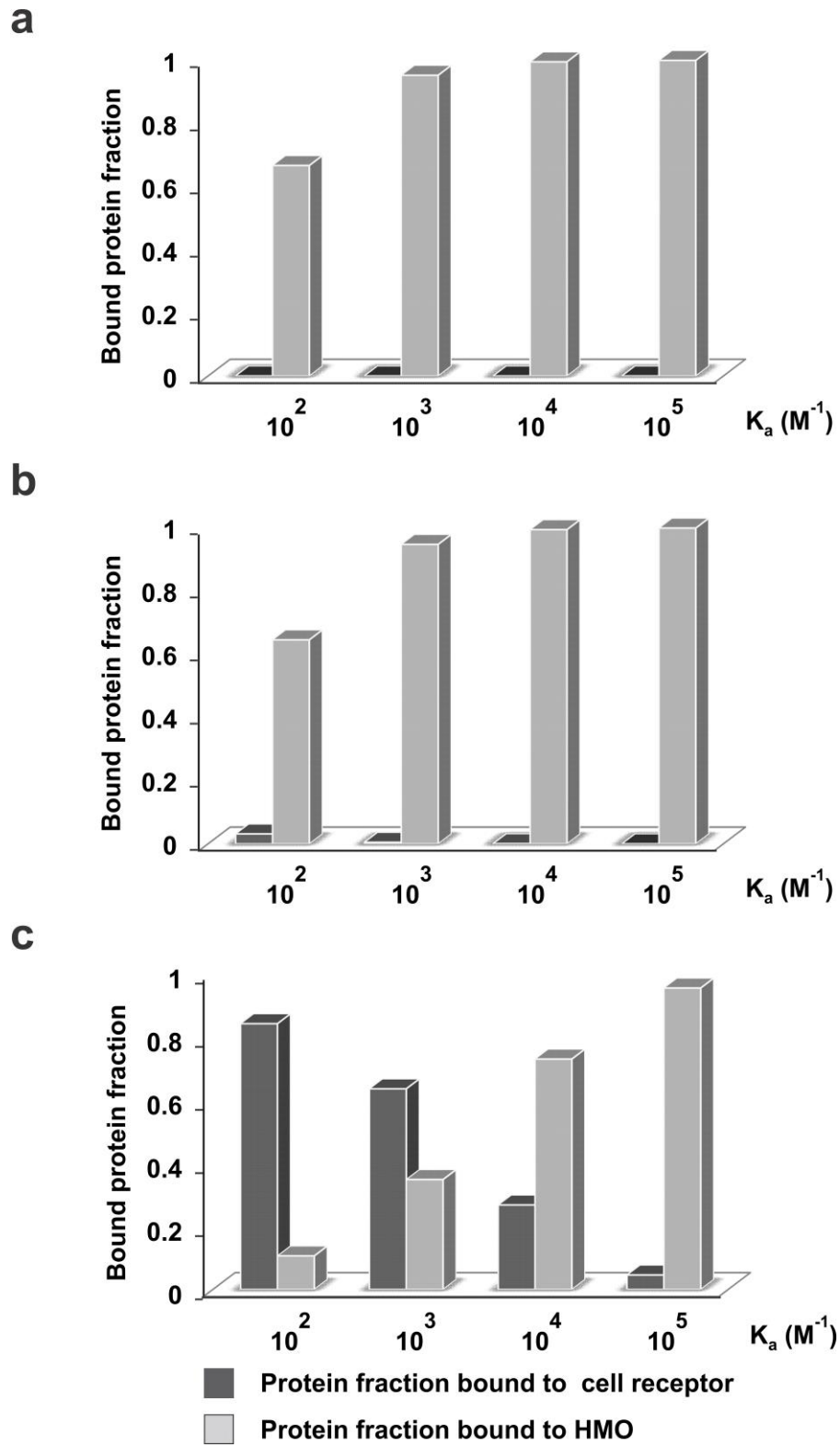


Figure 3

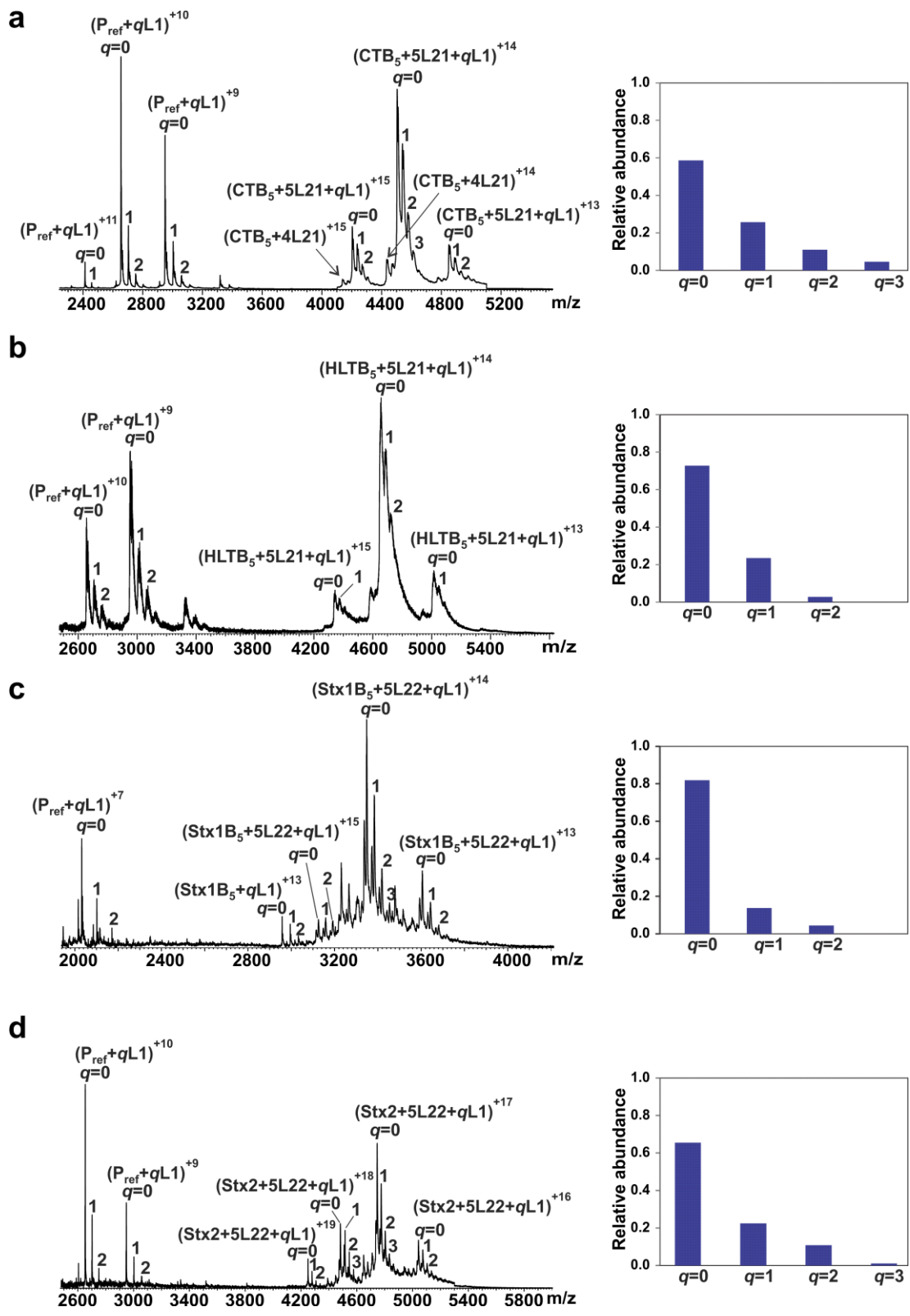


Figure 4

Supplementary Data For:

Recognition of Human Milk Oligosaccharides by Bacterial Exotoxins

Amr El-Hawiet, Elena N. Kitova, and John S. Klassen

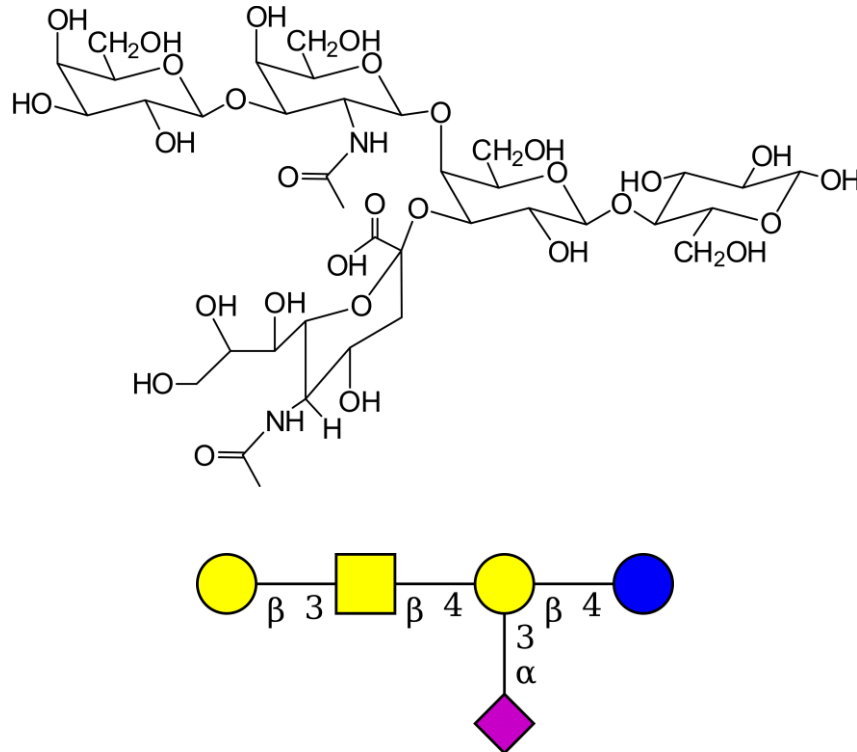







Figure S1. Structure of the **L21** (β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc-(1 \rightarrow 4)-[α -D-Neu5Ac-(2 \rightarrow 3)]- β -D-Gal-(1 \rightarrow 4)- β -D-Glc)  glucose:  lactose:  , N-acetylgalactosamine:  , sialic acid:  .

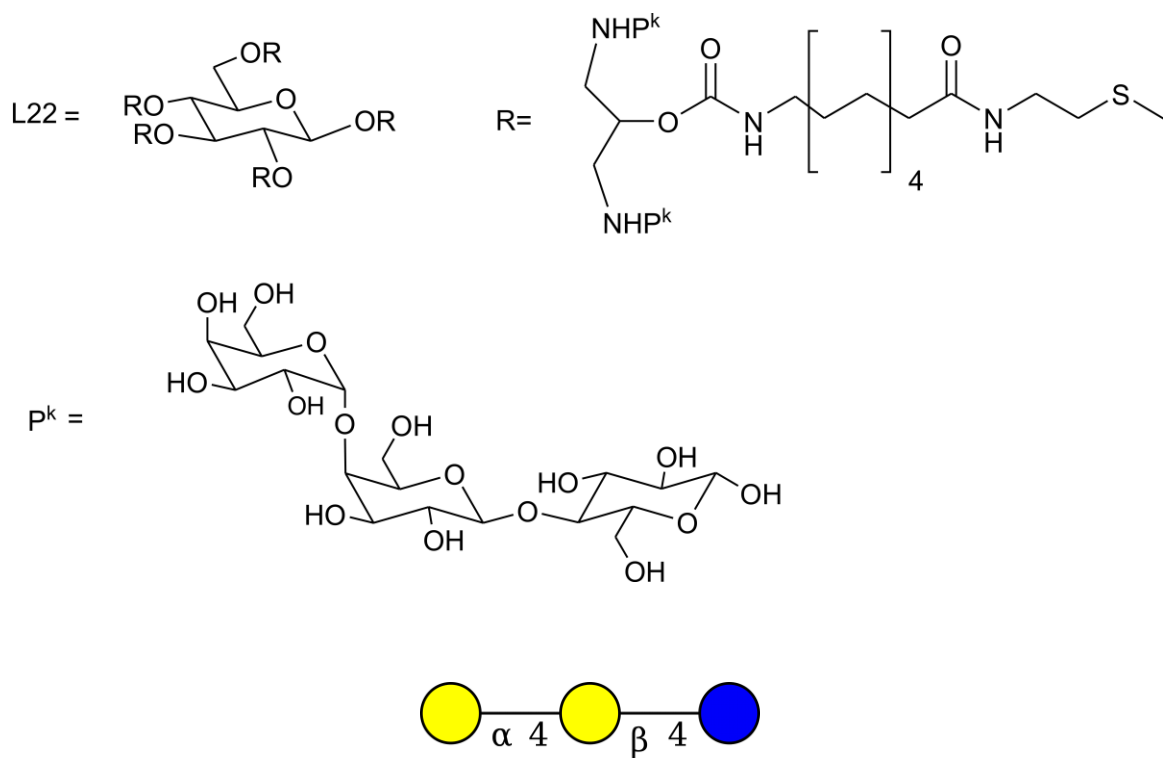




Figure S2. Structure of the **L22** (α -D-Gal(1 \rightarrow 4)- β -D-Gal-(1 \rightarrow 4)- β -D-Glc); glucose: , galactose .

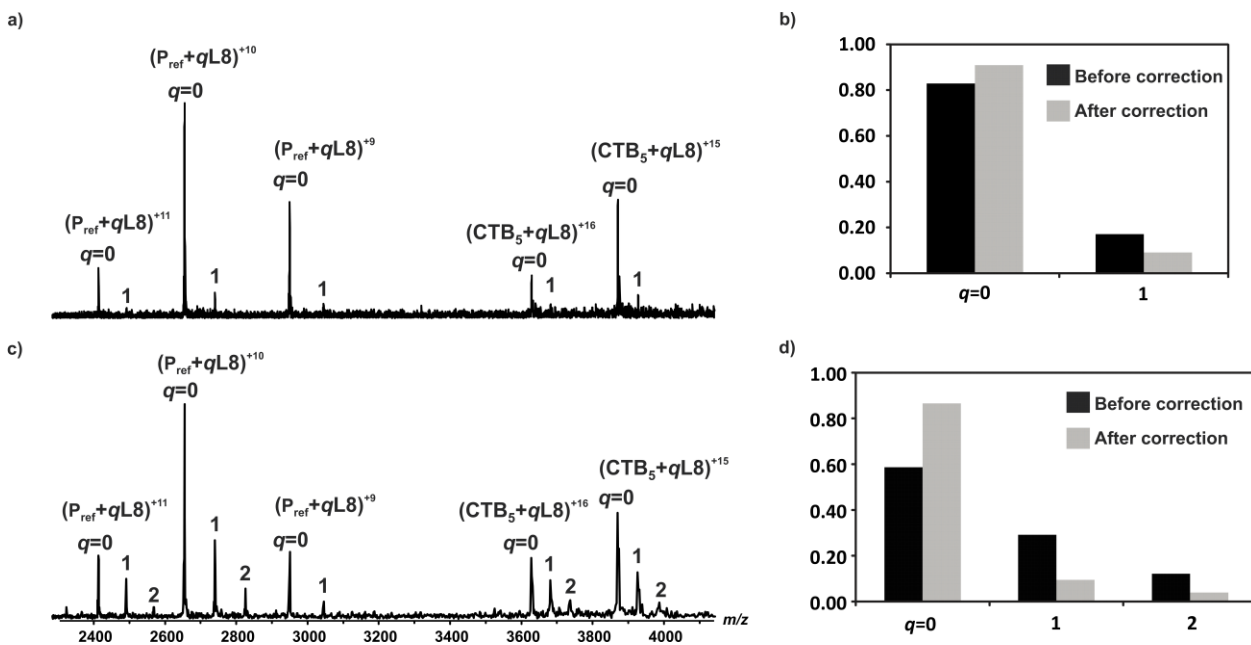


Figure S3. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (100 mM, pH 7 and 25°C) of CTB₅ (15 μM) and (a) **L8** (75 μM) or (c) **L8** (100 μM). A P_{ref} (8 μM) was added to each solution to quantify the extent of nonspecific protein-ligand binding during the ESI process. (b) and (d) Normalized distributions of **L8** bound to CTB₅, as determined from ESI mass spectra shown in (a) and (c), respectively.

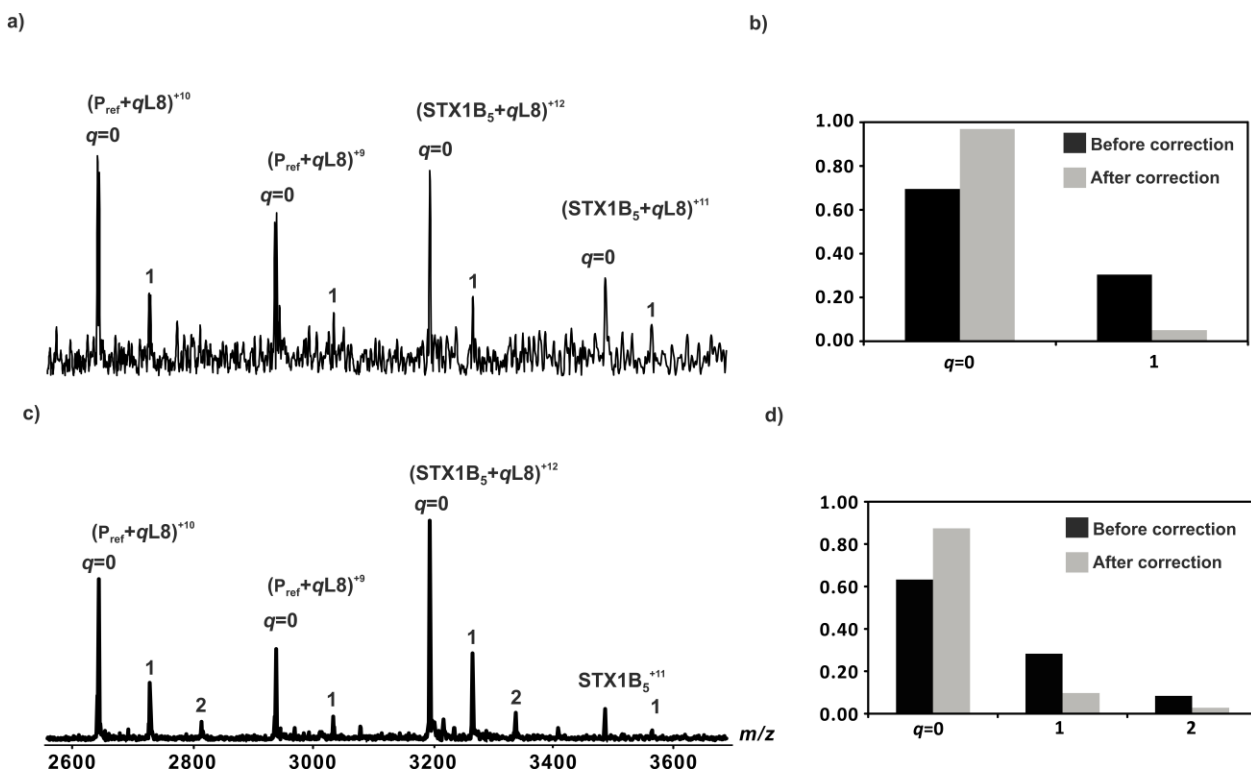


Figure S4. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (100 mM, pH 7 and 25°C) of Stx1B₅ (12 μM) and (a) **L8** (75 μM) or (c) **L8** (100 μM). A P_{ref} (1 μM) was added to each solution to quantify the extent of nonspecific protein-ligand binding during the ESI process. (b) and (d) Normalized distributions of **L8** bound to Stx1B₅, as determined from ESI mass spectra shown in (a) and (c), respectively, before and after correction for nonspecific binding.

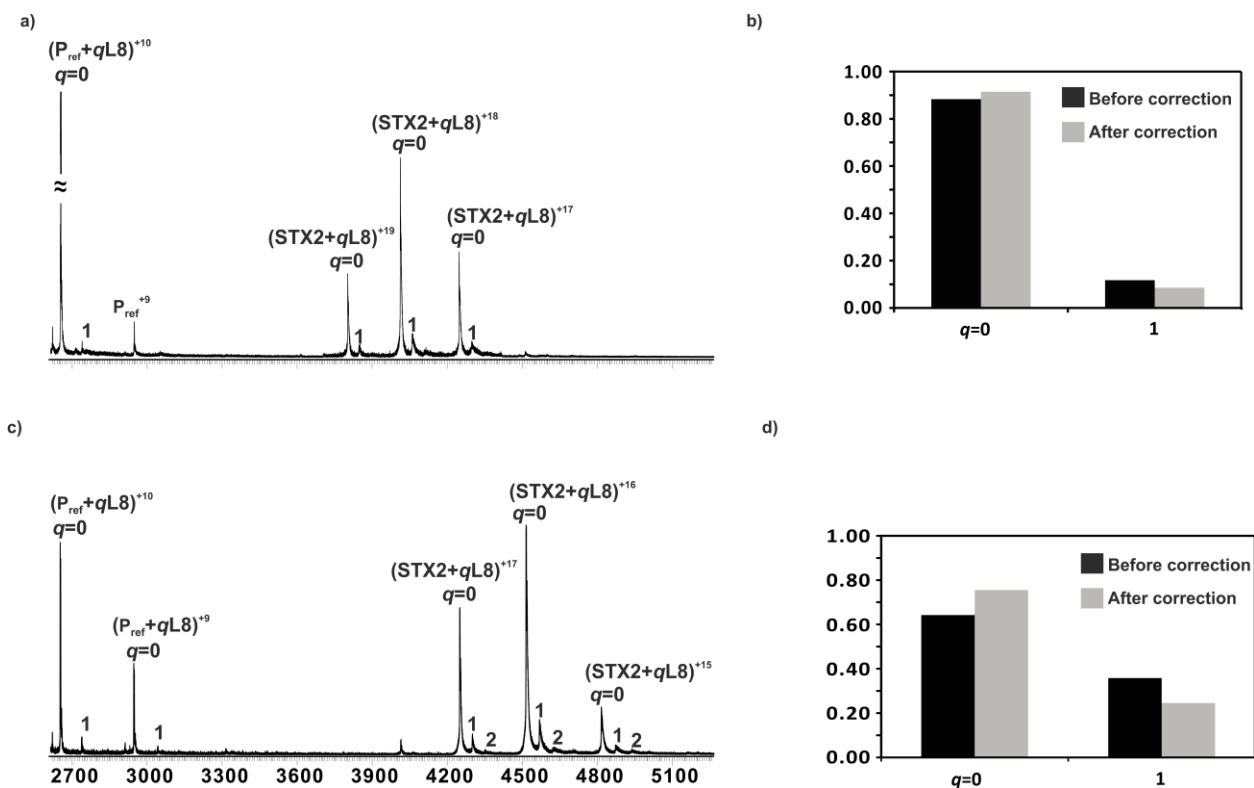


Figure S5. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (100 mM, pH 7 and 25°C) and Stx2 (10 μM) and (a) **L8** (25 μM) or (c) **L8** (60 μM). A P_{ref} (2 μM) was added to each solution to quantify the extent of nonspecific protein-ligand binding during the ESI process. (b) and (d) Normalized distributions of **L8** bound to Stx2, as determined from ESI mass spectra shown in (a) and (c), respectively, before and after correction for nonspecific binding.

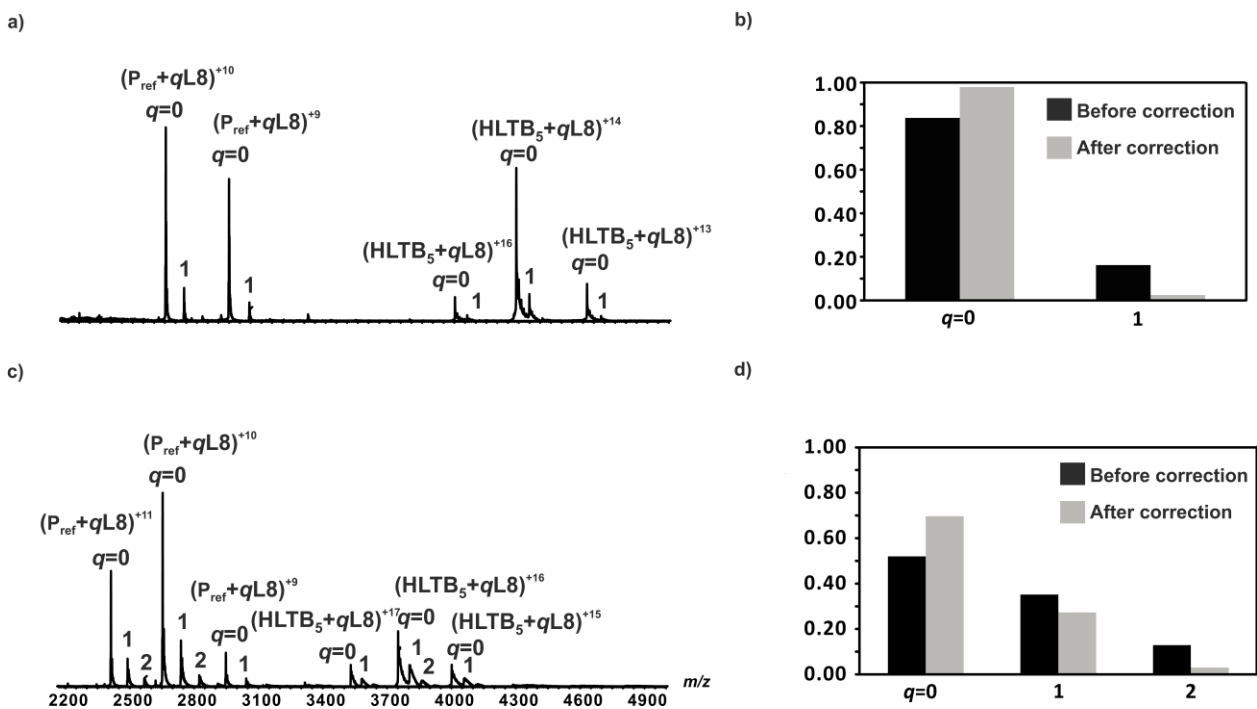


Figure S6 ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (100 mM, pH 7 and 25°C) and HLTB₅ (20 μM) and (a) L8 (25 μM) or (c) L8 (40 μM). A P_{ref} (1 μM) was added to each solution to quantify the extent of nonspecific protein-ligand binding during the ESI process. (b) and (d) Normalized distributions of L8 bound to HLTB₅, as determined from ESI mass spectra shown in (a) and (c), respectively, before and after correction for nonspecific binding.

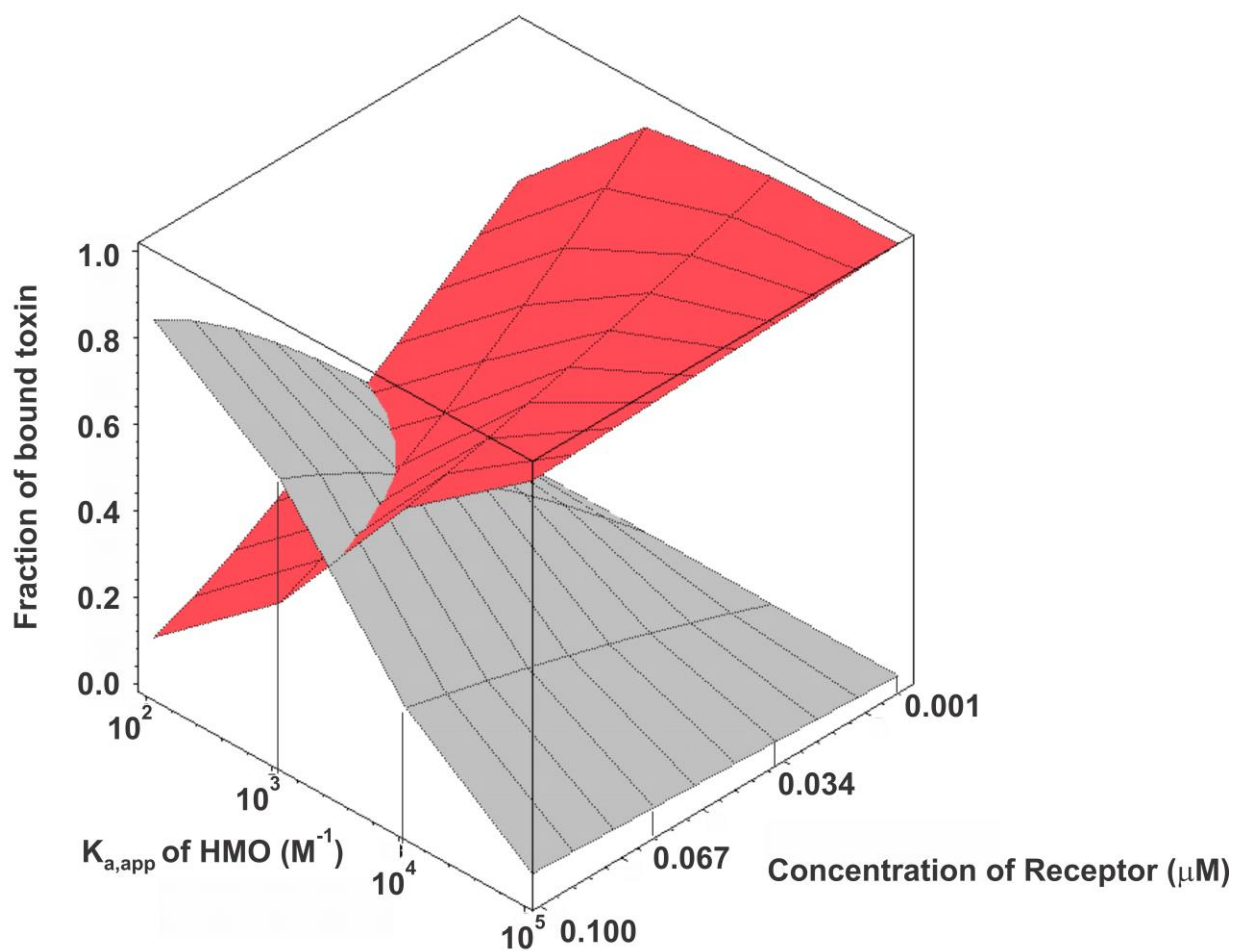


Figure S7. Fraction of ligand-bound protein calculated using a 1:1 protein-ligand binding model with two competing ligands - an HMO with affinities ranging from 10^2 to 10^5 M^{-1} (red surface) and a high affinity (10^9 M^{-1}) cellular receptor (gray surface) with concentrations ranging from 0.001 μM to 0.1 μM . The protein and HMO concentrations were 0.1 μM and 0.02 M, respectively.

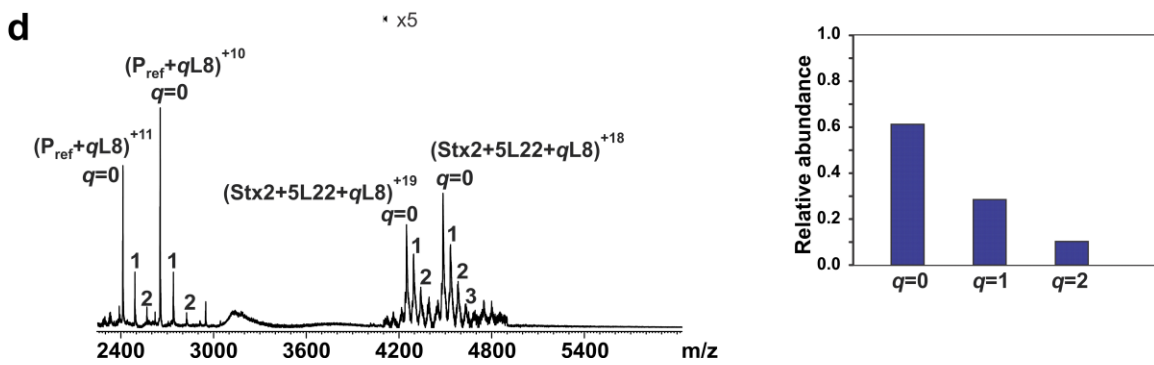
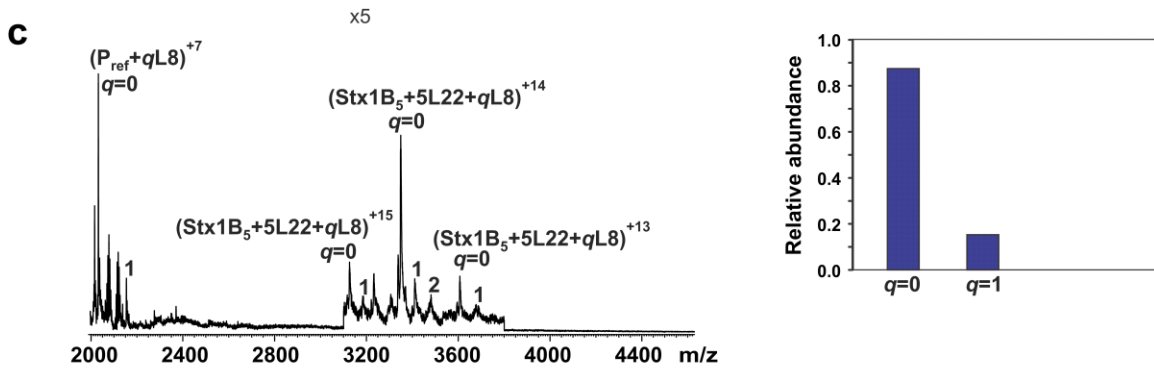
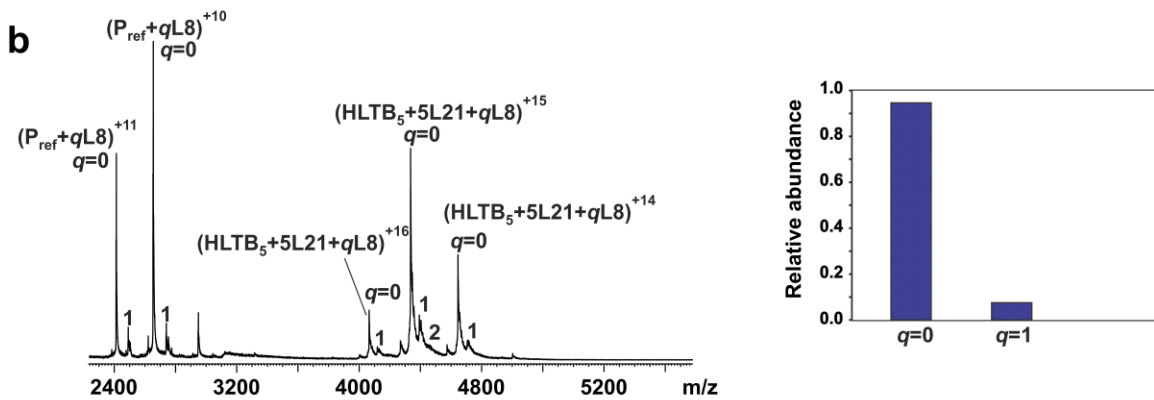
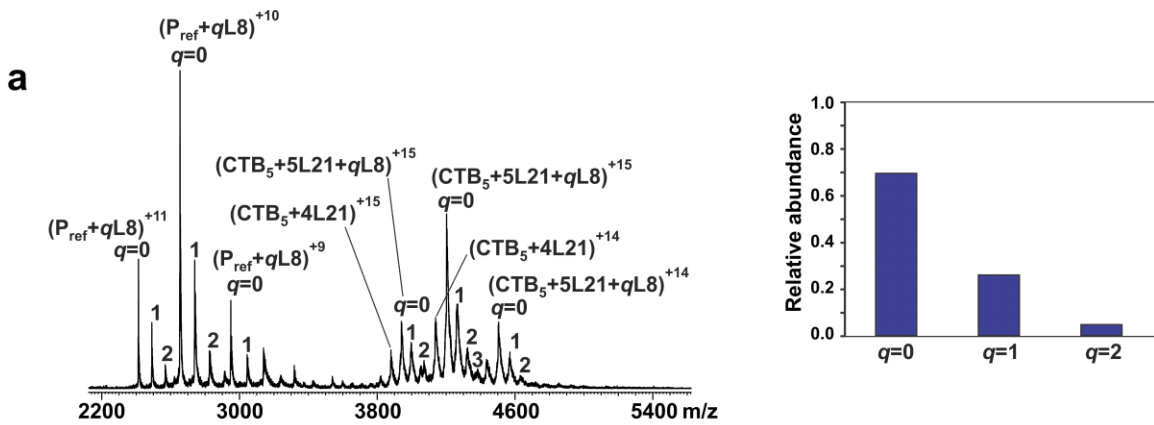


Figure S8. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate (100 mM, pH 7.2 and 25 °C) solutions containing (a) CTB₅ (4 μM), **L21** (22 μM), **L8** (200 μM) and P_{ref} (3 μM); the normalized distribution of **L8** bound to (CTB₅ + 5**L21**) complex after correction for nonspecific ligand binding is shown in the inset, (b) HLTB₅ (10 μM), **L21** (43 μM), **L8** (200 μM) and P_{ref} (3 μM); the normalized distribution of **L8** bound to (HLTB₅ + 5**L21**) complex after correction for nonspecific ligand binding is shown in the inset, (c) Stx1B₅ (4 μM), **L22** (4 μM), **L8** (200 μM) and P_{ref} (0.5 μM); the normalized distribution of **L8** bound to (Stx1B₅ + **L22**) complex after correction for nonspecific ligand binding is shown in the inset, (d) Stx2 (10 μM), **L22** (10 μM), **L8** (400 μM) and P_{ref} (3 μM); the normalized distribution of **L8** bound to (Stx2 + **L22**) complex after correction for nonspecific ligand binding is shown in the inset.