Affinities of Human Histo-Blood Group Antigens for Norovirus Capsid Protein Complexes

Ling Han[1], Elena N. Kitova[1], Ming Tan[2,3], Xi Jiang[2,3], Benjamin Pluvinage[4], Alisdair B. Boraston[4], and John S. Klassen[1]*

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

[2]*Division of Infectious Diseases, Cincinnati Children's Hospital Medical Center and*

[3]*Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA*

[4]*Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, Canada V8W 3P6*

*Corresponding Author's address:

Department of Chemistry

University of Alberta

Edmonton, AB CANADA T6G 2G2

Email: john.klassen@ualberta.ca

Telephone: (780) 492-3501

Abstract

The binding profiles of many human noroviruses (huNoVs) for human histo-blood group antigens (HBGAs) have been characterized. However, quantitative binding data for these important virus-host interactions are lacking. Here, we report on the intrinsic (per binding site) affinities of HBGA oligosaccharides for the huNoV VA387 virus-like particles (VLPs) and the associated subviral P particles measured using electrospray ionization mass spectrometry. The affinities of thirteen HBGA oligosaccharides, containing A, B and H epitopes, with variable sizes (disaccharide to tetrasaccharide) and different precursor chain types (type 1, 2, 3, 5 and 6), were measured for the P particle, while the affinities of the A and B trisaccharides and A and B type 6 tetrasaccharides for the VLP were determined. The intrinsic affinities of the HBGA oligosaccharides for the P particle range from 500 to 2300 M^{-1}; while those of the A and B trisaccharides and the A and B type 6 tetrasaccharides for the VLP range from 1000 to 4000 M^{-1}. Comparison of these binding data with those measured previously for the corresponding P dimer reveals that the HBGA oligosaccharides tested exhibit similar intrinsic affinities for the P dimer and P particle. The intrinsic affinities for the VLP are consistently higher than those measured for the P particle, but within a factor of three. While the cause of the subtle differences in HBGA oligosaccharide affinities for the P dimer and P particle and those for the VLP remains unknown, the present data support the use of P dimers or P particles as surrogates to the VLP for huNoV-receptor binding studies.

Introduction

Human noroviruses (huNoVs), which are the predominant cause of acute gastroenteritis outbreaks worldwide (Patel et al. 2008), is a genus of non-enveloped, single-stranded +RNA viruses in the *Caliciviridae* family. As there are no *in vitro* cell culture systems or suitable animal models available for huNoVs, the characterization of their structures and receptor interactions has relied on recombinant forms of their major capsid protein (VP1). For example, recombinant VP1, expressed using the baculovirus system, spontaneously assembles into a virus-like particle (VLP) that is devoid of genomic RNA for infection and replication but is structurally and antigenically indistinguishable from the authentic huNoV (Jiang et al. 1992). X-ray crystallography performed on the Norwalk VLP revealed that the intact particle is composed of 180 copies of VP1, which form a $T = 3$ icosahedral virion (Prasad et al. 1999). The formation of a smaller $T = 1$ viral capsid, consisting of 60 copies of VP1, has also been reported (White et al. 1997).

Production of VLPs using the baculovirus cell system is expensive and time consuming. Consequently, alternative protein complexes that can act as surrogates to VLPs are desirable. VP1 consists of two major domains bound by a flexible peptide linker, the N-terminal shell (S) domain and the protrusion (P) domain at the C-terminus (Prasad et al. 1999). The interior S domain is critical to maintaining the icosahedral structure of the virion, whereas the P domain forms a dimeric structure that is located on the outer surface and is implicated in the virus-receptor recognition process and, thus, cell entry. Expression of the P domain in *E. coli* has been shown to produce homodimers, called P dimers (Tan et al. 2004). The P dimers can also

assemble into larger complexes, a 12-mer small P particle (Tan et al. 2011) and a 24-mer P particle (Tan and Jiang 2005b, Tan et al. 2008). A recent native electrospray ionization mass spectrometry (ESI-MS) study revealed that, in 100 mM ammonium acetate (pH 7.4), the P particle is made up of approximately 85% of 24-mer and 15% of 18-mer (Bereszczak et al. 2012). Importantly, both the sub-viral particles and the P dimer are believed to retain the authentic antigenicity and receptor binding capability of the VLP (Tan et al. 2004, Tan and Jiang 2005b, Tamminen et al. 2012) and are, therefore, seen as attractive substitutes to VLPs for investigating the nature of huNoV-host cell interactions and discovering potential inhibitors. However, to our knowledge, a quantitative comparison of the receptor binding properties of a huNoV VLP and its corresponding P particle and P dimer has yet to be carried out.

It is well established that many huNoVs recognize human histo-blood group antigens (HBGAs), which are found on the surfaces of red blood cells and mucosal epithelial cells in the form of glycoproteins and glycolipids (Oriol 1990, Ravn and Dabelsteen 2000), as cellular receptors or attachment factors (Hutson et al. 2002, Hutson et al. 2003, Huang et al. 2005, Tan and Jiang 2005a). The HBGAs are divided into four types, namely A, B, H and Lewis, based on the carbohydrate structure at the non-reducing end. Additionally, each HBGA is further divided into six subtypes (type 1 to 6) based on the carbohydrates structure at the reducing end. To date, there have been few quantitative binding studies performed on the capsid proteins of huNoVs. Using saturation-transfer difference nuclear magnetic resonance (STD-NMR) spectroscopy, Peters and coworkers (Fiege et al. 2012) estimated the apparent association constants ($K_{a,app}$) of VLP from huNoV Ast6139 (GII.4 strain) for a variety of HBGA oligosaccharides (including H

disaccharide, A and B trisaccharides, H type 1, 2 and 6 trisaccharides, Lewis a, b, x, y and sialyl-Lewis a and x oligosaccharides) to be in the ~10^4 M^{-1} range. Based on these results, the intrinsic (per binding site) association constants ($K_{a,int}$) are predicted to be in the ~10^2 M^{-1} range. Using surface plasmon resonance spectroscopy, Belliot and colleagues (de Rougemont et al. 2011) analyzed BSA-conjugated type H, A and B carbohydrates binding to VLPs isolated from six GII.4 strains of huNoV. However, affinities could not be accurately determined due to the uncertainty in the number of active binding sites in each VLP, as well as the number of immobilized (on the chip) carbohydrates that participated in interactions. Quantitative $K_{a,app}$ values have also been reported for huNoV P dimers. Using STD-NMR spectroscopy, Kwong and co-workers (Hansmann et al. 2012) measured $K_{a,app}$ values of 2200 M^{-1} and 2600 M^{-1} for fucose and H type 2 trisaccharide, respectively, binding to a GII.10 huNoV P dimer. More recently, the $K_{a,int}$ values of a library of 42 HBGA oligosaccharides for the huNoV VA387 P dimer were measured using the direct ESI-MS assay (Han et al. 2013). The results of this study revealed that the P dimer exhibits broad specificity and binds to A, B, H and Lewis type antigens, although with low affinities ($\leq$3000 M^{-1}).

Here, we report $K_{a,int}$ values of variety of HBGA oligosaccharides for the P particle and VLP of huNoV VA387. Application of the direct ESI-MS assay, which was used for the P dimer measurements, to these large capsid protein complexes is not feasible due to the difficulty in resolving the free and ligand-bound forms of the protein complexes. Consequently, an adaptation of the *proxy protein* ESI-MS method (El-Hawiet et al. 2012) was used to carry out the measurements. The *proxy protein* ESI-MS method combines direct ESI-MS binding

measurements and competitive protein binding to evaluate protein-ligand affinities. Specifically, a proxy protein (P_{proxy}), which binds to the ligand of interest with known affinity and can be detected directly by ESI-MS, is used to quantitatively monitor the extent of ligand binding to the protein of interest – the P particle or VLP in the present case. Using this method the $K_{a,int}$ values of thirteen HBGA oligosaccharides, containing A, B and H epitopes, with variable sizes (disaccharide to tetrasaccharide) and different precursor chain types (type 1, 2, 3, 5 and 6), for the P particle were measured; affinities for the A and B trisaccharides and A and B type 6 tetrasaccharides for the VLP were also quantified. Comparison of these values with binding data recently reported for the P dimer of huNoV VA387 provides a unique opportunity to assess the similarity of the HBGA binding sites in the P dimer, P particle and VLP.

Experimental Section

Proteins

For VLP production, the gene encoding the capsid protein (VP1) of huNoV VA387 (GII.4, GenBank accession no. AY038600) was cloned and expressed through the Bac-to-Bac baculovirus expression system (Invitrogen Life Technologies, Grand Island, NY) as described elsewhere (Jiang et al. 2002, Huang et al. 2003). VLPs, comprised of 180 copies of VP1 (which has a molecular weight (MW) of 58,887 Da), assembled spontaneously. For P particle and P dimer production, the gene fragments encoding the P domain (residues 222-539) of VP1 with and without a C-terminus-fused peptide CDCRGDCFC, respectively, were cloned and expressed in bacteria through the GST-Gene Fusion System (GE-Healthcare Life Sciences, Piscataway, NJ) following a protocol described previously (Tan et al. 2004, Tan et al. 2005b). The resulting P

dimer (MW 69,312 Da) or P particle (24-mer, MW 865,036 Da) that assembled spontaneously were purified using glutathione affinity chromatography, followed by gel filtration chromatography.

The N-terminal family 51 carbohydrate-binding module (CBM, expected MW 20,735 Da) from *Streptococcus pneumonia* SP3-BS71 GH98 was recombinantly produced in *Escherichia coli* and purified by Ni^{2+} immobilized metal affinity chromatography (GE-Healthcare Life Sciences, Piscataway, NJ) using procedures described elsewhere (Higgins et al. 2011). A recombinant fragment of the C-terminus of human galectin-3 (Gal-3C, MW 16,330 Da) was a gift from Prof. C. Cairo (University of Alberta). A recombinant soluble fragment of α-(1→3)N-acetyl galactosaminyltransferase (GTA, homodimer, MW 69,040 Da), which contains a full C-terminal and catalytic domain, as well as a truncated N-terminal domain, was expressed in bacterial cells and purified by ion exchange chromatography using a SP-Sepharose FF resin (GE-Healthcare Life Sciences, Piscataway, NJ), followed by affinity purification using a UDP-hexanolamine resin (Seto et al. 1997). Bovine ubiquitin (Ubq, MW 8,565 Da), which was purchased from Sigma-Aldrich Canada (Oakville, Canada), and recombinant single chain fragment (scFv, MW 26,539 Da) of monoclonal antibody Se155-4, which was produced using procedures described before (Zdanov et al. 1994), served as reference proteins (P_{ref}) for the binding measurements. Each protein was dialyzed and concentrated against 50 mM aqueous ammonium acetate (pH 7) using Vivaspin 0.5 mL centrifugal filters (Sartorius Stedim Biotech, Göttingen, Germany) with a 10 kDa MW cutoff. The concentrations were measured by UV spectroscopy. Protein stock solutions were stored at −80 °C until used.

Carbohydrates

The oligosaccharides (**L1** – **L14**) were a gift from Prof. T. Lowary (University of Alberta) (Meloncelli et al. 2009, 2010, 2011). Their structures are shown in Figure S1 (Supplementary Materials). To prepare stock solutions, solid sample of each compound was weighed and dissolved in a known volume of ultrafiltered water (Milli-Q, Millipore, Billerica, MA) to yield a final concentration of 1 mM. These solutions were stored at -20 °C until needed.

Mass spectrometry

The ESI-MS measurements were carried out in positive ion mode using either a Synapt G2S quadrupole-ion mobility separation-time of flight mass spectrometer (Waters, Manchester, UK) or a 9.4T ApexQe Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker-Daltonics, Billerica, MA). The Synapt G2S mass spectrometer, with its high mass capabilities, was used for the ESI-MS analysis of the huNoV P particle and VLP. Both the G2S and the ApexQe mass spectrometers were used to carry out the direct and *proxy protein* ESI-MS measurements. Both mass spectrometers were equipped with nanoflow ESI (nanoESI) sources. To carry out ESI, the sample solution was loaded into a nanoESI tip pulled from a borosilicate capillary (1.0 mm o.d., 0.68 mm i.d.) using a micropipette puller (P-1000, Sutter Instruments, Novato, CA). To initiate ESI, a voltage of 1.0 – 1.4 kV was applied to a platinum wire inserted into the nanoESI tip. A detailed description of the instrumental conditions used to implement the direct and *proxy protein* ESI-MS binding measurements can be found elsewhere (Han et al. 2013, El-Hawiet et al. 2012). Unless otherwise indicated, the ESI-MS measurements performed on the ApexQe mass spectrometer were carried out using 50 mM aqueous ammonium acetate solutions

(pH 7 and 25 °C) containing protein and ligand of interest, while those on the Synapt G2S mass spectrometer were carried out using 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing protein and ligand of interest. To carry out the *proxy protein* ESI-MS measurements, a P_{proxy} was also added to the solutions. In all cases, the reference protein method, which involves the addition of a P_{ref} to the solution, was used to correct the mass spectra for the occurrence of nonspecific protein-carbohydrate binding during the ESI process. A complete description of the correction method can be found elsewhere (Sun et al. 2006, Kitova et al. 2012).

Gel filtration Chromatography

Gel filtration chromatography was performed using a Superdex 200 size exclusion column (HiLoad 16/60, 120 mL bed volume, GE Healthcare Life Sciences, Piscataway, NJ) powered by an AKTA fast-performance liquid chromatography (FPLC) system (model 920, GE Healthcare Life Sciences, Piscataway, NJ). The column was equilibrated and run in 1X phosphate-buffered saline (PBS, pH 7.4) at a flow rate of 1.0 mL min^{-1}; 1.0 mL of the VLP sample (5 mg mL^{-1}) was loaded onto the column using a manual injector. The MW of the proteins in each elution volume was calibrated with the Gel Filtration Calibration kit (GE Healthcare Life Sciences, Piscataway, NJ) and purified P particle as described elsewhere (Wang et al. 2013 and 2014).

Determination of K_a values

Direct ESI-MS assay. The direct ESI-MS assay was used to quantify oligosaccharide affinities for CBM and Gal-3C, two of four P_{proxy}'s used in this study, and the affinity of **L14** for the P dimer. For a 1:1 protein-ligand complex the association constant (K_a) is calculated from the

abundance (*Ab*) ratio (*R*) of the ligand-bound (PL) to free protein (P) ions measured by ESI-MS and the initial concentrations of protein ($[P]_0$) and ligand ($[L]_0$), eq 1 (Kitova et al. 2012):

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

where *R* is taken to be equal to the equilibrium concentration ratio in the solution, eq 2:

$$R = \frac{\sum Ab(PL)}{\sum Ab(P)} = \frac{[PL]}{[P]} \quad (2)$$

The abundances of free and ligand-bound proteins were calculated as the sum of the peak areas for all of the charge states detected for each species. In cases where ligand binding was weak (K_a $<10^4$ M^{-1}), a titration approach was employed, whereby the protein concentration was kept constant and the ligand concentration was varied (Daniel et al. 2002). Nonlinear regression analysis of the concentration-dependence of the fraction of ligand-bound protein, [*R*/(*R*+1)] was used to determine K_a, eq 3 (Kitova et al. 2012):

$$\frac{R}{R+1} = \frac{1 + K_a[L]_0 + K_a[P]_0 - \sqrt{(1 - K_a[L]_0 + K_a[P]_0)^2 + 4K_a[L]_0}}{2K_a[P]_0} \quad (3)$$

Proxy protein ESI-MS assay. The *proxy protein* ESI-MS assay was recently developed to quantify the affinities of protein–ligand complexes that could not be directly measured by ESI-MS (El-Hawiet et al. 2012). Briefly, the method involves the use of the direct ESI-MS assay to monitor the extent of ligand binding to a proxy protein (P_{proxy}) in the presence of the target protein (P) (El-Hawiet et al. 2012). The general expression relating the intrinsic association constant of P ($K_{a,int,P}$), with *h* identical and independent binding sites, to R_{proxy}, the abundance ratio of ligand-bound to free P_{proxy} ions, for a P_{proxy} with a single binding site is given by eq 4a

(El-Hawiet et al. 2012):

$$K_{a,int,P} = \frac{K_{a,P_{proxy}} / R_{proxy}}{\frac{h[P]_0}{[L]_0 - \frac{[P_{proxy}]_0 R_{proxy}}{R_{proxy}+1} - \frac{R_{proxy}}{K_{a,P_{proxy}}}} - 1} \quad (4a)$$

As described in more detail below, the huNoV P particle and VLP investigated in the present study exist in multiple forms in solution. The P particle exists predominately as a 24-mer, however, both the 18-mer and 36-mer have also been detected (Bereszczak et al. 2012). Similarly, although VLP exists predominantly as a 180-mer, the 60-mer and 80-mer, as well as dimer, are also present (Shoemaker et al. 2010). Consequently, it is more appropriate to rewrite eq 4a in terms of the total concentration of ligand binding sites ($[P]_{m,0} = \Sigma h_i[P_i]_0$), which is equal to the number of protomers that make up the assembly, eq 4b:

$$K_{a,P,int} = \frac{K_{a,P_{proxy}} / R_{proxy}}{\frac{[P]_{m,0}}{[L]_0 - \frac{[P_{proxy}]_0 R_{proxy}}{R_{proxy}+1} - \frac{R_{proxy}}{K_{a,P_{proxy}}}} - 1} \quad (4b)$$

It must be stressed that the $K_{a,int,P}$ values determined in this way represent the weighted average of the affinities of the different assemblies present in solution.

The implementation of the *proxy protein* method was previously demonstrated using a P_{proxy} with a single binding site. Two of the P_{proxy} used in the present study possess multiple (two) ligand binding sites. In this case, $K_{a,P,int}$ can be found using eq 5:

$$K_{a,P,int} = \frac{K_{a,P_{proxy},int}/\theta}{\frac{[P]_{m,0}}{[L_0] - \frac{g[P_{proxy}]_0\theta}{\theta+1} - \frac{\theta}{K_{a,P_{proxy},int}}} - 1} \quad (5)$$

where $K_{a,proxy,int}$ and g are the intrinsic affinity of P_{proxy} and the number of binding sites in P_{proxy}, respectively. The ratio of occupied-to-free binding sites in the P_{proxy} (θ) is calculated using eq 6:

$$\theta = \frac{[\text{occupied binding sites}]_{proxy}}{[\text{free binding sites}]_{proxy}} = \frac{\sum_{j=1}^{g} jR_{proxy,j}}{g(1+\sum_{j=1}^{g} R_{proxy,j})-\sum_{j=1}^{g} jR_{proxy,j}} \quad (6)$$

where $R_{proxy,j}$ corresponds to the concentration ratio of ligand-bound (to j molecules of L) to free P_{proxy} and is taken to be equal to the abundance (Ab) ratio of ligand-bound (to j molecules of L) to free P_{proxy} gas-phase ions, eq 7:

$$R_{proxy,j} = \frac{\sum Ab\left(P_{proxy}L_j\right)}{\sum Ab\left(P_{proxy}\right)} = \frac{\left[P_{proxy}L_j\right]}{\left[P_{proxy}\right]} \quad (7)$$

In the case of a P_{proxy} with two ligand binding sites (i.e., g = 2), θ is given by eq 8:

$$\theta = \frac{R_{proxy,1} + 2R_{proxy,2}}{2 + R_{proxy,1}} \quad (8)$$

Given in the Supplementary Materials is a derivation of the equations relevant to the implementation of the *proxy protein* method using a P_{proxy} with multiple ligand binding sites.

Results and Discussion

ESI-MS analysis of huNoV VA387 P particle and VLP

Representative ESI mass spectra measured for a 200 mM ammonium acetate aqueous solution (pH 7 and 25 °C) containing 3 μM of P particle (which corresponds to 72 μM of monomer) or 0.2 μM VLP (corresponding to 36 μM VP1) are shown in Figure 1. As seen in Figure 1a, the P particle is present predominantly as a 24-mer (with a charge state distribution centred around +67), along with the 18-mer (charge state distribution centred around +56) at lower abundance.

These observations are consistent with those reported previously (Bereszczak et al, 2012). The MWs of the 24-mer and 18-mer, 865,000 ± 540 Da and 648,900 ± 400 Da, respectively, are in reasonably good agreement with the expected values calculated from the protein sequence, 865,036 Da and 648,782 Da, respectively (Tan and Jiang 2005b). As seen in Figure 1b, the VP1 monomer (MW 58,887 Da), 60-mer (~3.5 MDa), 80-mer (~4.7 MDa) and 180-mer (~10 MDa) are all present in solution. Notably, the distribution of VP1 species is reproducible over a period of weeks, Figure S2 (Supplementary Materials). The smaller oligomers, which are believed to exist in dynamic equilibrium with the 180-mer, were previously observed by native ESI-MS, ion mobility separation MS, as well as atomic force atomic force microscopy (Uetrecht et al. 2011, Shoemaker et al. 2010). Due to the high MWs of the oligomers, it was not possible to resolve individual charge states and, thus, their identification was based on previously reported ESI-MS results (Shoemaker et al. 2010). Moreover, since the ESI-MS ionization/detection efficiencies of high MW oligomers are expected to be significantly different than those of VP1 monomer, the relative abundances of the monomer and oligomers measured by ESI-MS likely do not accurately reflect their relative concentrations in solution. In fact, gel-filtration chromatography performed on 5 mg mL^{-1} HuNoV VA387 VLP samples (in 1X PBS, pH 7.4, 25 °C) produced a single peak (Figure S3, Supplementary Materials), which corresponds to the void volume, suggesting that the capsid protein assembles predominantly into large complexes (>800 kDa) in solution.

Affinities of HBGA oligosaccharides for the huNoV VA387 P particle and VLP

The *proxy protein* ESI-MS assay was used to evaluate the affinities of the thirteen HBGA

oligosaccharides, **L1 – L13**, for the huNoV VA387 P particle. Due to the limited availability of huNoV VA387 VLP, measurements were restricted to four oligosaccharides, **L1**, **L2**, **L7** and **L8**. In order to implement the assay, a suitable P_{proxy} (one that exhibits moderate/high affinity for the ligand) was required for each oligosaccharide tested. Four different P_{proxy} were used for these measurements, the P dimer of huNoV VA387, a truncated recombinant form of the human blood group glycosyltransferase α-(1→3)-N-acetylgalactosaminyltransferase (GTA), the family 51 carbohydrate-binding module (CBM) and a recombinant fragment of the C-terminus of human galectin-3 (Gal-3C). Representative ESI mass spectra acquired for each of the P_{proxy} are shown in Figure S4 (Supplementary Materials). Although the VA387 P dimer binds to a broad range of HBGAs, the interactions are uniformly weak (Tan et al. 2004, Han et al. 2013). Consequently, in the present study the P dimer was only used to quantify the interaction between the P particle and B type 3 tetrasaccharide (**L12**), which is the highest affinity HBGA ligand ($K_{a,int}$ = 1500 M^{-1}) identified for the P dimer (Han et al. 2013). Truncated recombinant GTA, which forms a homodimer in aqueous solution at neutral pH, possesses two thermodynamically equivalent and independent acceptor substrate binding sites (Patenaude et al. 2002; Shoemaker et al. 2008). Recent ESI-MS measurements revealed that GTA exhibits modest intrinsic affinities for B trisaccharide (**L1**, 1.6×10^4 M^{-1}) and H disaccharide (**L13**, 3.2×10^4 M^{-1}) (Shoemaker et al. 2008, Soya et al. 2010), which enabled the use of GTA as a P_{proxy} to quantify the interactions of **L1** and **L13** with the P particle. Three distinct CBM species (labeled as I, II and III) with MWs of 20,738 Da, 20,798 Da and 20,916 Da, respectively, are evident from the mass spectrum shown in Figure S4c (Supplementary Materials). The MW measured for CBM I agrees well with the theoretical

MW of 20,735, the nature of the modifications giving rise to the other two forms of CBM were not established. However, all three forms bind to HBGA oligosaccharides with identical affinities (data not shown). Consequently, for the direct and *proxy protein* ESI-MS binding measurements, the abundances of all three forms were summed together to calculate the ratio of ligand-bound to free CBM. Glycan array screening carried out on CBM revealed binding to the A and B trisaccharides, as well as A and B type 2 and type 6 tetrasaccharides (Higgins et al. 2011). According to isothermal titration calorimetry (ITC), binding between CBM and the A and B type 2 tetrasaccharides is quite strong, with K_a values in the range of 10^4 – 10^5 M^{-1} (Higgins et al. 2011). To extend the utility of CBM as a P_{proxy} for the current study, the direct ESI-MS assay was used to measure the affinities of **L1** – **L14** for CBM (Table S1, Supplementary Materials). For a majority of the ligands (**L1** – **L8** and **L12**), ESI-MS titration experiments were performed and the K_a values obtained by fitting eq 3 to the fraction of ligand bound protein measured experimentally (Figure S5, Supplementary Materials). The K_a values for the other ligands (**L9** – **L11** and **L13** and **L14**) were determined from ESI-MS measurements carried out at ≥3 different ligand concentrations. The results show that CBM only binds to A and B oligosaccharides, with the following trend in affinities: A/B trisaccharides > type 2 ~ type 6 > type 5 > type 3 > type 1 tetrasaccharides. It should be noted that the binding data for **L3** ($(5.3 \pm 0.3)\times10^4$ M^{-1}) and **L4** ($(7.4 \pm 0.3)\times10^4$ M^{-1}) measured by ESI-MS are in reasonable agreement with the K_a values obtained using ITC, 7.8×10^4 M^{-1} and 6.6×10^4 M^{-1}, respectively (Higgins et al. 2011). Recombinant Gal-3C contains a carbohydrate recognition domain that interacts with the β-galactoside motif (Hirabayashi et al. 2002, Patrick et al. 2014). The affinities of the HBGA

oligosaccharides (**L1** – **L14**) for Gal-3C were measured by the direct ESI-MS assay (Table S1, Supplementary Materials). For **L9** – **L11** and **L14** the K_a values were obtained from ESI-MS titration experiments (Figure S6, Supplementary Materials). For the remaining oligosaccharides the K_a values were determined from ESI-MS measurements carried out at ≥3 different ligand concentrations. Notably, Gal-3C binds strongly (K_a ~10^5 M^{-1}) to A/B type 1, 2 and 6 tetrasaccharides; exhibits moderately strong binding (K_a ~10^4 M^{-1}) to B type 3, A type 5, B type 5 tetrasaccharides, H type 6 and type 2 trisaccharides, and weak binding (K_a <10^3 M^{-1}) to A/B trisaccharides and H disaccharide. It can also be seen from these data that Gal-3C binding to B type oligosaccharides is consistently stronger than to the corresponding A type oligosaccharides. It should also be pointed out that the K_a values for **L7** [$(1.02 \pm 0.04)\times10^5$ M^{-1}], **L8** [$(6.4 \pm 0.6)\times10^4$ M^{-1}], **L10** [$(1.34 \pm 0.02)\times10^5$ M^{-1}] and **L11** [$(1.38 \pm 0.06)\times10^4$ M^{-1}] agree reasonably well with values measured using frontal affinity chromatography for B type 6 tetrasaccharide (1.7×10^5 M^{-1}), A type 6 tetrasaccharide (7.1×10^4 M^{-1}), A type 1 hexasaccharide (5.6×10^5 M^{-1}) and H type 6 trisaccharide (1.1×10^4 M^{-1}) (Hirabayashi et al. 2002).

To quantify the affinities of **L1** – **L13** for the P particle, the *proxy protein* ESI-MS method was implemented using a titration format, whereby the concentrations of P_{proxy}, HBGA oligosaccharide and P_{ref} were fixed, while the concentration of the target protein (i.e. P particle or VLP) was varied. From the dependence of R_{proxy} (which corresponds to the abundance ratio of ligand-bound to free P_{proxy}) on target protein concentration, $K_{a,int}$ for the target protein could be determined. Shown in Figure 2 are representative ESI mas spectra acquired for aqueous 50 mM ammonium acetate solutions (pH 7, 25 °C) containing CBM (12 μM), B trisaccharide (**L1**, 35

μM), Ubq (8 μM) with 0, 6 and 12 μM P particle (corresponding to 0, 144 and 288 μM of monomer, respectively). Visual inspection reveals that the relative abundance of **L1**-bound P_{proxy} decreased with increasing P particle concentration, indicating that **L1** binds to the P particle. In Figure 2d the measured R_{proxy} values are plotted versus P particle concentration. Fitting eq 4b to these data gives a $K_{a,int}$ of 940 ± 90 M^{-1}. Analogous measurements were carried out using GTA as P_{proxy}. Representative mass spectra acquired for aqueous 50 mM ammonium acetate solutions (pH 7, 25 °C) containing GTA (10 μM), **L1** (60 μM), scFv (8 μM, P_{ref}) with 0, 6 and 12 μM P particle are shown in Figure 3. Clearly, the relative abundance of **L1**-bound P_{proxy} decreased with the addition of P particle to the solution, consistent with the occurrence of binding of **L1** to the P particle. From the ratios $R_{proxy,1}$ and $R_{proxy,2}$, which correspond to abundance ratios of **L1**-bound GTA (to one or two **L1**, respectively) to free GTA, the magnitude θ was calculated (eq 8). Shown in Figure 3d is a plot of the calculated values of θ versus P particle concentration. Fitting eq 5 to these data gives a $K_{a,int}$ of 1100 ± 100 M^{-1}. Importantly, these two values of $K_{a,int}$ (which were determined using different P_{proxy}) are indistinguishable, within experimental error. Using an analogous strategy (Figures S7 – S11), Supplementary Materials), $K_{a,int}$ values for the interactions of **L2** – **L13** with the P particle were determined, Table 1.

The aforementioned measurements were carried out using 50 mM aqueous ammonium acetate solutions. To ensure that the affinity measurements were not sensitive to ionic strength, ESI-MS measurements were repeated using substantially higher concentrations of ammonium acetate. Representative mass spectra acquired for solutions of CBM (12 μM), **L1** (35 μM), Ubq (8 μM) and 6 μM P particle in either 200 mM or 800 mM aqueous ammonium acetate solution

(pH 7, 25 °C) are shown in (Figure S12a and 12b, respectively, Supplementary Materials). Notably, the $K_{a,int}$ values of 940 ± 60 M^{-1} and 900 ± 110 M^{-1}, respectively, are indistinguishable from the value measured in 50 mM aqueous ammonium acetate (940 ± 90 M^{-1}). From these results it is concluded that the affinities of the HBGA ligands for the P particle and, presumably, VLP are relatively insensitive to the ionic strength of the solution.

Control experiments were also carried out to rule out the possibility that the aglycone of the HBGA oligosaccharides used in the study, $–(CH_2)_8COOC_2H_5$ for **L1**, **L2** and **L13** and $–(CH_2)_6CH=CH_2$ for **L3** – **L12**, promotes nonspecific binding to the P particle or VLP. The affinities of a type 2 trisaccharide (**L14**), which also has $–(CH_2)_6CH=CH_2$ at the reducing end but does not bind to GII.4 huNoVs (Fiege et al. 2012) for both the VA387 P dimer and the P particle were measured using direct and the *proxy protein* ESI-MS assay, respectively. Shown in Figure S13 (Supplementary Materials) is a representative ESI mass spectrum acquired for an aqueous 200 mM ammonium acetate solutions (pH 7, 25 °C) of P dimer (12 μM), **L14** (80 μM) and scFv (8 μM). It can be seen from the normalized distributions of **L14** bound P dimer (after correction for nonspecific ligand binding) that **L14** does not exhibit any detectable binding to the P dimer. Shown in Figures S14a–c are representative ESI mass spectra acquired for aqueous 200 mM ammonium acetate solutions (pH 7, 25 °C) containing Gal-3C (5 μM), **L14** (40 μM), Ubq (3 μM) with 0, 4 and 8 μM P particle (corresponding to 0, 96 and 192 μM of monomer, respectively). Notably, it can be seen from Figure S14d (Supplementary Materials) that R_{proxy} is independent of P particle concentration. This result confirms that the $–(CH_2)_6CH=CH_2$ aglycone does not bind nonspecifically to the P dimer, the P particle or, presumably, the VLP.

According to the results of the *proxy protein* ESI-MS measurements, the thirteen HBGA oligosaccharides investigated (**L1** – **L13**) exhibit binding to the P particle, with $K_{a,int}$ values ranging from 500 M^{-1} to 2300 M^{-1}. From these data, the apparent affinities of these HBGA oligosaccharides for the P particle (24-mer) can be estimated to be between 1×10^4 and 6×10^4 M^{-1}. Notably, the trend in measured affinities of **L1** – **L13** for the P particle mirrors that found for the P dimer, with the B type 3 tetrasaccharide and H disaccharide being the strongest and weakest binders, respectively. Additionally, the H type 6 trisaccharide (**L11**, which is 2'-fucosyllactose), which is abundant in human milk and was recently shown to inhibit VA387 P particles from binding to other HBGAs (Shang et al. 2013), exhibits a $K_{a,int}$ of 880 ± 50 for the P particle. Overall, the $K_{a,int}$ values for the P dimer and P particle are within a factor of two. This finding suggests that the binding sites of the P dimer and P particle are structurally identical, or nearly so.

The affinities of four HBGA oligosaccharides (**L1**, **L2**, **L7** and **L8**) for the VLP were also measured using CBM as the P_{proxy}. Shown in Figures 4a – 4c are representative ESI mass spectra acquired for aqueous 200 mM ammonium acetate solutions (pH 7, 25 °C) containing CBM (12 μM), **L1** (25 μM) and Ubq (4 μM) with 0, 380 and 760 nM VLP (corresponding to 0, 68 and 136 μM of monomer, respectively). From visual inspection it is seen that the relative abundance of **L1**–bound P_{proxy} decreased with the addition of the VLP to the solution, consistent with binding between **L1** and the VLP. Shown in Figure 5 are plots of R_{proxy} versus VLP concentrations determined for **L1**, **L2**, **L7** and **L8**. By fitting eq 4b to these data, the corresponding $K_{a,int}$ values were determined, Table 1. The $K_{a,int}$ values, which range from 1000 M^{-1} to 4000 M^{-1}, are

consistently larger than those measured for the P particle, but are within a factor of three. From the measured $K_{a,int}$ values, the apparent affinities of these ligands for the VLP (180-mer form) are estimated to be between 2×10^5 and 7×10^5 M^{-1}.

Taken together, the binding data measured in the present study and those reported previously for the P dimer indicate that the interactions between the HBGA ligands and the P dimer, P particle and VLP of huNoV VA387, while similar, are not identical and exhibit the following trend: P dimer ≈ P particle < VLP. This finding suggests that there exist subtle differences in the structure of the carbohydrate binding sites presented by the P dimer and P particle and those of the VLP. It is interesting to note, however, that the crystal structures of the VLP (Prasad et al. 1999) and P dimer (Choi et al. 2008, Bu et al. 2008) of Norwalk virus (GI.1) do not reveal a distinct structural difference between the P dimer of VLP and the one formed from the isolated P domain. Therefore, our data may imply that such subtle structural difference may not be easily recognized from crystal structures or, alternatively, that such subtle difference may occur between the VLP and the P dimers/P particle of VA387 (GII.4) but not between those of Norwalk virus. Nevertheless, the differences in $K_{a,int}$ values for the VA387 P dimer, P particle and VLP are small and support the use of P dimers and P particles as surrogates to the VLP.

Conclusion

In summary, the interactions between the huNoV VA387 P particle and VLP and a series of HBGA oligosaccharides were quantified for the first time. The measured $K_{a,int}$ values of the 13 HBGA oligosaccharides for the P particle range from 500 to 2300 M^{-1}; those of the A and B trisaccharides and the A and B type 6 tetrasaccharides for the VLP range from 1000 M^{-1} to 4000

M^{-1}. Comparison of the binding data with those measured previously for the P dimer reveals that the HBGA oligosaccharides tested exhibit similar intrinsic affinities for the P dimer and P particle. The intrinsic affinities exhibited for the VLP are consistently higher than those measured for the P particle, by a factor of three. Based on these data, the apparent affinities of the HBGA oligosaccharides tested for the P particle and VLP were estimated to be in the $10^4 – 10^5$ M^{-1} and $10^5 – 10^6$ M^{-1} range, respectively. While the cause of the subtle differences in HBGA oligosaccharide affinities for the P dimer and P particle and those for the VLP remains unknown, the present data support the use of P dimers and P particles as substitutes to the VLP for NoV-receptor binding studies.

Acknowledgements

The authors thank the Alberta Glycomics Centre (JSK), the Natural Sciences and Engineering Research Council (JSK) and the Canadian Institutes for Health Research (ABB; MOP 130305) for supporting this research and Prof. T. Lowary (University of Alberta) for generously providing carbohydrate ligands and Prof. C. Cairo (University of Alberta) for providing Gal-3C used in this work. LH also acknowledges an Alberta Innovates Graduate Student Scholarship and XJ and MT acknowledge support from the National Institutes of Health of the Unites States of America. ABB was supported by an EWR Steacie Memorial Fellowship, a Canada Research Chair in Molecular Interactions, and a Michael Smith Foundation for Health Research Career Scholar Award.

References

Bereszczak JZ, Barbu IM, Tan M, Xia M, Jiang X, van Duijn E, Heck AJR. 2012. Structure, stability and dynamics of norovirus P domain derived protein complexes studied by native mass spectrometry. *J. Struct. Biol.*, 177:273-282.

Bu WM, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS. 2008. Structural basis for the receptor binding specificity of norwalk virus. *J. Virol.*, 82:5340-5347.

Choi JM, Hutson AM, Estes MK, Prasad BVV. 2008. Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proc. Natl. Acad. Sci. U. S. A.*, 105:9175-9180.

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.

de Rougemont A, Ruvoen-Clouet N, Simon B, Estienney M, Elie-Caille C, Aho S, Pothier P, Le Pendu J, Boireau W, Belliot G. 2011. Qualitative and quantitative analysis of the binding of gii.4 norovirus variants onto human blood group antigens. *J. Virol.*, 85:4057-4070.

El-Hawiet A, Kitova EN, Arutyunov D, Simpson DJ, Szymanski CM, Klassen JS. 2012. Quantifying ligand binding to large protein complexes using electrospray ionization mass spectrometry. *Anal. Chem.*, 84:3867-3870.

Fiege B, Rademacher C, Cartmell J, Kitov PI, Parra F, Peters T. 2012. Molecular details of the recognition of blood group antigens by a human norovirus as determined by STD NMR spectroscopy. *Angew. Chem. Int. Ed.*, 51:928-932.

Han L, Kitov PI, Kitova EN, Tan M, Wang L, Xia M, Jiang X, Klassen JS. 2013. Affinities of

recombinant norovirus P dimers for human blood group antigens. *Glycobiology*, 23:276-285.

Hansman GS, Shahzad-ul-Hussan S, McLellan JS, Chuang G-Y, Georgiev I, Shimoike T, Katayama K, Bewley CA, Kwong PD. 2012. Structural basis for norovirus inhibition and fucose mimicry by citrate. *J. Virol.*, 86:284-292.

Higgins MA, Ficko-Blean E, Meloncelli PJ, Lowary TL, Boraston AB. 2011. The Overall architecture and receptor binding of pneumococcal carbohydrate-antigen-hydrolyzing enzymes. *J. Mol. Biol.*, 411:1017-1036.

Hirabayashi J, Hashidate T, Arata Y, Nishi N, Nakamura T, Hirashima M, Urashima T, Oka T, Futai M, Muller WEG, *et al.* 2002. Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim. Biophys. Acta*, 1572:232-254.

Huang P, Farkas T, Zhong W, Thornton S, Morrow AL, Jiang X. 2005. Norovirus and histo-blood group antigens: Demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. *J. Virol.*, 79:6714-6722.

Hutson AM, Atmar RL, Graham DY, Estes MK. 2002. Norwalk virus infection and disease is associated with ABO histo-blood group type. *J. Infect. Dis.*, 185:1335-1337.

Hutson AM, Atmar RL, Marcus DM, Estes MK. 2003. Norwalk virus-like particle hemagglutination by binding to H histo-blood group antigens. *J. Virol.*, 77:405-415.

Jiang X, Wang M, Graham DY, Estes MK. 1992. Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. *J. Virol.*, 66:6527-6532.

Jiang X, Zhong W, Farkas T, Huang P, Wilton N, Barrett E, Fulton D, Morrow R, Matson DO.

2002. Baculovirus expression and antigenic characterization of the capsid proteins of three Norwalk-like viruses. *Arch. Virol*, 147:119-130.

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.

Meloncelli PJ, Lowary TL. 2009. Synthesis of ABO histo-blood group type V and VI antigens. *Aust. J. Chem.*, 62:558-574.

Meloncelli PJ, Lowary TL. 2010. Synthesis of ABO histo-blood group type I and II antigens. *Carbohydr. Res.*, 345:2305-2322.

Meloncelli PJ, West LJ, Lowary TL. 2011. Synthesis and NMR studies on the ABO histo-blood group antigens: synthesis of type III and IV structures and NMR characterization of type I-VI antigens. *Carbohydr. Res.*, 346:1406-1426.

Oriol R. 1990. Genetic-control of the fucosylation of abh precursor chains - evidence for new epistatic interactions in different cells and tissues. *J. Immunogenet.*, 17:235-245.

Patel MM, Widdowson M-A, Glass RI, Akazawa K, Vinje J, Parashar UD. 2008. Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg. Infect. Dis.*, 14:1224-1231.

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.

Prasad BVV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK. 1999. X-ray

crystallographic structure of the Norwalk virus capsid. *Science*, 286:287-290.

Ravn V, Dabelsteen E. 2000. Tissue distribution of histo-blood group antigens. *APMIS*, 108:1-28.

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.

Shang J, Piskarev VE, Xia M, Huang PW, Jiang X, Likhosherstov LM, Novikova OS, Newburg DS, Ratner DM. 2013. Identifying human milk glycans that inhibit norovirus binding using surface plasmon resonance. *Glycobiology*, 23:1491-1498.

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.

Shoemaker GK, van Duijn E, Crawford SE, Uetrecht C, Baclayon M, Roos WH, Wuite GJL, Estes MK, Prasad BVV, Heck AJR. 2010. Norwalk virus assembly and stability monitored by mass spectrometry. *Mol. Cell. Proteomics*, 9:1742-1751.

Soya N, Shoemaker GK, Palcic MM, Klassen JS. 2009. Comparative study of substrate and product binding to the human ABO(H) blood group glycosyltransferases. *Glycobiology*, 19:1224-1234.

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.

Tamminen K, Huhti L, Koho T, Lappalainen S, Hytonen VP, Vesikari T, Blazevic V. 2012. A comparison of immunogenicity of norovirus GII-4 virus-like particles and P-particles. *Immunology*, 135:89-99.

Tan M, Fang P, Chachiyo T, Xia M, Huang P, Fang Z, Jiang W, Jiang X. 2008. Noroviral P particle: Structure, function and applications in virus-host interaction. *Virology*, 382:115-123.

Tan M, Fang P, Xia M, Chachiyo T, Jiang W, Jiang X. 2011. Terminal modifications of norovirus P domain resulted in a new type of subviral particles, the small P particles. *Virology*, 410:345-352.

Tan M, Hegde RS, Jiang X. 2004. The P domain of norovirus capsid protein forms dimer and binds to histo-blood group antigen receptors. *J. Virol.*, 78:6233-6242.

Tan M, Jiang X. 2005a. Norovirus and its histo-blood group antigen receptors: an answer to a historical puzzle. *Trends Microbiol.*, 13:285-293.

Tan M, Jiang X. 2005b. The P domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors. *J. Virol.*, 79:14017-14030.

Uetrecht C, Barbu IM, Shoemaker GK, van Duijn E, Heck AJR. 2011. Interrogating viral capsid assembly with ion mobility-mass spectrometry. *Nat. Chem.*, 3:126-132.

Wang L, Huang P, Fang H, Xia M, Zhong W, McNeal MM, Jiang X, Tan M. 2013. Polyvalent complexes for vaccine development. *Biomaterials*, 34:4480-4492.

Wang L, Xia M, Huang P, Fang H, Cao D, Meng X-J, McNeal M, Jiang X, Tan M. 2014. Branched-linear and agglomerate protein polymers as vaccine platforms. *Biomaterials*,

35:8427-8438.

White LJ, Hardy ME, Estes HK. 1997. Biochemical characterization of a smaller form of recombinant Norwalk virus capsids assembled in insect cells. *J. Virol.*, 71:8066-8072.

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-angstrom resolution. *Proc. Natl. Acad. Sci. U. S. A.*, 91:6423-6427.

Table 1. Intrinsic association constants ($K_{a,int}$) for HBGA oligosaccharides (**L1 – L14**) binding to huNoV VA387 P dimer, P particle and VLP, measured at 25 °C and pH 7 using the ESI-MS *proxy protein* assay.[a]

	HBGA	**Structures**	$K_{a,int}$ (M^{-1}) P dimer [b]	$K_{a,int}$ (M^{-1}) P particle	$K_{a,int}$ (M^{-1}) VLP
L1	B trisaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-$O(CH_2)_8COOC_2H_5$	800 ± 100	940 ± 90 [c] 1100 ± 100 [d]	2300 ± 250 [c]
L2	A trisaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-$O(CH_2)_8COOC_2H_5$	500 ± 50	840 ± 90 [c]	1400 ± 150 [c]
L3	B type 2 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-$O(CH_2)_6CH=CH_2$	410 ± 45	870 ± 60 [c]	n.d. [h]
L4	A type 2 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-$O(CH_2)_6CH=CH_2$	290 ± 30	710 ± 90 [c]	n.d. [h]
L5	B type 5 tetrasaccharide	α-D-Gal-(1→3)- [α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-Gal-$O(CH_2)_6CH=CH_2$	700 ± 100	930 ± 80 [c]	n.d. [h]
L6	A type 5 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal- (1→3)-β-D-Gal-$O(CH_2)_6CH=CH_2$	560 ± 40	900 ± 100 [c]	n.d. [h]
L7	B type 6 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-$O(CH_2)_6CH=CH_2$	600 ± 45	650 ± 30 [c]	1000 ± 160 [c]
L8	A type 6 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-$O(CH_2)_6CH=CH_2$	1200 ± 50	1400 ± 120 [c]	3900 ± 260 [c]
L9	B type 1 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-$O(CH_2)_6CH=CH_2$	700 ± 100	1530 ± 40 [e]	n.d. [h]
L10	A type 1 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-$O(CH_2)_6CH=CH_2$	600 ± 65	1500 ± 100 [e]	n.d. [h]
L11	H type 6 trisaccharide	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc-$O(CH_2)_6CH=CH_2$	330 ± 25	880 ± 50 [e]	n.d. [h]
L12	B type 3 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-α-D-GalNAc-$O(CH_2)_6CH=CH_2$	1500 ± 150	2300 ± 110 [f]	n.d. [h]
L13	H disaccharide	α-L-Fuc-(1→2)-β-D-Gal-$O(CH_2)_8COOC_2H_5$	240 ± 40	520 ± 100 [d]	n.d. [h]
L14	type 2 trisaccharide	α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-$O(CH_2)_6CH=CH_2$	NB [g]	NB [g]	n.d. [h]

a. Uncertainties correspond to one standard deviation. b. Values are adapted from Han et al. 2013. c, d, e, f. Values are measured using CBM, GTA, Gal-3C and P dimer as P_{proxy}, respectively. g. NB ≡ no binding detected. h. n.d. ≡ not determined.

Figure captions

Figure 1. (a) ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate (200 mM, pH 7 and 25 °C) solution with (a) 3 μM P particle (corresponding to 72 μM monomer) and (b) 0.2 μM VLP (corresponding to 36 μM VP1) of huNoV VA387. The measurements were carried out using a Waters Synapt G2S mass spectrometer.

Figure 2. Representative ESI mass spectra obtained in positive ion mode for a 50 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12 μM CBM, 8 μM Ubq (P_{ref}), 35 μM **L1** (B trisaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of huNoV VA387. (d) Plot of R_{proxy} versus concentration of monomer in the P particle. The solution condition for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to the best fit of eq 4b for the experimental data. The error bars correspond to one standard deviation. The measurements were carried out using a Bruker ApexQe FT-ICR mass spectrometer.

Figure 3. Representative ESI mass spectra obtained in positive ion mode for a 50 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 10 μM GTA, 10 μM scFv (P_{ref}), 60 μM **L1** (B trisaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of huNoV VA387. (d) Plot of θ versus concentration of monomer in the P particle. The solution condition for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to

the best fit of eq 5 to the experimental data. The error bars correspond to one standard deviation. The measurements were carried out using a Bruker ApexQe FT-ICR mass spectrometer.

Figure 4. Representative ESI mass spectra obtained in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12 μM CBM, 4 μM Ubq (P_{ref}), 25 μM **L1** (B trisaccharide) with (a) 0 nM, (b) 380 nM and (c) 760 nM VLP (180-mer) of huNoV VA387. The measurements were carried out using a Waters Synapt G2S mass spectrometer.

Figure 5. Plots of R_{proxy} versus concentration of monomer in the VLP measured for aqueous ammonium acetate solutions (200 mM, pH 7 and 25 °C) containing CBM (12 μM), Ubq (P_{ref}, 4 μM), VLP (180-mer) of huNoV VA387 (0 – 760 nM) with (a) **L1** (B trisaccharide, 25 μM), (b) **L2** (A trisaccharide, 20 μM), (c) **L7** (B type 6 tetrasaccharide, 40 μM) and (d) **L8** (A type 6 tetrasaccharide, 25 μM). The solid curves correspond to the best fit of eq 4b to the experimental data for each ligand. The error bars correspond to one standard deviation. The measurements were carried out using a Waters Synapt G2S mass spectrometer.

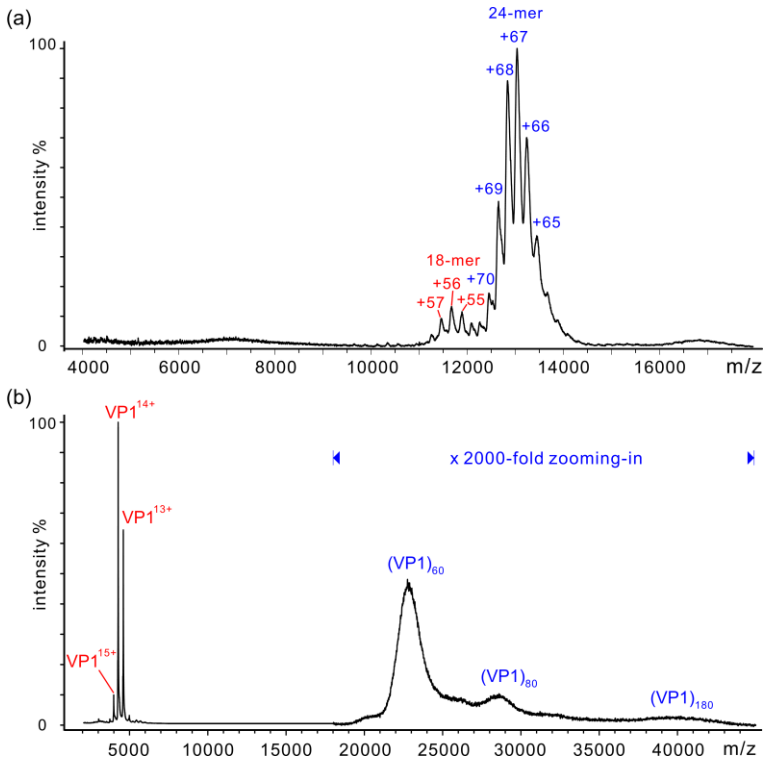

(a)
24-mer
+67
+68
+66
+69
+65
18-mer
+56
+70
+57
+55
intensity %
100
0
4000
6000
8000
10000
12000
14000
16000
m/z
(b)
VP1^14+
x 2000-fold zooming-in
VP1^13+
(VP1)60
VP1^15+
(VP1)80
(VP1)180
intensity %
100
0
5000
10000
15000
20000
25000
30000
35000
40000
m/z

Figure 1

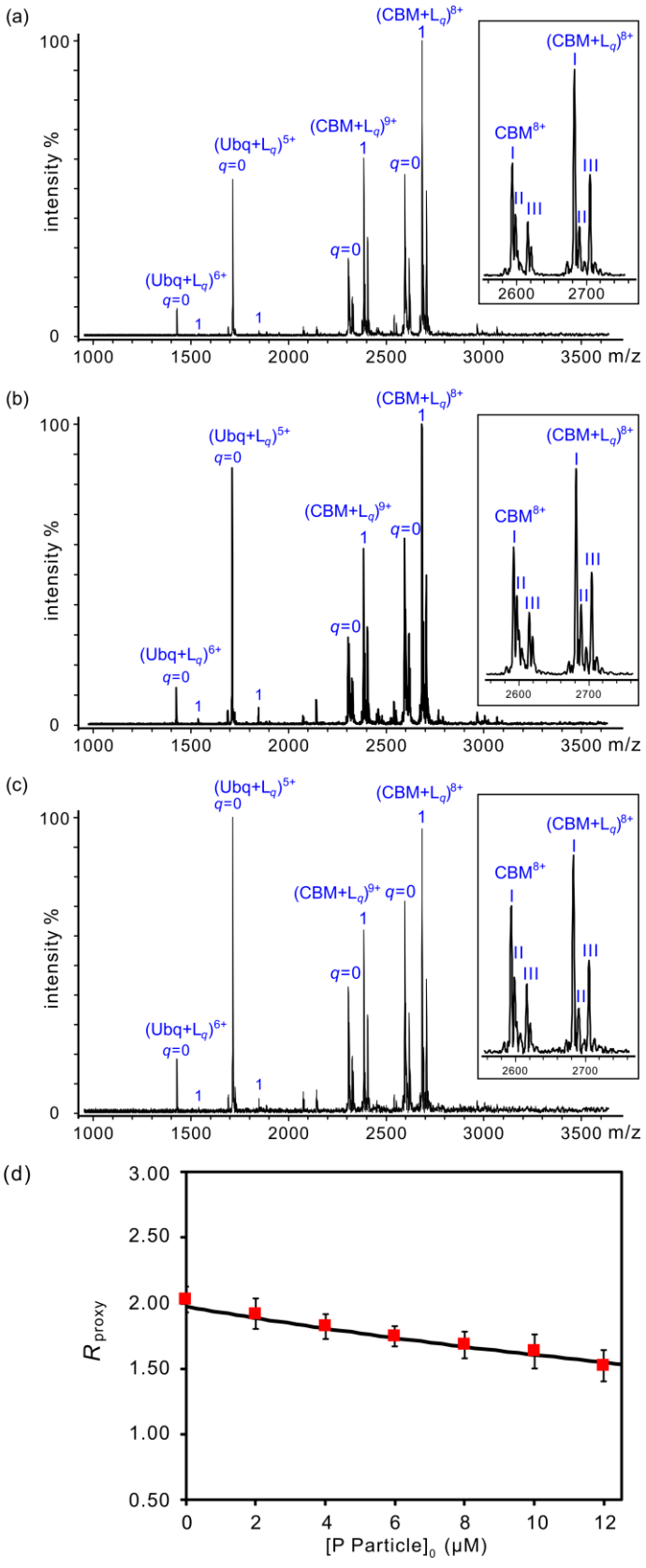

(a)
(b)
(c)
(d)
intensity %
m/z
(CBM+L$_q$)$^{8+}$
(Ubq+L$_q$)$^{5+}$
(CBM+L$_q$)$^{9+}$
(Ubq+L$_q$)$^{6+}$
CBM^{8+}
q=0
1
I
II
III
R_{proxy}
[P Particle]$_0$ (μM)

Figure 2

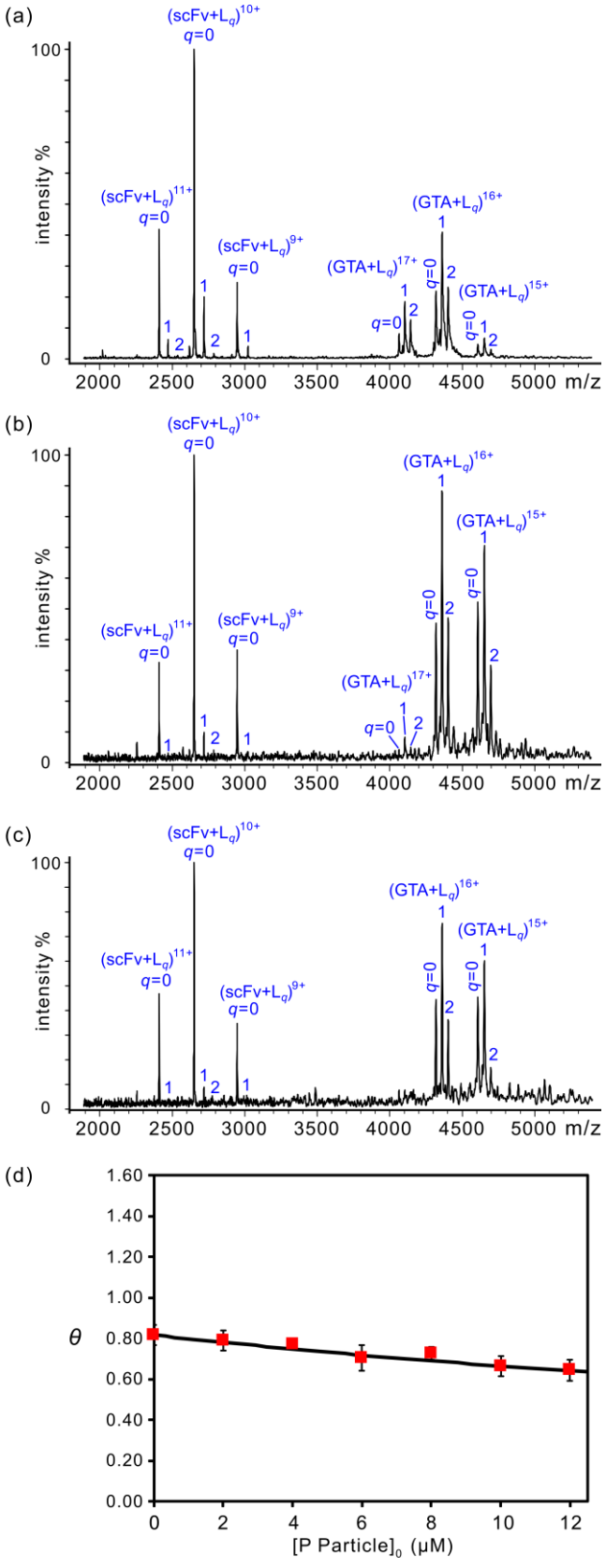

(a)
(scFv+L$_q$)$^{10+}$ q=0
(scFv+L$_q$)$^{11+}$ q=0
(scFv+L$_q$)$^{9+}$ q=0
(GTA+L$_q$)$^{16+}$
(GTA+L$_q$)$^{17+}$
(GTA+L$_q$)$^{15+}$
intensity %
100
0
2000 2500 3000 3500 4000 4500 5000 m/z
(b)
(scFv+L$_q$)$^{10+}$ q=0
(GTA+L$_q$)$^{16+}$
(GTA+L$_q$)$^{15+}$
(scFv+L$_q$)$^{11+}$ q=0
(scFv+L$_q$)$^{9+}$ q=0
(GTA+L$_q$)$^{17+}$
intensity %
100
0
2000 2500 3000 3500 4000 4500 5000 m/z
(c)
(scFv+L$_q$)$^{10+}$ q=0
(GTA+L$_q$)$^{16+}$
(GTA+L$_q$)$^{15+}$
(scFv+L$_q$)$^{11+}$ q=0
(scFv+L$_q$)$^{9+}$ q=0
intensity %
100
0
2000 2500 3000 3500 4000 4500 5000 m/z
(d)
θ
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0 2 4 6 8 10 12
[P Particle]$_0$ (μM)

Figure 3

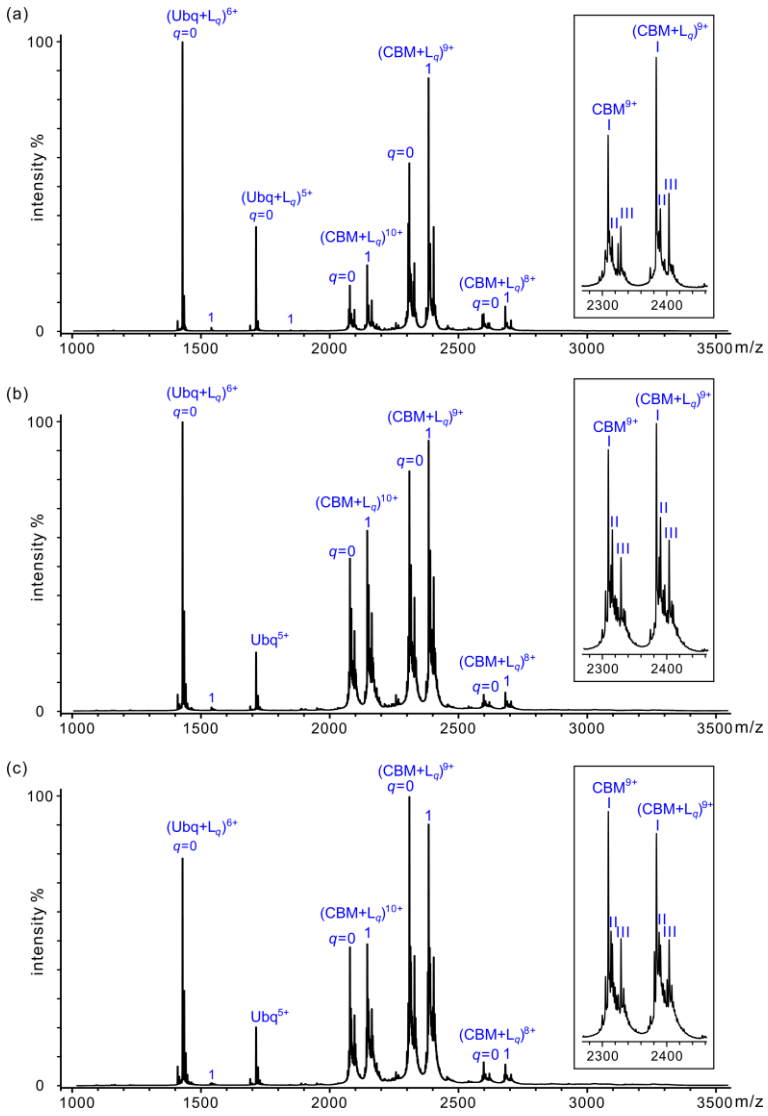

(a)
(Ubq+L$_q$)$^{6+}$
q=0
(CBM+L$_q$)$^{9+}$
(Ubq+L$_q$)$^{5+}$
(CBM+L$_q$)$^{10+}$
(CBM+L$_q$)$^{8+}$
CBM^{9+}
intensity %
m/z
(b)
Ubq^{5+}
(c)

Figure 4

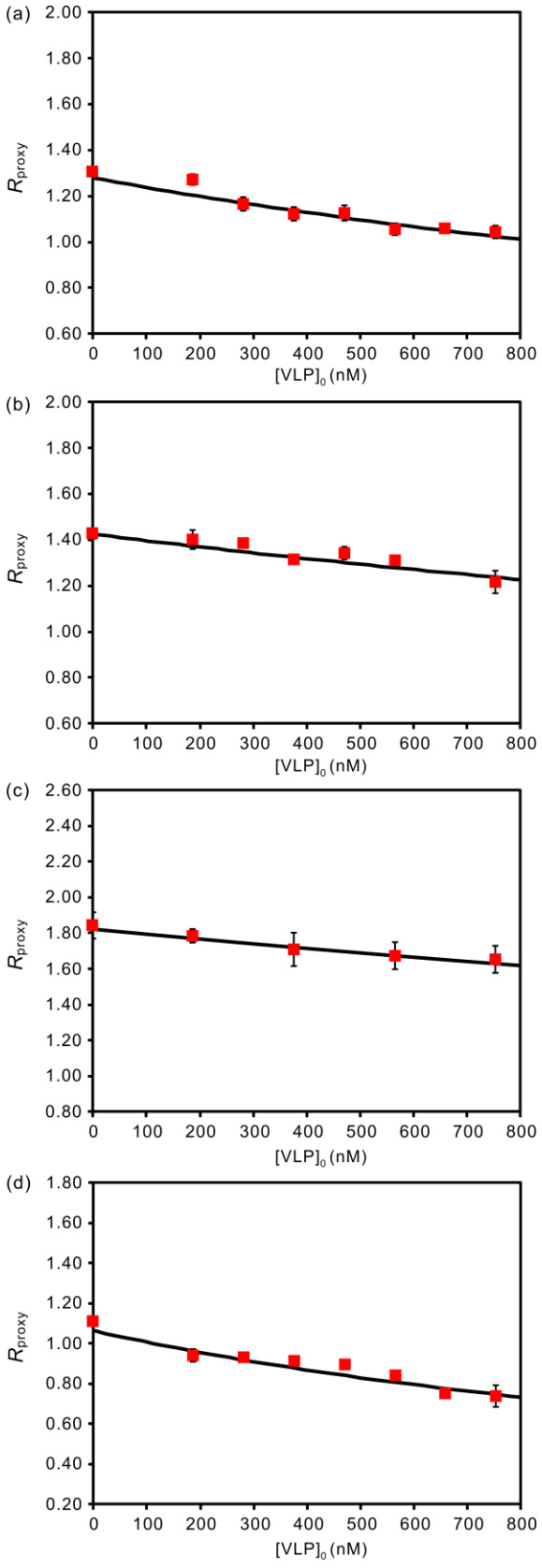


Figure 5

Supplementary Materials for:

Affinities of Human Histo-Blood Group Antigens for Norovirus Capsid Protein Complexes

Ling Han, Elena N. Kitova, Ming Tan, Xi Jiang, Benjamin Pluvinage, Alisdair B. Boraston, and John S. Klassen

Methods

Application of the proxy protein ESI-MS method using P_{proxy} that possesses multiple ligand binding sites.

A special extension of the *proxy protein* ESI-MS method, in which both the target protein (P) and proxy protein (P_{proxy}) possesses multiple ligand (L) binding sites, was used in the current study. The relevant equations of mass balance for the situation where P and P_{proxy} possess *h* and *g* identical L binding sites, respectively, are given by eqs S1a-S1c:

$$[P]_0 = [P] + \sum_{i=1}^{h} [PL_i] \quad \text{(S1a)}$$

$$[P_{proxy}]_0 = [P_{proxy}] + \sum_{j=1}^{g} [P_{proxy}L_j] \quad \text{(S1b)}$$

$$[L]_0 = [L] + \sum_{i=1}^{h} i[PL_i] + \sum_{j=1}^{g} j[P_{proxy}L_j] \quad \text{(S1c)}$$

where $[P]_0$, $[P_{proxy}]_0$ and $[L]_0$ are the initial concentrations of P, P_{proxy} and L, respectively, [P], $[P_{proxy}]$ and [L] are the equilibrium concentrations of P, P_{proxy} and L, respectively, and $[PL_i]$ and

$[P_{proxy}L_j]$ are the equilibrium concentrations of P and P_{proxy} bound to *i* and *j* molecules of L, respectively.

The values of $[P_{proxy}]$ and $[P_{proxy}L_j]$ can be calculated from eqs S2a and S2b, respectively:

$$[P_{proxy}] = \frac{[P_{proxy}]_0}{1+\sum_{j=1}^{g}[P_{proxy}L_j]/[P_{proxy}]} = \frac{[P_{proxy}]_0}{1+\sum_{j=1}^{g}R_{proxy,j}} \quad \text{(S2a)}$$

$$[P_{proxy}L_j] = [P_{proxy}]R_{proxy,j} = \frac{[P_{proxy}]_0 R_{proxy,j}}{1+\sum_{j=1}^{g}R_{proxy,j}} \quad \text{(S2b)}$$

where $R_{proxy,j}$ corresponds to the concentration ratio of ligand-bound (bound to *j* molecules of L) to free P_{proxy} and is taken to be equal to the abundance (*Ab*) ratio of ligand-bound (bound to *j* molecules of L) to free P_{proxy} gas phase ions, eq S3:

$$R_{proxy,j} = \frac{\left[P_{proxy}L_j\right]}{\left[P_{proxy}\right]} = \frac{\sum Ab\left(P_{proxy}L_j\right)}{\sum Ab\left(P_{proxy}\right)} \quad \text{(S3)}$$

The ratio of occupied-to-free binding sites for P_{proxy} (θ) can be calculated using eq S4:

$$\theta = \frac{[\text{occupied binding sites}]_{P_{proxy}}}{[\text{free binding sites}]_{P_{proxy}}} = \frac{\sum_{j=1}^{g} j[P_{proxy}L_j]}{g[P_{proxy}]_0 - \sum_{j=1}^{g} j[P_{proxy}L_j]} = \frac{\sum_{j=1}^{g} jR_{proxy,j}}{g(1+\sum_{j=1}^{g}R_{proxy,j}) - \sum_{j=1}^{g} jR_{proxy,j}} \quad \text{(S4)}$$

and the intrinsic association constant for L binding to P_{proxy} ($K_{a,Pproxy,int}$) can be expressed by eq S5:

$$K_{a,P_{proxy},int} = \frac{[\text{occupied binding sites}]_{P_{proxy}}}{[\text{free binding sites}]_{P_{proxy}}[L]} = \frac{\theta}{[L]} \quad \text{(S5)}$$

It follows that [L] can be found using eq S6:

$$[\mathrm{L}] = \frac{\theta}{\mathrm{K}_{\mathrm{a,P_{proxy},int}}} \tag{S6}$$

The intrinsic association constant of P ($K_{a,P,int}$) can be expressed by eq S7:

$$\mathrm{K}_{\mathrm{a,P,int}} = \frac{[\text{occupied binding sites}]_{\mathrm{P}}}{[\text{free binding sites}]_{\mathrm{P}}[\mathrm{L}]} \tag{S7}$$

and the concentrations of occupied and free binding sites are given by eqs S8a and S8b:

$$[\text{occupied binding sites}]_{\mathrm{P}} = \sum_{i=1}^{h} i[\mathrm{PL}_i] \tag{S8a}$$

$$[\text{free binding sites}]_{\mathrm{P}} = [\mathrm{P}]_{\mathrm{m},0} - \sum_{i=1}^{h} i[\mathrm{PL}_i] \tag{S8b}$$

where the total concentration of binding sites can be written as eq S9:

$$[\text{total binding sites}]_{\mathrm{P}} \equiv [\mathrm{P}]_{\mathrm{m},0} = h[\mathrm{P}]_0 \tag{S9}$$

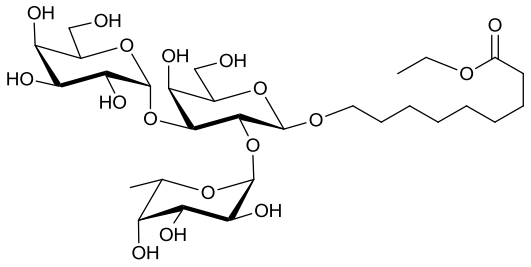
From eqs S1c, S4 and S6, the following expression can be derived, S10:

$$\sum_{i=1}^{h} i[\mathrm{PL}_i] = [\mathrm{L}]_0 - \frac{\theta}{\mathrm{K}_{\mathrm{a,P_{proxy},int}}} - \frac{g[\mathrm{P_{proxy}}]_0\theta}{\theta+1} \tag{S10}$$

Finally, the value of $K_{a,P,int}$ can be calculated from eq S11:

$$K_{\mathrm{a,P,int}} = \frac{K_{a,\mathrm{P_{proxy},int}}/\theta}{\dfrac{[\mathrm{P}]_{\mathrm{m},0}}{[\mathrm{L}]_0 - \dfrac{g[\mathrm{P_{proxy}}]_0\theta}{\theta+1} - \dfrac{\theta}{K_{a,\mathrm{P_{proxy},int}}}} - 1} \tag{S11}$$

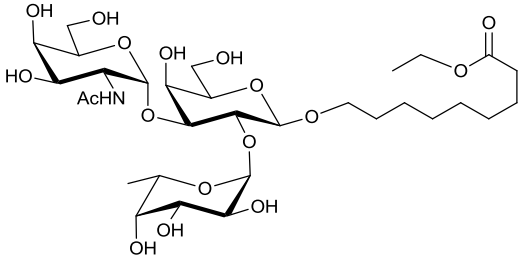
It can be shown that, in the case where P_{proxy} possesses a single binding site (i.e., g = 1), eq S11 reduces to eq 4b.



L1

B trisaccharide (MW 672.32 Da)

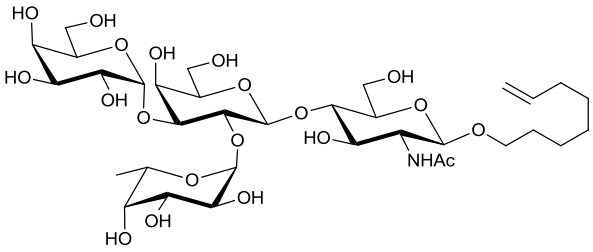
α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-



L2

A trisaccharide (MW 713.35 Da)

α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-

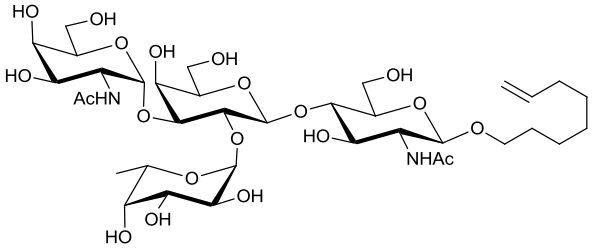


L3

B type 2 tetrasaccharide (MW 801.36 Da)

α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-

(1→4)-β-D-GlcNAc-O$(CH_2)_6CH=CH_2$

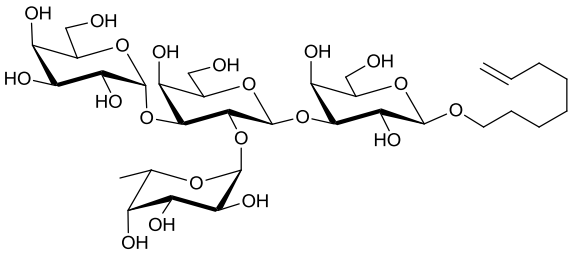


L4

A type 2 tetrasaccharide (MW 842.39 Da)

α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-

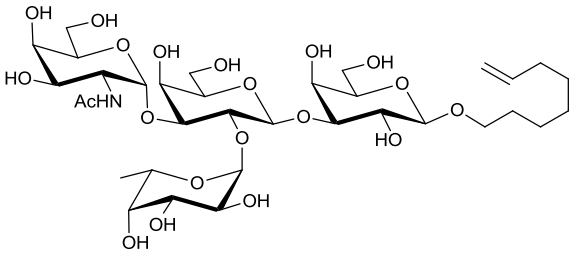
Gal-(1→4)-β-D-GlcNAc-O$(CH_2)_6CH=CH_2$



L5

B type 5 tetrasaccharide (MW 760.34 Da)

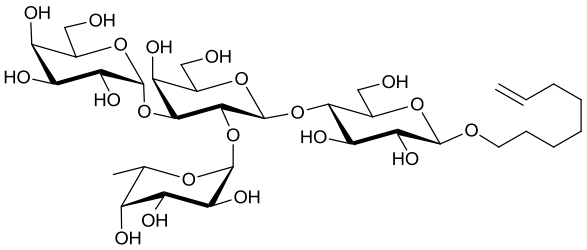
α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-



L6

A type 5 tetrasaccharide (MW 801.36 Da)

α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-

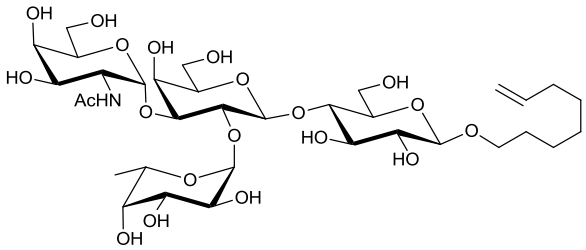


L7

B type 6 tetrasaccharide (MW 760.34 Da)

α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-

(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$

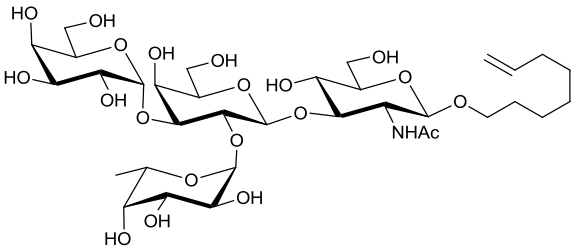


L8

A type 6 tetrasaccharide (MW 801.36 Da)

α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-

Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$

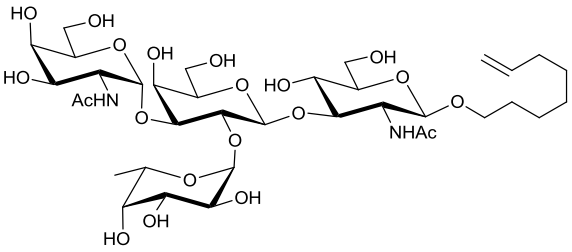


L9

B tetrasaccharide type 1 (MW 801.36 Da)

α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-

(1→3)-β-D-GlcNAc-OC_8H_{15}

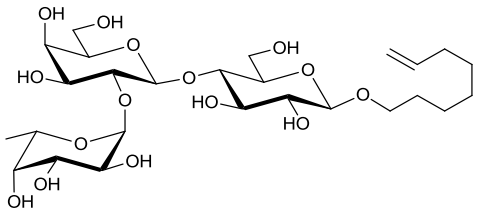


L10

A tetrasaccharide type 1 (MW 842.39 Da)

α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-

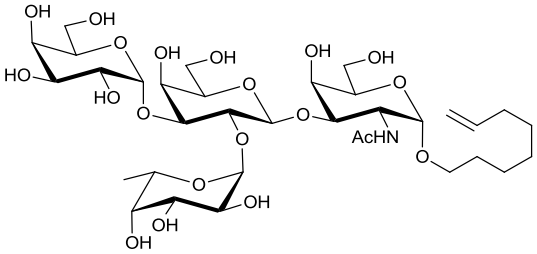
Gal-(1→3)-β-D-GlcNAc-OC_8H_{15}



L11

H type 6 trisaccharide (MW 598.28 Da)

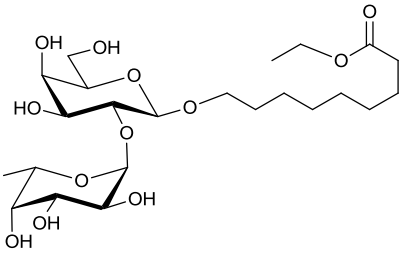
α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc-



L12

B type 3 tetrasaccharide (MW 801.36 Da)

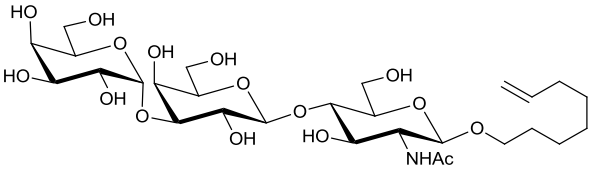
α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-



L13

H disaccharide (MW 510.57 Da)

α-L-Fuc-(1→2)-β-D-Gal-



L14

Type 2 trisaccharide (MW 655.69 Da)

α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-

Figure S1. Structures of the HBGA oligosaccharides **L1** – **L14**.

Table S1. Association constants (K_a) for binding of the HBGA oligosaccharides (**L1 – L14**) with CBM, Gal-3C and GTA measured at 25 °C and pH 7 by the ESI-MS assay.[a]

	HBGA	**Structures**	K_a ($\times 10^4$ M^{-1}) CBM	K_a ($\times 10^4$ M^{-1}) Gal-3C	$K_{a,int}$ [b] ($\times 10^4$ M^{-1}) GTA
L1	B trisaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-O$(CH_2)_8COOC_2H_5$	(7.3 ± 0.4)	(0.08 ± 0.01)	1.6 [c]
L2	A trisaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-O$(CH_2)_8COOC_2H_5$	(11.3 ± 0.7)	(0.056± 0.002)	n.d. [d]
L3	B type 2 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-O$(CH_2)_6CH=CH_2$	(5.3 ± 0.3)	(22.3 ± 1.7)	n.d. [d]
L4	A type 2 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-O$(CH_2)_6CH=CH_2$	(7.4 ± 0.3)	(17.4 ± 1.4)	n.d. [d]
L5	B type 5 tetrasaccharide	α-D-Gal-(1→3)- [α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-Gal-O$(CH_2)_6CH=CH_2$	(3.1 ± 0.1)	(3.02 ± 0.06)	n.d. [d]
L6	A type 5 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal- (1→3)-β-D-Gal-O$(CH_2)_6CH=CH_2$	(3.3 ± 0.6)	(1.05 ± 0.04)	n.d. [d]
L7	B type 6 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$	(5.6 ± 0.3)	(10.2 ± 0.4)	n.d. [d]
L8	A type 6 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$	(5.8 ± 0.2)	(6.4 ± 0.6)	n.d. [d]
L9	B type 1 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-O$(CH_2)_6CH=CH_2$	(0.36 ± 0.05)	(16.4 ± 0.9)	n.d. [d]
L10	A type 1 tetrasaccharide	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-O$(CH_2)_6CH=CH_2$	(0.31 ± 0.05)	(13.4 ± 0.2)	n.d. [d]
L11	H type 6 trisaccharide	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$	< 0.02	(1.38 ± 0.06)	n.d. [d]

L12	B type 3 tetrasaccharide	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-α-D-GalNAc-O(CH_2)$_6$CH=CH_2	(0.44 ± 0.03)	(0.70 ± 0.05)	n.d. [d]
L13	H disaccharide	α-L-Fuc-(1→2)-β-D-Gal-O(CH_2)$_8$COOC_2H_5	< 0.02	NB [e]	3.2 [c]
L14	type 2 trisaccharide	α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-O(CH_2)$_6$CH=CH_2	< 0.02	(1.20 ± 0.04)	n.d. [d]

a. Uncertainties correspond to one standard deviation. b. $K_{a,int}$ values corresponds to the intrinsic (per binding site) association constants. c. Values are adapted from Shoemaker et al. 2008 and Soya et al. 2010. d. n.d. ≡ not determined. e. NB ≡ no binding detected.

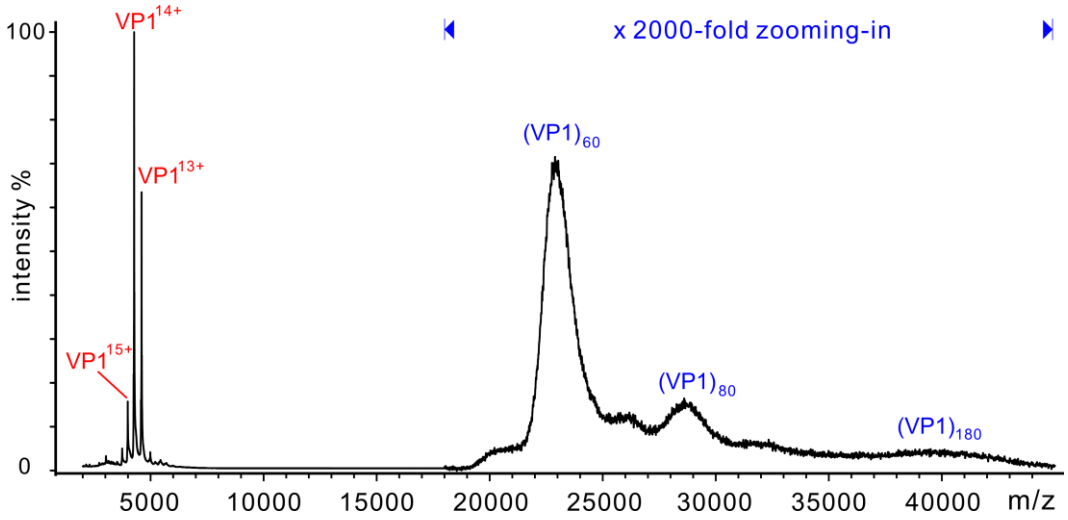


Figure S2. ESI mass spectrum acquired in positive ion mode for a 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 0.2 μM (corresponding to 36 μM VP1) huNoV VA387 VLP. This measurement was carried out on a different day than that shown in Figure 1a. The mass spectrum was acquired using a Waters Synapt G2S mass spectrometer.

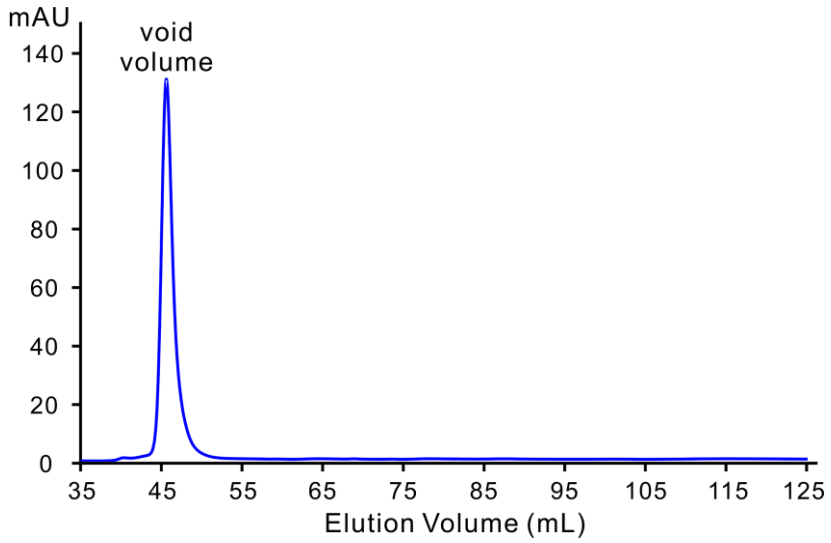


Figure S3. Chromatograph of gel filtration of the HuNoV VA387 VLP (5 mg mL^{-1}) in 1X PBS (pH 7.4, 25°C) buffer. The observation of only a single peak (corresponding to the void volume at ~45 mL) indicates the capsid protein of HuNoV VA387 predominantly assembles into large complexes, with molecular weights ≥800 kDa. The Superdex 200 gel filtration column was calibrated as described previously (Tan and Jiang 2005).

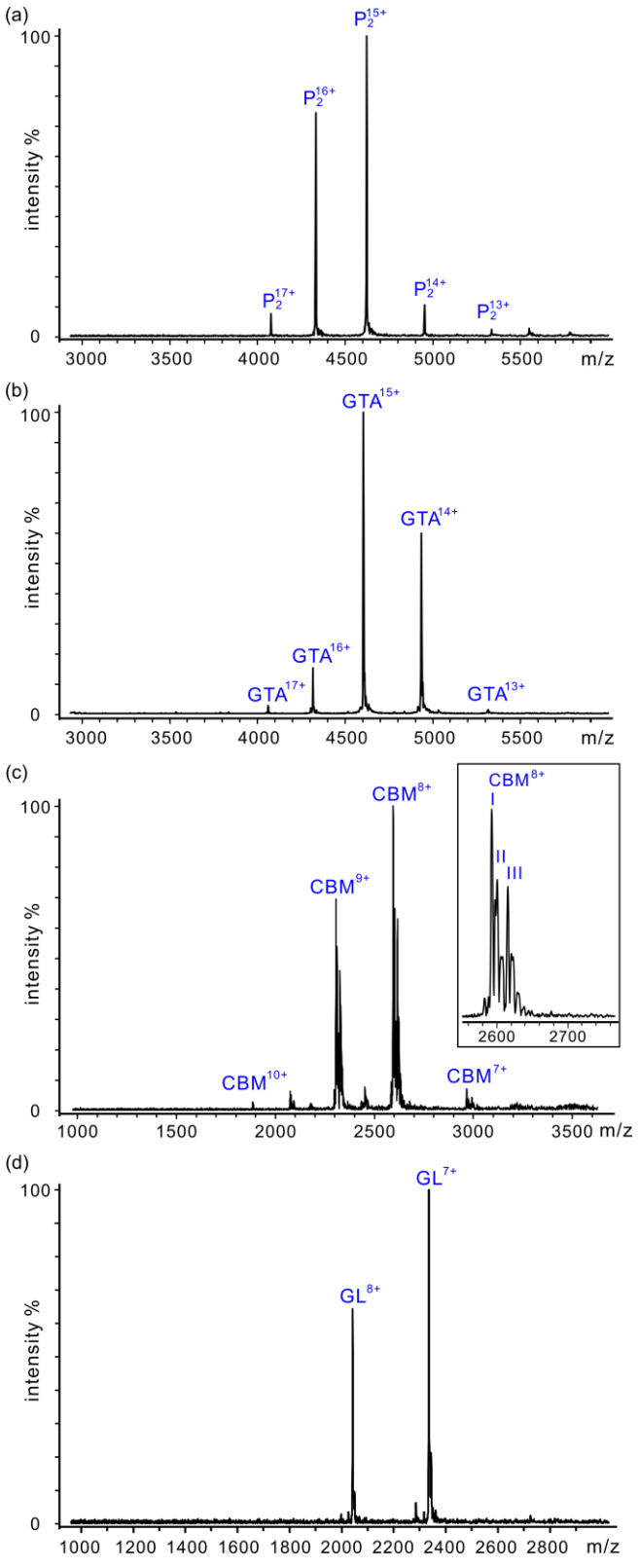
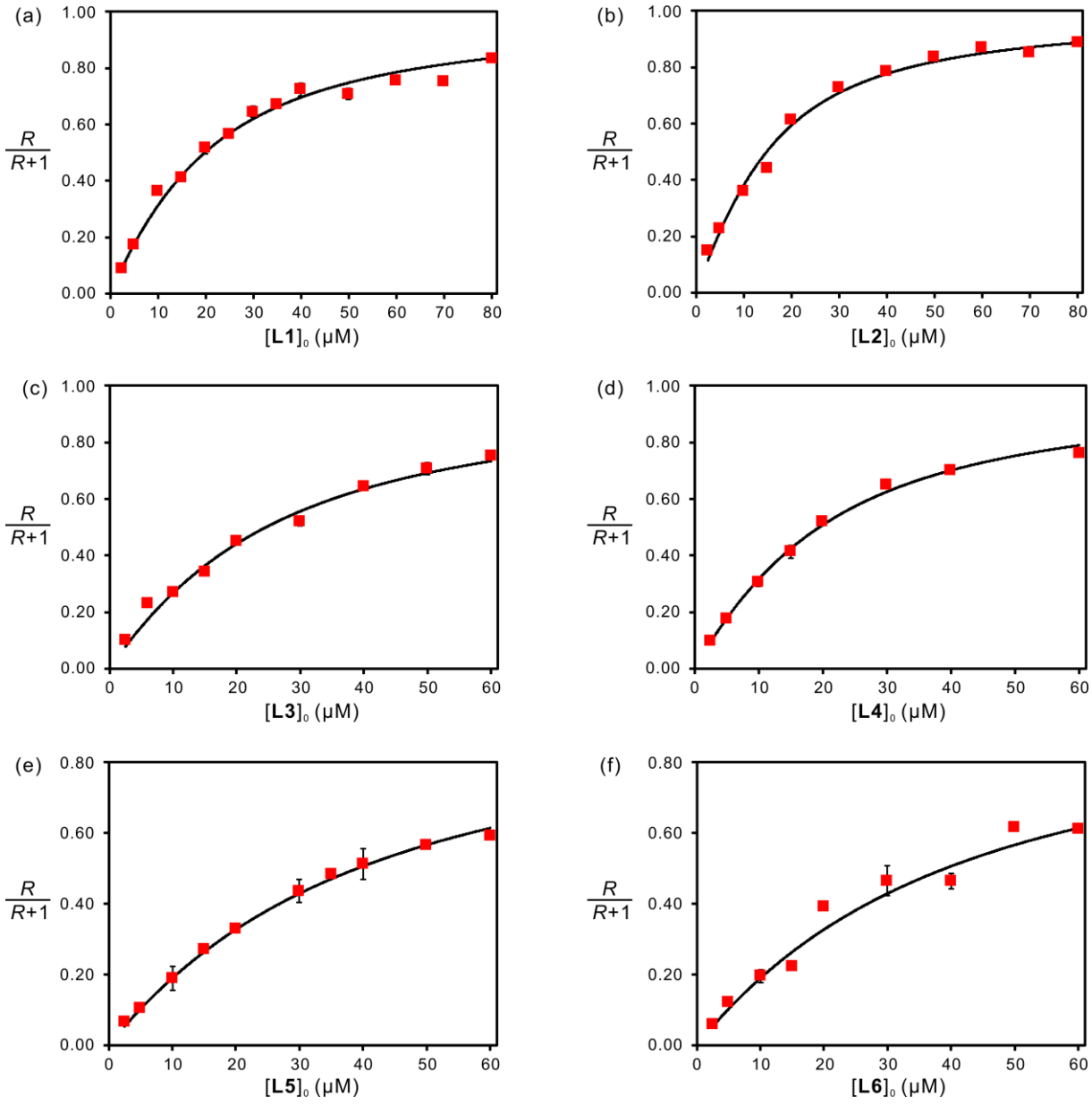

(a)
100
intensity %
0
P_2^{15+}
P_2^{16+}
P_2^{17+}
P_2^{14+}
P_2^{13+}
3000 3500 4000 4500 5000 5500 m/z
(b)
100
intensity %
0
GTA^{15+}
GTA^{14+}
GTA^{16+}
GTA^{17+}
GTA^{13+}
3000 3500 4000 4500 5000 5500 m/z
(c)
100
intensity %
0
CBM^{8+}
CBM^{9+}
CBM^{10+}
CBM^{7+}
1000 1500 2000 2500 3000 3500 m/z
CBM^{8+}
I
II
III
2600 2700
(d)
100
intensity %
0
GL^{7+}
GL^{8+}
1000 1200 1400 1600 1800 2000 2200 2400 2600 2800 m/z

Figure S4. ESI mass spectra obtained in positive ion mode for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing P dimer (P_2, 12 μM), (b) GTA (10 μM), (c) CBM (12 μM) and (d) Gal-3C (GL, 5 μM). The measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.


(a)
(b)
(c)
(d)
(e)
(f)
R / R+1
[L1]0 (μM)
[L2]0 (μM)
[L3]0 (μM)
[L4]0 (μM)
[L5]0 (μM)
[L6]0 (μM)

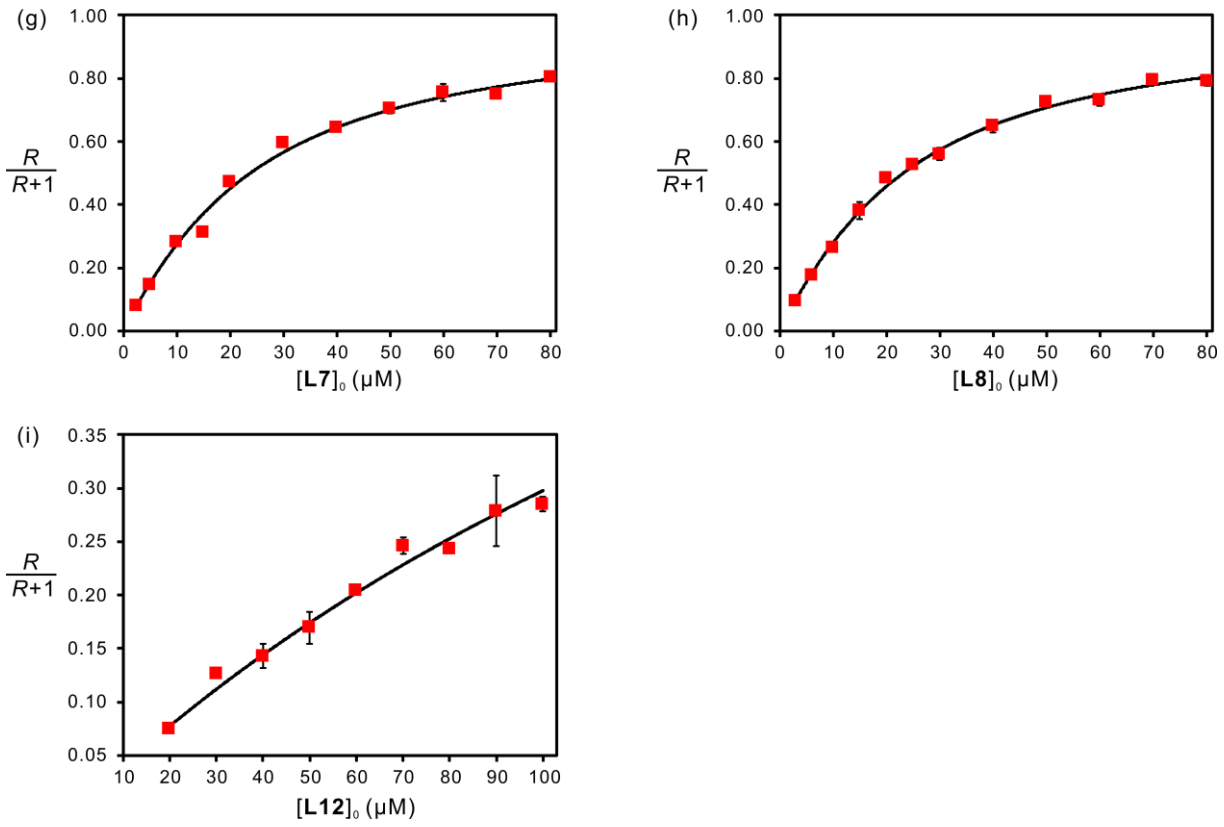


Figure S5. Plot of fraction of ligand-bound CBM versus ligand concentration measured for (a) **L1** (B trisaccharide), (b) **L2** (A trisaccharide), (c) **L3** (B type 2 tetrasaccharide), (d) **L4** (A type 2 tetrasaccharide), (e) **L5** (B type 5 tetrasaccharide), (f) **L6** (A type 5 tetrasaccharide), (g) **L7** (B type 6 tetrasaccharide), (h) **L8** (A type 6 tetrasaccharide), and (i) **L12** (B type 3 tetrasaccharide). The ESI-MS binding measurements were carried out on 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing CBM (12 μM), Ubq (P_{ref}, 8 μM) and each ligand at a minimum of eight different concentrations ranging from 2.5 to 100 μM. The solid curves correspond to the best fit of eq 3 to the experimental data and the error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.

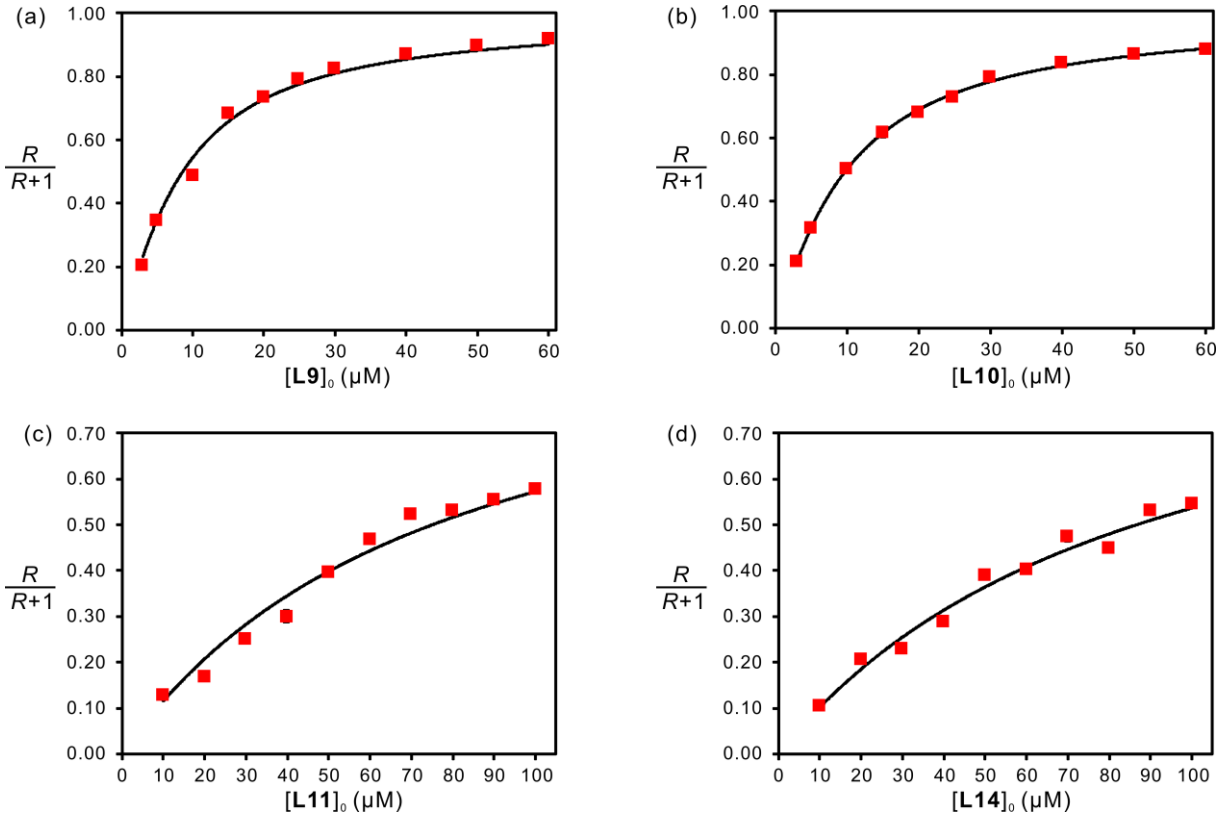
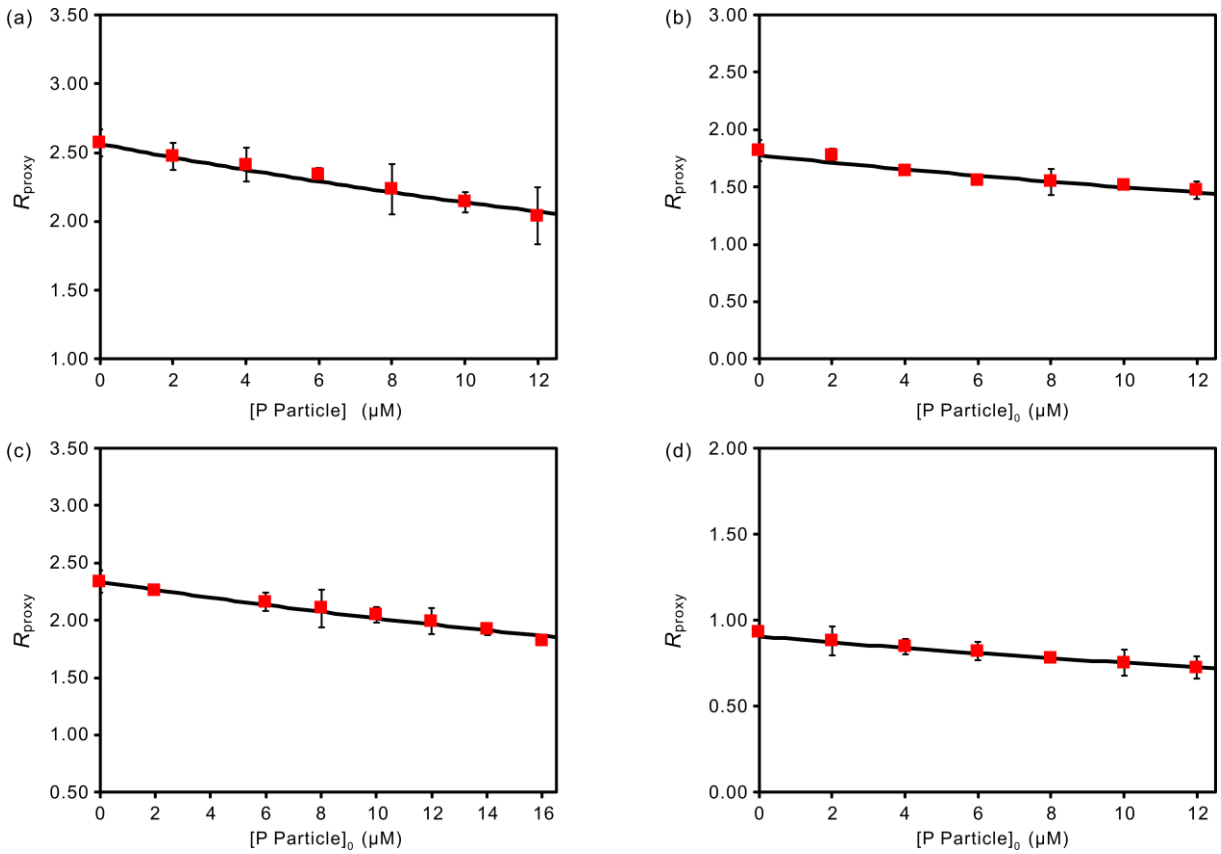


Figure S6. Plot of fraction of ligand-bound Gal3C versus ligand concentration measured for (a) **L9** (B type 1 tetrasaccharide), (b) **L10** (A type 1 tetrasaccharide), (c) **L11** (H type 6 trisaccharide), and (d) **L14** (type 2 trisaccharide). The ESI-MS binding measurements were carried out on 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing Gal-3C (5 μM), Ubq (P_{ref}, 3 μM) and each ligand at a minimum of nine different concentrations ranging from 2.5 to 100 μM. The solid curves correspond to the best fit of eq 3 to the experimental data and the error bars correspond to one standard derivation. These measurements were performed on Waters Synapt G2S mass spectrometer.


(a)
3.50
3.00
2.50
2.00
1.50
1.00
R_{proxy}
0
2
4
6
8
10
12
[P Particle] (μM)
(b)
3.00
2.50
2.00
1.50
1.00
0.50
0.00
R_{proxy}
0
2
4
6
8
10
12
[P Particle]$_0$ (μM)
(c)
3.50
3.00
2.50
2.00
1.50
1.00
0.50
R_{proxy}
0
2
4
6
8
10
12
14
16
[P Particle]$_0$ (μM)
(d)
2.00
1.50
1.00
0.50
0.00
R_{proxy}
0
2
4
6
8
10
12
[P Particle]$_0$ (μM)

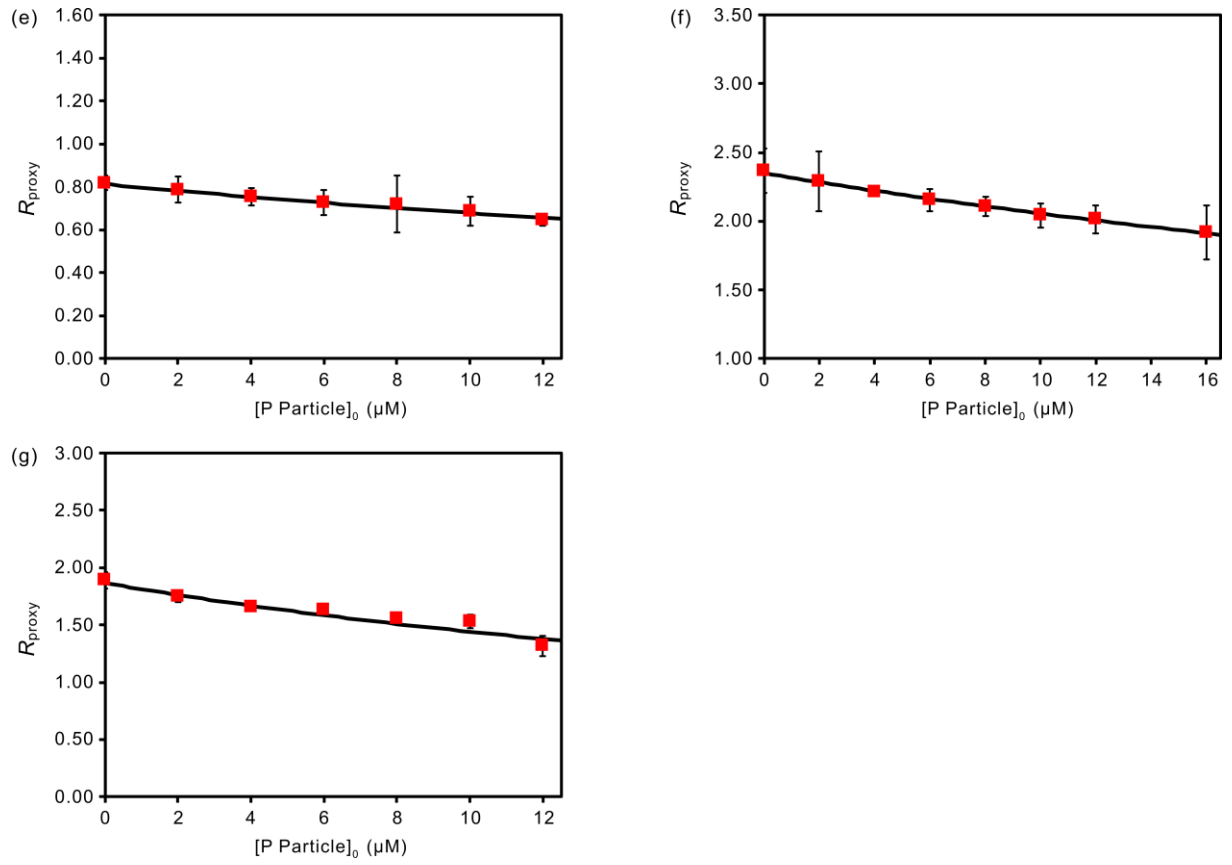


Figure S7. Plots of R_{proxy} versus P particle concentration measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing CBM (12 μM), Ubq (P_{ref}, 8 μM), (a) **L2** (A trisaccharide, 30 μM), (b) **L3** (B type 2 tetrasaccharide, 40 μM), (c) **L4** (A type 2 tetrasaccharide, 40 μM), (d) **L5** (B type 5 tetrasaccharide, 35 μM), (e) **L6** (A type 5 tetrasaccharide, 30 μM), (f) **L7** (B type 6 tetrasaccharide, 50 μM) and (g) **L8** (A type 6 tetrasaccharide, 40 μM), and P particle (0 – 16 μM, which corresponds to 0 – 384 μM of monomer). The solid curves correspond to the best fit of eq 4b to the experimental data for each ligand. The error bars correspond to one standard derivation. These measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.

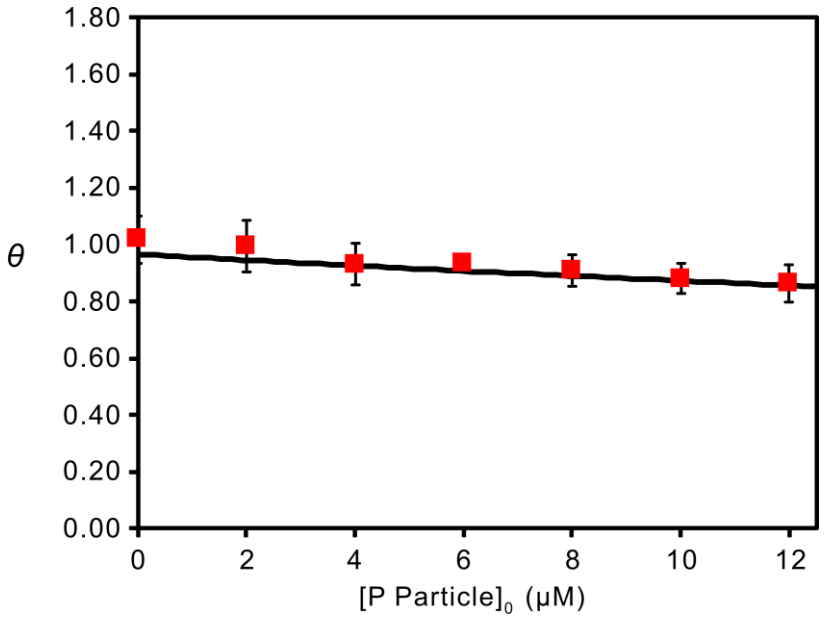


Figure S8. Plot of the θ versus P particle concentration measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing GTA (10 μM), scFv (P_{ref}, 10 μM), **L13** (H disaccharide, 40 μM) and huNoV VA387 P particle (0 – 12 μM, which corresponds to 0 – 288 μM of monomer). The solid curve corresponds to the best fit of eq 5 to the experimental data. The error bars correspond to one standard derivation. These measurements were performed on Bruker ApexQe FT-ICR mass spectrometer.

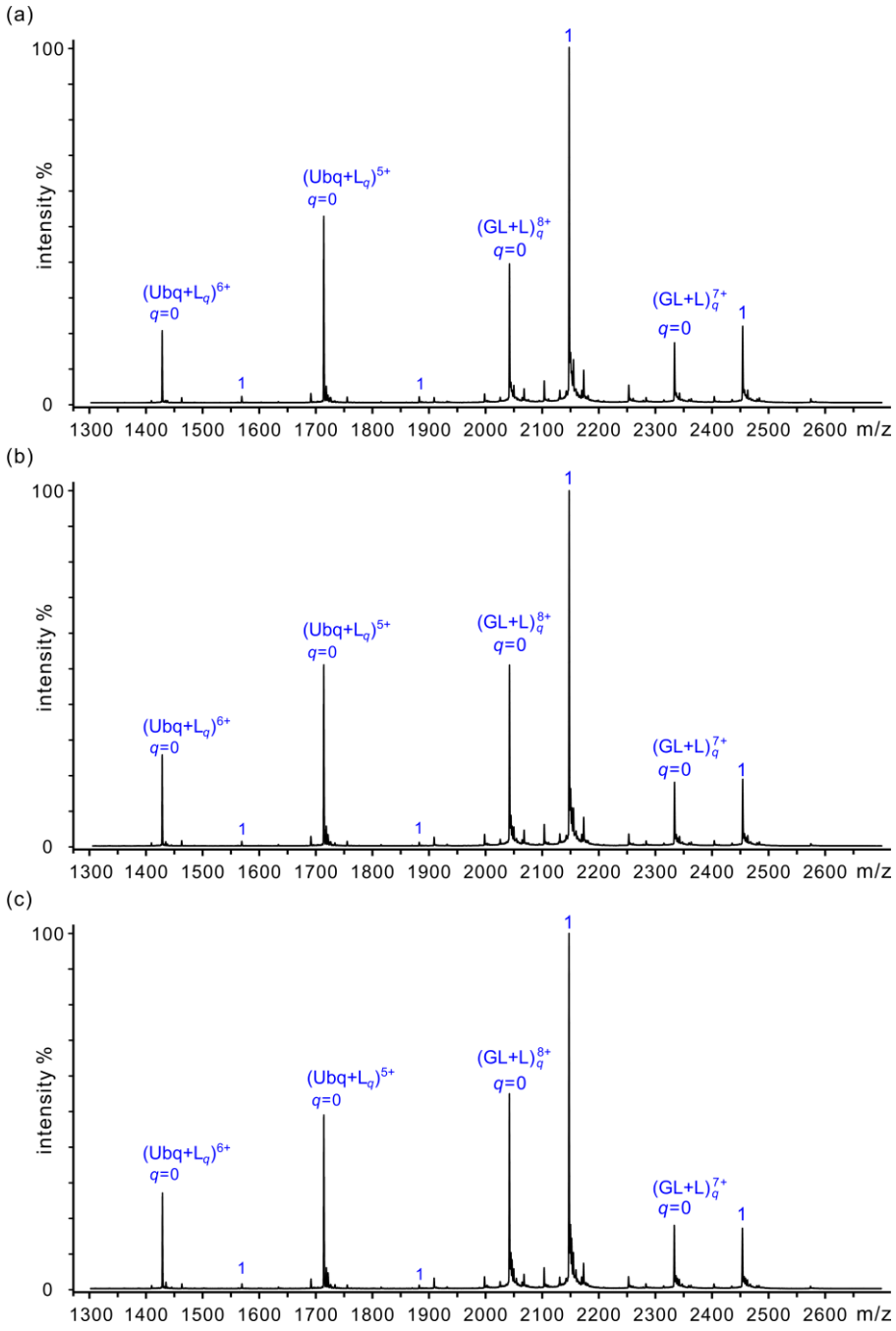

(a)
100
intensity %
0
1
(Ubq+L$_q$)$^{5+}$
q=0
(GL+L)$_q^{8+}$
q=0
(Ubq+L$_q$)$^{6+}$
q=0
(GL+L)$_q^{7+}$
q=0
1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 2400 2500 2600 m/z
(b)
100
intensity %
0
1
(Ubq+L$_q$)$^{5+}$
q=0
(GL+L)$_q^{8+}$
q=0
(Ubq+L$_q$)$^{6+}$
q=0
(GL+L)$_q^{7+}$
q=0
1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 2400 2500 2600 m/z
(c)
100
intensity %
0
1
(Ubq+L$_q$)$^{5+}$
q=0
(GL+L)$_q^{8+}$
q=0
(Ubq+L$_q$)$^{6+}$
q=0
(GL+L)$_q^{7+}$
q=0
1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 2400 2500 2600 m/z

Figure S9. Representative ESI mass spectra measured for 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 5 μM Gal-3C (GL), 3 μM Ubiquitin (Ubq, P_{ref}), 20 μM **L10** (A type 1 tetrasaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of huNoV VA387. The measurements were performed on a Waters Synapt G2S mass spectrometer.

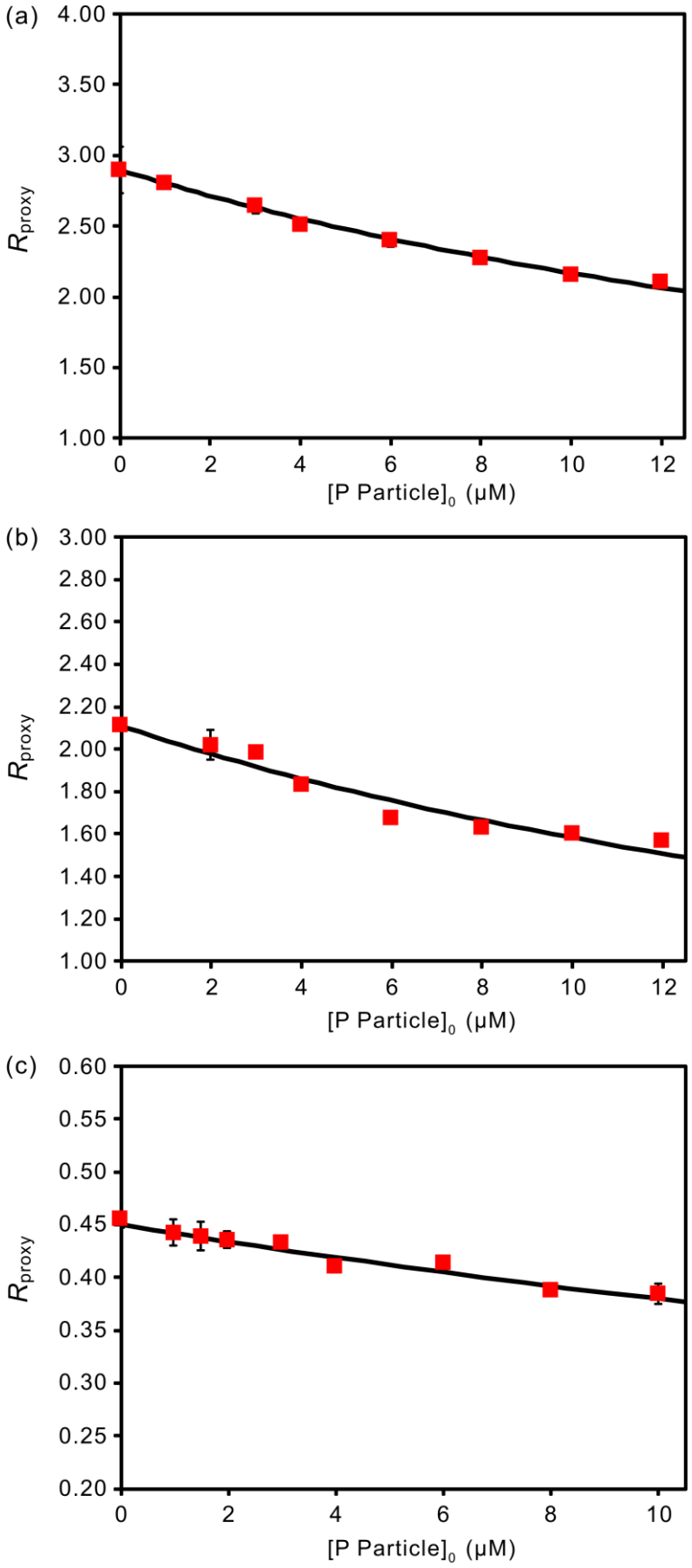

(a)
4.00
3.50
3.00
2.50
2.00
1.50
1.00
R_{proxy}
0
2
4
6
8
10
12
[P Particle]$_0$ (μM)
(b)
3.00
2.80
2.60
2.40
2.20
2.00
1.80
1.60
1.40
1.20
1.00
R_{proxy}
0
2
4
6
8
10
12
[P Particle]$_0$ (μM)
(c)
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
R_{proxy}
0
2
4
6
8
10
[P Particle]$_0$ (μM)

Figure S10. Plots of R_{proxy} versus P particle concentration measured for 200 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing Gal-3C (5 μM), Ubq (P_{ref}, 3 μM), (a) **L9** (B type 1 tetrasaccharide, 20 μM), (b) **L10** (A type 1 tetrasaccharide, 20 μM), and (c) **L11** (H type 6 trisaccharide, 40 μM), and P particle (0 – 12 μM, which corresponds to 0 – 288 μM of monomer). The solid curves correspond to the best fit of eq 4b to the experimental data for each ligand. The error bars correspond to one standard derivation. These measurements were performed on a Waters Synapt G2S mass spectrometer.

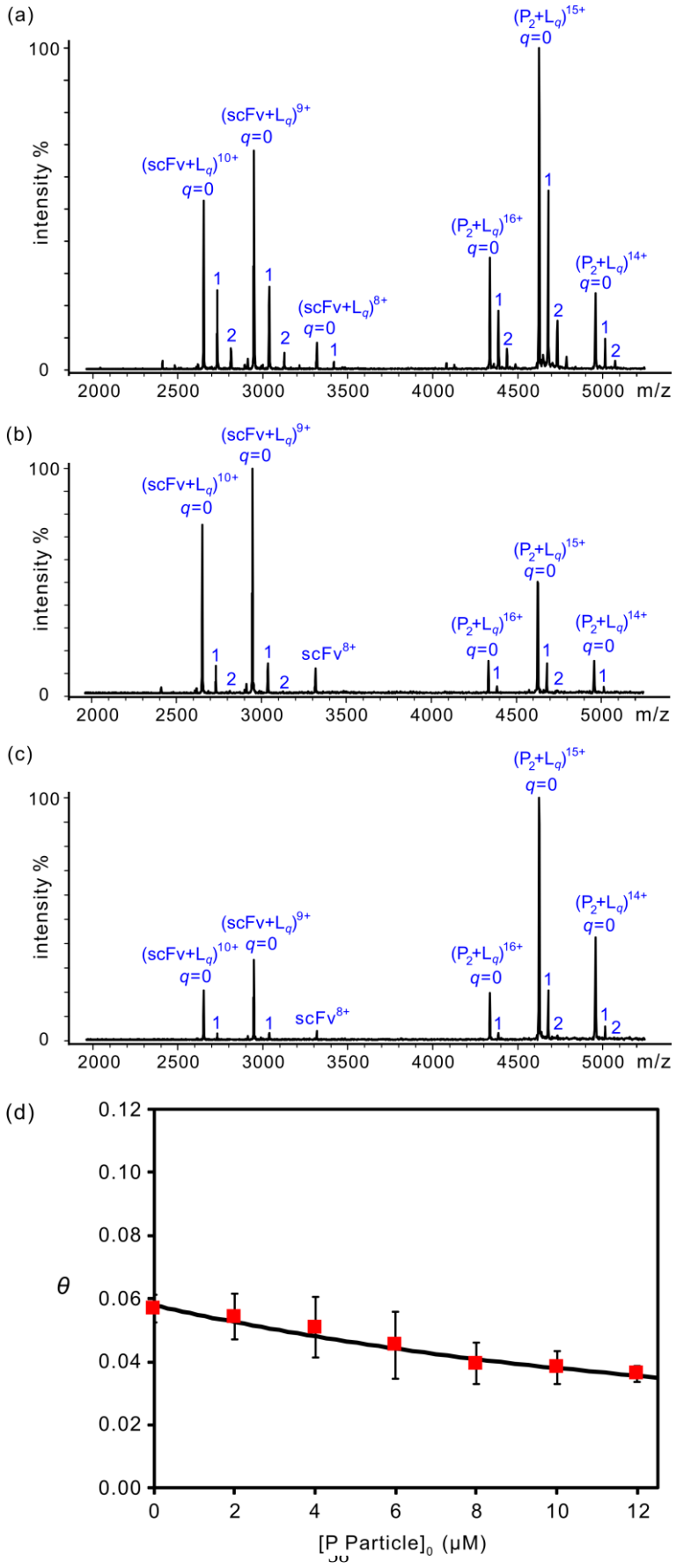

(a)
100
intensity %
0
2000
2500
3000
3500
4000
4500
5000
m/z
(scFv+L$_q$)$^{10+}$ q=0
(scFv+L$_q$)$^{9+}$ q=0
(scFv+L$_q$)$^{8+}$ q=0
(P$_2$+L$_q$)$^{16+}$ q=0
(P$_2$+L$_q$)$^{15+}$ q=0
(P$_2$+L$_q$)$^{14+}$ q=0
1
2
(b)
scFv^{8+}
(c)
(d)
θ
0.12
0.10
0.08
0.06
0.04
0.02
0.00
[P Particle]$_0$ (μM)

Figure S11. ESI mass spectra measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 12 μM P dimer (P_2), 10 μM scFv (P_{ref}), 40 μM **L12** (B type 3 tetrasaccharide) with (a) 0 μM, (b) 6 μM and (c) 12 μM P particle (24-mer) of huNoV VA387. (d) Plot of the occupancy function θ versus P particle concentration. The solution conditions for each measurement was same as (a), but with the addition of 0 to 12 μM P particle. The solid curve corresponds to the best fit of eq 5 to the experimental data. The error bars correspond to one standard derivation. The measurements were performed on a Bruker ApexQe FT-ICR mass spectrometer.

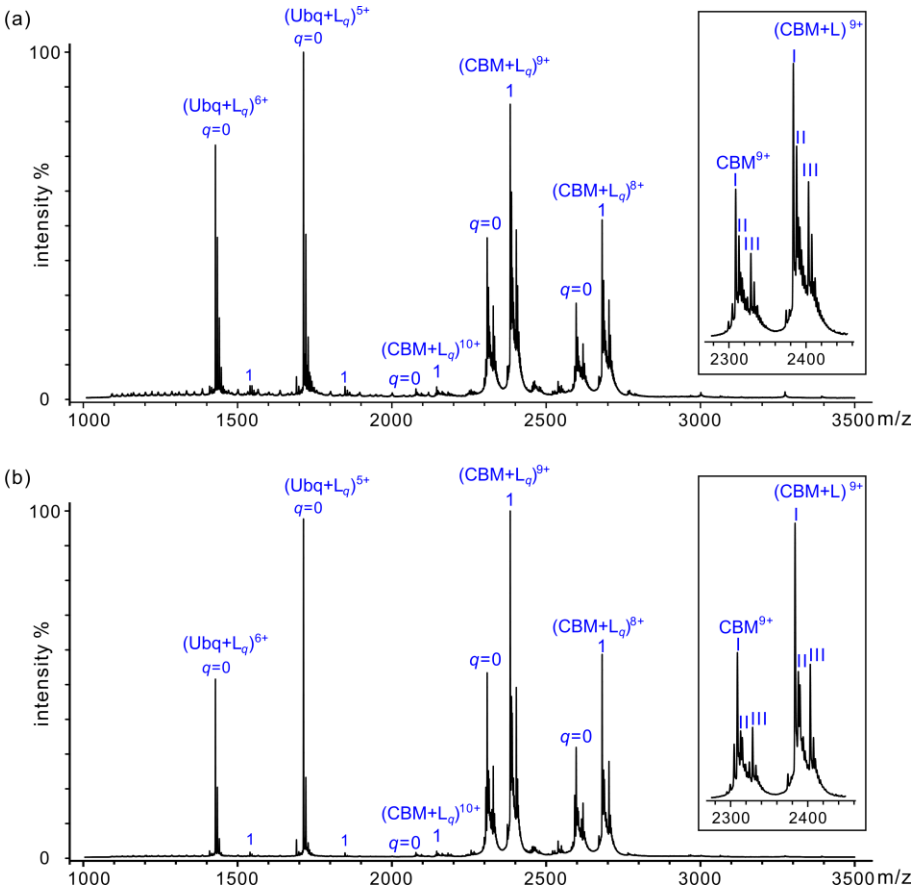


Figure S12. ESI mass spectra obtained in positive ion mode for aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12 μM CBM, 8 μM Ubq (P_{ref}), 35 μM **L1** (B trisaccharide) with (a) 0 μM and (b) 6 μM P particle (24-mer) of huNoV VA387. The concentration of ammonium acetate in (a) was 200 mM and in (b) 800 mM. These measurements were performed on a Waters Synapt G2S mass spectrometer.

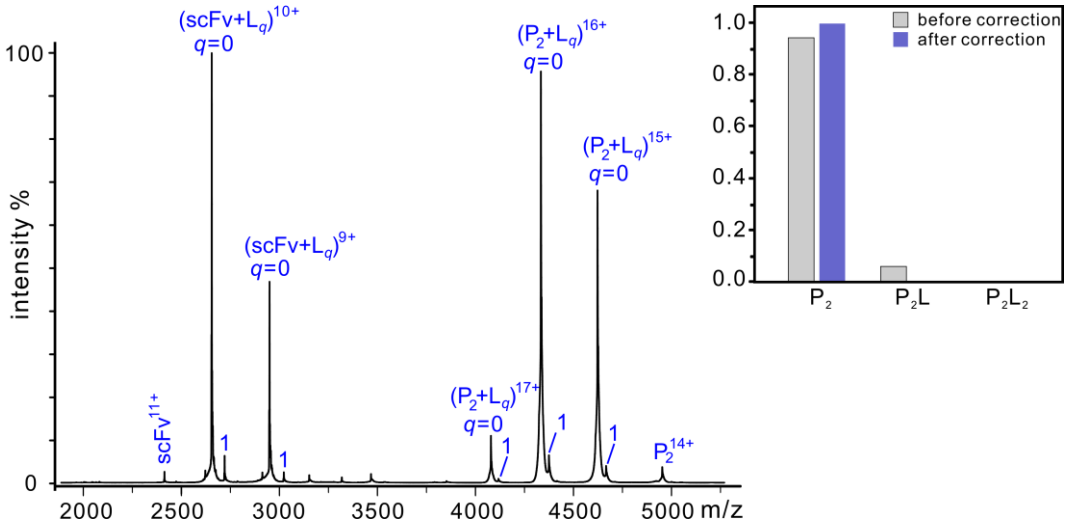


Figure S13. A representative ESI mass spectrum measured in positive ion mode for 200 mM aqueous ammonium acetate solution (pH 7 and 25 °C) containing 12 μM VA387 P dimer (P_2), 8 μM scFv (P_{ref}) and 80 μM **L14** (type 2 trisaccharide). Inset, normalized distribution of **L14** bound P dimer before and after correction for nonspecific ligand binding. The measurement was performed on a Waters Synapt G2S mass spectrometer.

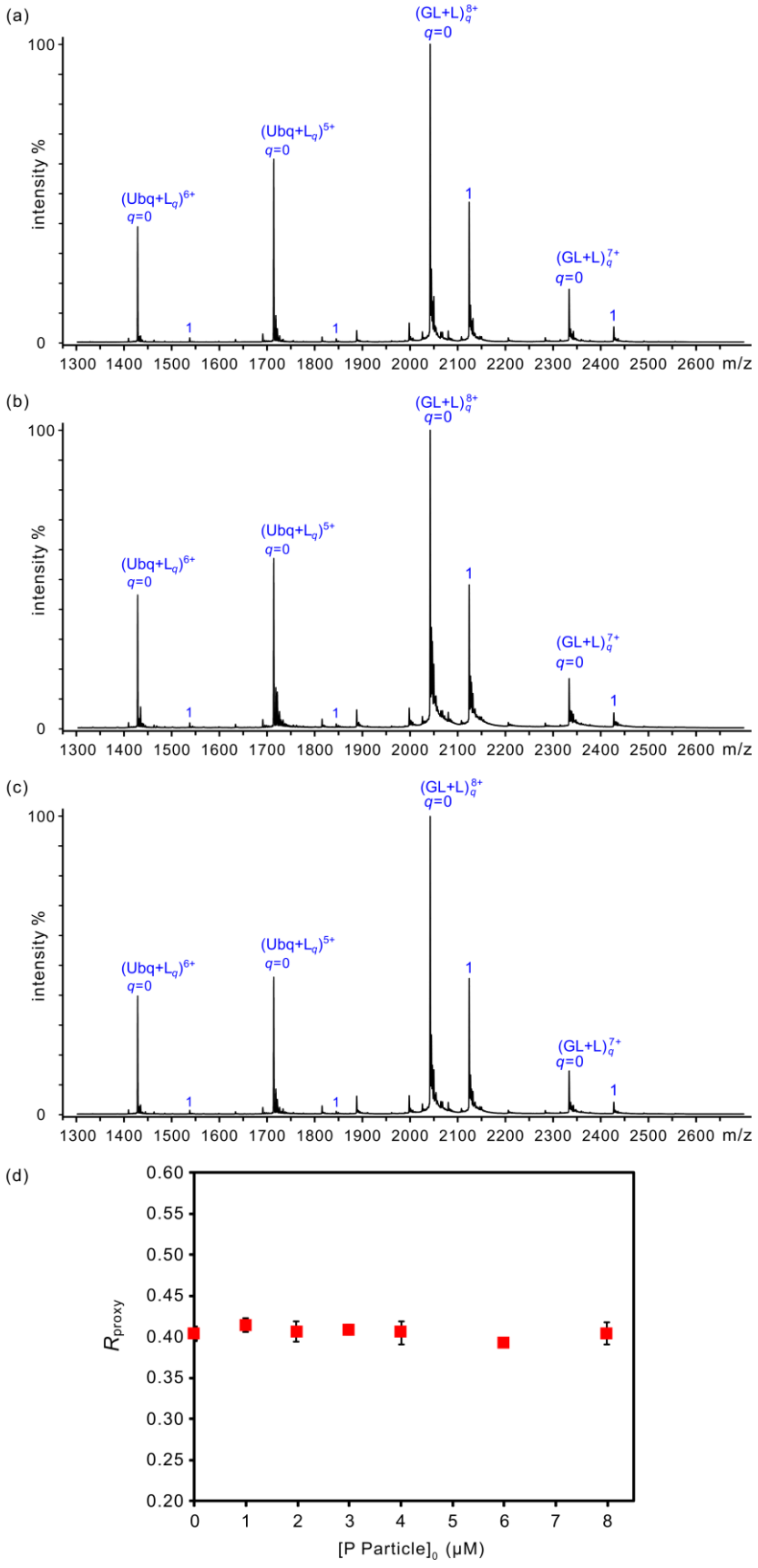

(a)
(b)
(c)
(d)
intensity %
100
0
1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 2400 2500 2600 m/z
(GL+L)$_q^{8+}$ q=0
(Ubq+L$_q$)$^{5+}$ q=0
(Ubq+L$_q$)$^{6+}$ q=0
(GL+L)$_q^{7+}$ q=0
1
R_{proxy}
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0 1 2 3 4 5 6 7 8
[P Particle]$_0$ (μM)

Figure S14. ESI mass spectra measured for 50 mM aqueous ammonium acetate solutions (pH 7 and 25 °C) containing 5 μM Gal-3C (GL), 10 μM Ubiquitin (Ubq, P_{ref}), 40 μM **L14** (type 2 trisaccharide) with (a) 0 μM, (b) 4 μM and (c) 8 μM P particle (24-mer) of huNoV VA387. (d) Plot of R_{proxy} versus P particle concentration. The solution conditions for each measurement was same as (a), but with the addition of 0 to 8 μM P particle. The error bars correspond to one standard derivation. The measurements were performed on a Waters Synapt G2S mass spectrometer.

References

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.

Soya N, Shoemaker GK, Palcic MM, Klassen JS. 2009. Comparative study of substrate and product binding to the human ABO(H) blood group glycosyltransferases. *Glycobiology*, 19:1224-1234.

Tan M, Jiang X. 2005. The P domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors. *J. Virol.*, 79:14017-14030.