Identifying Carbohydrate Ligands of a Norovirus P Particle using a Catch and Release Electrospray Ionization Mass Spectrometry Assay

Ling Han[1], Elena N. Kitova[1], Ming Tan[2,3], Xi Jiang[2,3] 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*

Abstract

Noroviruses (NoVs), the major cause of epidemic acute gastroenteritis, recognize human histo-blood group antigens (HBGAs), which are present as free oligosaccharides in bodily fluid or glycolipids and glycoproteins on the surfaces of cells. The subviral P particle formed by the protruding (P) domain of the NoV capsid protein serves as a useful model for the study NoV-HBGA interactions. Here, we demonstrate the application of a catch-and-release electrospray ionization mass spectrometry (CaR-ESI-MS) assay for screening carbohydrate libraries against the P particle to rapidly identify NoV ligands and potential inhibitors. Carbohydrate libraries of 50 and 146 compounds, which included 18 and 24 analogs of HBGA receptors, respectively, were screened against the P particle of VA387, a member of the predominant GII.4 NoVs. Deprotonated ions corresponding to the P particle bound to carbohydrates were isolated and subjected to collision-induced dissociation to release the ligands in their deprotonated forms. The released ligands were identified by ion mobility separation followed by mass analysis. All 13 and 16 HBGA ligands with intrinsic affinities >500 M^{-1} were identified in the 50 and the 146 compound libraries, respectively. Furthermore, screening revealed interactions with a series of oligosaccharides with structures found in the cell wall of mycobacteria and human milk. The affinities of these newly discovered ligands are comparable to those of the HBGA receptors, as estimated from the relative abundance of released ligand ions.

Introduction

Noroviruses (NoVs), a group of non-enveloped, single–stranded positive-sense RNA viruses in the *Calciviridae* family, are the most common viral pathogens causing acute gastroenteritis [1]. Human histo-blood group antigens (HBGAs) have been shown to be the receptors or attachment factors of human NoVs that control the host susceptibility of NoV infection [2 – 9]. The HBGAs are either present as free oligosaccharides in bodily fluids, such as blood, saliva, milk and the intestinal contents, or as complex carbohydrates covalently linked to proteins or lipids on red blood cells or mucosal epithelial cells [10,11]. The structures of the carbohydrates at the non-reducing end determine the type of HBGAs, including A, B, H or Lewis antigens. In addition, each HBGA can be further divided into six subtypes (Types 1 - 6), based on their detailed carbohydrate structures at the reducing ends. The binding specificity and affinity of NoVs to HBGAs are strain or genotype dependent and different NoV-HBGA binding patterns are known [4]. For example, VA387, a member of the widely circulated GII.4 NoVs, has been shown to bind to a variety of HBGAs, including all A, B and H antigens and some Lewis antigens [3,4,12]. On the other hand, the MOH, a strain of GII.5 genotype, binds only saliva samples from type A and B individuals, while the VA207 (GII.9) shows a preference for Le^X and Le^Y antigens [3, 4].

Currently, there is no *in vitro* cell culture system or a suitable animal model for human NoVs, which has hindered the characterization of NoV receptors. The 7.7 kb RNA genome of NoVs has three open reading frames (ORFs), in which ORF2 encodes the capsid protein (VP1). VP1 possesses two major domains, the N-terminal shell (S) domain and the C-terminal protrusion (P)

domain, linked by a short flexible hinge [13]. The S domain forms the interior capsid shell, which maintains the icosahedral structure of the virion, while the P domain forms an exterior dimeric structure that is important for the virus-receptor interactions and host immune response [14,15]. Recombinant VP1 can spontaneously assemble into a virus-like particle (VLP) *in vitro* [16]. X-ray crystallography of Norwalk virus (GI.1) VLPs revealed that the icosahedral NoV capsid is composed of 180 VP1s that organize into 90 homodimers [13]. The results of a recent electrospray ionization mass spectrometry (ESI-MS) study of Norwalk virus VLPs [17] support this observation.

The P domains of NoV VP1 can be structurally and functionally independent. Expression of the P domain alone produced P dimers with HBGA binding function [14,18,19]. Crystallography of NoV P dimers in complex with HBGAs demonstrated that the structure of the recombinant P dimers is identical to that of the native NoV capsids and revealed two symmetric HBGA binding sites on the tops of the P dimers [20–26], which corresponds to the outermost surface of the capsid. Two other P domain complexes with authentic HBGA binding function, the 24-mer P particles [27–29] and the 12-mer small P particles [30], were also made through end-modifications of the P domain. These P domain complexes provide multiple tools for the study NoV-host interactions [12,19,31–34], while the P particles have been shown as useful platforms for vaccine development against NoVs and other pathogens [29, 35–40].

Identification of inhibitors as potential antivirals against NoVs represents an active area of NoV research. Using VLPs as a model, Jiang and coworkers screened a library comprising 5000 drug-like small molecules and identified 14 compounds that efficiently inhibited binding of

VA387 VLP to HBGAs in a saliva-binding assay [41]. Screening a library of 340 compounds using saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy and spin-lock filtered NMR spectroscopy, Peters and coworkers identified 26 compounds that bound to the HBGA binding sites of a GII.4 VLP [42]. In a more recent study using in silico screening of a large online library comprising over two million compounds followed by a validation of blocking assays on P dimer-saliva interaction, five compounds that shared a common structure of cyclopenta [a] dimethyl phenanthren with an IC50 < 10.0 μM were identified [43]. Glycan array screening has also provided insights into the carbohydrate binding specificities of some human and murine VLPs. For example, the VLP of VA207 (GII.9) was shown to bind strongly to oligosaccharides with Lewis epitopes (1,3/4 linked α-L-fucose residue), while the VLP of Norwalk virus (GI.1) was found to bind to a variety of structures not found in the HBGAs (Consortium for Functional Glycomics, http://www.functionalglycomics.org/).

Recently, ESI-MS has emerged as a promising tool for identifying and quantifying protein-carbohydrate interactions *in vitro*. In particular, the catch-and-release (CaR) ESI-MS assay enables the rapid screening of carbohydrate libraries against target proteins [44–46]. The assay involves incubating the target protein with a library of compounds, followed by direct ESI-MS analysis of intact protein-ligand complexes. In principle, the identity of ligands "caught" by the protein can be found from the molecular weight (MW) of the corresponding protein-ligand complex. In cases where MW cannot be accurately determined or when dealing with isomeric species, the ligands are "released" (as ions) from the protein using collision-induced dissociation (CID), followed by accurate mass analysis, alone or in

combination with ion mobility separation (IMS) or another stage of CID [46]. Because carbohydrates have relatively low gas-phase acidities and are able to effectively compete with the protein for negative charge, the assay is normally carried out in negative ion mode [46]. It has been shown that moderate-to-high affinity ligands (K_a >10^4 M^{-1}) can be identified from libraries containing over 200 carbohydrates in a single CaR-ESI-MS measurement, which is typically completed within 1–2 min [46].

Here, we report the application of the CaR ESI-MS assay for screening carbohydrate libraries against the NoV P particle. A series of control experiments, including the screening of a 50 compound library containing multiple HBGA oligosaccharides with known affinities for the corresponding P dimer [12], were performed to confirm the reliability of the assay. The validated assay was then used to screen a carbohydrate library of 146 compounds to identify new carbohydrate ligands and potential inhibitors against NoV–HBGA interaction.

Experimental

Proteins

The P particle (24-mer, MW 865 253 Da) of NoV strain VA387 (GII.4) was produced from the P domain (residues 222-539) of VP1. A cysteine rich peptide CDCRGDCFC was linked to the C terminus of the P domain to enhance the stability of the P particle [27,28]. The procedures used for the production and purification of the P particle have been described previously [27,28]. Prior to ESI-MS analysis, the P particle was concentrated (to a final concentration of 20 μM) and exchanged into aqueous 200 mM ammonium acetate (pH 7) using Vivaspin 0.5 mL centrifugal filters (Sartorius Stedim Biotech, Germany) with a MW cutoff of 10 kDa and stored at −20 °C

until needed. The concentration of the P particle (24-mer) was measured using a Pierce BCA assay kit (Thermo Scientific, Ottawa, Canada) according to the manufacturer's instructions.

Carbohydrates

A full list of the carbohydrates used in the present study is shown in Table S1 (Supplementary Data), along with their MWs. Compounds **L1**–**L6**, **L12**–**L13**, **L55**–**L63**, **L68**, **L83**–**L113** were gifts from Prof. T. Lowary (University of Alberta); **L16, L37**–**L38**, **L114**–**L115** were gifts from Alberta Innovates Technology Futures (Alberta, Canada); **L39**–**L40**, **L69**–**L70**, **L72**–**L82**, **L116**–**L117** were gifts from Prof. C-C. Ling (University of Calgary); **L14**–**L15**, **L28**–**30**, **L50, L128**–**L140** were purchased from Dextra (Reading, UK); **L41**–**L49** were purchased from Sigma-Aldrich Canada (Oakville, Canada); **L64**–**L67**, **L141**–**L143** were purchased from IsoSep AB (Sweden) and **L7**–**L11**, **L17**–**L27**, **L31**–**L36**, **L51**–**L54**, **L71**, **L118**–**L127** and **L144**–**L146** were purchased from Elicityl SA (Crolles, France). For each compound, an aqueous 2.5 mM stock solution was prepared by dissolving a known mass of the solid sample into ultrafiltered Milli-Q water (Millipore, MA) and was stored at -20 °C until needed. To apply the CaR-ESI-MS assay, solutions of P particle (5 μM) with one or more carbohydrates (at a concentration of 10 μM each) in 200 mM ammonium acetate (25 °C, pH 7) were prepared.

Mass spectrometry

All CaR-ESI-MS measurements were carried out in negative ion mode using a Synapt G2S quadrupole-ion mobility separation-time of flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK) equipped with a nanoflow ESI (nanoESI) source. NanoESI tips were produced from borosilicate capillaries (1.0 mm o.d., 0.68 mm i.d.) pulled to ~5 μm using a P-1000

micropipette puller (Sutter Instruments, Novato, CA). A platinum wire was inserted into the nanoESI tip and a capillary voltage of 0.80–1.00 kV was applied. The source parameters were: source temperature 60 °C, cone voltage 50 V, Trap voltage 5 V, and Transfer voltage 2 V. To identify carbohydrate ligands for the P particle, ions corresponding to ligand-bound P particle were isolated using the quadrupole mass filter. The quadrupole was set to transmit a broad m/z window (approximately 200 m/z units), which allows for the simultaneous passage of free and ligand-bound P particle complexes at a given charge state. Protein-ligand complexes were subjected to CID in the Trap region of the Synapt G2S by increasing the trap voltage from 5 V to 180 V. Argon (1.42×10^{-2} mbar) was used for CID in the Trap region. In most instances, the deprotonated ligands released from the complexes could be identified from their MWs. Where required, IMS was used to separate the released isomeric ligands. For IMS separation a wave height of 35 V was used while ramping the wave velocity from 2000 to 500 m s^{-1}. In all cases a helium flow rate of 150 mL min^{-1} and a nitrogen flow rate of 40 mL min^{-1} were used. The arrival time distributions (ATDs) for the released ligands were compared to reference ATDs, which were measured for the deprotonated carbohydrates produced directly from solution. Data acquisition and processing were performed using MassLynx software (version 4.1).

Results and Discussion

Validating the CaR-ESI-MS assay

Shown in Figure 1a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM, pH 7) solution of P particle (5 μM) at 25°C. Under these solution conditions, the recombinant NoV VA387 P particle exists predominantly as a 24-mer,

with a charge state distribution from -61 to -70. The 18-mer was also detected, but with much lower abundance, with a charge state distribution ranging from -53 to -56, consistent with observation of a recent ESI-MS study of the P particle [47]. The identity of the broad, unresolved peak centred at m/z ~7000 is not known; a similar spectral feature was also reported in the previous ESI-MS study [47]. The peaks corresponding to the different charge states of the P particle are broadened due to adduct formation during the ESI process. Therefore, in order to more precisely establish the MW of the 24- and 18-mers, modest collisional heating in the Trap region was applied to strip away labile adducts. Shown in Figure S1 (Supplemental Data) is a representative ESI mass spectrum obtained using a Trap voltage of 80 V. From the mass spectrum, MWs of 864 700 ± 80 Da and 648 300 ± 110 Da are found for the 24-mer and 18-mer, respectively. These values are in reasonable agreement with the theoretical values of 865 036 Da (24-mer) and 648 782 Da (18-mer) [27].

As a starting point for establishing the reliability of the CaR-ESI-MS assay for detecting specific interactions between the P particle and carbohydrate ligands, the assay was applied to a solution of P particle and the B type 3 tetrasaccharide (**L1**), which has been shown to bind to norovirus VA387 P dimer with an intrinsic association constant ($K_{a,\mathrm{int}}$) of 1.5×10^3 M^{-1} [12]. Shown in Figure 1b is a representative ESI mass spectrum measured in negative ion mode for an aqueous ammonium acetate solution of P particle (5 μM) and **L1** (10 μM). Based on the reported $K_{a,\mathrm{int}}$ value and the assuming all 24 HBGA binding sites of the P particle are equivalent and independent, the P particle is expected to bind a maximum of two molecules of **L1** under these solution conditions. However, it was not possible to resolve the signal corresponding to the

individual complexes, i.e., the (P particle + *i***L1**) complexes where $i = 0 - 2$, at any of the observed charge states. To establish that **L1** was bound to the P particle in solution, the quadrupole mass filter was set to pass a broad m/z window centred at m/z 13 420 (which corresponds to the -65 charge state) and the transmitted ions were subjected to CID in order to release bound ligands. The resulting CID mass spectrum reveals strong signal at m/z 800.4, which corresponds to deprotonated **L1** (Figure 1c). This result demonstrates that the CaR-ESI-MS assay, as implemented here, can detect specific HBGA ligands for the P particle.

Measurements were also carried out on solutions of P particle with three different HBGA ligands, **L1**, **L2** and **L3** (Figure 2a). The corresponding $K_{a,\mathrm{int}}$ values for **L2** and **L3**, determined from binding measurements performed on the P dimer, are 8.0×10^2 M^{-1} and 6.5×10^2 M^{-1}, respectively [12]. Collision-induced dissociation of the -65 charge state, at a Trap voltage of 180 V (which was used in order to maximize the release of ligands) produced abundant signal for the deprotonated ions of **L1**, **L2** and **L3** (Figure 2b). Under these conditions, the P particle also releases protein monomers. Similar results were obtained when CID was performed on other charge states (Figure S2). These results demonstrate that multiple carbohydrate ligands can be identified, simultaneously, using the CaR-ESI-MS assay. Moreover, the relative abundances of the three ligands are qualitatively consistent with the trends in $K_{a,\mathrm{int}}$ values determined for the P dimer [12]. This finding suggests that the release efficiencies for the three ligands are similar, despite the differences in their size and structure, and raises the possibility of using the CaR-ESI-MS assay to, not only identify carbohydrate ligands from mixtures, but also establish their relative affinities.

Given the propensity for carbohydrates to bind nonspecifically to proteins during the ESI process, leading to false positives [48, 49], it was important to establish that ligands identified by the CaR-ESI-MS assay originate from specific binding in solution. To test for the occurrence of nonspecific binding, the CaR-ESI-MS assay was applied to solutions of P particle and the type 1 tetrasaccharide **L20**, which lacks the minimal recognition moiety (α-L-fucose residue) and is not expected to specifically bind to the VA387 P particle in the solution [6,12], at varying concentrations (10 – 35 μM). Shown in Figure S3a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM) solution of P particle (5 μM) and **L20** (10 μM). The CID measurements, which were carried out in a manner analogous to that described above, failed to produce any signal corresponding to the deprotonated **L20** ion (Figure S3b). Similar results were also obtained at the higher concentrations investigated (Figure S3c). These results suggest that the CaR-ESI-MS assay is not prone to false positives, at least over the range of carbohydrate concentrations used here.

Screening carbohydrate libraries against the NoV P particle

The CaR ESI-MS was used to screen two carbohydrate libraries against the NoV P particle. One of the libraries (*Library1*) was composed of 50 carbohydrates and included 18 HBGA oligosaccharides (**L1** – **L18**) known to bind to the NoV VA387 P dimer, with intrinsic affinities ranging from 200 to 1500 M^{-1} [12]. The rest of the library was made up from human and plant oligosaccharides (**L19** – **L50**) that were not expected to bind to the NoV P particle. The second library (*Library2*) was composed of 146 compounds (**L1** – **L146**) and included 24 HBGA oligosaccharides (**L1** – **L18, L51** – **L53, L69** – **L71**), other human, as well as plant and bacterial

oligosaccharides.

Library1

This library served as an additional control to validate the CaR ESI-MS assay for screening carbohydrates against the NoV P particle. Shown in Figure 3a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM) solution of P particle (5 μM) and *Library1* (each 10 μM). The -65 charge state was selected for CID (at 150 V) and the resulting mass spectrum is shown in Figure 3b. Inspection of the CID mass spectrum reveals signal corresponding to 12 different HBGA ligand MWs. Eleven of the HBGA ligands released from the P particle have unique MWs and can be identified simply from the measured m/z of their deprotonated ions ((**L2** (m/z 841.4), **L3** (m/z 638.3), **L4** (m/z 597.3), **L6** (m/z 759.3), **L7** (m/z 690.3), **L8** (m/z 893.4), **L9** (m/z 852.3), **L10** (m/z 1055.4), **L11** (m/z 1014.5), **L12** (m/z 712.4) and **L13** (m/z 671.3)). Overall, the relative abundances of the released ligands agree qualitatively with their relative affinities. The one notable exception is **L4**, which is as abundant as some of the higher affinity ligands, such as **L9** and **L12**. While there is no definitive explanation for the unusually high abundance of **L4** in the CID spectrum, it is possible that the relative HBGA affinities measured for the P particle do not reflect their affinities for the P dimer. Efforts to quantify HBGA ligand binding to the P particle are now underway in our laboratory. *Library1* contained two isomeric HBGA ligands, **L1** and **L5,** and the signal at m/z 800.4 could correspond to either or both of them. To confirm that both **L1** and **L5** bind to the P particle, the released ions were subjected to IMS (Figure 3c). Inspection of the ATD measured for the released ions reveals two features (at 7.4 ms and 8.0 ms), which indicates that there are at least

two structures present. Comparison of the ATD measured for the released ions with those of the individual **L1** and **L5** ions (Figure 3c) confirms that both ligands were released from the P particle. Furthermore, the relative areas of the ATDs for the two ligands are consistent with **L1** having a higher affinity than **L5** for the P particle [12].

Analysis of the CaR-ESI-MS data obtained for *Library1* reveals that 13 of the 18 HBGA ligands could be identified in a single measurement. Furthermore, all ligands with $K_{a,\mathrm{int}}$ >500 M^{-1} (as determined for the P dimer) were successfully detected. Interestingly, the five HBGAs that were not detected (**L14** – **L18**) correspond to Lewis oligosaccharides that have very low affinity ($K_{a,\mathrm{int}}$ <300 M^{-1} for the P dimer. Measurements carried out on solutions of the P particle and **L14** – **L18** at higher concentrations (each 20 μM) failed to identify binding of any of these ligands (Figure S4a and b). These results may suggest that the Lewis antigens have affinities that are lower than expected based on the results obtained for the P dimer.

Library2

With the goal of identifying new carbohydrate ligands and potential inhibitors for NoVs, the CaR-ESI-MS was used to screen a library containing 146 carbohydrates against the VA387 P particle. Shown in Figure 4a is a representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate (200 mM) solution of P particle (5 μM) and *Library2* (10 μM each component). The -65 charge state was selected for CID (at 150 V) and the resulting mass spectrum is shown in Figure 4b. A total of 28 “hits” were identified by CaR-ESI-MS, including the 13 ABH oligosaccharides that were detected from *Library1*. Signals for three other HBGA oligosaccharides not previously tested against the corresponding P dimer, Globo A

heptasaccharide (**L52**, m/z 1217.4), Globo B heptasaccharide (**L53**, m/z 1176.4) and Globo H hexasaccharide (**L51**, m/z 1014.5), were detected (Figure 4b). The deprotonated **L51** ion is isobaric with deprotonated **L11** ion; however, these two ions can be distinguished by IMS (Figure 4c). Moreover, comparison of the relative area of the ATDs for the two compounds indicates that **L11** has higher affinity to P particle than **L51**. The fucosyl GM1 hexasaccharide (**L54**, m/z 1143.5), which has a 1,2-linked α-L-fucose residue, was also detected and its abundance is similar to that of the H type oligosaccharides (e.g. **L3**, **L4** and **L7**). Of the 17 human milk oligosaccharides (HMOs) present in *Library2* (**L4**, **L18** – **L23**, **L28**, **L64** – **L67**, **L86** – **L87**, **L141** – **L143**), four were found to bind to the P particle, **L4** (m/z 597.3), **L64** (m/z 487.2), **L66** (m/z 778.2) and **L67** (m/z 1363.9). Each of these possesses an α-L-fucose residue, which has been shown to be an important recognition element for NoV VA387. Of the four HMOs, **L66** (α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-D-Glc) exhibited the highest affinity, comparable to the strongest HBGA binders in the library, for the P particle based on the intensity of the released HMOs.

Surprisingly, a number of bacterial oligosaccharides that are based on structures found in the cell wall of mycobacteria and contain α-L-rhamnose (**L60** (m/z 429.1), **L61** (m/z 478.2), **L62** (m/z 603.2) and **L63** (m/z 617.2)) or α-D-arabinofuranose residues (**L56** (m/z 427.0), **L57** (m/z 559.1), **L58** (m/z 698.2) and **L59** (m/z 768.2)) were detected; their abundances are comparable to those of some of the HBGA oligosaccharides tested. Notably, **L56** – **L61** do not possess α-L-fucose. This finding is consistent with earlier reports of HBGA binding NoVs that also interact with compounds that do not possess α-L-fucose residues [34,41,42,50]. It has been

suggested that these ligands may either mimic the fucose moiety and interact with the NoV in the same binding site, or bind at a different site [42,50]. Both **L60** and **L61** possess an α-L-rhamnose residue, which may mimic the structure of α-L-fucose and interact with P particle through the same binding site. To our knowledge, interactions between NoVs and arabinofuranose containing glycans have not been previously reported and the nature of these interactions is not known. Efforts to localize the binding sites of these bacterial oligosaccharides using hydrogen-deuterium exchange MS are currently underway in our laboratory. Curiously, 4',6'-O-benzylidene maltose (**L55**) was also detected. Given that none of the unmodified maltooligosaccharides present in the library (**L41** – **L46**) were found to interact with the P particle, it seems likely that the benzylidene group is responsible for binding.

Conclusions

In summary, a CaR-ESI-MS assay was used to screen carbohydrate libraries against the P particle of NoV VA387 to identify new carbohydrate ligands. To our knowledge this is the first reported example of the application of a CaR-ESI-MS assay to a large protein assembly. The results of control experiments demonstrated the reliability of the assay for rapidly (1 – 2 min) identifying multiple HBGA ligands present in mixtures carbohydrates. Isomeric ligands could be distinguished by performing IMS on the released ligands (in their deprotonated form). Moreover, the relative abundances of the released ligands provides a qualitatively measure of their relative affinities for the P particle. Application of the CaR-ESI-MS assay to a library of 146 carbohydrates identified all 16 ABH type ligands present. Furthermore, screening revealed interactions with a series of oligosaccharides with structures found in the cell wall of

mycobacteria and human milk. The affinities of these newly discovered ligands are comparable to those of the HBGA receptors, as estimated from the relative abundance of released ligand ions.

Acknowledgements

The authors thank the Alberta Glycomics Centre and the National Institutes of Health of the Unites States of America for supporting this research and Professors T. Lowary (University of Alberta) and C-C. Ling (University of Calgary), as well as Alberta Innovates Technology Futures for generously providing some of the carbohydrates used in this work. L.H. also acknowledges an Alberta Innovates Graduate Student Scholarship.

References

1. Patel, M. M., Widdowson, M.-A., Glass, R. I., Akazawa, K., Vinje, J., Parashar, U. D.: Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg. Infect. Dis.* **14**, 1224-1231 (2008)

2. Hutson, A. M., Atmar, R. L., Marcus, D. M., Estes, M. K.: Norwalk virus-like particle hemagglutination by binding to H histo-blood group antigens. *Journal of Virology* **77**, 405-415 (2003)

3. Huang, P., Farkas, T., Marionneau, S., Zhong, W., Ruvoen-Clouet, N., Morrow, A. L., Altaye, M., Pickering, L. K., Newburg, D. S., LePendu, J., Jiang, X.: Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: Identification of 4 distinct strain-specific patterns. *J. Infect. Dis.* **188**, 19-31 (2003)

4. Huang, P., Farkas, T., Zhong, W., Thornton, S., Morrow, A. L., Jiang, X.: 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 (2005)

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

6. Fiege, B., Rademacher, C., Cartmell, J., Kitov, P. I., Parra, F., Peters, T.: 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 (2012)

7. Tan, M., Jiang, X.: Norovirus-host interaction: implications for disease control and

prevention. *Expert reviews in molecular medicine* **9**, 1-22 (2007)

8. Tan, M., Jiang, X.: Norovirus gastroenteritis, carbohydrate receptors, and animal models. *PLoS pathogens* **6**, e1000983 (2010)

9. Tan, M., Jiang, X.: Norovirus-host interaction: Multi-selections by human histo-blood group antigens. *Trends in microbiology* **19**, 382-388 (2011)

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

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

12. Han, L., Kitov, P. I., Kitova, E. N., Tan, M., Wang, L., Xia, M., Jiang, X., Klassen, J. S.: Affinities of recombinant norovirus P dimers for human blood group antigens. *Glycobiology* **23**, 276-285 (2013)

13. Prasad, B. V. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G., Estes, M. K.: X-ray crystallographic structure of the Norwalk virus capsid. *Science* **286**, 287-290 (1999)

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

15. Tan, M., Huang, P., Meller, J., Zhong, W., Farkas, T., Jiang, X.: Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: Evidence for a binding pocket. *J. Virol.* **77**, 12562-12571 (2003)

16. Jiang, X., Wang, M., Graham, D. Y., Estes, M. K.: Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. *J. Virol.* **66**, 6527-6532 (1992)

17. Shoemaker, G. K., van Duijn, E., Crawford, S. E., Uetrecht, C., Baclayon, M., Roos, W. H., Wuite, G. J. L., Estes, M. K., Prasad, B. V. V., Heck, A. J. R.: Norwalk virus assembly and stability monitored by mass spectrometry. *Mol. Cell. Proteomics* **9**, 1742-1751 (2010)

18. Tan, M., Meller, J., Jiang, X.: C-terminal arginine cluster is essential for receptor binding of norovirus capsid protein. *Journal of Virology* **80**, 7322-7331 (2006)

19. Tan, M., Xia, M., Cao, S., Huang, P., Farkas, T., Meller, J., Hegde, R. S., Li, X., Rao, Z., Jiang, X.: Elucidation of strain-specific interaction of a GII-4 norovirus with HBGA receptors by site-directed mutagenesis study. *Virology* **379**, 324-334 (2008)

20. Bu, W., Mamedova, A., Tan, M., Xia, M., Jiang, X., Hegde, R. S.: Structural basis for the receptor binding specificity of Norwalk virus. *J Virol* **82**, 5340-5347 (2008)

native mass spectrometry. *J. Struct. Biol.* **177**, 273-282 (2012)

21. Cao, S., Lou, Z., Tan, M., Chen, Y., Liu, Y., Zhang, Z., Zhang, X. C., Jiang, X., Li, X., Rao, Z.: Structural basis for the recognition of blood group trisaccharides by norovirus. *J. Virol.* **81**, 5949-5957 (2007)

22. Choi, J. M., Hutson, A. M., Estes, M. K., Prasad, B. V. V.: Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9175-9180 (2008)

23. Chen, Y., Tan, M., Xia, M., Hao, N., Zhang, X. C., Huang, P., Jiang, X., Li, X., Rao, Z.: Crystallography of a Lewis-binding norovirus, elucidation of strain-specificity to the Polymorphic Human Histo-Blood Group Antigens. *PLoS Pathog.* **7**, e1002152 (2011)

24. Hansman, G. S., Biertumpfel, C., Georgiev, I., McLellan, J. S., Chen, L., Zhou, T. Q.,

Katayama, K., Kwong, P. D.: Crystal structures of GII.10 and GII.12 norovirus protruding domains in complex with histo-blood group antigens reveal details for a potential site of vulnerability. *J. Virol.* **85**, 6687-6701 (2011)

25. Shanker, S., Choi, J. M., Sankaran, B., Atmar, R. L., Estes, M. K., Prasad, B. V. V.: Structural analysis of histo-blood group antigen binding specificity in a norovirus GII.4 epidemic variant: Implications for epochal evolution. *J. Virol.* **85**, 8635-8645 (2011)

26. Kubota, T., Kumagai, A., Ito, H., Furukawa, S., Someya, Y., Takeda, N., Ishii, K., Wakita, T., Narimatsu, H., Shirato, H.: Structural basis for the recognition of Lewis antigens by genogroup I norovirus. *J Virol* **86**, 11138-11150 (2012)

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

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

29. Tan, M., Huang, P., Xia, M., Fang, P. A., Zhong, W., McNeal, M., Wei, C., Jiang, W., Jiang, X.: Norovirus P particle, a novel platform for vaccine development and antibody production. *J Virol* **85**, 753-764 (2011)

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

31. Tan, M., Zhong, W., Song, D., Thornton, S., Jiang, X.: E. coli-expressed recombinant

norovirus capsid proteins maintain authentic antigenicity and receptor binding capability. *J. Med. Virol.* **74**, 641-649 (2004)

32. Tan, M., Xia, M., Chen, Y., Bu, W., Hegde, R. S., Meller, J., Li, X., Jiang, X.: Conservation of carbohydrate binding interfaces: evidence of human HBGA selection in norovirus evolution. *PLoS One* **4**, e5058 (2009)

33. Shang, J., Cheng, F., Dubey, M., Kaplan, J. M., Rawal, M., Jiang, X., Newburg, D. S., Sullivan, P. A., Andrade, R. B., Ratner, D. M.: An organophosphonate strategy for functionalizing silicon photonic biosensors. *Langmuir* **28**, 3338-3344 (2012)

34. Zhang, X., Dai, Y., Zhong, W., Tan, M., Lv, Z., Zhou, Y., Jiang, X.: Tannic acid inhibited norovirus binding to HBGA receptors, a study of 50 Chinese medicinal herbs. *Bioorg. Med. Chem.* **20**, 1616-1623 (2012)

35. Tan, M., Jiang, X.: Norovirus P particle, a subviral nanoparticle for vaccine development against norovirus, rotavirus and influenza virus. *Nanomedicine* **7**, 1-9 (2012)

36. Xia, M., Tan, M., Wei, C., Zhong, W., Wang, L., McNeal, M., Jiang, X.: A candidate dual vaccine against influenza and noroviruses. *Vaccine* **29**, 7670-7677 (2011)

37. Fang, H., Tan, M., Xia, M., Wang, L., Jiang, X.: Norovirus P particle efficiently elicits innate, humoral and cellular immunity. *PLoS One* **8**, e63269 (2013)

38. Dai, Y., Wang, Y., Zhang, X., Tan, M., Xia, M., Wu, X., Jiang, X., Nie, J.: Evaluation of anti-norovirus IgY from egg yolk of chickens immunized with norovirus P particles. *Journal of virological methods* **186**, 126-131 (2012)

39. Dai, Y., Zhang, X., Tan, M., Huang, P., Lei, W., Fang, H., Zhong, W., Jiang, X.: A dual

chicken IgY against rotavirus and norovirus. *Antivir Res*. **97**, 293-300 (2013)

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

41. Feng, X., Jiang, X.: Library screen for inhibitors targeting norovirus binding to histo-blood group antigen receptors. *Antimicrob. Agents Chemother.* **51**, 324-331 (2007)

42. Rademacher, C., Guiard, J., Kitov, P. I., Fiege, B., Dalton, K. P., Parra, F., Bundle, D. R., Peters, T.: Targeting norovirus infection-multivalent entry inhibitor design based on NMR experiments. *Chem. Eur. J.* **17**, 7442-7453 (2011)

43. Zhang, X., Tan, M., Chhabra, M., Dai, Y., Meller, J., Jiang, X.: Inhibition of histo-blood group antigen binding as a novel strategy to block norovirus infections. *PLoS One* **8**, e69379 (2013)

44. Cederkvist, F., Zamfir, A. D., Bahrke, S., Eijsink, V. G. H., Sorlie, M., Peter-Katalinic, J., Peter, M. G.: Identification of a high-affinity-binding oligosaccharide by (+) nanoelectrospray quadrupole time-of-flight tandem mass spectrometry of a noncovalent enzyme-ligand complex. *Angew. Chem. Int. Ed.* **45**, 2429-2434 (2006)

45. Abzalimov, R. R., Dubin, P. L., Kaltashov, I. A.: Glycosaminoglycans as naturally occuring combinatorial libraries: Spectrometry-based strategy for characterization of anti-thrombin interaction strategy with low molecular weight heparin and heparin oligomers. *Anal. Chem.* **79**, 6055-6063 (2007)

46. El-Hawiet, A., Shoemaker, G. K., Daneshfar, R., Kitova, E. N., Klassen, J. S.: Applications of a catch and release electrospray ionization mass spectrometry assay for carbohydrate library

screening. *Anal. Chem.* **84**, 50-58 (2012)

47. Bereszczak, J. Z., Barbu, I. M., Tan, M., Xia, M., Jiang, X., van Duijn, E., Heck, A. J. R.: Structure, stability and dynamics of norovirus P domain derived protein complexes studied by native mass spectrometry. *J. Struct. Biol.* **177**, 273-282 (2012)

48. Sun, J. X., Kitova, E. N., Wang, W. J., Klassen, J. S.: Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal. Chem.* **78**, 3010-3018 (2006)

49. Kitova, E. N., El-Hawiet, A., Schnier, P. D., Klassen, J. S.: Reliable determinations of protein-ligand interactions by direct ESI-MS measurements. Are we there yet? *J. Am. Soc. Mass Spectrom.* **23**, 431-441 (2012)

50. Hansman, G. S., Shahzad-ul-Hussan, S., McLellan, J. S., Chuang, G.-Y., Georgiev, I., Shimoike, T., Katayama, K., Bewley, C. A., Kwong, P. D.: Structural basis for norovirus inhibition and fucose mimicry by citrate. *J. Virol.* **86**, 284-292 (2012)

Figure captions

Figure 1. ESI mass spectra acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of (a) P particle (5 μM) and (b) P particle (5 μM) and **L1** (10 μM), at pH 7 and 25 °C. (c) CID mass spectrum measured for the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 150 V.

Figure 2. (a) ESI mass spectrum obtained in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5 μM) and **L1**, **L2** and **L3** (10 μM each), at pH 7 and 25 °C. (b) CID mass spectrum measured for the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 180 V.

Figure 3. (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5 μM) and *Library1* (10 μM each), at pH 7 and 25 °C. (b) CID mass spectrum of the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 150 V. (c) Arrival time distributions measured for the deprotonated **L1** and **L5** ions (m/z 800.3) following release from the P particle (*post-release*) and the deprotonated **L1** and **L5** ions obtained directly from solution (*reference*).

Figure 4. (a) Representative ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) solution of P particle (5 μM) and *Library2* (10 μM each), at pH 7 and 25 °C. (b) CID mass spectrum of the P particle and its carbohydrate complexes at the -65 charge state acquired at a Trap voltage of 150 V.

(c) Arrival time distributions measured for the deprotonated **L11** and **L51** ions (m/z 1014.4) following release from the P particle (*post-release*) and the deprotonated **L11** and **L51** ions obtained directly from solution (*reference*).

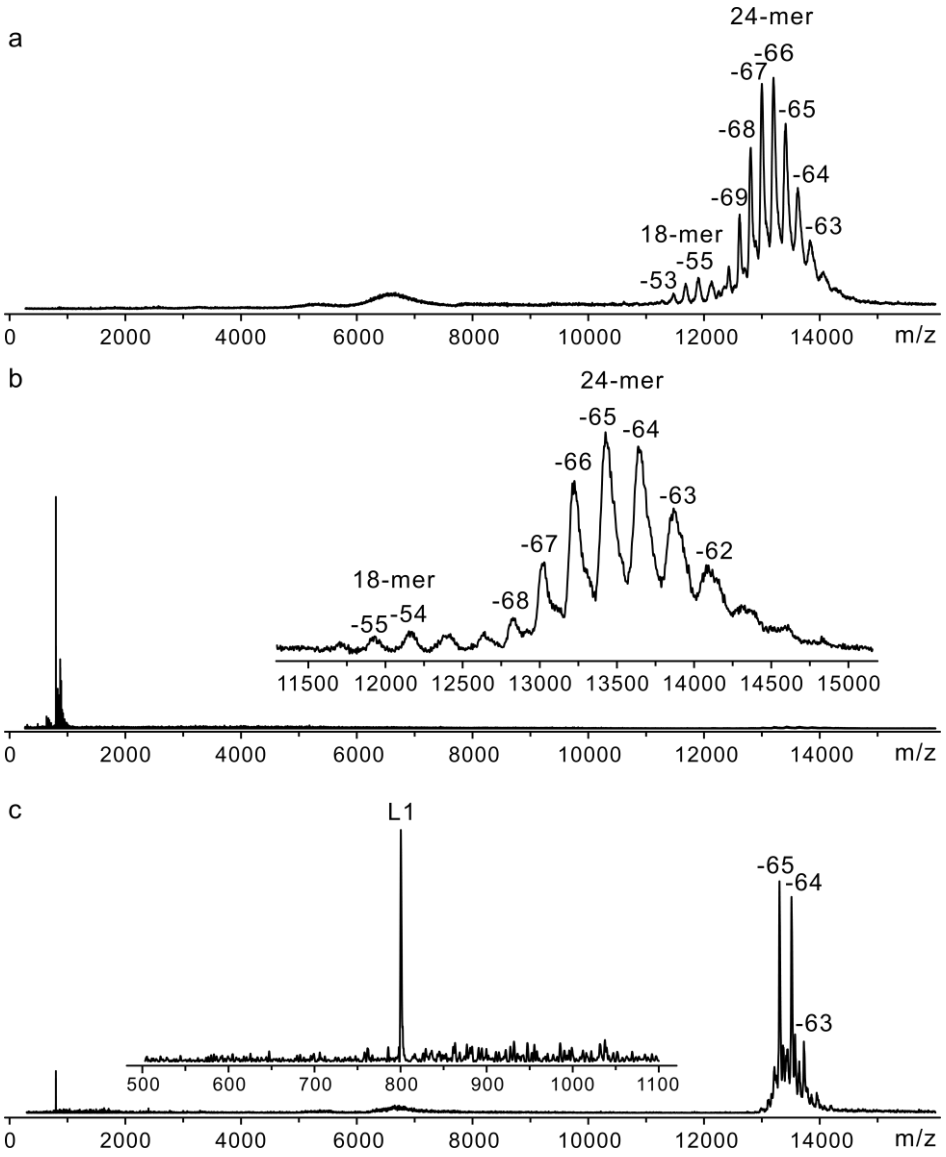
a
24-mer
-66
-67
-65
-68
-64
-69
-63
18-mer
-55
-53
0
2000
4000
6000
8000
10000
12000
14000
m/z
b
24-mer
-65
-64
-66
-63
-67
-62
18-mer
-55
-54
-68
11500
12000
12500
13000
13500
14000
14500
15000
0
2000
4000
6000
8000
10000
12000
14000
m/z
c
L1
-65
-64
-63
500
600
700
800
900
1000
1100
0
2000
4000
6000
8000
10000
12000
14000
m/z

Figure 1

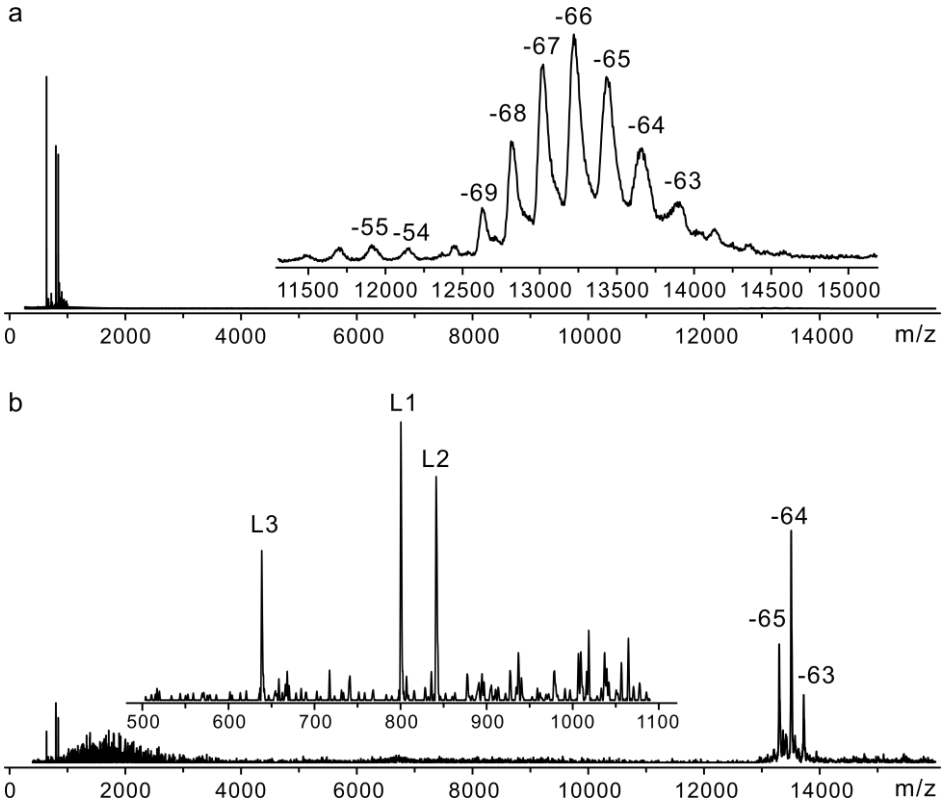
a
-66
-67
-65
-68
-64
-63
-69
-55
-54
11500 12000 12500 13000 13500 14000 14500 15000
0 2000 4000 6000 8000 10000 12000 14000 m/z
b
L1
L2
L3
-64
-65
-63
500 600 700 800 900 1000 1100
0 2000 4000 6000 8000 10000 12000 14000 m/z

Figure 2

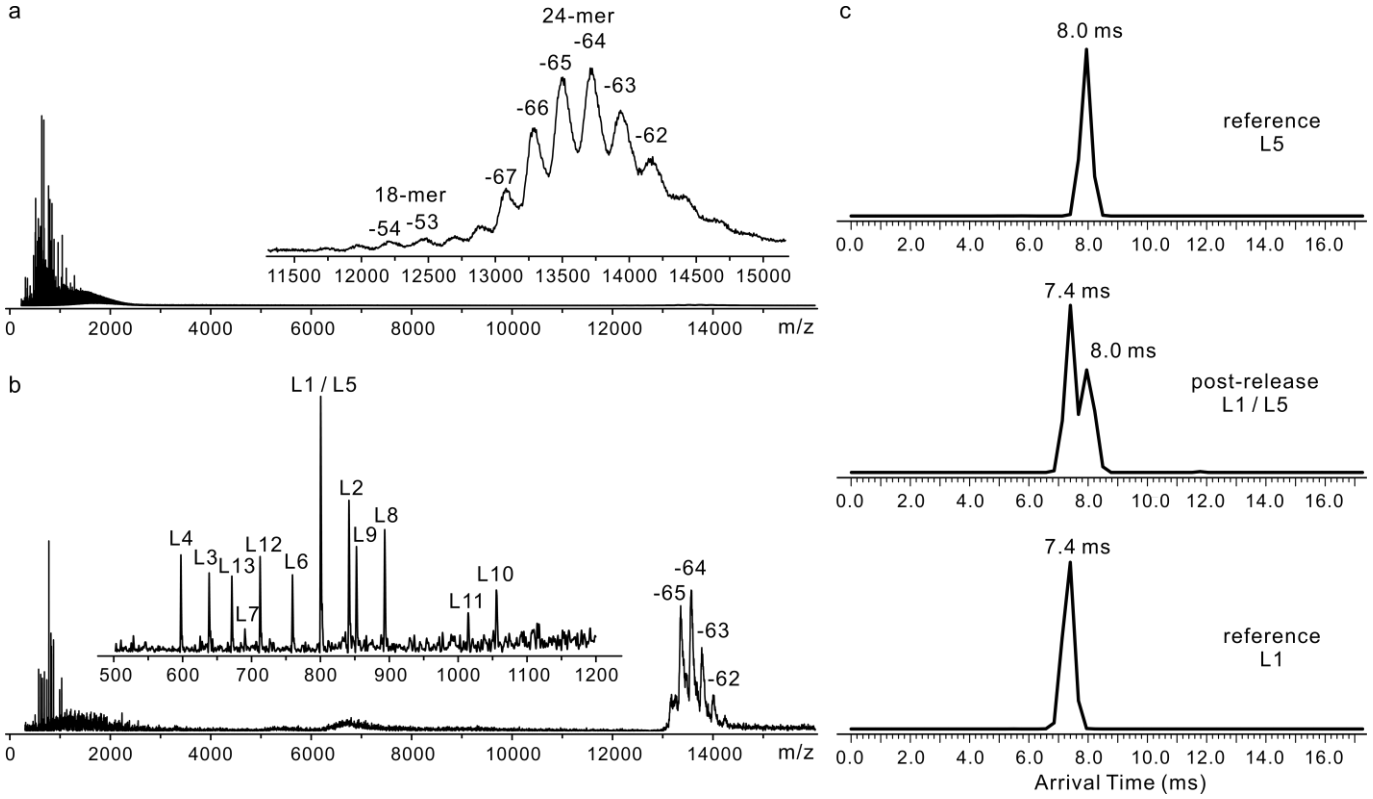
a
24-mer
-64
-65
-63
-66
-62
-67
18-mer
-54
-53
11500 12000 12500 13000 13500 14000 14500 15000
0 2000 4000 6000 8000 10000 12000 14000 m/z
b
L1 / L5
L2
L8
L9
L4
L3
L12
L13
L6
L7
L10
L11
-64
-65
-63
-62
500 600 700 800 900 1000 1100 1200
0 2000 4000 6000 8000 10000 12000 14000 m/z
c
8.0 ms
reference
L5
0.0 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0
7.4 ms
8.0 ms
post-release
L1 / L5
0.0 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0
7.4 ms
reference
L1
0.0 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0
Arrival Time (ms)

Figure 3

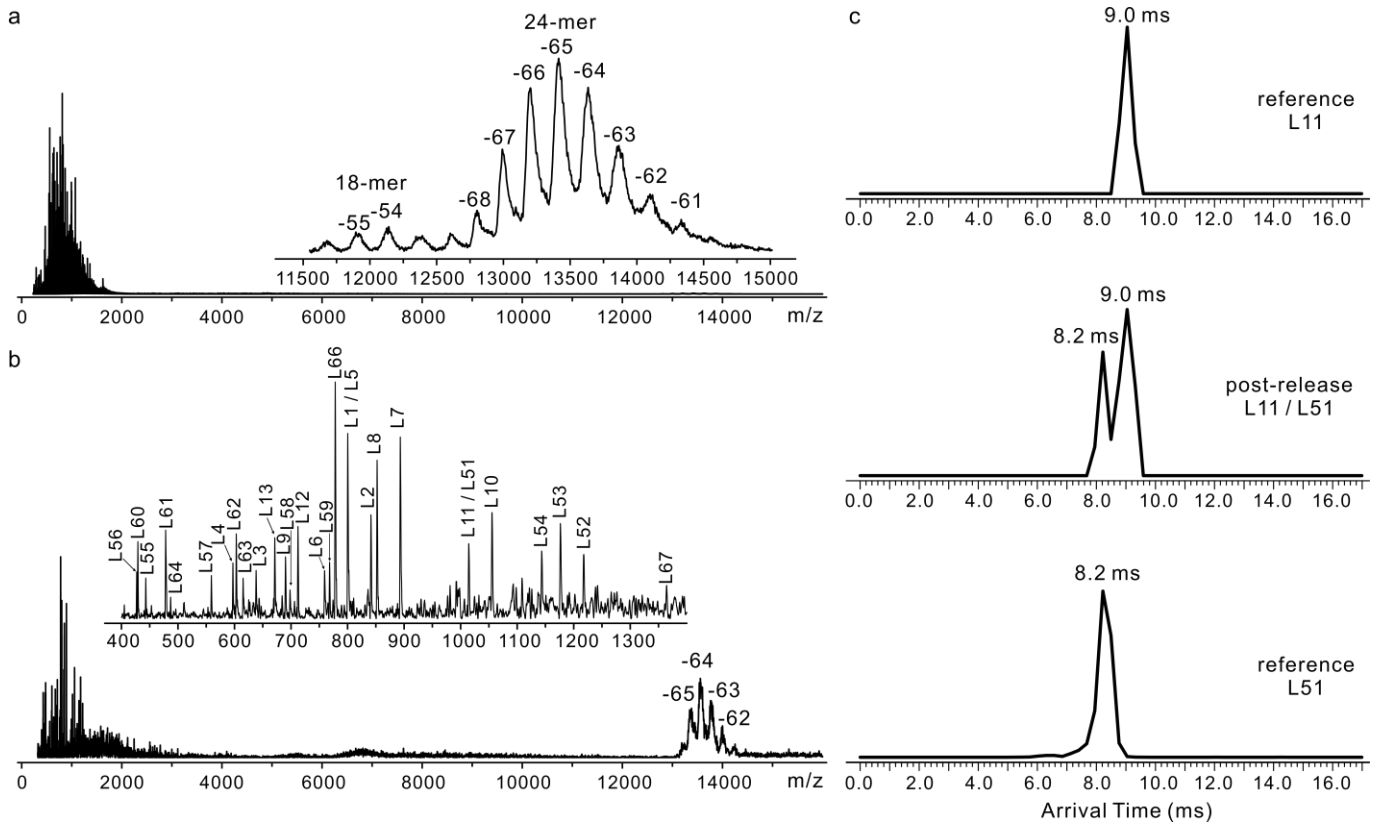
a
24-mer
-65
-66
-64
-67
-63
-62
-61
18-mer
-55
-54
-68
11500 12000 12500 13000 13500 14000 14500 15000
0 2000 4000 6000 8000 10000 12000 14000 m/z
b
L66
L1 / L5
L8
L7
L11 / L51
L10
L2
L54
L53
L52
L67
L56
L60
L55
L61
L64
L57
L4
L62
L63
L3
L13
L9
L58
L12
L6
L59
400 500 600 700 800 900 1000 1100 1200 1300
-64
-65
-63
-62
0 2000 4000 6000 8000 10000 12000 14000 m/z
c
9.0 ms
reference
L11
0.0 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0
9.0 ms
8.2 ms
post-release
L11 / L51
0.0 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0
8.2 ms
reference
L51
0.0 2.0 4.0 6.0 8.0 10.0 12.0 14.0 16.0
Arrival Time (ms)

Figure 4

Supplementary data for:

Identifying Carbohydrate Receptors of a Norovirus P Particle using a Catch and Release Electrospray Ionization Mass Spectrometry Assay

Ling Han, Elena N. Kitova, Ming Tan, Xi Jiang and John S. Klassen

Table S1. Structures and molecular weights (MWs) of components of carbohydrate library.[a]

Carbohydrate	Structure	MW (Da)
L1	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-α-D-GalNAc-O$(CH_2)_6CH=CH_2$	801.36
L2	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-α-D-GalNAc-O$(CH_2)_6CH=CH_2$	842.39
L3	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-α-D-GalNAc-O$(CH_2)_6CH=CH_2$	639.31
L4	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$	598.28
L5	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$	801.36
L6	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-O$(CH_2)_6CH=CH_2$	760.34
L7	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	691.25
L8	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	894.33
L9	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	853.31
L10	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	1056.39
L11	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	1015.36
L12	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-O$(CH_2)_8COOC_2H_5$	713.35
L13	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)-]-β-D-Gal-O$(CH_2)_8COOC_2H_5$	672.32
L14	β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-D-GlcNAc	529.20
L15	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-D-GlcNAc	675.26
L16	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-O$(CH_2)_8CONH(CH_2)_2NH_2$	727.37
L17	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-D-Gal	837.31
L18	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-D-Glc	999.36
L19	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	545.20
L20	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	707.25

L21	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	707.25
L22	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	1072.38
L23	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)- β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)- β-D-Gal-(1→4)-D-Glc	1437.51
L24	β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	545.20
L25	β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-β-D-Gal-(1→4)-D-Glc	707.25
L26	β-D-GalNAc-(1→4)-β-D-Gal-(1→4)-D-Glc	545.20
L27	α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	504.17
L28	β-D-Gal-(1→4)-β-D-GlcNAc-(1→6)-[β-D-Gal-(1→4)- β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-D-Glc	1072.38
L29	β-D-GlcNAc-(1→4)-β-D-GlcNAc-(1→4)-β-D-GlcNAc-(1→4)-D-GlcNAc	830.33
L30	β-D-GlcNAc-(1→4)-β-D-GlcNAc-(1→4)-β-D-GlcNAc-(1→4)- β-D-GlcNAc-(1→4)-β-D-GlcNAc-(1→4)-D-GlcNAc	1236.49
L31	α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)- β-D-Gal-(1→4)-D-Glc	869.30
L32	β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	869.30
L33	β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→3)-β-D-Gal-(1→4)-D-Glc	869.30
L34	α-D-GalNAc-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)- D-Glc	910.33
L35	α-D-GalNAc-(1→3)-β-D-GalNAc-(1→3)-D-Gal	586.22
L36	α-D-GalNAc-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→3)-β-D-Gal-(1→4)- D-Glc	910.33
L37	α-D-Glc-(1→3)-α-D-Man-(1→2)-α-D-Man-O$(CH_2)_8COOCH_3$	674.30
L38	α-D-Glc-(1→3)-α-D-Man-(1→2)-α-D-Man-O$(CH_2)_2CH_3$	546.22
L39	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc- O$(CH_2)_6NH_2$	847.38
L40	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc- O$(CH_2)_6NH_2$	847.38
L41	α-D-Glc-(1→4)-α-D-Glc-(1→4)-D-Glc	504.17
L42	α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-D-Glc	666.22
L43	α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-D-Glc	828.27
L44	α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)- α-D-Glc-(1→4)-D-Glc	990.33
L45	α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)- α-D-Glc-(1→4)-α-D-Glc-(1→4)-D-Glc	1152.38
L46	α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)- α-D-Glc-(1→4)-α-D-Glc-(1→4)-α-D-Glc-(1→4)-D-Glc	1314.43
L47	β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-D-Glc	666.22

L48	β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-D-Glc	990.33
L49	α-D-Man-(1→6)-[α-D-Man-(1→3)]-α-D-Man-(1→6)-[α-D-Man-(1→3)]-D-Man	828.27
L50	α-D-Man-(1→6)-[α-D-Man-(1→3)]-α-D-Man-(1→6)-D-Man	666.22
L51	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	1015.36
L52	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	1218.44
L53	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	1177.41
L54	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	1144.40
L55	4',6'O-benzylidene-α-D-Glc-(1→4)-α-D-Glc-OCH_3	444.16
L56	α-D-Ara*f*-(1→3)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OCH_3	428.15
L57	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-[α-D-Ara*f*-(1→3)]-α-D-Ara*f*-OCH_3	560.20
L58	α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-$O(CH_2)_8N_3$	699.31
L59	α-D-Ara*f*-(1→3)-[α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-$O(CH_2)_8NHCOCF_3$	769.30
L60	α-L-Rha(1→3)-2-OCH_3-α-L-Rha-O(p-OCH_3Ph)	430.18
L61	α-L-Rha-(1→3)-α-D-GlcNAc-$O(CH_2)_7CH_3$	479.27
L62	2,4-di-OCH_3-α-L-Fuc-(1→3)-α-L-Rha-(1→3)-2-OCH_3-α-L-Rha-O(p-OCH_3Ph)	604.27
L63	2,3,4-tri-OCH_3-α-L-Fuc-(1→3)-α-L-Rha-(1→3)-2-OCH_3-α-L-Rha-O(p-OCH_3Ph)	618.29
L64	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-D-Glc	488.17
L65	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-D-Glc	634.23
L66	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-D-Glc	779.27
L67	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→6)-[α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)]-β-D-Gal-(1→4)-D-Glc	1364.50
L68	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-D-GlcNAc-OCH_2COOH	733.26
L69	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcNAc-(1→3)-$O(CH_2)_2N_3$	963.37
L70	β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-β-D-GalNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-$O(CH_2)_2N_3$	1109.42
L71	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-D-Gal	982.35
L72	β-D-Gal-(1→4)-[α-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-$O(CH_2)_6NH_2$	993.44
L73	α-D-Gal-(1→3)-β-D-Gal-(1→4)-[α-Fuc-(1→3)]-β-D-GlcNAc-	790.36

	$O(CH_2)_6NH_2$	
L74	β-D-Gal-(1→4)-[α-Fuc-(1→3)]-β-D-GlcNAc-$O(CH_2)_6NH_2$	628.31
L75	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-Fuc-(1→3)]-β-D-GlcNAc-$O(CH_2)_6NH_2$	993.44
L76	β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-$O(CH_2)_6NH_2$	1139.50
L77	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-$O(CH_2)_6NH_2$	831.38
L78	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-$O(CH_2)_2NH_2$	775.32
L79	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc-$O(CH_2)_6N_3$	670.29
L80	β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glc-$O(CH_2)_6NH_2$	806.35
L81	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-$O(CH_2)_6NH_2$	685.33
L82	β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-$O(CH_2)_2N_3$	817.31
L83	α-D-GlcNAc-(1→3)-D-Glc	383.14
L84	α-D-Tal-(1→3)-D-Glc	342.12
L85	β-D-Gal-(1→3)-β-D-$GlcNH_2$ x AcOH-OCH_3	415.17
L86	β-D-Gal-(1→4)-β-D-Glc-OCH_3	356.13
L87	β-D-Gal-(1→4)-α-D-Glc-OCH_2COOH	400.12
L88	β-D-GlcNAc-(1→4)-β-D-GlcNAc-OCH_2CCl_3	554.08
L89	α-D-Ara*f*-(1→5)-(2,3)-anhydro-α-D-Ara*f*-$O(CH_2)_7CH_3$	376.21
L90	3,6-di-OCH_3-β-D-Glc-(1→4)-2,3-di-OCH_3-α-L-Rha-(1→2)-3-OCH_3-α-L-Rha-O(p-OCH_3Ph)	648.30
L91	(6-OCH_3)-β-D-Glc-(1→4)-(2,3-di-OCH_3)-α-L-Rha(1→2)-(3-OCH_3)-α-L-Rha-O(p-OCH_3Ph)	634.28
L92	β-D-Gal*f*-(1→5)-β-D-Gal*f*-(1→6)-β-D-Gal*f*-$O(CH_2)_7CH_3$	616.29
L93	β-D-Gal*f*-(1→6)-β-D-Gal*f*-(1→5)-β-D-Gal*f*-$O(CH_2)_7CH_3$	616.29
L94	α-D-Ara*f*-(1→5)-[β-D-Gal*f*-(1→5)-β-D-Gal*f*-(1→6)]-β-D-Gal*f*-$O(CH_2)_7CH_3$	748.34
L95	β-D-Gal*f*-(1→6)-[α-D-Ara*f*-(1→5)]-β-D-Gal*f*-(1→5)-β-D-Gal*f*-$O(CH_2)_7CH_3$	748.34
L96	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-[α-D-Ara*f*-(1→3)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OCH_3	692.24
L97	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OCH_3	692.24
L98	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OCH_3	824.28
L99	5-SCH_3-α-D-Xyl*f*-(1→4)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→5)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-$O(CH_2)_8NH_2$	1321.51
L100	5-SCH_3-α-D-Xyl*f*-(1→4)-α-D-Man*p*-(1→5)-β-D-Ara*f*-(1→2)-	1261.49

	α-D-Ara*f*-(1→5)-[β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$NH_2$	
L101	5-SCH_3-α-D-Xyl*f*-(1→4)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→5)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$NH_2$	1159.46
L102	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$N_3$	1227.48
L103	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$N_3$	1227.48
L104	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$N_3$	1095.43
L105	α-D-Man*p*-(1→2)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[α-D-Man*p*-(1→2)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$NHCOCF_3	1357.49
L106	α-D-Man*p*-(1→5)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$NHCOCF_3	931.35
L107	α-D-Man*p*-(1→2)-α-D-Man*p*-(1→5)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$NHCOCF_3	1093.40
L108	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$NHCOCF_3	1033.38
L109	α-D-Man*p*-(1→2)-α-D-Man*p*-(1→2)-α-D-Man*p*-(1→5)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$NHCOCF_3	1255.46
L110	α-D-Man*p*-(1→2)-α-D-Man*p*-(1→2)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[α-D-Man*p*-(1→2)-α-D-Man*p*-(1→2)-β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$NHCOCF_3	1681.59
L111	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-[β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$N_3$	1623.60
L112	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-[β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→3)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O(CH_2)$_8$$N_3$	1623.60
L113	α-D-Tal(1→2)-[α-D-Abe-(1→3)]-α-D-Man-OCH_3	486.19
L114	α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-O(CH_2)$_8$COOCH_3	715.33
L115	α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-Glc-O(CH_2)$_8$CONHNH_2	674.31
L116	β-D-GlcNHis-(1→3)-β-D-Gal-(1→4)-β-D-Glc-O(CH_2)$_6$$N_3$	765.34
L117	β-D-GlcNArg-(1→3)-β-D-Gal-(1→4)-β-D-Glc-O(CH_2)$_6$$N_3$	784.38
L118	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-D-Gal	836.29
L119	α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc	924.31
L120	α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc	1215.40
L121	β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)]-	1127.39

	β-D-Gal-(1→4)-D-Glc	
L122	β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	1418.48
L123	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-β-D-Gal-(1→4)-D-Glc	998.34
L124	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	1289.44
L125	α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	1580.53
L126	β-D-Gal-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	795.26
L127	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-α-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	1160.40
L128	β-D-GlcA(2S)-(1→3)- β-D-GalNAc(4S,6S)	707.92
L129	β-D-GlcA(2S)-(1→3)-D-GalNAc(6S)	605.98
L130	β-D-GlcA-(1→3)-D-GalNAc(4S,6S)	605.98
L131	β-D-GlcA-(1→3)-D-GalNAc(6S)	504.04
L132	β-D-GlcA-(1→3)-β-D-GalNAc(4S)-(1→3)-β-D-GlcA-(1→3)-β-D-GalNAc(4S)-$OCH_2CH=CH_2$	1064.11
L133	β-D-Gal-(1→3)-D-GlcNAc	383.14
L134	β-D-Gal-(1→4)-D-GlcNAc	383.14
L135	β-D-Gal-(1→6)-D-GlcNAc	383.14
L136	β-D-Gal-(1→3)-D-GalNAc	383.14
L137	α-D-Gal-(1→3)-D-Gal	342.12
L138	β-D-Gal-(1→4)-D-Gal	342.12
L139	α-D-Gal-(1→3)-β-D-Gal-(1→4)-D-GlcNAc	545.20
L140	α-D-Gal-(1→3)-β-D-Gal-(1→4)-α-D-Gal-(1→3)-D-Gal	666.22
L141	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)-D-Glc	1289.44
L142	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	998.34
L143	α-D-Neu5Ac-(2→6)-[β-D-Gal-(1→3)]-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	998.34
L144	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	836.29
L145	β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	1289.44
L146	β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)-α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc	1580.53

a. *f* : furanose ring; *p* : pyranose ring. Oligosaccharide residues in pyranose form unless otherwise indicated.

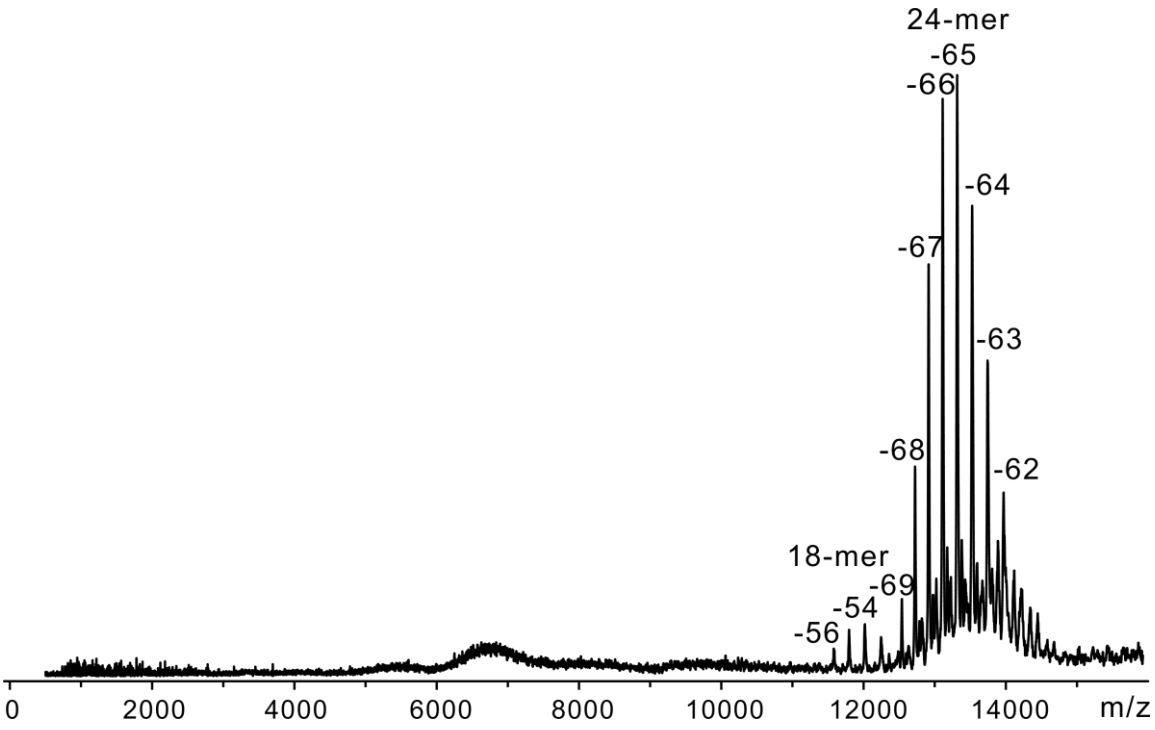

Figure S1. ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) of P particle (5 μM) under mild CID conditions (80 V Trap voltage).

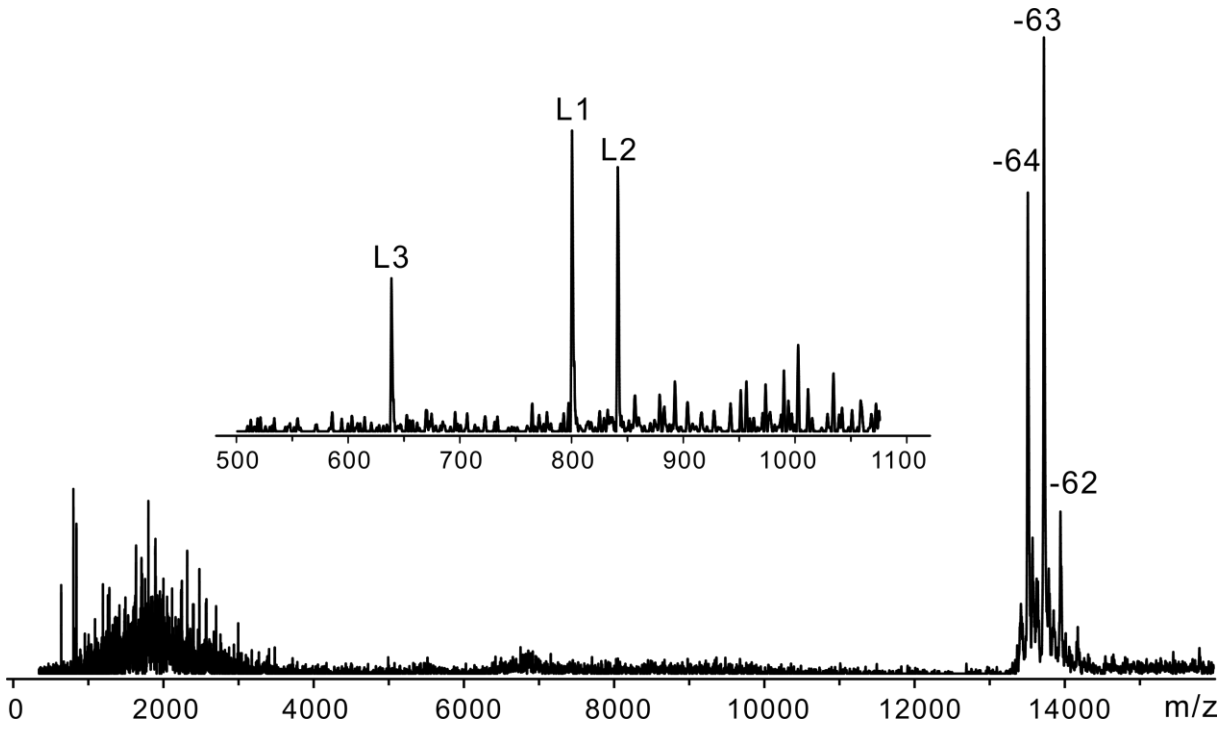

Figure S2. CID mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) of P particle (5 μM) and **L1**, **L2** and **L3** (10 μM each) using a broad (m/z 200 units) quadrupole isolation window centered at m/z 13650. A Trap voltage of 180 V was used.

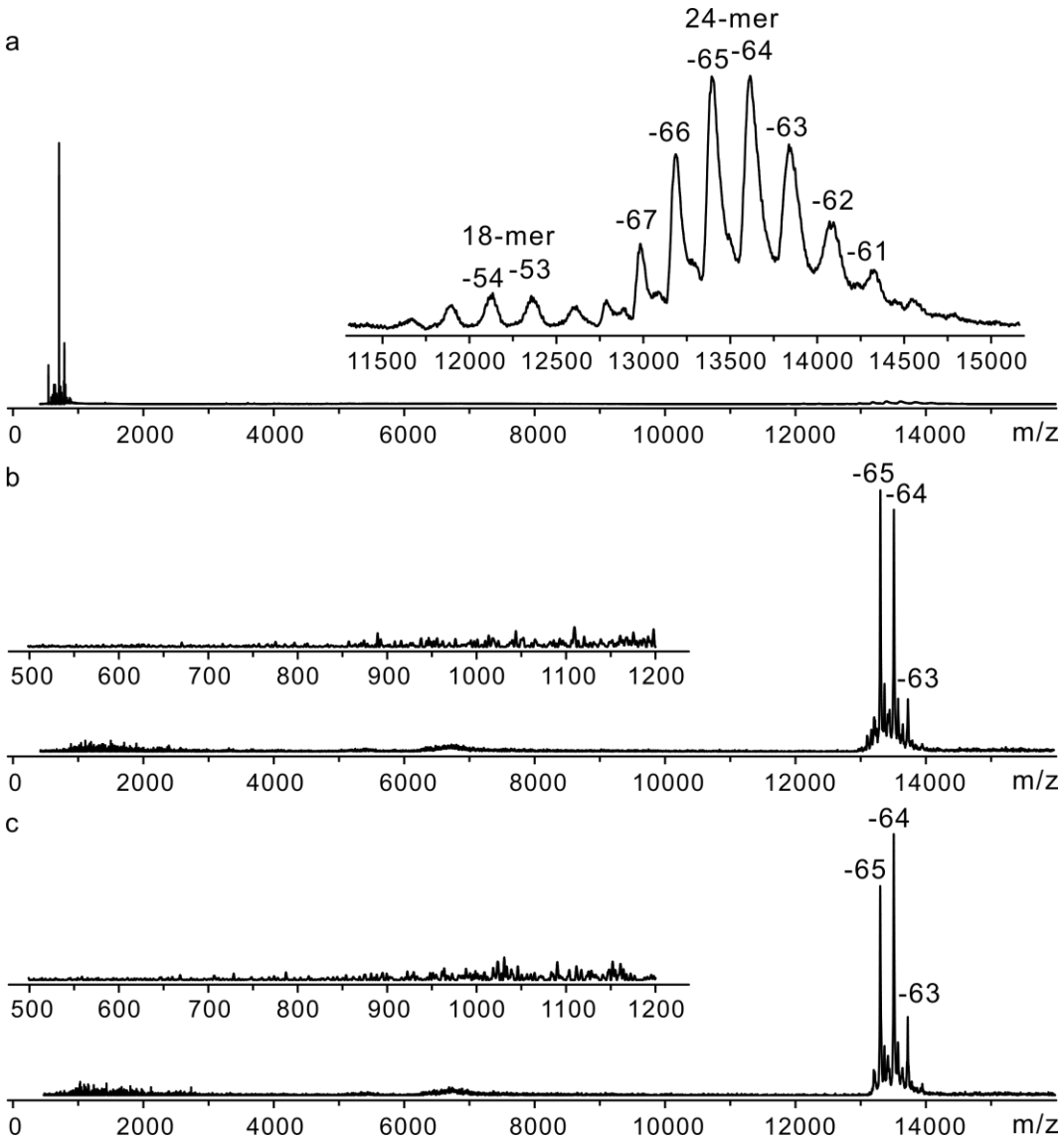

Figure S3. (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) of P particle (5 μM) and **L20** (10 μM). (b) CID mass spectrum acquired for solution in part (a) using a broad (m/z 200 units) quadrupole isolation window centered at m/z 13420. (c) CID mass spectrum acquired for an aqueous ammonium acetate solution (200 mM) of P particle (5 μM) and **L20** (10 μM) using a broad (m/z 200 units) quadrupole isolation window centered at m/z 13420. For (b) and (c), a Trap voltage of 150 V was used.

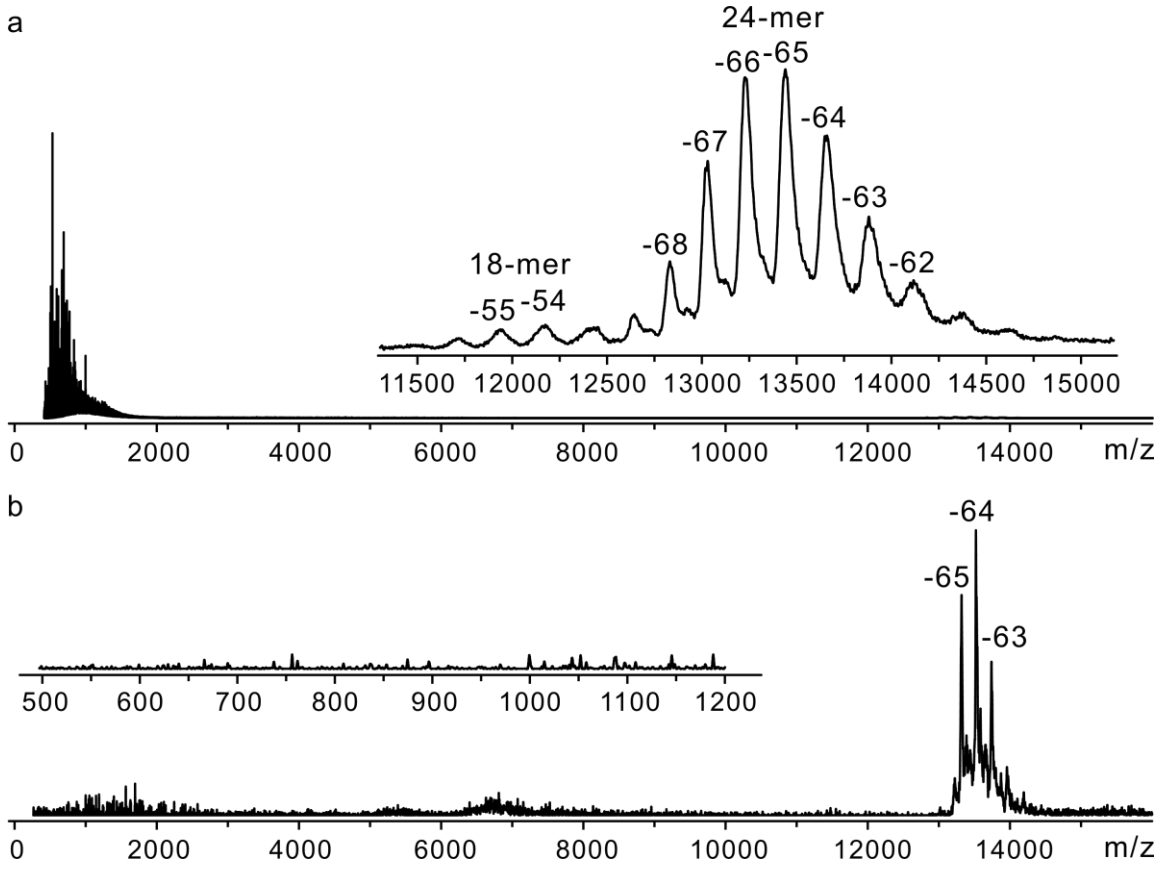

Figure S4. (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM) of P particle (5 μM) and **L14**, **L15**, **L16**, **L17** and **L18** (20 μM each). (b) CID mass spectrum acquired for solution in part (a) using a broad (m/z 200 units) quadrupole isolation window centered at m/z 13420 and a Trap voltage of 150 V.