Screening Carbohydrate Libraries for Protein Interactions using the Direct ESI-MS Assay. Applications to Libraries of Unknown Concentration

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

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

[*] Current address: Faculty of Pharmacy, University of Alexandria, Egypt

Abstract

A semi-quantitative electrospray ionization mass spectrometry (ESI-MS) binding assay suitable for analyzing mixtures of oligosaccharides, at unknown concentrations, for interactions with target proteins is described. The assay relies on the differences in the ratio of the relative abundances of the ligand-bound and free protein ions measured by ESI-MS at two or more initial protein concentrations to distinguish low affinity ($\leq 10^3$ M^{-1}) ligands from moderate and high affinity ($>10^5$ M^{-1}) ligands present in the library and to rank their affinities. Control experiments were performed on solutions of a single chain antibody and a mixture of synthetic oligosaccharides, with known affinities, in the absence and presence of a forty-component carbohydrate library to demonstrate the implementation and reliability of the assay. The application of the assay for screening natural libraries of carbohydrates against proteins is also demonstrated using mixtures of human milk oligosaccharides, isolated from breast milk, and fragments of a bacterial toxin and human galectin 3.

Introduction

It is well recognized that carbohydrate-protein interactions play vital roles in a wide range of physiological and pathological cellular processes, e.g., inflammation, cell-cell interactions, signal transduction, fertility, bacterial and viral infections and the immune response [1,2]. It is, therefore, important to identify and characterize carbohydrate-protein interactions in order to fully understand cellular processes. This understanding also helps to guide the discovery and design of new drugs to treat a variety of diseases and infections [3,4]. There are a number of analytical methods available to detect carbohydrate-protein interactions *in vitro* and to quantify their kinetic and thermodynamic parameters, e.g. isothermal titration microcalorimetry, surface plasmon resonance spectroscopy, frontal affinity chromatography and enzyme linked immunosorbent assays [5]. Glycan microarrays are now routinely used to screen oligosaccharide libraries against proteins and their use has dramatically accelerated the discovery of new and biologically important carbohydrate-protein interactions [3]. However, it is known that glycan array screening may produce false negatives, particularly for low affinity interactions, is sensitive to the nature of the linker used to immobilize the glycans and does not provide a quantitative measure of the affinities of the identified interactions [6,7].

Recently, electrospray ionization mass spectrometry (ESI-MS) has emerged as a promising tool for identifying carbohydrate-protein interactions in solution and quantifying their affinities [8]. The direct ESI-MS assay is gaining popularity as a convenient and versatile method for quantifying the association constants (K_a) of carbohydrate-protein complexes, as well as for other types of protein-ligand interactions

[9-19]. The determination of K_a values for a given protein-ligand (PL) interaction using the direct ESI-MS assay is based on the ratio (*R*) of total abundance (*Ab*) of ligand-bound and free protein ions measured for solutions of known initial concentrations of protein ($[P]_o$) and ligand ($[L]_o$). For a 1:1 PL complex (eq 1), K_a is calculated using eq 2:

$$P + L \rightleftharpoons PL \tag{1}$$

$$K_a = \frac{R}{[L]_o - \frac{R}{1+R}[P]_o} \tag{2}$$

where *R* is given by eq 3:

$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]}{[P]} \tag{3}$$

An assumption underlying the implementation of this method is that PL and P have similar ionization and detection efficiencies (i.e. similar ESI-MS response factors) such that gas-phase abundance ratio is equal to the concentration ratio in solution [8, 9].

The direct ESI-MS assay can also be used to measure multiple binding equilibria, simultaneously, and is uniquely suited to quantify stepwise ligand binding [19] and the binding of multiple ligands, with distinct molecular weights [9,20]. Implemented in a catch-and-release (CaR) format, direct ESI-MS analysis allows for the rapid screening of libraries of hundreds of soluble oligosaccharides against target proteins [20-24]. The CaR-ESI-MS assay involves incubating the target protein with the library, followed by direct ESI-MS analysis to detect protein binding to the highest affinity ligands. In cases where the mass of the complex can't be accurately determined or when dealing with isomeric ligands, bound ligands are released, as ions, from the protein using collisional activation, followed by accurate mass measurement, alone or in combination with

fragmentation or ion mobility separation of the released ligands for ligand identification [20-22]. The application of the CaR-ESI-MS assay to carbohydrate library screening has been demonstrated for libraries of synthetic oligosaccharides, as well as mixtures of oligosaccharides extracted from natural sources [20-24].

While the direct ESI-MS and CaR-ESI-MS assays have proven to be well suited for identifying carbohydrate-protein interactions *in vitro*, their use in quantifying affinities requires that the protein and ligand concentrations are known. This requirement hinders their application to mixtures of carbohydrates extracted or derived from natural sources, such as human milk, for which the concentrations of the individual components are generally not known. Here, we describe a semi-quantitative ESI-MS binding assay suitable for analyzing mixtures of oligosaccharides at unknown concentrations. The assay relies on the differences in the ratio of the abundances of the ligand-bound and free protein ions measured by ESI-MS at two or more initial protein concentrations to distinguish low affinity ($\leq 10^3$ M^{-1}) ligands from moderate and high affinity ($>10^5$ M^{-1}) ligands present in the library and to rank ligand affinities. Control experiments were performed on solutions of a single chain antibody and a mixture of synthetic oligosaccharides, with known affinities, in the absence and presence of a carbohydrate library to demonstrate the implementation and reliability of the assay. The application of the assay for screening natural libraries of carbohydrates against proteins was also demonstrated using mixtures of human milk oligosaccharides (HMOs), isolated from breast milk, and fragments of bacterial and human proteins.

Experimental

Proteins and ligands

A single chain fragment (scFv, MW 26 539 Da) of the monoclonal antibody (mAb) Se155-4 was produced and purified, as described previously [25]. A fragment of the carboxy-terminus of the toxin TcdB from *Clostridium difficile* strain 630 (TcdB-B3, MW 30 241 Da) was a gift of Prof. K. Ng (University of Calgary). Human galectin 3 C-terminal fragment (Gal-3C, MW 16 325 Da) was a gift from Prof. C. Cairo (University of Alberta). α-Lactalbumin (MW 14 210 Da), which served as a reference protein (P_{ref}), was purchased from Sigma–Aldrich Canada (Oakville, ON). Each protein was concentrated and dialyzed against aqueous 100 mM ammonium acetate using microconcentrators (Millipore Corp., Bedford, MA) with a MW cut-off of 10 kDa and stored at –20 ^{o}C, if not used immediately. The carbohydrates α-D-Abe-(1→3)-2-*O*-CH_3-α-D-Man*p*-(1→3)-α-D-Glc*p*-(1→4)-β-D-Glc*p*-OCH_3 (**L1**), β-D-Glc*p*-(1→2)-[α-D-Abe*p*-(1→3)]-α-D-Man*p*-OCH_3 (**L2**) and α-D-Abe*p*-(1→3)-α-D-Tal*p*-OCH_3 (**L3**) were gifts from Prof. D. Bundle (University of Alberta). Stock solutions of individual carbohydrates were prepared by dissolving a known mass of each solid compound in ultrafiltered water (Milli-Q, Millipore) to give a final concentration of 1 mM. The stock solutions were stored at -20 ^{o}C until needed. A forty-component library (*Library 1*) consisting of human blood antigen oligosaccharides, HMOs and bacterial and plant oligosaccharides was prepared. The composition of the library is given in Table S1, Supplementary Information. Two other libraries (*Library 2* and *Library 3*), consisting of mixtures of HMOs extracted from pooled human milk, were prepared using a procedure described previously [26]. As described below, the composition of *Library 2* and *Library 3* (Table S2 and S3, Supplementary Information) was assessed from the molecular weights of the major

components identified by ESI-MS analysis and previously determined HMO structures [27-30].

Mass spectrometry

All carbohydrate-protein binding measurements were carried out in positive ion mode using a Synapt G2 quadrupole-ion mobility separation-time of flight (Q-IMS-TOF) mass spectrometer (Waters, UK) equipped with nanoflow ESI (nanoESI) source. Cesium iodide (concentration 30 ng μL^{-1}) was used for calibration. NanoESI tips were produced from borosilicate tubes (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). A platinum wire was inserted into the nanoESI tip, and a ~1.0 kV voltage was applied to the wire to carry out ESI. A cone voltage of 5 V was used and the source block temperature was maintained at 60 ºC. Injection voltages into the Trap and Transfer ion guides were 3 V and 0 V, respectively. Argon was used in the Trap and Transfer ion guides at pressures of 2.22 x 10^{-2} mbar and 3.36 x 10^{-2} mbar, respectively. The helium chamber preceding the traveling wave ion mobility separation device was maintained at 7.72 mbar. Data acquisition and processing were carried out using MassLynx (v 4.1).

ESI-MS analysis of *Library 2* and *Library 3* was carried out in negative ion mode using the Synapt G2 mass spectrometer. Representative ESI mass spectra acquired for the two libraries are shown in Figures S1 and S2 (Supplementary Information). For *Library 2*, ion signals corresponding to at least fifteen distinct molecular weights were measured, while for *Library 3*, signals for four different molecular weights were measured. The measured molecular weights and proposed composition of the HMOs, tested in screening experiments, are listed in Tables S2 and S3 (Supplementary Information).

Results and discussion

Estimating K_a from ESI-MS data

Generally, the initial ligand and protein concentrations must be known in order to extract reliable K_a values from ESI-MS binding data [8,9]. It is, nevertheless, possible to distinguish between low affinity ($\leq 10^3$ M^{-1}) and moderate and high affinity ($>10^5$ M^{-1}) ligands and to rank ligand affinities from ESI-MS data acquired for ligands of unknown concentration. The approach is based on the differential dependence of *R* on protein concentration for ligands with different affinities. Shown in Figure 1 are theoretical plots of *R* versus $[P]_o$ for a 1:1 PL complex with a K_a of 1.4 x 10^5 M^{-1}, 9.5 x 10^4 M^{-1}, 8.0 x 10^4 M^{-1}, 4.0 x 10^4 M^{-1}, 2.5 x 10^4 M^{-1} and 5 x 10^3 M^{-1}. The *R* values were calculated using eq 4 and $[L]_o$ of 50 μM.

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

For $[P]_o$ values in the μM range (which is typical for ESI-MS binding measurements), the selected K_a values produce *R* values that can be accurately measured (i.e., are in the ~0.2 to ~10 range). Similar plots can be constructed for other values of K_a. Inspection of Figure 1 reveals that there is a dramatic difference in the dependence *R* with $[P]_o$ depending on the magnitude of K_a. For example, for a K_a of 5 x 10^3 M^{-1} (i.e., a low affinity interaction), *R* decreases by only ~0.04 (from 0.25 to 0.21) as $[P]_o$ increases from 5 to 50 μM. In contrast, for a K_a of 1.4 x 10^5 M^{-1} (a relatively high carbohydrate affinity), *R* decreases by a factor of three, from 6.4 to 2.1. For K_a of 2.5 x 10^4 M^{-1} (a moderate affinity) *R* changes from 1.2 to 0.7. It follows that, at this value of $[L]_o$, low, moderate

and high affinity carbohydrate-interactions can be distinguished from the dependence of the *R* value on $[P]_o$.

To demonstrate that this approach can be applied to a wide range of $[L]_o$ values, theoretical *R* values were calculated for a range of ligand affinities for $[L]_o$ values ranging from 2 to 100 μM (Figure 2). The surface corresponding to K_a of 5 x 10^3 M^{-1} is essentially flat (i.e., is parallel to $[P]_o - [L]_o$ plane), indicating that the *R* values do not change significantly with protein concentration. In contrast, the surface corresponding to K_a of 1.4 x 10^5 M^{-1} exhibits a significant dependence on protein concentration, with *R* decreasing dramatically with increasing $[P]_o$ at the higher $[L]_o$ values considered. Based on this analysis it is concluded that moderate and high affinity ligands can be distinguished from low affinity ligands based on the magnitude of the changes in their corresponding *R* values with protein concentration (known).

Distinguishing low affinity from moderate/high affinity carbohydrate ligands

The interactions between the Se155-4 scFv and three oligosaccharide ligands, **L1**, **L2** and **L3**, served as a model to demonstrate the implementation of the assay. The affinities of the three ligands for scFv ((1.6 ± 0.1) x 10^5 (**L1**) [20], (5.0 ± 1.0) x 10^3 (**L2**), and (1.4 ± 0.1) x 10^4 M^{-1} (**L3**)) in aqueous ammonium acetate (50 mM) at pH 7 and 25 °C were determined in separate experiments using the direct ESI-MS assay [8,9]. Shown in Figure 3 are representative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions of **L1** (15 μM), **L2** (50 μM), **L3** (50 μM) and scFv at 5 μM (Figure 3a) or 60 μM (Figure 3b). A reference protein (P_{ref}) was added to both solutions in order to correct the mass spectra for the formation of nonspecific carbohydrate-protein complexes during the ESI process [31]. Ion signal corresponding to free (i.e., $scFv^{n+}$) and

ligand-bound scFv (i.e., (scFv + **L1**)$^{n+}$) , (scFv + **L2**)$^{n+}$) and (scFv + **L3**)$^{n+}$), at n = 9 - 11, was detected. Free (i.e., P_{ref}^{n+}) and carbohydrate-bound P_{ref} ions (i.e., (P_{ref} + **L2**)$^{n+}$ and (P_{ref} + **L3**)$^{n+}$ ions at charge states n = 7 and 8 were also detected, indicating the occurrence of nonspecific binding of **L2** and **L3** to scFv. The distributions of free and ligand-bound scFv, as measured by ESI-MS and following correction of the mass spectra for nonspecific binding, are shown in Figures 3c (for $[scFv]_o$ = 5 μM) and 3d (for $[scFv]_o$ = 60 μM). Notably, the abundance ratio *Ab*(scFv + **L1**)/*Ab*(scFv) (i.e., R_{L1}) decreased dramatically, from 4.4 ± 0.1 to 0.4 ± 0.1, as the scFv concentration increased from 5 to 60 μM. In contrast, the *Ab*(scFv + **L2**)/*Ab*(scFv) ratio (R_{L2}) was unchanged, with a value of 0.20 ± 0.05, while the ratio *Ab*(scFv + **L3**)/*Ab*(scFv) (R_{L3}) decreased slightly, from 0.78 ± 0.02 to 0.45 ± 0.01. These results confirm that, for ligands with K_a values in the 10^3 M^{-1} range (e.g. **L2**), the dependence of *R* on protein concentration is negligible, while for moderate and high affinity ligands (**L3** and **L1**, respectively) modest changes in protein concentration lead to measurable changes in *R*.

Analogous ESI-MS measurements experiments were carried on solutions of **L1**, **L2**, scFv and a library of 40 oligosaccharides (*Library 1*) that do not bind specifically to scFv (data not shown). Shown in Figure 4 are representative ESI mass spectra obtained for solutions of *Library 1* (with each library component at 5 μM concentration), **L1** (15 μM), **L2** (50 μM) and scFv at 9 μM (Figure 4a) or 45 μM (Figure 4b). For the 9 μM scFv solution, the R_{L1} and R_{L2} values are 1.6 ± 0.1 and 0.3 ± 0.1, respectively; for the 45 μM scFv solution they are 0.4 ± 0.1 and 0.3 ± 0.1, respectively. The significant decrease in R_{L1} (by factor of 4) is consistent with the relatively high affinity of **L1**, while the invariance in R_{L2} is with the low affinity of the **L2**.

Taken together, the aforementioned results demonstrate that the newly developed assay can be used to screen complex mixtures of oligosaccharides. To further illustrate the implementation of the assay, mixtures of HMOs, extracted from breast milk, were screened against recombinant TcdB-B3 and Gal-3C proteins. It was recently shown that recombinant fragments of TcdB bind to a variety of commercially available neutral and acidic HMOs, with affinities of 10^2 – 10^3 M^{-1} per binding site [26]. Here, *Library 2*, which consists of at least 15 different species, including a number of relatively high molecular weight HMOs, was screened against TcdB-B3. Because neither the exact composition nor the concentration of HMOs in the mixture were known, preliminary ESI-MS binding measurements were carried out to establish an appropriate volume ratio of *Library 2* and TcdB-B3 stock solutions that would produce detectable concentrations of protein-HMO complexes. Shown in Figure 5 are representative ESI mass spectra acquired for solutions containing TcdB-B3 (10 μM or 50 μM), P_{ref} (2 μM) and 4 μL of *Library 2* (total solution volume was 10 μL). This amount of *Library 2* solution was chosen because TcdB-B3 and P_{ref}, as well as their HMO complexes (specific and nonspecific,) could be detected by ESI-MS with reasonable signal-to-noise ratios. Signal corresponding to the protonated ions of free TcdB-B3 (i.e., TcdB-$B3^{n+}$) and HMO-bound TcdB-B3 (i.e., $(TcdB\text{-}B3 + HMO)^{n+}$), at charge states +10 and +11, was measured for both solutions. Signal corresponding to free P_{ref} and HMO-bound P_{ref}, at charge states +6, +7 and +8, was also detected. After correction for nonspecific HMO binding [31], eleven molecular weight "hits" were detected (Table 1), indicating that at least eleven HMOs (**L4 – L14**) in the mixture bind to TcdB-B3. Notably, with the exception of the ligand with molecular weight 1364.5 Da (**L5**, $Hex_4HexNAc_2Fuc_2$), none of the ligands identified

from the present screening measurements correspond to HMOs previously reported to bind to TcdB. Inspection of the *R* values measured at the two TcdB-B3 concentrations investigated reveals that increasing the concentration of TcdB-B3 by a factor of five had little effect on the abundance ratios of the HMO-bound and free TcdB-B3 ions, with the change in *R* within 0.01 in all cases (Table 1). Given the absence of a dependence of the *R* values on protein concentration it can be concluded that TcdB-B3 binding to the HMO ligands present in the mixture is uniformly weak, with affinities in the 10^3 M^{-1} range. These results are in agreement with previously measured K_a values for the interaction of TcdB-B1 fragment with individual HMOs determined by direct ESI-MS measurements [26].

Gal-3C is known to bind to a variety of HMOs including lactose, lactotetraose, lacto*neo*tetraose, lactofucopentaoses 1, 2 and 3 and lacto-N-difucohexaose [32]. Here, *Library 3,* which contains primarily Hex_2NeuAc (**L15**), $Hex_3HexNAc$ (**L16**), $Hex_3HexNAcFuc$ (**L17**) and $Hex_3HexNAcFuc_2$ (**L18**) HMOs, was screened against Gal-3C. Representative mass spectra acquired for solutions of Gal-3C (at concentrations of 3 and 15 μM) and *Library 3* are shown in Figure 6. At both protein concentrations, signal for protonated ions of free Gal-3C and Gal-3C bound to HMO ligands with molecular weights corresponding to **L15** – **L18,** at charge states +8 and +9, was detected. The abundance of the complex of Gal-3C with **L15**, (Hex_2NeuAc) was significantly lower than the abundances of the Gal-3C complexes with **L16** – **L18**. This could reflect a lower concentration of **L15**, compared to **L16** – **L18**, or the lower affinity compared to the other HMOs. To provide insight into the relative affinities, the R_{L15}, R_{L16}, R_{L17} and R_{L18} values (after correction for nonspecific binding) determined at the two different Gal-3C

concentrations were compared (Table S4, Supplementary Information). The data show that *R* values measured for **L16**, **L17** and **L18** decreased with increasing Gal-3C concentration, while *R* values measured for **L15** did not change significantly. These findings suggest that **L15** has a low affinity for the Gal-3C fragment. It is interesting to note that **L15** is likely a mixture of 3´-sialyllactose and 6´-sialyllactose and it was previously shown that 3´-sialyllactose binds weakly to Gal-3C (4×10^3 M^{-1}) while 6´-sialyllactose does not bind [32]. The magnitude of the changes in *R* values for **L16**, **L17** and **L18** (0.11 ± 0.02, 0.21 ± 0.02 and 0.11 ± 0.02, respectively) upon increasing the Gal-3C concentration from 3 to 15 μM are similar to value measured for (scFv + **L3**) complex, which has an affinity of 1.4×10^4 M^{-1}. **L16** is likely a mixture of lactotetraose and lactoneotetraose, while **L17** is a mixture of lactofucopentaoses and **L18** is assumed to correspond to lacto-N-difucohexaose isomers. The reported affinities of lacto-fucopentaoses 1, 2, 3 and for lacto-N-difucohexaose for Gal-3C are in the 10^4 or 10^5 M^{-1} range, respectively [32]. Consequently, the detectable decreases in the *R* values for **L16**, **L17** and **L18** with increasing protein concentration are consistent with the previously reported affinity data. These results provide further evidence that the newly developed assay can distinguish between low and moderate affinity ligands present in carbohydrate mixtures.

Conclusions

In summary, a semi-quantitative ESI-MS binding assay suitable for analyzing mixtures of oligosaccharides, at unknown concentrations, for interactions with target proteins has been developed. The assay relies on the differences in the ratio of the relative abundances of the ligand-bound and free protein ions measured by ESI-MS at two or more initial

protein concentrations to distinguish low affinity ligands from moderate and high affinity ligands present in the library and to rank their affinities. The implementation of the assay was demonstrated using three carbohydrate-binding proteins. Control experiments performed on scFv and oligosaccharide ligands with known affinities, in the absence and presence of a carbohydrate library, demonstrated that the assay can reliably distinguish between low and moderate/high affinity carbohydrate ligands present in mixtures. The application of the assay for screening mixtures of HMOs, isolated directly from breast milk, and fragments of a bacterial toxin and human galectin 3 demonstrated the use of the assay for screening natural libraries of carbohydrates with unknown concentrations.

Acknowledgements

The authors acknowledge the Alberta Glycomics Centre and the Natural Sciences and Engineering Research Council of Canada for funding and Prof. D. Bundle (University of Alberta) for generously providing some of the oligosaccharides used in this study, Prof. K. Ng (University of Calgary) for the gift of TcdB fragment and Prof. C. Cairo (University of Alberta) for providing the fragment of human galectin 3.

References

1. Arnaud, J., Audfray, A., Imberty, A. Binding sugars: from natural lectins to synthetic receptors and engineered neolectins. Chem. Soc. Rev. **42**, 4798-4813 (2013)
2. Moran, A.P., Gupta, A., Joshi, L. Sweet-talk: role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract. Gut **60**, 1412-1425 (2011)
3. Rillahan, C.D., Paulson, J.C. Glycan microarrays for decoding the glycome. Annu. Rev. Biochem. **80**, 797 – 823 (2011)
4. Yu, Y., Mishra, S., Song, X.Z., Lasanajak, Y., Bradley, K.C., Tappert, M.M., Air, G.M., Steinhauer, D.A., Halder, S., Cotmore, S., Tattersall, P., Agbandje-McKenna, M., Cummings, R.D., Smith, D.F. Functional glycomic analysis of human milk glycans reveals the presence of virus receptors and embryonic stem cell biomarkers. J. Biol. Chem. **287**, 44784- 44799 (2012)
5. Recognition of carbohydrates in biological systems. Part A: General procedures. In Methods of enzymology, 362, Lee, Y.C., Lee, R.T. Eds.; Academic Press: Amsterdam, 2003.
6. He, X.G., Gerona-Navarro, G., Jaffrey, S.R. Ligand discovery using small-molecule microarrays. J. Pharmacol. Exp. Ther. **313**, 1-7 (2005)
7. Grant, O.C., Smith, H.M., Firsova, D., Fadda, E., Woods, R.J. Presentation, presentation, presentation! Molecular-level insight into linker effects on glycan array screening data. Glycobiology **24**, 17-25 (2014)

8. 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)
9. Wang, W., Kitova, E.N., Klassen, J.S. Influence of solution and gas phase processes on protein-carbohydrate binding affinities determined by nanoelectrospray fourier transform ion cyclotron resonance mass spectrometry. Anal. Chem. **75**, 4945-4955 (2003)
10. Shoemaker, G.K., Soya, N., Palcic, M.M., Klassen, J.S. Temperature-dependent cooperativity in donor-acceptor substrate binding to the human blood group glycosyltransferases. Glycobiology **18**, 587-592 (2008)
11. Soya, N., Shoemaker, G.K., Palcic, M.M., Klassen, J.S. Comparative Study of Substrate and Product Binding to the Human ABO(H) Blood Group Glycosyltransferases. Glycobiology **19**, 1224-1234 (2009)
12. Rademacher, C., Shoemaker, G.K., Kim, H.-S., Zheng, R.B., Taha, H., Liu, C., Nacario, R.C., Schriemer, D.C., Klassen, J.S., Peters T., Lowary, T.L. Ligand specificity of CS-35, a monoclonal antibody that recognizes mycobacterial lipoarabinomannan: a model system for oligofuranoside - protein recognition. J. Am. Chem. Soc. **129**, 10489-10502 (2007)
13. Jorgensen, T.J.D., Roepstorff, P., Heck, A.J.R. Direct determination of solution binding constants for noncovalent complexes between bacterial cell wall peptide analogues and vancomycin group antibiotics by electrospray ionization mass spectrometry. Anal. Chem. **70**, 4427-4432 (1998).

14. Jecklin, M.C., Touboul, D., Bovet, C., Wortmann, A., Zenobi, R. Which electrospray-based ionization method best reflects protein-ligand interactions found in solution? A comparison of ESI, NanoESI, and ESSI for the determination of dissociation constants with mass spectrometry. J. Am. Soc. Mass Spectrom. **19**, 332-343 (2008)
15. Yu, Y. H., Kirkup, C.E., Pi, N., Leary, J.A. Characterization of noncovalent protein-ligand complexes and associated enzyme intermediates of GlcNAc-6-O-sulfotransferase by electrospray ionization FT-ICR mass spectrometry. J. Am. Soc. Mass Spectrom. **15**, 1400-1407 (2004)
16. Hofstadler, S.A., Sannes-Lowery, K.A. Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. Nat. Rev. Drug Discov. **5**, 585-595 (2006)
17. Ganem, B., Henion, J.D. Going gently into flight: analyzing noncovalent interactions by mass spectrometry. Bioorg. Med. Chem. **11**, 311-314 (2003)
18. Liu, L., Kitova, E.N., Klassen, J.S. Quantifying protein-fatty acid interactions using electrospray ionization mass spectrometry. J. Am. Soc. Mass Spectrom. **22**, 310-318 (2011)
19. Lin, H., Kitova, E.N., Klassen, J.S. Measuring positive cooperativity using the direct ESI-MS assay. Cholera toxin B subunit homopentamer binding to GM1 pentasaccharide. J. Am. Soc. Mass Spectrom. **25**, 104-110 (2014)
20. El-Hawiet, A., Shoemaker, G., 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)

21. El-Hawiet, A., Kitova, E.N., Klassen, J.S. Quantifying protein interactions with isomeric carbohydrate ligands using a catch and release electrospray ionization-mass spectrometry assay. Anal. Chem. **85**, 7637-7644 (2013)
22. Han, L., Kitova, E.N., Tan, M., Jiang, X., Klassen, J.S. Identifying carbohydrate ligands of a norovirus P particle using a catch and release electrospray ionization mass spectrometry assay. J. Am. Soc. Mass Spectrom. **25**, 111-119 (2014)
23. Abzalimov, R.R., Dubin, P. L., Kaltashov, I.A. Glycosaminoglycans as naturally occurring 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)
24. Cederkvist, F., Zamfir, A.D., Bahrke, S., Eijsink, V.G.H., Sørlie, 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)
25. Zdanov, A., Li, Y., Bundle, D. R., Deng, S.-J., MacKenzie, C. R., Narang, S. A., Young, N. M., Cygler, M. Proc. Natl. Acad. Sci. U.S.A. **91** , 6423-6427 (1994)
26. El-Hawiet, A., Kitova, E.N., Kitov, P.I., Eugenio, L., Ng, K.K., Mulvey, G.L., Dingle, T.C., Szpacenko, A., Armstrong, G.D., Klassen, J.S. Binding of *Clostridium difficile* toxins to human milk oligosaccharides. Glycobiology **21**, 1217-1227 (2011)

27. Wu, S., Tao, N., German, J.B., Grimm, R., Lebrilla, C.B. Development of an annotated library of neutral human milk oligosaccharides. J. Proteome Res. **9**, 4138-4151 (2010)

28. Wu, S., Grimm, R., German, J.B., Lebrilla, C.B. Annotation and structural analysis of sialylated human milk oligosaccharides. J. Proteome Res. **10**, 856-868 (2011)

29. De Leoz, M.L., Wu, S., Strum, J.S., Niñonuevo, M.R., Gaerlan, S.C., Mirmiran, M., German, J.B., Mills, D.A., Lebrilla, C.B., Underwood, M.A. A quantitative and comprehensive method to analyze human milk oligosaccharide structures in the urine and feces of infants. Anal. Bioanal. Chem. **405**, 4089-4105 (2013)

30. Ninonuevo, M.R., Park, Y., Yin, H., Zhang, J., Ward, R.E., Clowers, B.H., German, J.B., Freeman, S.L., Killeen, K., Grimm, R., Lebrilla, C.B. A Strategy for annotating the human milk glycome. J. Agric. Food Chem. **54**, 7471-7480 (2006)

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

32. Hirabayashi, J.1., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W.E., Yagi, F., Kasai, K. Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. Biochim. Biophys. Acta **1572**, 232-254 (2002)

Table 1. Values of *R* measured for HMO ligands in *Library 2* from ESI-MS screening measurements performed using two different concentrations of TcdB-B3.[a]

HMO	***R*** ($[TcdB\text{-}B3]_o$ = 10 μM)	***R*** ($[TcdB\text{-}B3]_o$ = 50 μM)
L4	0.03	0.03
L5	0.03	0.02
L6	0.05	0.03
L7	0.04	0.04
L8	0.05	0.04
L9	0.05	0.04
L10	0.05	0.04
L11	0.04	0.03
L12	0.04	0.03
L13	0.03	0.03
L14	0.03	0.02

a. Standard deviation for *R* values was ≤ 0.01 in all cases.

Figure captions

Figure 1. Plot of *R* versus $[P]_o$ calculated for $[L]_o = 50$ μM, K_a of 1.4 x 10^5 M^{-1} (red), 9.5 x 10^4 M^{-1} (light blue), 8.0 x 10^4 M^{-1} (green), 4.0 x 10^4 M^{-1} (purple), 2.5 x 10^4 M^{-1} (orange), and 5 x 10^3 M^{-1} (dark blue). The *R* values represent the ratio of ligand-bound to free protein at equilibrium.

Figure 2. Plot of *R* versus $[P]_o$ and $[L]_o$ calculated for K_a of 5 x 10^3 M^{-1} (blue), 2.5 x 10^4 M^{-1} (pale brown), 4.0 x 10^4 M^{-1} (purple), 8.0 x 10^4 M^{-1} (green), 9.5 x 10^4 M^{-1} (cyan), and 1.4 x 10^5 M^{-1} (red). The *R* values represent the ratio of ligand-bound to free protein at equilibrium.

Figure 3. ESI mass spectra acquired for aqueous ammonium acetate (50 mM) solutions of (a) scFv (5 μM), **L1** (15 μM), **L2** (50 μM), **L3** (50 μM); (b) scFv (60 μM), **L1** (15 μM), **L2** (50 μM), **L3** (50 μM). Lysozyme (P_{ref}) (4 μM (a) or 10 μM (b) was added as P_{ref}. Distributions of the relative abundances of the solution species, calculated from mass spectra shown in (a) and (b), are shown in (c) and (d), correspondingly.

Figure 4. ESI mass spectra acquired for aqueous ammonium acetate (50 mM) solutions of (a) scFv (≡ P) (9 μM), **L1** (15 μM), **L2** (50 μM); (b) scFv (45 μM), **L1** (15 μM), **L2** (50 μM). Lysozyme (P_{ref}) (4 μM (a) or 8 μM (b) was added to correct for nonspecific carbohydrate binding

Figure 5. ESI mass spectra acquired for aqueous ammonium acetate (50 mM) solutions TcdB-B3 (≡ P) ((a) 10 μM, (b) 50 μM) and 4 μL HMO *Library 2*. α-lactalbumin (P_{ref}) (4 μM) was added to the solution to correct for

nonspecific HMO binding. Inset (a) and (b): +11 charge state of free and HMO-bound TcdB-B3 ions.

Figure 6. ESI mass spectra acquired for aqueous ammonium acetate (100 mM) solutions Gal-3C (≡ P) ((a) 3 μM and (b) 15 μM) containing 1 μL of HMO *Library 3*. Ubiquitin (2 μM) was added as P_{ref}.

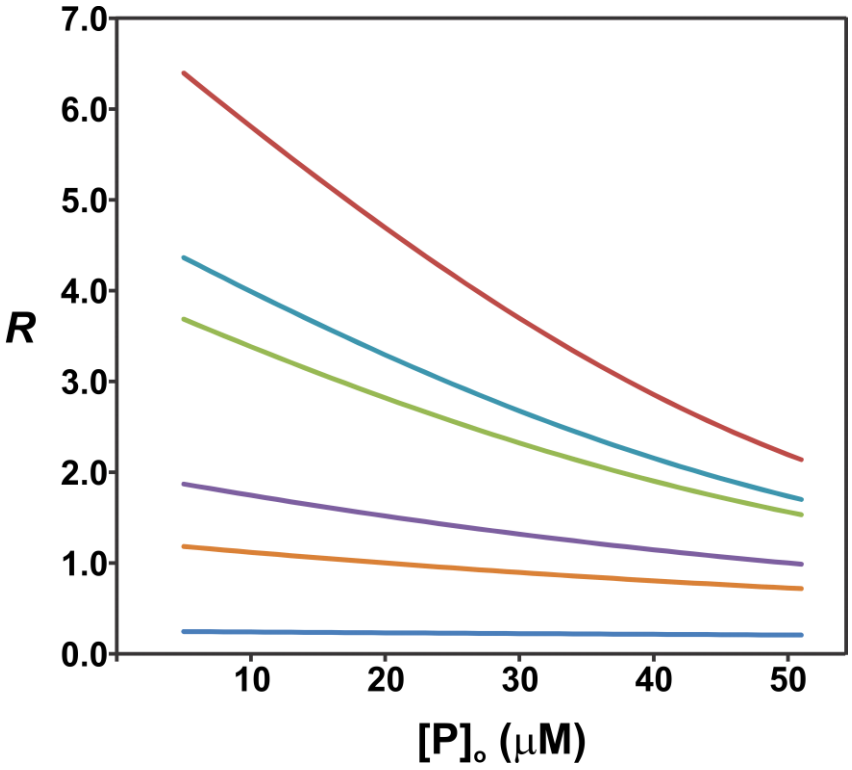


Figure 1

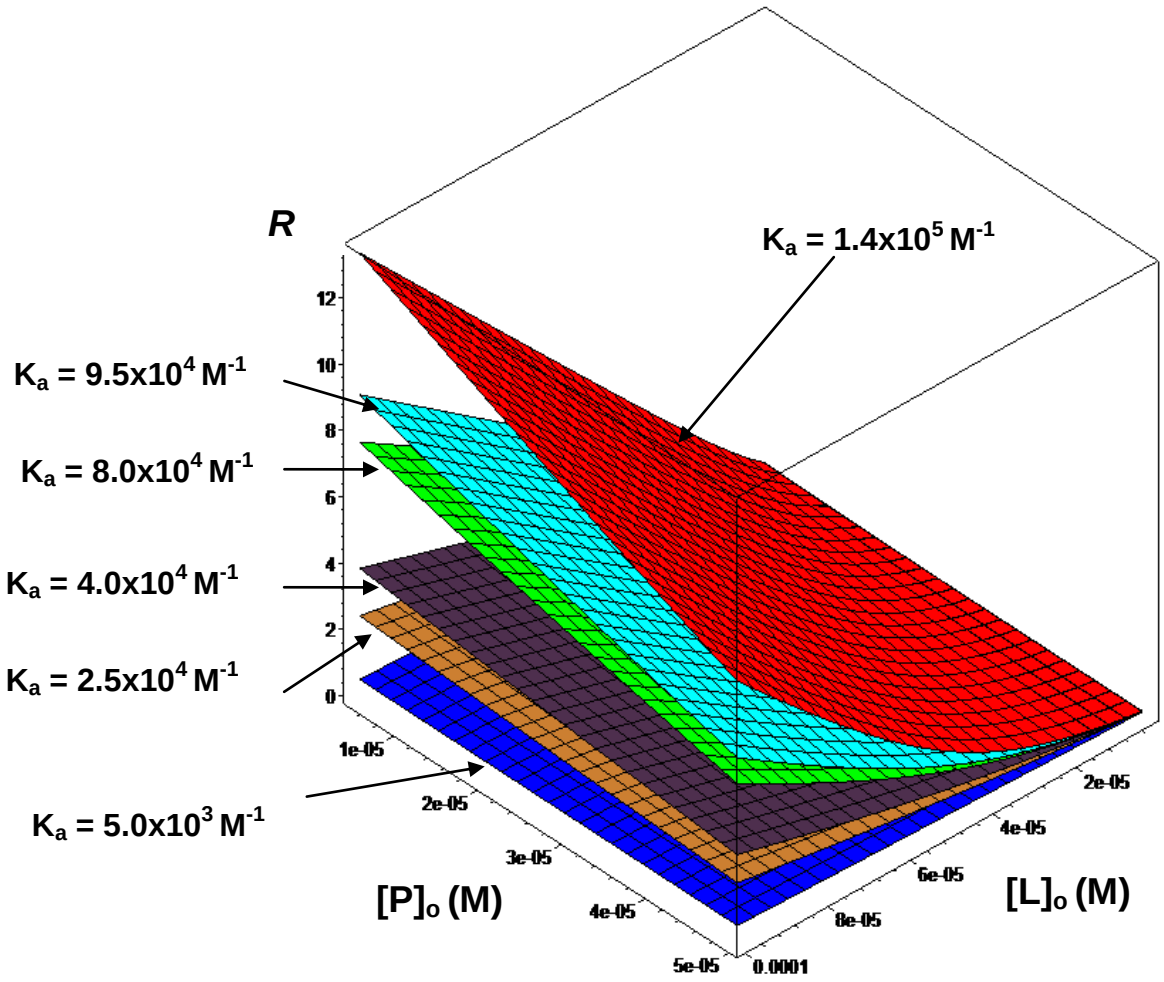

R
$K_a = 1.4x10^5\ M^{-1}$
$K_a = 9.5x10^4\ M^{-1}$
$K_a = 8.0x10^4\ M^{-1}$
$K_a = 4.0x10^4\ M^{-1}$
$K_a = 2.5x10^4\ M^{-1}$
$K_a = 5.0x10^3\ M^{-1}$
12
10
8
6
4
2
0
1e-05
2e-05
3e-05
4e-05
5e-05
$[P]_o$ (M)
0.0001
8e-05
6e-05
4e-05
2e-05
$[L]_o$ (M)

Figure 2

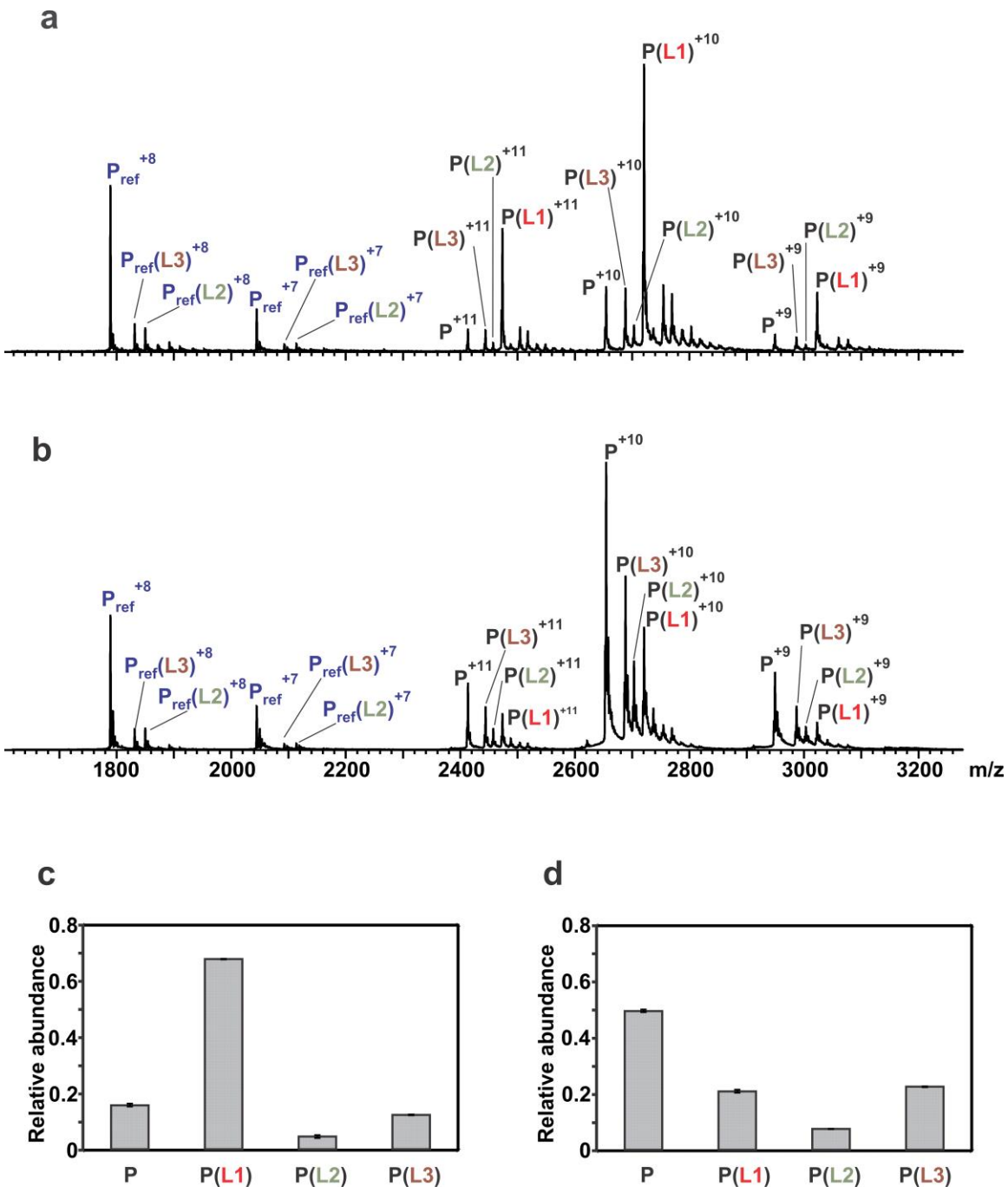
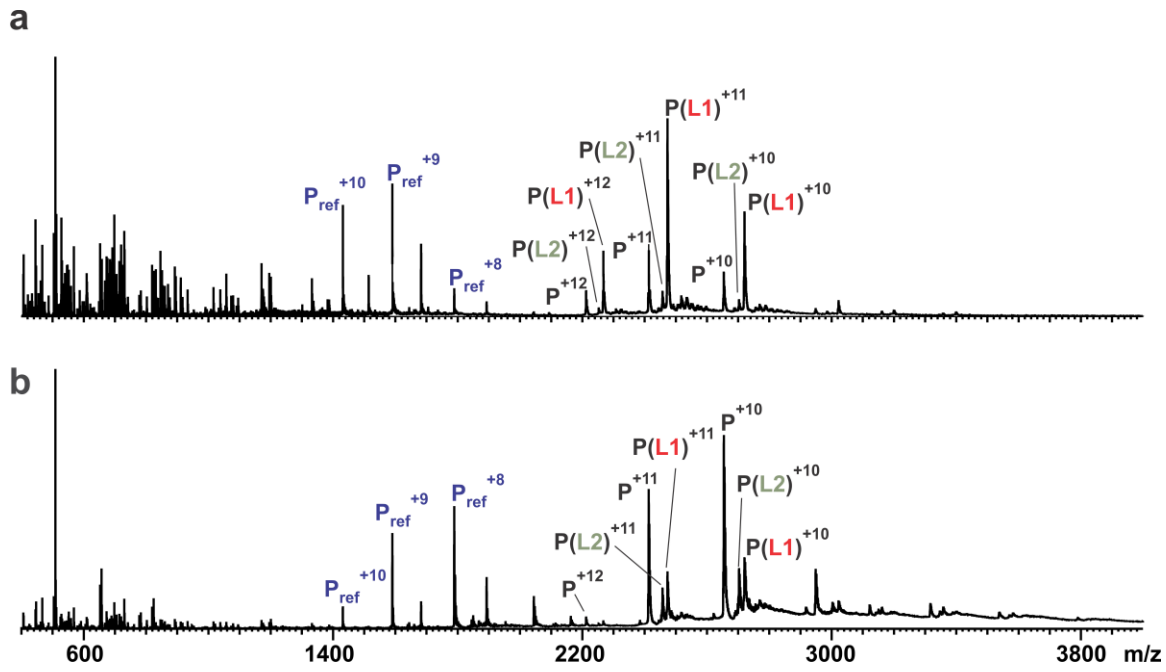

a
b
c
d
Relative abundance
P
P(L1)
P(L2)
P(L3)
m/z
1800
2000
2200
2400
2600
2800
3000
3200
0
0.2
0.4
0.6
0.8

Figure 3


a
P(L1)+11
P(L2)+11
P(L2)+10
P(L1)+12
P(L1)+10
P ref +10
P ref +9
P(L2)+12
P+11
P ref +8
P+10
P+12
b
P+10
P(L1)+11
P+11
P ref +9
P ref +8
P(L2)+10
P(L2)+11
P(L1)+10
P ref +10
P+12
600
1400
2200
3000
3800
m/z

Figures 4

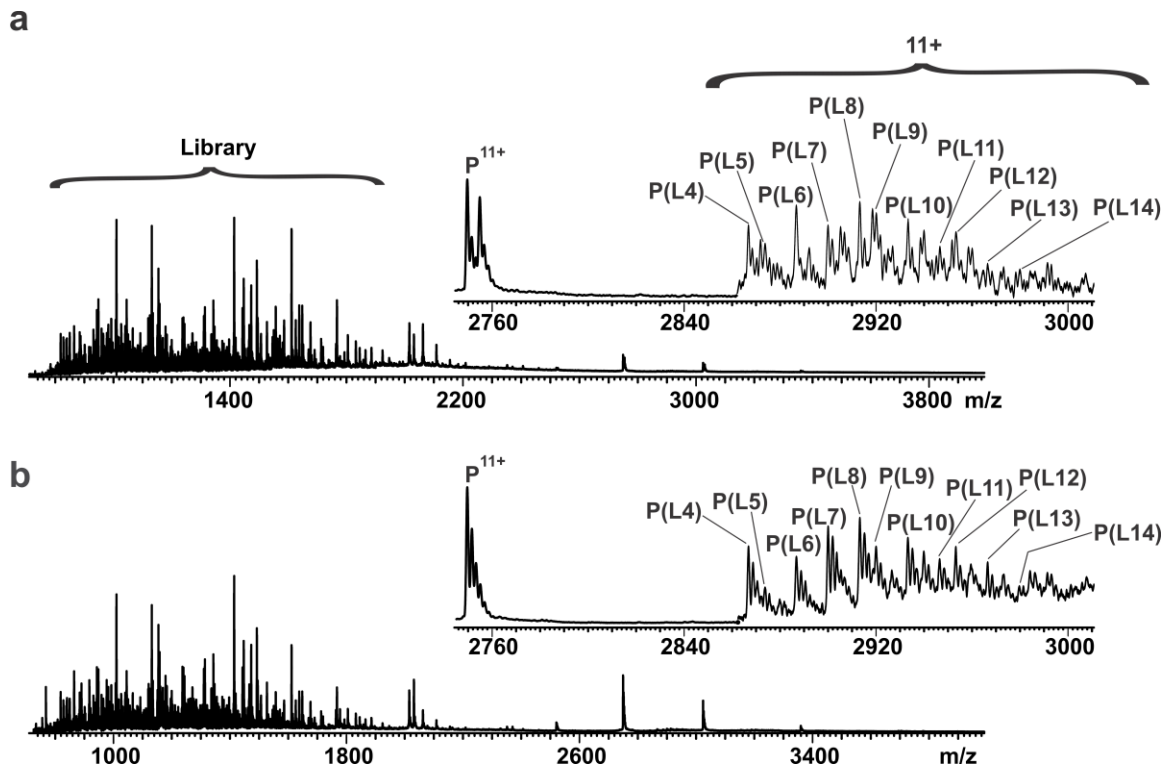

a
11+
Library
P11+
P(L4)
P(L5)
P(L6)
P(L7)
P(L8)
P(L9)
P(L10)
P(L11)
P(L12)
P(L13)
P(L14)
2760
2840
2920
3000
1400
2200
3000
3800 m/z
b
P11+
P(L4)
P(L5)
P(L6)
P(L7)
P(L8)
P(L9)
P(L10)
P(L11)
P(L12)
P(L13)
P(L14)
2760
2840
2920
3000
1000
1800
2600
3400
m/z

Figure 5

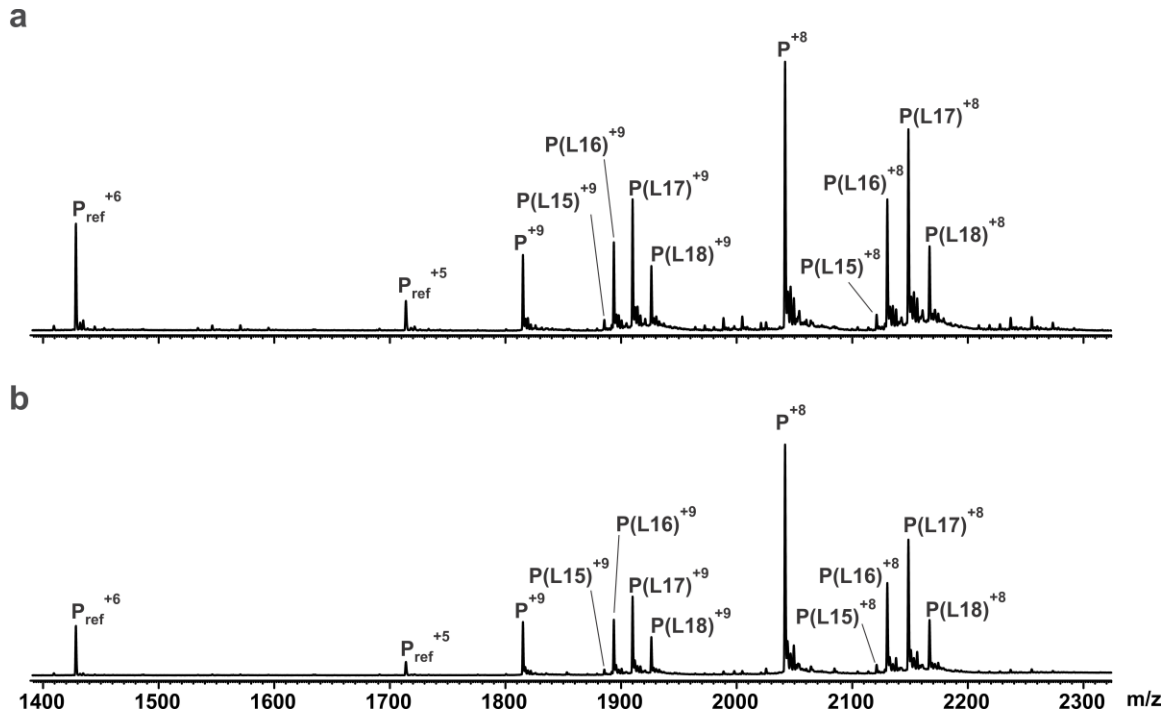

a
b
P+8
P(L17)+8
P(L16)+9
P(L16)+8
P(L15)+9
P(L17)+9
P+9
P(L18)+8
P(L18)+9
P(L15)+8
Pref+6
Pref+5
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
m/z

Figure 6

Supplementary Information for:

Screening Carbohydrate Libraries for Protein Interactions using the Direct ESI-MS Assay. Applications to Libraries of Unknown Concentration

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

Table S1. Structures and molecular weights (MWs) of carbohydrates in *Library 1*.

Carbohydrate	Structure	MW (Da)
1	α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	894.33
2	α-D-GalNAc-(1→3)-β-D-GalNAc-(1→3)-α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc	910.33
3	α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc	924.31
4	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→3)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	998.34
5	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
6	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
7	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
8	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
9	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-Glc-O(CH2)6CH=CH2	760.34
10	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
11	α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→8)-α-D-Neu5Ac-(2→3)- β-D-Gal-(1→4)-D-Glc	1215.40
12	GlcNAc-β–(1-4)-GlcNAc- β-(1-4)-GlcNAc- β-(1-4)-GlcNAc- β-(1-4)-GlcNAc- β-(1-4)-GlcNAc	1237.17
13	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
14	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
15	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
16	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
17	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)-	1437.51

	β-D-Gal-(1→4)-D-Glc	
18	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
19	β-D-Gal-(1→4)-β-D-Glc-OCH_3	356.13
20	α-D-GlcNAc-(1→3)-D-Glc	383.14
21	β-D-Gal-(1→3)-β-D-GlcNH_2 x AcOH-OCH_3	415.17
22	4',6'O-benzylidene-α-D-Glc-(1→4)-α-D-Glc-OCH_3	444.16
23	β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-D-GlcNAc	529.20
24	β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	545.20
25	α-D-GalNAc-(1→3)-β-D-GalNAc-(1→3)-D-Gal	586.22
26	6-*O*-CH_3-β-D-Glc*p*(1-4)-2,3-di-*O*-CH_3-α-D-Rha*p*(1-2)-3-OCH_3-α-D-Rah*p*O*p*OCH_3Ph	620.60
27	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-Glc	633.55
28	α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-Gal-(1→4)-Glc	650.58
29	β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-D-Glc	666.22
30	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-[α-L-Fuc-(1→4)]-D-GlcNAc	675.26
31	α-L-Fuc-(1→2)-β-D-Gal-(1→3)-β-D-GalNAc-(1→3)-D-Gal	691.25
32	β-D-Gal-(1→3)-β-D-GlcNAc-(1→4)-β-D-Gal-(1→4)-D-Glc	707.63
33	α-D-Neu5Ac-(2→3)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-D-Glc	779.27
34	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
35	α-D-Man-(1→6)-[α-D-Man-(1→3)]-α-D-Man-(1→6)-[α-D-Man-(1→3)]-D-Man	828.27
36	α-L-Fuc-(1→2)-β-D-Gal-(1→4)-[α-L-Fuc-(1→3)]-β-D-GlcNAc-(1→3)-D-Gal	837.31
37	α-D-Gal-(1→3)-β-D-Gal-(1→4)-β-D-GlcNAc-(1→3)-β-D-Gal-(1→4)-D-Glc	869.30
38	β-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→3)-[β-D-Ara*f*-(1→2)-α-D-	1095.43

	Ara*f*-(1→5)]-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-$O(CH_2)_8N_3$	
39	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
40	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*-$O(CH_2)_8N_3$	1623.60

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

Table S2. Molecular weights and proposed composition of HMO ligands in *Library 2*.

HMO	m/z of doubly deprotonated HMO ions	MW (experimental) (Da)	MW (theoretical) (Da)	Composition of HMO
L4	643.71	1289.43	1289.44	$Hex_3HexNAcNeuAc_2$
L5	681.24	1364.48	1364.50	$Hex_4HexNAc_2Fuc_2$
L6	753.76	1509.52	1509.53	$Hex_4HexNAc_2FucNeuAc$
L7	826.28	1654.56	1654.57	$Hex_4HexNAc_2Fuc_2NeuAc$
L8	899.31	1800.62	1800.63	$Hex_6HexNAc_4$
L9	936.33	1874.65	1874.67	$Hex_5HexNAc_3FucNeuAc$
L10	1009.35	2020.70	2020.72	$Hex_5HexNAc_3Fuc_2NeuAc$
L11	1082.38	2166.75	2166.78	$Hex_7HexNAc_5$
L12	1119.41	2240.82	2240.82	$Hex_6HexNAc_4FucNeuAc$
L13	1192.43	2386.86	2386.87	$Hex_6HexNAc_4Fuc_2NeuAc$
L14	1264.96	2531.92	2531.91	$Hex_6HexNAc_4Fuc_3NeuAc$

Table S3. Molecular weights and proposed composition of HMO ligands in *Library 3*.

HMO	m/z of singly deprotonated HMO ions	MW (experimental) (Da)	MW (theoretical) (Da)	Composition of HMO
L15	632.20	633.20	633.21	Hex_2NeuAc
L16	706.23	707.23	707.24	$Hex_3HexNAc$
L17	852.29	853.29	853.30	$Hex_3HexNAcFuc$
L18	998.34	999.34	999.36	$Hex_3HexNAcFuc_2$

Table S4. Values of *R* measured for HMO ligands in *Library 3* from ESI-MS screening measurements performed using two different concentrations of Gal–3C.

HMO	***R*** ($[Gal\text{-}3C]_o$ = 3 μM)	***R*** ($[Gal\text{-}3C]_o$ = 15 μM)
L15	0.05 ± 0.02	0.06 ± 0.01
L16	0.59 ± 0.01	0.48 ± 0.01
L17	0.89 ± 0.01	0.68 ± 0.01
L18	0.45 ± 0.01	0.34 ± 0.01

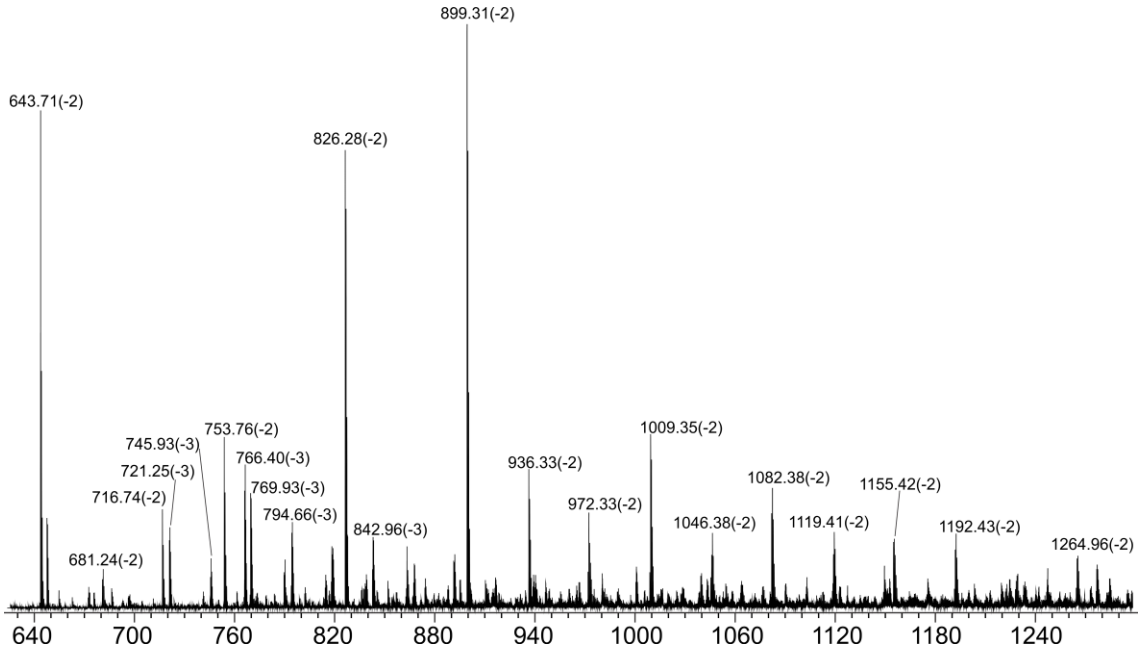


Figure S1. ESI mass spectrum acquired in negative mode for an aqueous solution of *Library 2* (total HMO concentration was 0.1 mg/mL).

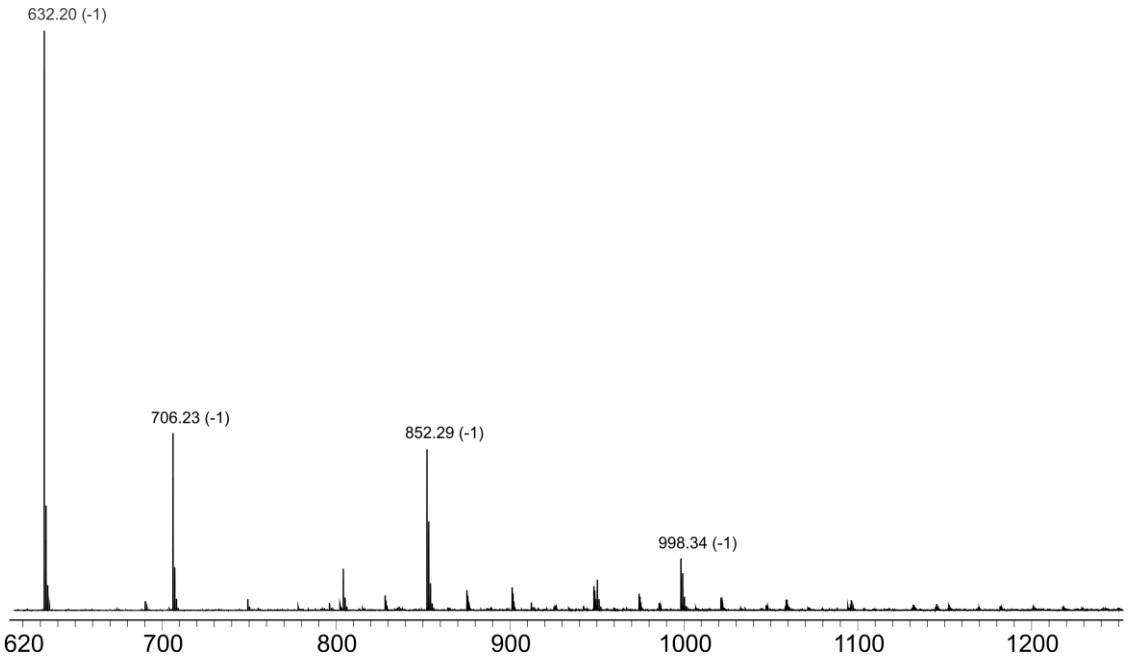


Figure S2. ESI mass spectrum acquired in negative mode for an aqueous solution of *Library 3* (total HMO concentration was 0.1 mg/mL).