Protein-Glycolipid Interactions Studied in vitro using ESI-MS and Nanodiscs.

Insights into the Mechanisms and Energetics of Binding

Ling Han[1], Elena N. Kitova[1], Jun Li[1], Sanaz Nikjah[1], Hong Lin[1], Benjamin Pluvinage[2], Alisdair B. Boraston[2] and John S. Klassen[1]*

[1]*Alberta Glycomics Centre and Department of Chemistry,*

University of Alberta, Edmonton, Alberta, Canada T6G 2G2

[2]*Department of Biochemistry and Microbiology, University of Victoria,*

Victoria, British Columbia, Canada V8W 3P6

*Corresponding Author's address:

Department of Chemistry, University of Alberta

Edmonton, AB CANADA T6G 2G2

Email: john.klassen@ualberta.ca

Telephone: (780) 492-3501

Abstract

Electrospray ionization-mass spectrometry (ESI-MS) analysis combined with the use of nanodiscs (NDs) to solubilize glycolipids (GLs) has recently emerged as a promising analytical method for detecting protein-GL interactions *in vitro* and, when applied to libraries of GLs, ranking their affinities. However, there is uncertainty regarding the mechanism(s) of complex formation in solution and the extent to which the relative abundances of protein-glycolipid complexes observed by ESI-MS reflect the relative concentrations in solution. Here, we describe the results of a systematic ESI-MS study aimed at elucidating the processes that influence binding of water soluble proteins to GLs incorporated into NDs and to exploit these insights to quantify the binding energetics. The interactions between the cholera toxin B subunit homopentamer (CTB_5) and its native ganglioside receptor, β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-β-D-Glc-ceramide (GM1), and between a recombinant fragment of family 51 carbohydrate-binding module (CBM), originating from *S. pneumoniae*, with a synthetic B type 2 neoglycolipid, α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-1,2-di-O-dodecyl-sn-glycero ($B2_{NGL}$) served as model protein-GL complexes for this study. The results of the ESI-MS measurements reveal that proteins bind reversibly to ND-bound GLs and that proteins possessing multiple ligand binding sites are able to interact with GLs originating from different NDs. Experimental evidence suggests that the diffusion of GLs between NDs is rapid and influences the nature of the protein-GL complexes that are detected. Using a newly developed ESI-MS assay, the *proxy ligand* method, the association constants for the CBM-$B2_{NGL}$ and CTB_5-GM1 interactions were quantified and found to be slightly smaller than those for the corresponding oligosaccharides in solution.

Introduction

Glycolipids (GLs) on the surfaces of cells serve a number of important roles. They function as receptors in signaling, pathogen recognition and cellular adhesion processes and convey immunological identity.[1-3] Due to the poor solubility of GL receptors, together with the low affinities that are typical of individual protein-carbohydrate interactions (K_a $<10^4$ M^{-1}),[4,5] the direct quantification of interactions between water-soluble proteins and GL ligands *in vitro* is generally not possible using conventional binding assays, such as isothermal titration calorimetry. Moreover, the structural and functional properties of the receptors may be significantly altered upon removal from a membrane environment.[6,7] Indeed, it is increasingly recognized that protein-GL binding is context dependent (e.g. cell versus model membrane and membrane composition) and is sensitive to GL concentration and fatty acid/ceramide content.[8-10] At present, quantitative binding data are typically obtained using spectroscopy- or microscopy-based measurements and GL that are solubilized by model membranes (e.g. supported lipid bilayer and tethered bilayer lipid membranes and vesicles).[11-15] However, the heterogeneous nature and limited stability of these model membranes make protein-GL interactions difficult to study experimentally and the interpretation of the binding data is not always straightforward.

Recently, the use of nanodiscs (NDs), which are water soluble discoidal phospholipid bilayers, has emerged as a promising method for studying protein interactions with GLs in a lipid environment.[16-18] Glycolipids are readily incorporated into NDs allowing their interactions with water-soluble proteins to be investigated in aqueous solutions using a variety of biophysical methods, including surface plasmon resonance (SPR) spectroscopy,[19] electrospray ionization-mass spectrometry (ESI-MS)[20,21] and silicon photonic sensors.[22] However, while it is possible to detect protein binding to GLs in NDs, interpretation of the binding data is challenging owing to a

lack of mechanistic insights into the association processes. The goal of the present study was to probe, primarily through the use of ESI-MS measurements, the mechanism(s) of protein binding to GLs contained in NDs and to quantify the thermodynamic stabilities of the resulting protein-GL complexes. The interactions between the cholera toxin B subunit homopentamer (CTB_5) and its native ganglioside receptor, β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-β-D-Glc-ceramide (GM1),[23,24] and between a recombinant family 51 carbohydrate binding module (CBM) originating from *S. pneumoniae*, a gram-positive bacterium responsible for a variety of life-threatening diseases including pneumonia, meningitis, and septicemia,[25] with a synthetic B type 2 neoglycolipid, α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-1,2-di-O-dodecyl-sn-glycero ($B2_{NGL}$), served as model protein-GL complexes for this study.

Experimental

Materials and Methods

Proteins

Cholera toxin B subunit homopentamer (CTB_5, molecular weight (MW) 58,040 Da) from *Vibrio cholerae* was purchased from Sigma-Aldrich Canada (Oakville, Canada). A gene fragment encoding a family 51 carbohydrate-binding module (CBM, MW 20,735 Da) was recombinantly produced in *Escherichia coli* and purified as described elsewhere.[26] The ESI-MS analysis of an aqueous solution of CBM (Figure S1a, Supporting Information) revealed the presence of three isoforms (referred to as CBM-I (MW 20,738 ± 2 Da), CBM-II (MW 20,798 ± 5 Da) and CBM-III (MW 20,916 ± 5 Da)). The origin of the structural heterogeneity is unknown, but the MW of the major form of CBM detected (CBM-I) is consistent with the theoretical value (MW 20,735 Da) obtained from the amino acid sequence (Figure S2, Supporting Information). Notably, the

three CBM forms exhibit similar affinities for A and B blood group oligosaccharides (Figure S1b, Supporting Information). Bovine ubiquitin (Ubq, MW 8,565 Da) purchased from Sigma-Aldrich Canada (Oakville, Canada) was used as reference protein (P_{ref}) for the binding measurements.[27] The recombinant membrane scaffold protein (MSP) MSP1E1 (MW 27,494 Da) used for ND preparation was expressed from the plasmid pMSP1E1 (Addgene, Cambridge, MA) and purified using a reported protocol.[28] Saposin A, used for the preparation of the lipoprotein discs (picodiscs), was a gift from Prof. G. Privé (University of Toronto).[29] Stock solutions of CTB_5 and CBM were concentrated and dialyzed into an aqueous 200 mM ammonium acetate solution (pH 6.8) using Amicon 0.5 mL microconcentrator (EMD Millipore, Billerica, MA) with a MW cutoff of 10 kDa. The concentrations of CTB_5 and CBM stock solutions were determined using a Pierce BCA assay kit (Thermo Scientific, Ottawa, Canada) following the manufacturer's instructions, whereas the concentration of Ubq, MSP1E1 and saposin A stock solutions were estimated by UV absorption at 280 nm. All the protein stock solutions were stored at −80 °C until used.

Phospholipids, glycolipids and oligosaccharides

1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC, MW 677.9 Da) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC, MW 760.1 Da) were purchased from Avanti Polar Lipids (Alabaster, AL). The ganglioside GM1, purified from bovine brain, was purchased from Axxora LLC (Farmingdale, NY). Two isoforms of GM1, i.e., d18:1-18:0 (MW 1545.9 Da) and d20:1-18:0 (MW 1573.9 Da), were identified in the GM1 sample. Blood group B type 2 tetrasaccharide neoglycolipid ($B2_{NGL}$, MW 1101.7 Da) and A type 2 tetrasaccharide neoglycolipid ($A2_{NGL}$, MW 1142.7 Da) were purchased from Dextra (Reading, UK). The structures of these phospholipids and GLs are shown in Figure S3 (Supporting Information). The GM1 pentasaccharide ($GM1_{os}$,

MW 998.34 Da) was purchased from Elicityl SA (Crolles, France). The blood group B trisaccharide (B-tri) was a gift from Prof T. Lowary (University of Alberta). The structures of $GM1_{os}$ and B-tri are also included in Figure S3 (Supporting Information). DMPC, POPC, GM1, $B2_{NGL}$ and $A2_{NGL}$ samples were dissolved in HPLC grade methanol/chloroform (1:1 v/v, Thermo Fisher, Ottawa, Canada) to prepare stock solutions of known concentrations. The $GM1_{os}$ and B-tri solid samples were weighed and dissolved in ultrafiltered Milli-Q water (EMD Millipore, Billerica, MA) to yield a stock solution at 1 mM concentration. All the stock solutions were stored at −20 °C until needed.

Preparation of nanodiscs and picodiscs

Nanodiscs containing DMPC alone or GL (GM1, $B2_{NGL}$, or $A2_{NGL}$) were prepared based on a protocol developed by Sligar and coworkers[16,17]; picodiscs containing SapA and POPC, alone or with GL, were prepared following a protocol described by Privé and coworkers.[29,30] Detailed descriptions of the procedures can be found in Supporting Information.

Mass spectrometry

All ESI-MS binding measurements were carried out in positive ion mode (unless otherwise indicated) 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. The *direct* ESI-MS assay[31] and the newly developed *proxy ligand* ESI-MS method were used to measure the affinities of the protein-GL interactions. A brief description of the assays is given below. All the ESI solutions were prepared using 200 mM aqueous ammonium acetate buffer (pH 6.8, 25°C) and allowed to equilibrate for 15 min at 25 °C prior to ESI-MS analysis, unless otherwise indicated. Additional details on the instrumental and experimental conditions used and the binding assays are given as Supporting Information.

***Direct* ESI-MS assay.** The *direct* ESI-MS assay was used to quantify protein (P)-ligand (L) binding.[31] The association constant (K_a) for a 1:1 PL complex is determined from the abundance (*Ab*) ratio (*R*) of the PL to P ions measured by ESI-MS, eq 1:

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

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

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

and $[P]_0$ and $[L]_0$ are the initial concentrations of P and L, respectively.

***Proxy ligand* ESI-MS assay.** The *proxy ligand* ESI-MS assay relies on a proxy ligand (L_{proxy}), which binds to P with known affinity ($K_{a,proxy}$) and competes with the GL ligand (L). In cases where P possesses a single ligand binding site, the extent of PL binding can be deduced by monitoring the relative abundance of PL_{proxy} using direct ESI-MS measurements. K_a for L binding to P can be calculated from eq 3:

$$K_a = \frac{1}{([L_{proxy}]_0 - \frac{R_{proxy}}{K_{a,proxy}})(\frac{[L]_0}{R_{proxy}[P]_0 - ([L_{proxy}]_0 - \frac{R_{proxy}}{K_{a,proxy}})(R_{proxy}+1)} - \frac{1}{R_{proxy}})} \quad (3)$$

where R_{proxy} corresponds to the abundance ratio of the PL_{proxy} to P ions measured by ESI-MS and is taken to be equal to the corresponding concentration ratio, eq 4:

$$R_{proxy} = \frac{\sum Ab(PL_{proxy})}{\sum Ab(P)} = \frac{[PL_{proxy}]}{[P]} \quad (4)$$

and $[P]_0$, $[L]_0$ and $[L_{proxy}]_0$ are the initial concentrations of P, L and L_{proxy}, respectively.

Ultracentrifugation and SDS-PAGE

Ultracentrifugation and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were used to analyze the species present in solutions containing proteins and GL NDs. Briefly, CTB_5 and NDs were incubated in a 200 mM ammonium acetate solution (pH 6.8, 25 °C) and placed in a microconcentrator (EMD Millipore, Billerica, MA) with a MW cutoff of 100 kDa and subjected to ultracentrifugation. The supernatant and filtrate were then analyzed by SDS-PAGE. Additional details can be found in Supporting Information.

Results and Discussion

a. CTB_5 binding to GM1 nanodiscs

The binding of GM1 to CTB_5 is one of the most extensively studied protein-glycosphingolipid interactions. CTB_5 can bind up to five molecules of GM1 and, according to crystal structures reported for the complex of CTB_5 with the water-soluble GM1 pentasaccharide ($GM1_{os}$), the β-D-Gal-(1→3)-β-D-GalNAc and α-D-Neu5Ac-(2→3) motifs in each $GM1_{os}$ interact primarily with a single B subunit of CTB_5 through eighteen direct or water mediated H-bonds.[24] The stepwise binding of $GM1_{os}$ to CTB_5 at neutral pH exhibits positive cooperativity, with intrinsic (per binding site) K_a values ranging from 10^6 to 10^7 M^{-1}.[33,34] The CTB_5-GM1 interaction serves as a useful model system for probing various aspects of protein binding to GLs incorporated into NDs. The measured distribution of GM1 bound to the five available CTB_5 binding sites can provide insights into the nature of the binding processes, such as the reversibility of the individual protein-GL interactions and, relatedly, the ability of CTB_5 to sample GM1 ligands from multiple NDs, as well as the diffusion of GM1 both within and between NDs. Moreover, because of the relatively high affinity of the interactions, the extent of GM1 binding can be used to quantify the concentration of available GM1 and, consequently, establish the efficiency of incorporation of GM1 into NDs.

CTB_5-GM1 nanodisc interactions revealed by ESI-MS. ESI-MS binding measurements were performed on solutions of CTB_5 and ND containing GM1 at percentages ranging from 0.5% to 10%; the corresponding average number of GM1 molecules per ND was estimated to be 1 (0.5%) to 20 (10%). Shown in Figure 1 are illustrative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, pH 6.8, 25 °C) containing CTB_5 (3 μM) with 3 μM and 24 μM 0.5% GM1 ND or 0.6 μM and 1.4 μM 10% GM1 ND. Inspection of the mass spectra reveals signal corresponding to the protonated ions of free and GM1-bound CTB_5, i.e., $(CTB_5 + qGM1)^{n+}$ with $q = 0 – 5$ at $n = 14 – 17$. Also shown in Figure 1 are the normalized distributions of $(CTB_5 + qGM1)$ species calculated from the corresponding mass spectra. Illustrative ESI mass spectra and distributions of bound GM1 measured for the 1%, 2.5% and 5% GM1 NDs are given in Figures S4 – S6 (Supporting Information). According to the ESI-MS data, the number of GM1 ligands bound to CTB_5 is sensitive to both the ND concentration, as well as the percentage of GM1 in the ND. For example, at low concentrations (e.g. 0.6 μM) of the 10% GM1 ND, CTB_5 exists predominantly as free protein with trace amounts of CTB_5 bound to between two and five GM1 (Figure 1c), whereas at higher concentrations (e.g. 1.4 μM), CTB_5 is bound predominantly to four and five GM1 (Figure 1d). Similarly, at low concentrations (e.g. 3 μM) for the 0.5% GM1 ND, the unbound form CTB_5 is the most abundant species (Figure 1a); at higher concentrations of ND (e.g. 24 μM), the distribution shifts to higher ligand occupancy, with the majority of CTB_5 bound to four GM1 (Figure 1b).

Notably, the distributions of bound GM1 measured using NDs with different percentages but with the same total concentration of GM1 are, in some cases, substantially different. As an example, CTB_5 is found to be bound predominantly to between three and five GM1 for solutions of 0.5% GM1 (12 μM) and 1% GM1 (6 μM) NDs (Figures S7a and S4a, Supporting

Information). In contrast, for solutions of higher percentage GM1 NDs, which also contain a total GM1 concentration of 12 μM, CTB_5 is found to be primarily in its free form and the fraction of GM1-bound CTB_5 decreases with the increase of GM1 percentage (Figure 1c and Figures S5a and S6a, Supporting Information). The observed differences in the measured distributions are less noticeable at higher GM1 concentrations. For example, for solutions containing GM1 NDs of different GM1 percentages but all with ~20 μM GM1, CTB_5 is found bound to between three to five GM1 in all cases (Figures 1b and 1d and Figures S4b, S5b and S6b, Supporting Information).

Comparing the measured distributions of bound GM1 to those expected based on the reported equilibrium constants for stepwise binding of $GM1_{os}$ to CTB_5 reveals that, under solution conditions that promote extensive GM1 binding (up to four or five GM1), the measured and theoretical distributions are similar, although the extent of GM1 binding measured by ESI-MS is generally less than expected (Figures 1b and 1d and Figures S4b, S5b and S6b, Supporting Information). In contrast, for solutions containing low concentrations of GM1 NDs, there are marked differences between the measured and theoretical distributions. For example, for solutions of 3 μM CTB_5 with 2.1 μM 2.5% GM1 ND, 1.2 μM 5% GM1 ND or 0.6 μM 10% GM1 ND, where the total GM1 concentration is 12 μM, free CTB_5 dominates the ESI mass spectra. However, based on the concentration of GM1 present in solution and the affinities reported for $GM1_{os}$, CTB_5 is expected to be nearly fully bound (Figure 1c and Figures S5a and S6a, Supporting Information). As described in more detail below, the apparent disagreement between the measured and expected distributions for solutions containing low concentrations of GM1 NDs can be explained in terms of differential ESI-MS response factors for free CTB_5 and

the (CTB_5 + *q*GM1) complexes, which are produced by gas-phase dissociation of ND-(CTB_5 + *q*GM1) complexes originating from solution.

Reversibility of CTB_5-GM1 nanodisc interactions. To test the reversibility of the CTB_5 interactions with GM1 contained in the NDs, the influence of adding free CTB_5 to a solution containing CTB_5 and GM1 ND was investigated. Shown in Figure S7a (Supporting Information) is an ESI mass spectrum acquired for a 200 mM ammonium acetate aqueous solution containing 3 μM CTB_5 and 12 μM 0.5% GM1 ND (incubated for 15 min). Under these conditions, CTB_5 is predominantly bound to between three and five GM1. However, upon addition of 3 μM CTB_5 to this solution, free CTB_5, as well as CTB_5 bound to between one and five GM1 are detected (Figure S7b, Supporting Information). This distribution is nearly identical to that observed for a solution initially containing 6 μM CTB_5 and 12 μM 0.5% GM1 ND (Figures S7c and 7d, Supporting Information). These results confirm that the CTB_5 interactions with GM1 (in NDs) in solution are reversible and that GM1 can be readily redistributed among the CTB_5 binding sites.

Release of CTB_5-GM1 complexes from nanodiscs in the gas phase. From the ESI-MS data acquired for the solutions of 0.5%, 1%, 2.5%, 5% and 10% GM1 ND, plots of the fraction (*f*) of occupied CTB_5 binding sites versus GM1 concentration were calculated (Figures S8a–8e, Supporting Information). Although most noticeable for the low % GM1 ND data, all of the plots are sigmoidal in appearance, which, on its own, is suggestive of positive cooperativity, and reach a maximum *f* of between 85% and 94%. Also plotted is the dependence of *f* expected assuming complete (stoichiometric) binding. Notably, the experimental values approach the theoretical values, at least at certain concentrations, indicating that the amount of GM1 incorporated into the NDs does not differ significantly from the value expected based on the molar ratios of GM1 to DMPC used the prepare the NDs. To our knowledge, this is the first experimental evidence that

the incorporation efficiency of GLs, such as GM1, into NDs is close to 100%. For comparison purposes the corresponding plot of *f* versus $GM1_{os}$ concentration measured by ESI-MS for solutions of CTB_5 (3 μM) and $GM1_{os}$ (1 – 60 μM) is also shown (Figure S8f, Supporting Information). Notably, the experimental data for $GM1_{os}$ binding are well described by the theoretical curve, which was calculated using the Homans' binding model[33] and the reported affinities.[34] Moreover, although $GM1_{os}$ binding to CTB_5 exhibits slight positive cooperativity[33,34] the binding isotherm increases nearly linearly with $GM1_{os}$ concentration until the binding sites are saturated, i.e., *f* (>99%). This latter result indicates that all five binding sites of CTB_5 are accessible for binding and that the *f* values <95% observed for GM1 binding are not due to structural effects related to ligand binding sites. Instead, it is proposed that a fraction of GM1 is retained by the ND upon release of the (CTB_5 + *q*GM1) ions in the gas phase, *vide infra*. The former result, the differences in the binding isotherms measured for the GM1 NDs and $GM1_{os}$, suggests that the origin of the apparent cooperative binding is different in the two cases, *vide infra*.

Previously, it was shown that, for solution of CTB_5 (5 μM) and 10% GM1 ND (10 μM), no free CTB_5 could be detected.[21] This finding led to the suggestion that the (CTB_5 + *q*GM1) ions measured by ESI-MS (under gentle sampling conditions) were the result of the kinetically facile dissociation of the (CTB_5 + *q*GM1) complexes from the NDs in the gas phase.[21] Analogous experiments were carried out in the present study to establish whether the (CTB_5 + *q*GM1) complexes present in solutions containing high and low concentrations of low % GM1 ND were associated with the NDs. For the high concentration case, an ammonium acetate solution (200 mM, pH 6.8, 25 °C) of CTB_5 (5 μM) and 0.5% GM1 ND (24 μM) was subjected to ultracentrifugation using a membrane with a 100 kDa MWCO and the filtrate and supernatant

solutions analyzed by SDS-PAGE (Figure S9a, Supporting Information). The results of this analysis failed to reveal the presence of CTB subunit in the filtrate, suggesting that the protein is predominantly bound to ND in solution. ESI-MS measurements were also carried out to identify the species present in the supernatant and filtrate. Notably, ions corresponding to CTB_5 bound to between three and five GM1, as well as MSP dimer, were detected in the supernatant (Figure S10a, Supporting Information). In contrast, no free or GM1-bound CTB_5 ions were detected in the filtrate (Figure S10b, Supporting Information). At lower concentration of GM1 ND (e.g. 3 μM), SDS-PAGE revealed bands corresponding to CTB subunit in both the supernatant and filtrate, similar to the results obtained for solutions of CTB_5 (5 μM) alone or with a ND containing no GM1 (Figure S9, Supporting Information). However, while free CTB_5 and (CTB_5 + *q*GM1) complexes were present in the supernatant (Figure S11a Supporting Information), only free CTB_5 was identified in the filtrate (Figure S11b, Supporting Information). To further confirm that no GM1-bound CTB_5 was present in the filtrate, CID was performed in negative ion mode on all ions with m/z >2500. The CID mass spectrum reveals signal corresponding to CTB subunit monomer and tetramer ions; no ions corresponding deprotonated GM1 were detected (Figure S11c, Supporting Information). Taken together, these results provide compelling evidence that the (CTB_5 + *q*GM1) ions detected by ESI-MS are the results of gas-phase dissociation of the ND complexes, which results in the release of intact (CTB_5 + *q*GM1) complexes.

Experimental support for the incomplete release of CTB_5-bound GM1 from the NDs in the gas phase can be found in the results of CID experiments performed on the ND ions. Shown in Figures S12b and S12c (Supporting Information) are CID mass spectra acquired in negative ion mode for ND ions produced from a 200 mM ammonium acetate solution (pH 6.8, 25 °C)

containing 14 μM 0.5% GM1 ND with and without 3 μM CTB_5, respectively. CID was carried out using an isolation window centered at m/z 11,000, which corresponds to the ND ions. A comparison of the CID mass spectra shows that the abundance ratio of GM1 to DMPC ions decreases after addition of CTB_5, which is consistent with a fraction of GM1 is extracted from ND, forming (CTB_5 + *q*GM1) complex ions. However, deprotonated GM1 ions were found to be released from the ND even in the presence of excess CTB_5.

The present binding data measured for solutions of CTB_5 and GM1 NDs reveal that the distributions of (CTB_5 + *q*GM1) complexes acquired by ESI-MS are sensitive to gas-phase processes. Two key conclusions are: the (CTB_5 + *q*GM1) ions detected by ESI-MS are produced by dissociation of the ND complexes in the gas phase and the dissociation process is not 100% efficient, with a small fraction of GM1 left behind in the NDs. Based on these finding, the apparent cooperative nature of CTB_5 binding to GM1 NDs, as suggested from the curvature in the plots of *f* versus GM1 concentration (Figure S8, Supporting Information), can be attributed to a higher ESI-MS response factor for free CTB_5, compared to the ND-associated (CTB_5 + *q*GM1) complexes, *vide supra*. Furthermore, the apparent inability to saturate the CTB_5 binding sites (i.e., *f* reaches a limiting value of <0.95) is attributed to the incomplete extraction of GM1 from the NDs by CTB_5 in the gas phase.

Mechanism of CTB_5-GM1 nanodisc binding. Although the distributions of (CTB_5 + *q*GM1) complexes measured by ESI-MS are sensitive to gas-phase reactions, the binding data provide new insight into how CTB_5 associates with GM1 in the NDs. Notably, the detection of (CTB_5 + 4GM1) and (CTB_5 + 5GM1) complexes in solutions with low percentage GM1 NDs (i.e., 0.5% and 1% GM1 NDs, which contain an average of 1 and 2 GM1, respectively) is consistent with a

stepwise binding model, in which CTB_5 sequentially binds to GM1 originating from multiple NDs. There are three possible mechanisms that could account for this observation.

i) *ND recruitment mechanism.* One possible mechanism would see CTB_5 binding irreversibly to GM1 from multiple NDs (Figure 2a). However, by overlaying the relative positions of the five ligand binding sites of CTB_5 [24] onto NDs with diameters of ~11 nm,[35,36] it can be concluded that one CTB_5 could bind simultaneously to at most two NDs. Even then, unfavourable steric effects are likely to be significant. Consequently, based on structural considerations it is unlikely that the simultaneous binding of CTB_5 to multiple NDs is responsible for the measured distributions of (CTB_5 + *q*GM1) complexes.

ii) *GL extraction mechanism*. A second possible mechanism would involve CTB_5 interacting with GM1 molecules in one ND, followed by dissociation of an intact (CTB_5 + *q*GM1) complex from the ND and rapid re-binding to GM1 in another ND (Figure 2b). The number of binding steps would depend on the number of GM1 per ND and, in the case of NDs containing a high numbers of GM1 (e.g. ≥5 per ND), CTB_5 would be expected to interact with a single ND. An argument against this mechanism comes from kinetic data measured by SPR spectroscopy for the dissociation of CTB_5 from immobilized NDs containing on average one or two GM1 at 25 °C in HEPES-buffered saline (pH 7.4).[19] Based on the measured rate constant, 0.028 min^{-1}, the lifetime of ND-bound (CTB_5 + *q*GM1) complexes will be >35 min, which is significantly longer than the time scale of the ESI-MS measurements. Moreover, the rate of dissociation from immobilized NDs containing >12 GM1 was too slow to be accurately measured.[19] Although it was not clear from these measurements whether free CTB_5 or (CTB_5 + *q*GM1) complexes were released from the NDs, the kinetic data suggest that the stepwise binding of CTB_5 to different NDs is too slow to account for the measured distributions of (CTB_5 + *q*GM1) complexes. The absence of

detectable amounts of (CTB_5 + *q*GM1) complexes in the filtrate from the ultracentrifugation experiments described above provides additional, although indirect, support for this conclusion.

iii) *GL diffusion mechanism*. A third possible mechanism would proceed through a rapid redistribution of GM1 between NDs such that, upon binding to one ND, CTB_5 can recruit additional GM1 from other NDs (Figure 2c). The exchange kinetics for DMPC between NDs have been quantified using small-angle neutron scattering and fluorescence methods.[37] These measurements, which support a monomeric lipid diffusion mechanism, yielded exchange rate constants (k_{ex}) of 0.0328 min^{-1} and 0.0378 min^{-1} for DMPC exchange at 27 °C and an activation Gibbs energy of 91.8 kJ mol^{-1}.[37] Using an average value of k_{ex} of 0.035 min^{-1}, the lifetime of DMPC in the ND is estimated to be ~29 min at 27 °C.

In an effort to evaluate the rate of exchange of GM1 between NDs, CID measurements were performed in negative ion mode on ions with a narrow range of m/z centred at 11,500 produced by ESI from four different solutions, one with 0.5% $A2_{NGL}$ ND and picodiscs[29,30] containing GM1 and POPC (in a 1:1:4 SapA:GM1:POPC ratio), one with GM1 picodiscs alone (1:1:4 SapA:GM1:POPC ratio), one with 0.5% $A2_{NGL}$ ND alone, and one with 0.5% $A2_{NGL}$ ND and 0.5% GM1 ND (Figure S13, Supporting Information). CID performed on solution of 12 μM 0.5% $A2_{NGL}$ ND and 54 μM GM1 picodisc produced negatively charged DMPC, $A2_{NGL}$ and GM1 ions (Figure S13b, Supporting Information). In contrast, in the absence of the GM1 picodiscs in solution, CID produced only DMPC and $A2_{NGL}$ ions. To rule out the possibility that the GM1 detected in the CID mass spectrum shown in Figure S13b (Supporting Information) originated from picodisc ions, analogous CID measurements were performed on ions with m/z centred at 11,500 produced from solutions of GM1 picodisc or 0.5% $A2_{NGL}$ ND. Notably, no GM1 ions were detected (Figures S13c and S13d, Supporting Information). Interestingly, the relative

abundances of GM1 and $A2_{NGL}$ ions detected in Figure S13b (Supporting Information) are similar to those measured by CID performed on ions (m/z ~11,500) produced from an equimolar mixture of 0.5% $A2_{NGL}$ ND and 0.5% GM1 ND (Figure S13e, Supporting Information). Taken together, these data suggest that GM1 readily transfers from the picodisc to the ND (on the min timescale), leading to NDs that have ~0.5% GM1.

The rapid transfer of GM1 from NDs to picodiscs was also demonstrated. As shown in Figure S14a (Supporting Information), GM1 ions were observed in the CID mass spectrum acquired for ions with m/z ~5,500 produced from an ammonium acetate solution (pH 6.8, 25 °C) of 16 μM 1% GM1 ND and 60 μM picodiscs (containing only POPC). CID was also performed on ions with m/z ~5,500 produced from solutions of either POPC-containing picodiscs (Figure S14b, Supporting Information) or 1% GM1 ND (Figure S14c, Supporting Information). In neither case were deprotonated GM1 ions detected; this finding suggests that the GM1 ions detected in Figure S14a (Supporting Information) arise from the transfer of GM1 from the NDs to the picodiscs. Taken together, these results establish the rapid exchange of GM1 between picodiscs and NDs and lend support to the hypothesis that GM1 diffusion between NDs influences, at least to some extent, the measured distributions of (CTB_5 + *q*GM1) complexes.

b. CBM binding to $B2_{NGL}$ in nanodiscs

The CBM-$B2_{NGL}$ interaction served as a second model system for investigating protein binding to GLs contained in NDs. CBM recognizes type A and B blood group oligosaccharides. Recent studies employing glycan array screening (Consortium for Functional Glycomics, http://www.functionalglycomics.org/), isothermal titration calorimetry (ITC),[26] as well as ESI-MS[38] revealed that CBM exhibits relatively strong binding for A/B trisaccharides and A/B type 2, 5 and 6 oligosaccharides (10^4 to 10^5 M^{-1}). Additionally, the X-ray crystal structure of CBM

bound to the B type 2 tetrasaccharide indicates that CBM possesses a single ligand binding site and forms a network of H-bonds with the α-L-Fuc, α-D-Gal and β-D-Gal residues.[26] The K_a of the histo-blood group B type 2 tetrasaccharide ($B2_{os}$) binding to CBM is reported to be 5×10^4 – 8×10^4 M^{-1}.[26,38]

ESI-MS measurements were performed on solutions of CBM and 2.5% and 10% $B2_{NGL}$ NDs. Shown in Figures 3a and 3b are representative ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate (200 mM, pH 6.8, 25 °C) solutions containing CBM (12 μM) with 8 μM and 31 μM 10% $B2_{NGL}$ ND, respectively. Notably, signal corresponding to both free and $B2_{NGL}$-bound CBM (all three CBM species) was detected, i.e., $(CBM + B2_{NGL})^{n+}$ at n = 8 – 10. Representative mass spectra acquired for solutions of CBM (12 μM) with 2.5% $B2_{NGL}$ NDs are shown in Figure S15 (Supporting Information). Plots of the fraction of ligand-bound CBM versus $B2_{NGL}$ concentration are shown in Figure 3c, along with the expected curve for $B2_{os}$ binding, based on the reported affinity.[38] Fitting eq 1 to the experimental data yields similar affinities, 3200 ± 100 M^{-1} (2.5% $B2_{NGL}$ ND) and 2900 ± 100 M^{-1} (10% $B2_{NGL}$ ND). These values are significantly smaller (by factor of 17 – 18) than the K_a reported for $B2_{os}$.[26,38] While this finding is, on its own, consistent with the reduced protein affinities reported for some surface immobilized glycans,[39] it is likely that measured affinities for $B2_{NGL}$ are influenced by non-uniform ESI response factors for the bound and unbound CBM species, *vide infra*.

To demonstrate that ligand-bound CBM remains associated with the NDs in solution, ultracentrifugation analysis using a membrane filter with a MW cutoff of 100 kDa was carried out on an ammonium acetate solution (pH 6.8, 25 °C) of CBM (12 μM) with 10% $B2_{NGL}$ ND (21 μM). Because CBM cannot be reliably distinguished from the MSP used for the NDs by SDS-PAGE, ESI-MS measurements were carried out to analyze the supernatant and filtrate

solutions. Shown in Figures S16a and 16b (Supporting Information) are mass spectra acquired for the supernatant and filtrate, respectively. It can be seen that free CBM is present in the filtrate, while (CBM + $B2_{NGL}$) is only detected in the supernatant. This result, which is consistent with those obtained for solutions of CTB_5 and GM1 NDs, suggests that $B2_{NGL}$-bound CBM is associated with the ND in solution and that the $(CBM + B2_{NGL})^{n+}$ ions detected by ESI-MS are the result of dissociation of the CBM-$B2_{NGL}$-ND complexes in the gas phase.

c. Protein affinities for glycolipids in nanodiscs – the *proxy ligand* ESI-MS assay

A weakness of the *direct* ESI-MS assay for quantifying protein-GL interactions involving NDs is that the detected protein-GL complexes result from dissociation of the protein-GL-ND complexes in the gas phase. Consequently, any differences in the ESI response factors for the free protein and the protein-GL complex ions will introduce errors into the affinity measurements. Given these limitations, a new ESI-MS binding assay, the *proxy ligand* method, was developed. This assay, which combines *direct* ESI-MS measurements with competitive ligand-protein binding, was used to quantify the affinities of CBM for NDs containing 10% and 15% $B2_{NGL}$. The B-tri ligand, which served as L_{proxy} for these measurements, has an affinity for CBM of 7.3×10^4 M^{-1}.[38] Shown in Figure S1b (Supporting Information) is a representative ESI mass spectrum acquired for the aqueous ammonium acetate solution (200 mM, pH 6.8, 25 °C) of 12 μM CBM and 40 μM B-tri. Ions corresponding to free CBM and CBM bound to B-tri were detected, i.e., CBM^{n+} and $(CBM + B\text{-tri})^{n+}$ at n = 8 to 10. The addition of 24 μM of 15% $B2_{NGL}$ ND to the solution resulted in the appearance of ions corresponding to CBM bound to $B2_{NGL}$, i.e., $(CBM + B2_{NGL})^{n+}$ at n = 8 to 10, also resulted in an increase in the abundance ratio of B-tri-bound to free CBM ions (i.e., R_{proxy}) (Figure 4a). The increase in R_{proxy} is consistent with a decrease in CBM available for binding to B-tri due to the competitive binding to $B2_{NGL}$. Shown

in Figure 4b is a plot of R_{proxy} versus $B2_{NGL}$ concentration. The data were analyzed according to the procedure described in the Experimental section and an affinity of $(1.4 \pm 0.1)\times10^4$ M^{-1} was obtained by fitting eq 3 to the experimental data. Measurements carried out using 10% $B2_{NGL}$ ND yielded an affinity of $(1.1 \pm 0.1)\times10^4$ M^{-1} (Figure 4b and Figure S17, Supporting Information). Notably, the $B2_{NGL}$ affinities measured using the *proxy ligand* ESI-MS assay are consistently higher (by a factor of ~5) than the values obtained by *direct* ESI-MS assays. The lower values measured directly by ESI-MS are attributed to non-uniform response factors for CBM and (CBM + $B2_{NGL}$) species, *vide supra*. That the K_a for the CBM-$B2_{NGL}$ interaction measured by *proxy ligand* method is lower (by a factor of ~5) than the value reported for $B2_{os}$ ($K_a = 5.3\times10^4$ M^{-1}) is also notable. This finding suggests that protein binding to GLs in NDs may be energetically less favorable than the interactions with the corresponding free oligosaccharides in solution.

The *proxy ligand* ESI-MS method was also extended to evaluate the affinities of CTB_5 for GM1 NDs. However, because of the presence of multiple binding sites, the cooperative nature of GM1 binding and the possibility of multivalent binding effects, interpretation of binding data is generally more complicated than in the case of CBM. To minimize the occurrence of multivalent binding, measurements were carried out on solutions of CTB_5 with low concentrations of low percentage (0.5% and 1%) GM1 NDs and high concentrations of $GM1_{os}$, which served as L_{proxy}. Under these conditions, it is expected that CTB_5 will bind preferentially $GM1_{os}$ and will not interact with multiple GM1. Shown in Figure 5a is a representative ESI mass spectrum acquired for an aqueous ammonium acetate solution (200 mM, pH 6.8, 25 °C) of 4.4 μM CTB_5 and 20 μM $GM1_{os}$. Ions corresponding to CTB_5 bound to between two and five $GM1_{os}$ were observed, with the (CTB_5 + $5GM1_{os}$) complex being the most

abundant. The addition of 2.5 μM 0.5% GM1 ND to the solution resulted in the appearance of $(CTB_5 + 4GM1_{os} + GM1)^{n+}$ ions, at n = 15 to 17, (Figure 5b) and a measurable increase of the abundance ratio of the $(CTB_5 + 5GM1_{os})$ to $(CTB_5 + 4GM1_{os})$ ions ($\equiv R_{proxy,5}$), which is consistent with CTB_5 binding to GM1 ND in solution. A plot of $R_{proxy,5}$ versus GM1 (in the ND) concentration is shown in Figure 5d. Using the binding model described in Supporting Information, which is an extension of the Homans' model,[33] the association constants $K_{a,1}$, $K_{a,2}$ and $K_{a,3}$, corresponding to GM1 binding to CTB_5 sites with zero, one or two occupied nearest neighbour subunits, respectively, which gave the closest agreement to the experimentally determined $R_{proxy,5}$ – $K_{a,1} = 2.8\times10^6$ M^{-1}, $K_{a,2} = 4.8\times10^6$ M^{-1} and $K_{a,3} = 8.2\times10^6$ M^{-1}. Shown in Figure 5c is a comparison of the theoretical distribution of bound $GM1_{os}$ and GM1 (calculated using these $K_{a,1}$, $K_{a,2}$ and $K_{a,3}$ values) with the experimentally-determined distribution determined from the mass spectrum in Figure 5b. Notably, there is excellent agreement in the distributions of bound $GM1_{os}$. In contrast, the predicted distribution for bound GM1 does not resemble the experimental distribution. However, this disagreement can be explained in terms of non-uniform ESI-MS response factors for the $(CTB_5 + qGM1_{os})$ and $(CTB_5 + qGM1_{os} + GM1)$ complexes, *vide supra*. Moreover, the concentration dependence of $R_{proxy,5}$ predicted theoretically agrees well with the experimental observations made over a range of concentrations (Figure 5d). Analogous measurements performed using 1% GM1 ND gave a similar affinities - $K_{a,1} = 1.2\times10^6$ M^{-1}, $K_{a,2} = 2.0\times10^6$ M^{-1} and $K_{a,3} = 3.5\times10^6$ M^{-1} (Figure S18, Supporting Information). Notably, the measured affinities are slightly smaller than the value obtained for the corresponding CTB_5-$GM1_{os}$ interactions - $K_{a,proxy,1} = 3.2\times10^6$ M^{-1}, $K_{a,proxy,2} = 5.5\times10^6$ M^{-1} and $K_{a,proxy,3} = 9.5\times10^6$ M^{-1}, a finding consistent with what was found for the CBM and $B2_{NGL}$ interaction.

Conclusions

The present study represents the first detailed investigation into the mechanisms and energetics of protein interactions with GLs in NDs. The results of ESI-MS measurements performed on solutions of CTB_5 and GM1 NDs reveal that proteins bind reversibly to ND-bound GLs and, in the case of proteins with multiple ligand binding sites, are able to interact with GLs originating from different NDs. The results of ESI-MS measurements performed on solutions of NDs and picodiscs provide direct evidence for rapid GL diffusion between picodiscs and NDs. Based on this finding it is proposed that diffusion of GLs between NDs influences the nature of the protein-GL complexes detected. While ESI-MS serves as a convenient method for detecting protein interactions with GLs in NDs, the measured abundances of free and GL-bound protein ions do not necessarily reflect solution composition. There is overwhelming evidence that, in solution, the GL-bound proteins remain associated with NDs and are only released (as protein-GL complexes) in the gas phase. Consequently, different ESI-MS response factors are expected for the free proteins and GL-bound proteins. Finally, using the newly developed *proxy ligand* ESI-MS assay, K_a values for CBM-$B2_{NGL}$ and CTB_5-GM1 interactions were quantified. A key finding of this study is that the affinities of the proteins for the GL ligands in the NDs are slightly lower (by a factor of ≤ 5) than those of the corresponding oligosaccharides in solution. Future efforts will exploit the *proxy ligand* ESI-MS method to study, in detail, the effects of ND composition on protein-GL binding.

Acknowledgement

The authors would like to acknowledge the Natural Sciences and Engineering Research Council of Canada and the Alberta Glycomics Centre for generous funding and T. Lowary (University of Alberta) for providing the blood group B trisaccharide used in this work. ABB thanks the

Canadian Institutes for Health Research for funding (MOP 130305). LH also thanks Alberta Innovates for a Graduate Student Scholarship.

Supporting Information Available

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Email: john.klassen@ualberta.ca

Notes

The authors declare no competing financial interests.

References

(1) Hakomori, S. *Curr. Opin. Hematol.* **2003**, *10*, 16.

(2) Sharon, N.; Lis, H. *Sci. Am.* **1993**, *268*, 82.

(3) Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Stanley, P.; Bertozzi, C.R.; Hart, G.W.; Etzler, M.E. Essentials of Glycobiology 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA 2009.

(4) Holgersson, J.; Gustafsson, A.; Breimer, M. E. *Immunol. Cell Biol.* **2005**, *83*, 694.

(5) Lopez P.H.; Schnaar, R.L. *Methods Enzymol*. **2006**, *417*, 205.

(6) DeMarco, M. L. *Biochemistry* **2012**, *51*, 5725.

(7) Evans, S. V.; MacKenzie, C. R. *J. Mol. Recognit.* **1999**, *12*, 155.

(8) Lingwood, C. A. *Glycoconjugate J.* **1996**, *13*, 495.

(9) Lingwood, C. A.; Manis, A.; Mahfoud, R.; Khan, F.; Binnington, B.; Mylvaganam, M. *Chem. Phys. Lipids* **2010**, *163*, 27.

(10) Lingwood, D.; Binnington, B.; Rog, T.; Vattulainen, I.; Grzybek, M.; Coskun, U.; Lingwood, C. A.; Simons, K. *Nat. Chem. Biol.* **2011**, *7*, 260.

(11) MacKenzie, C. R.; Hirama, T.; Lee, K. K.; Altman, E.; Young, N. M. *J. Biol. Chem.* **1997**, *272*, 5533.

(12) Lauer, S.; Goldstein, B.; Nolan, R. L.; Nolan, J. P. *Biochemistry* **2002**, *41*, 1742.

(13) Gallegos, K. M.; Conrady, D. G.; Karve, S. S.; Gunasekera, T. S.; Herr, A. B.; Weiss, A. A. *PLoS One* **2012**, *7*, 10.

(14) Vogel, J.; Bendas, G.; Bakowsky, U.; Hummel, G.; Schmidt, R. R.; Kettmann, U.; Rothe, U. *Biochim. Biophys. Acta.* **1998**, *1372*, 205.

(15) Shi, J. J.; Yang, T. L.; Kataoka, S.; Zhang, Y. J.; Diaz, A. J.; Cremer, P. S. *J. Am. Chem.*

Soc. **2007**, *129*, 5954.

(16) Nath, A.; Atkins, W. M.; Sligar, S. G. *Biochemistry* **2007**, *46*, 2059.

(17) Bayburt, T. H.; Sligar, S. G. *FEBS Lett.* **2010**, *584*, 1721.

(18) Jayaraman, N.; Maiti, K.; Naresh, K. *Chem. Soc. Rev.* **2013**, *42*, 4640.

(19) Borch, J.; Torta, F.; Sligar, S. G.; Roepstorff, P. *Anal. Chem.* **2008**, *80*, 6245.

(20) Zhang, Y. X.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C. S.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618.

(21) Leney, A. C.; Fan, X. X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271.

(22) Sloan, C. D. K.; Marty, M. T.; Sligar, S. G.; Bailey, R. C. *Anal. Chem.* **2013**, *85*, 2970.

(23) Vanheyni, S. *Science* **1974**, *183*, 656.

(24) Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Ho, W. G. J. *J. Mol. Biol.* **1998**, *282*, 1043.

(25) Kadioglu, A.; Andrew, P. W. *Trends Immunol.* **2004**, *25*, 143.

(26) Higgins, M. A.; Ficko-Blean, E.; Meloncelli, P. J.; Lowary, T. L.; Boraston, A. B. *J. Mol. Biol.* **2011**, *411*, 1017.

(27) Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.

(28) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Nano Lett.* **2002**, *2*, 853.

(29) Popovic, K.; Holyoake, J.; Pomès, R.; Privé, G. G. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 2908.

(30) Leney, A. C.; Rezaei Darestani, R.; Li, J.; Nikjah, S.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Xiong, Z. J.; Prive, G. G.; Klassen, J. S. *Anal. Chem.* **2015**, in press, DOI: 10.1021/acs.analchem.5b00170.

(31) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.*

2012, *23*, 431.

(32) Kitova, E. N.; Kitov, P. I.; PaszkiewiCz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.

(33) Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1047.

(34) Lin, H.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 104.

(35) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. *J. Am. Chem. Soc.* **2004**, *126*, 3477.

(36) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. *Arch. Biochem. Biophys.* **2006**, *450*, 215.

(37) Nakano, M.; Fukuda, M.; Kudo, T.; Miyazaki, M.; Wada, Y.; Matsuzaki, N.; Endo, H.; Handa, T. *J. Am. Chem. Soc.* **2009**, *131*, 8308.

(38) Han, L.; Kitova, E. N.; Tan, M.; Jiang, X.; Pluvinage, B.; Boraston, A. B.; Klassen, J. S. *Glycobiology* **2015**, *25*, 170.

(39) Moonens, K.; Bouckaert, J.; Coddens, A.; Tran, T.; Panjikar, S.; De Kerpel, M.; Cox, E.; Remaut, H.; De Greve, H. *Mol. Microbiol.* **2012**, *86*, 82.

Figure Captions

Figure 1. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3 μM CTB_5 with (a) 3 μM, (b) 24.4 μM 0.5% GM1 ND (corresponding to 3 and 24.4 μM GM1, respectively); (c) 0.6 μM and (d) 1.4 μM 10% GM1 ND (corresponding to 12 and 28 μM GM1, respectively). Insets show normalized distributions of free and GM1-bound CTB_5; theoretical distributions were calculated using association constants reported in reference 34 for the stepwise binding of $GM1_{os}$ to CTB_5.

Figure 2. Possible mechanisms for the stepwise binding of CTB_5 to GM1 ND. (a) *Nanodisc recruitment mechanism* - CTB_5 binds irreversibly to GM1 ligands from multiple NDs. (b) *Glycolipid extraction mechanism* - CTB_5 interacts with GM1 in one ND, followed by dissociation of the resulting (CTB_5 + *q*GM1) complex from the original ND and rapid re-binding to GM1 in another ND. (c) *Glycolipid diffusion mechanism* - GM1 rapidly redistribute between NDs and can be recruited by CTB_5. Note: to facilitate visualizing multivalent binding, a linear arrangement of subunits is used to represent the CTB_5 homopentamer.

Figure 3. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 μM) with (a) 8 μM and (b) 30.8 μM 10% $B2_{NGL}$ ND. (c) Plots of fraction of ligand-bound CBM (*f*) versus $B2_{NGL}$ concentration. The experimental conditions were the same as in (a) and (b), but with addition of 3.2 – 30.8 μM ND containing 2.5% (●) or 10% (■) $B2_{NGL}$. The dashed curve represents the theoretical plot calculated from the association

constant reported in reference 38 for CBM binding to $B2_{os}$. The error bars correspond to one standard deviation.

Figure 4. (a) ESI mass spectrum acquired in positive ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 12 μM CBM, 40 μM B-tri (L_{proxy}) with 24 μM 15% $B2_{NGL}$ ND (corresponding to 360 μM and 720 μM $B2_{NGL}$, respectively); 5 μM P_{ref} (Ubq) was added to the solution to correct for the nonspecific ligand binding during ESI process. (b) Plots of R_{proxy} (≡ *Ab*(CBM + B-tri)/*Ab*(CBM)) versus $B2_{NGL}$ concentration. The experimental conditions were the same as in (a), but with addition of 0 – 24 μM 15% $B2_{NGL}$ ND (●) or 0 – 28 μM 10% $B2_{NGL}$ ND (■). The error bars correspond to one standard deviation.

Figure 5. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 4.4 μM CTB_5 and 20 μM $GM1_{os}$ with (a) 0 μM and (b) 2.5 μM 0.5% GM1 ND. Inset shows the normalized distributions of free and $GM1_{os}$-bound CTB_5. (c) (■) Normalized distributions of free and ligand-bound CTB_5 measured from the mass spectrum in (b); (▨) theoretical distributions were calculated using association constants determined from the *proxy ligand* method and values reported in reference 34 for the stepwise binding of $GM1_{os}$ and GM1 to CTB_5. (d) Plot of $R_{proxy,5}$ (≡ *Ab*(CTB_5 + $5GM1_{os}$)/*Ab*(CTB_5 + $4GM1_{os}$)) versus GM1 concentration. The experimental conditions were the same as in (a) and (b), but with addition of 0 – 2.5 μM 0.5% GM1 ND. The error bars correspond to one standard deviation.

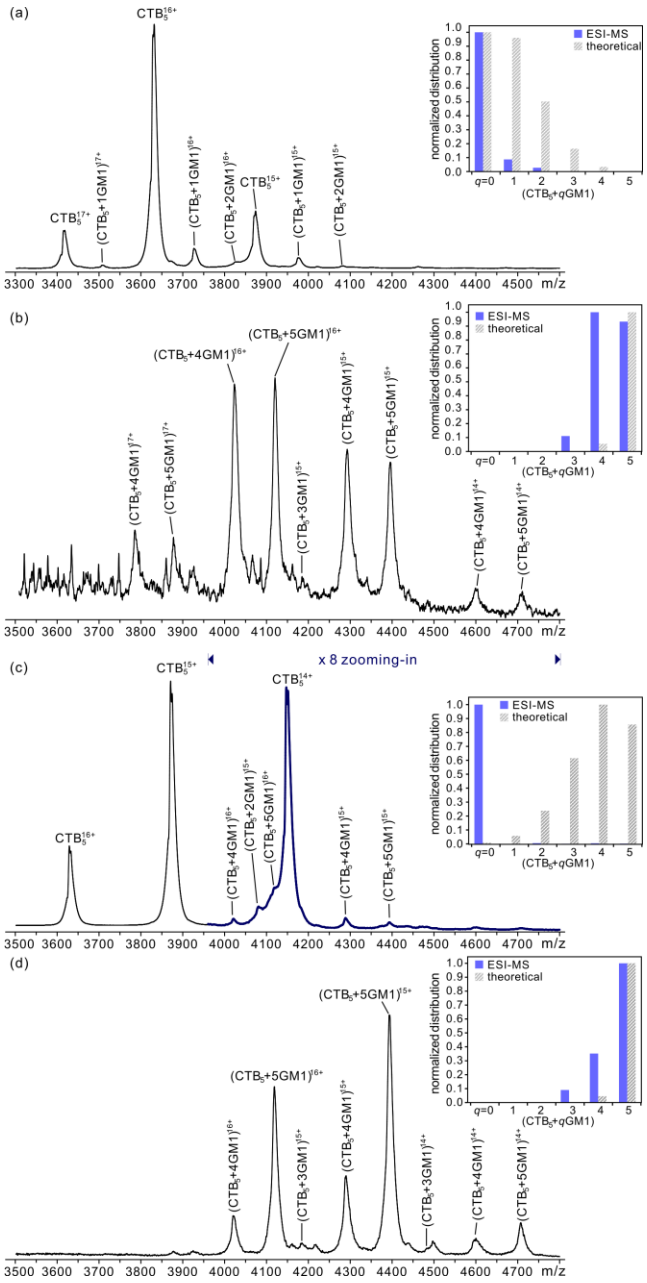

(a)
CTB_5^{16+}
CTB_5^{17+}
$(CTB_5+1GM1)^{17+}$
$(CTB_5+1GM1)^{16+}$
$(CTB_5+2GM1)^{16+}$
CTB_5^{15+}
$(CTB_5+1GM1)^{15+}$
$(CTB_5+2GM1)^{15+}$
ESI-MS
theoretical
normalized distribution
q=0 1 2 3 4 5
(CTB_5+qGM1)
3300 3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 m/z
(b)
$(CTB_5+4GM1)^{17+}$
$(CTB_5+5GM1)^{17+}$
$(CTB_5+4GM1)^{16+}$
$(CTB_5+5GM1)^{16+}$
$(CTB_5+3GM1)^{15+}$
$(CTB_5+4GM1)^{15+}$
$(CTB_5+5GM1)^{15+}$
$(CTB_5+4GM1)^{14+}$
$(CTB_5+5GM1)^{14+}$
3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 4700 m/z
x 8 zooming-in
(c)
CTB_5^{15+}
CTB_5^{14+}
CTB_5^{16+}
$(CTB_5+4GM1)^{16+}$
$(CTB_5+2GM1)^{15+}$
$(CTB_5+5GM1)^{16+}$
$(CTB_5+4GM1)^{15+}$
$(CTB_5+5GM1)^{15+}$
(d)
$(CTB_5+5GM1)^{15+}$
$(CTB_5+5GM1)^{16+}$
$(CTB_5+4GM1)^{16+}$
$(CTB_5+3GM1)^{15+}$
$(CTB_5+4GM1)^{15+}$
$(CTB_5+3GM1)^{14+}$
$(CTB_5+4GM1)^{14+}$
$(CTB_5+5GM1)^{14+}$

Figure 1

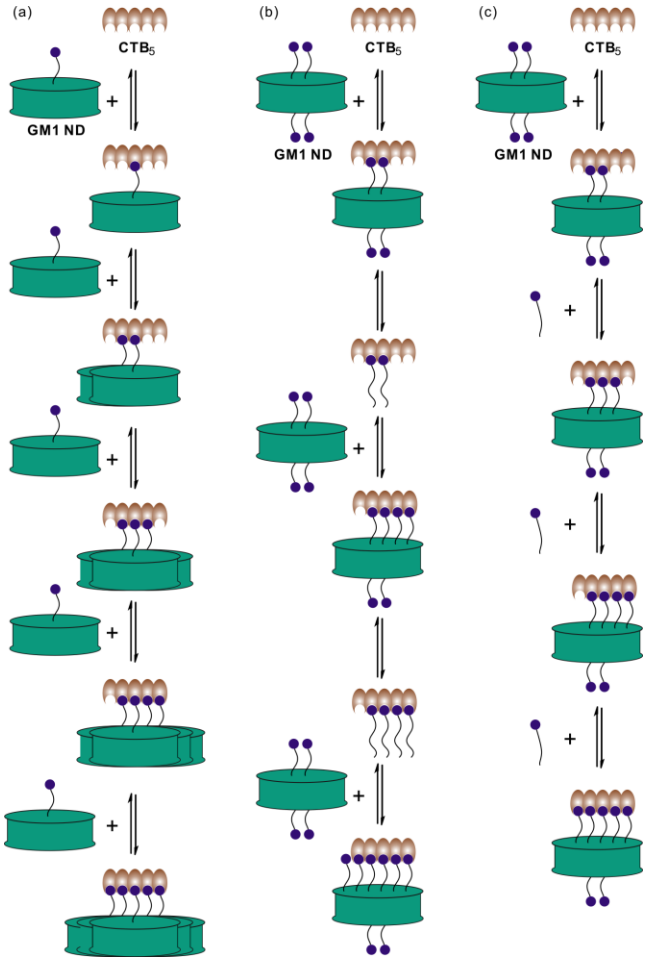


Figure 2

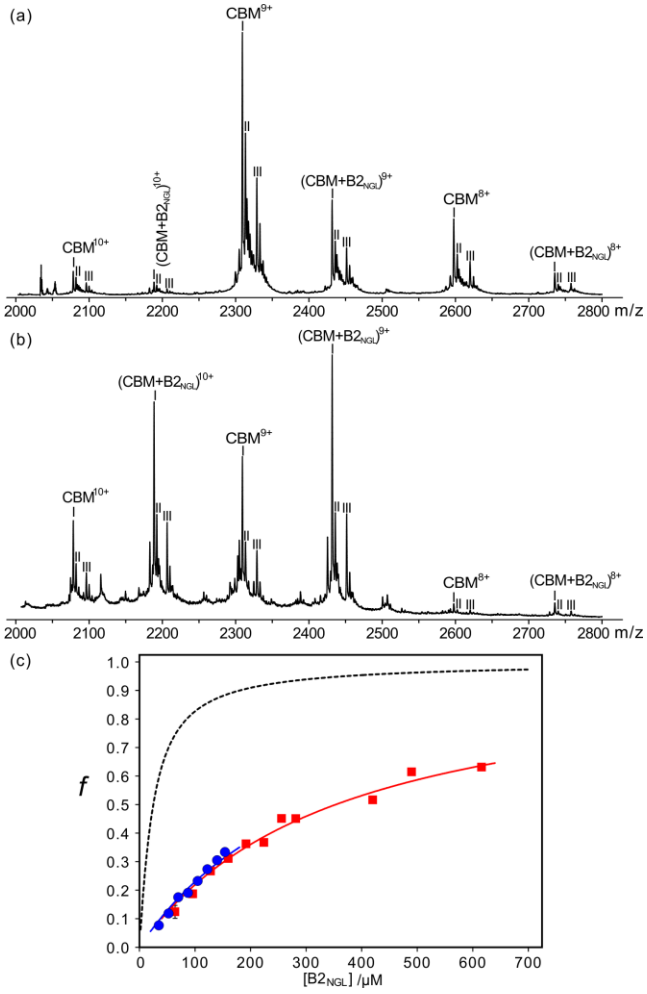

(a)
CBM9+
(CBM+B2NGL)10+
(CBM+B2NGL)9+
CBM8+
CBM10+
(CBM+B2NGL)8+
2000 2100 2200 2300 2400 2500 2600 2700 2800 m/z
(b)
(CBM+B2NGL)9+
(CBM+B2NGL)10+
CBM9+
CBM10+
CBM8+
(CBM+B2NGL)8+
2000 2100 2200 2300 2400 2500 2600 2700 2800 m/z
(c)
f
1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0
0 100 200 300 400 500 600 700
[B2NGL] /μM

Figure 3

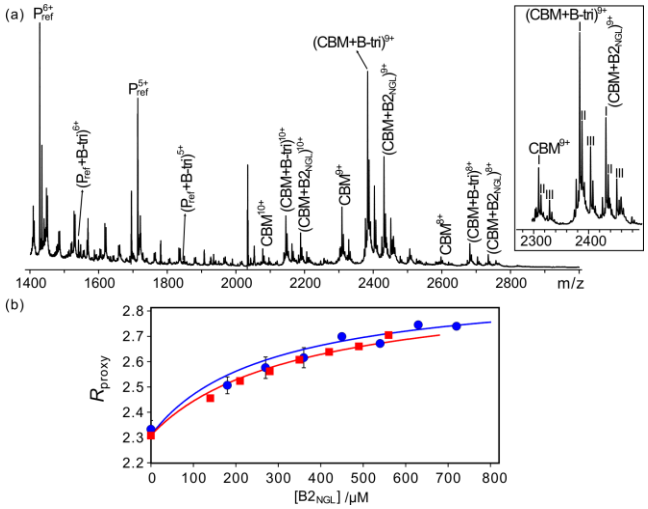

(a)
P_{ref}^{6+}
$(P_{ref}+B\text{-}tri)^{6+}$
P_{ref}^{5+}
$(P_{ref}+B\text{-}tri)^{5+}$
CBM^{10+}
$(CBM+B\text{-}tri)^{10+}$
$(CBM+B2_{NGL})^{10+}$
CBM^{9+}
$(CBM+B\text{-}tri)^{9+}$
$(CBM+B2_{NGL})^{9+}$
CBM^{8+}
$(CBM+B\text{-}tri)^{8+}$
$(CBM+B2_{NGL})^{8+}$
1400 1600 1800 2000 2200 2400 2600 2800 m/z
2300 2400
(b)
R_{proxy}
2.2 2.3 2.4 2.5 2.6 2.7 2.8
0 100 200 300 400 500 600 700 800
$[B2_{NGL}]$ /μM

Figure 4

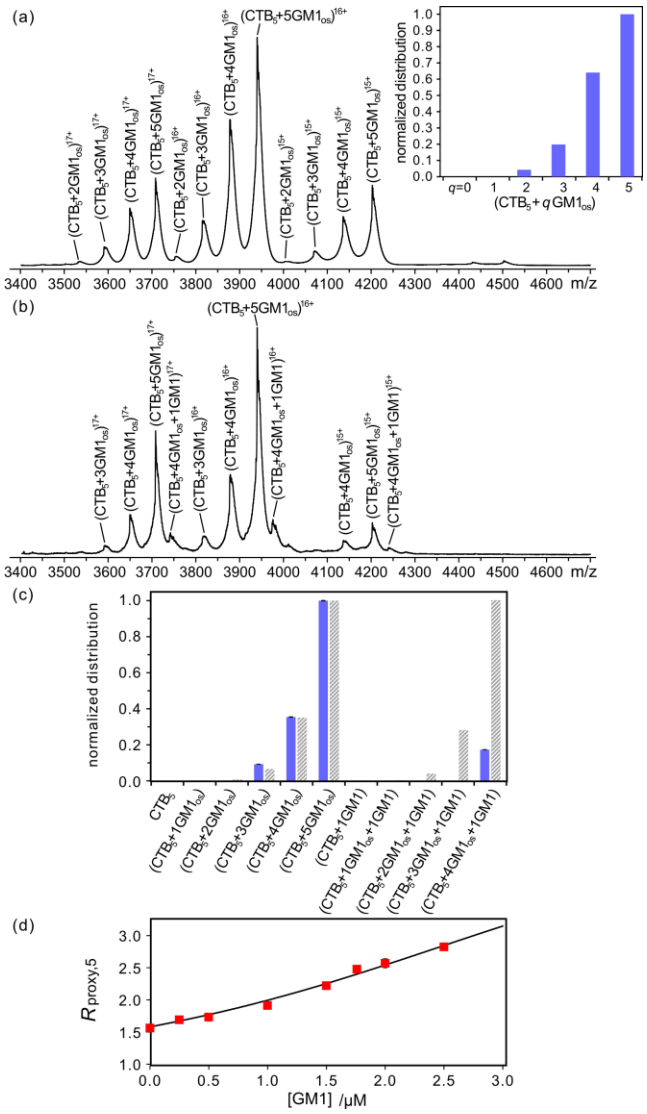
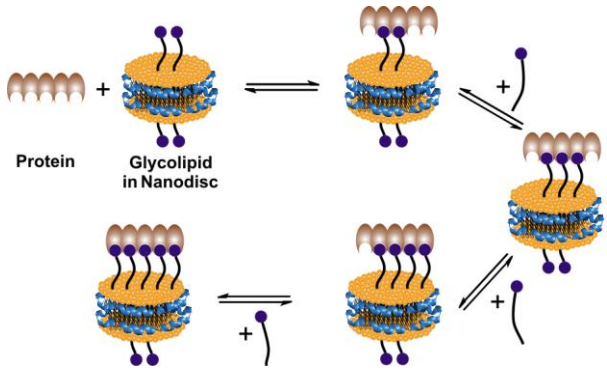

(a)
(CTB5+2GM1os)17+
(CTB5+3GM1os)17+
(CTB5+4GM1os)17+
(CTB5+5GM1os)17+
(CTB5+2GM1os)16+
(CTB5+3GM1os)16+
(CTB5+4GM1os)16+
(CTB5+5GM1os)16+
(CTB5+2GM1os)15+
(CTB5+3GM1os)15+
(CTB5+4GM1os)15+
(CTB5+5GM1os)15+
normalized distribution
q=0 1 2 3 4 5
(CTB5+qGM1os)
3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 m/z
(b)
(CTB5+3GM1os)17+
(CTB5+4GM1os)17+
(CTB5+5GM1os)17+
(CTB5+4GM1os+1GM1)17+
(CTB5+3GM1os)16+
(CTB5+4GM1os)16+
(CTB5+5GM1os)16+
(CTB5+4GM1os+1GM1)16+
(CTB5+4GM1os)15+
(CTB5+5GM1os)15+
(CTB5+4GM1os+1GM1)15+
3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 m/z
(c)
normalized distribution
CTB5
(CTB5+1GM1os)
(CTB5+2GM1os)
(CTB5+3GM1os)
(CTB5+4GM1os)
(CTB5+5GM1os)
(CTB5+1GM1)
(CTB5+1GM1os+1GM1)
(CTB5+2GM1os+1GM1)
(CTB5+3GM1os+1GM1)
(CTB5+4GM1os+1GM1)
(d)
Rproxy,5
[GM1] /μM

Figure 5

TOC Graphic



SUPPORTING INFORMATION FOR:

Protein-Glycolipid Interactions Studied in vitro using ESI-MS and Nanodiscs. Insights into the Mechanisms and Energetics of Binding

Ling Han, Elena N. Kitova, Jun Li, Sanaz Nikjah, Hong Lin, Benjamin Pluvinage, Alisdair B. Boraston and John S. Klassen

Experimental

Materials and Methods

Preparation of nanodiscs

Nanodiscs containing GL (GM1, $B2_{NGL}$, or $A2_{NGL}$) were prepared using a protocol developed by Sligar and coworkers[S1,S2] and only a brief description is given here. DMPC was mixed with GM1, $B2_{NGL}$ or $A2_{NGL}$ at the desired ratios. The lipids were dried under a gentle stream of nitrogen overnight at room temperature and re-dissolved in a Tris buffer containing 20 mM sodium cholate (Sigma-Aldrich Canada, Oakville, Canada) at neutral pH. The recombinant membrane scaffold protein MSP1E1 was added to the mixture to yield the an MSP1E1:lipid molar ratio of 1:100. To initiate the ND self-assembly process, an equal volume of pre-washed biobeads (Bio-Rad, Mississauga, Canada) were added and incubated with the mixture for 4 h at room temperature. The supernatant was recovered and then loaded onto the Superdex 200 10/300 size exclusion column (GE-Healthcare Life Sciences, Piscataway, NJ). Finally, the ND fraction was collected, concentrated and dialyzed against 200 mM ammonium acetate (pH 6.8) using an Amicon microconcentrator (EMD Millipore, Billerica, MA) with a 30 kDa MW cut off . The ND stock solutions were stored at -80 °C before use and the concentration was estimated based on the UV absorption of MSP1E1 at 280 nm. As the nominal molar ratio of MSP to total lipid is

1:100 and each ND possesses two MSPs, the number of GLs per ND is estimated to be two times the percentage of GL.

Preparation of picodiscs

Picodiscs, containing SapA and POPC, alone or with GL, were prepared following a protocol described elsewhere.[S3,S4] Briefly, GM1 and POPC (dissolved in 1:1 methanol: chloroform) were mixed in a 1:4 ratio and dried under flowing nitrogen overnight to form a lipid film. The lipid film was re-suspended in 50 mM sodium acetate and 150 mM NaCl (pH 4.8) followed by sonication and thaw cycles to form liposomes. Saposin A protein was then added into the liposomes at 1:10 molar ratio of SapA:(GM1+POPC) to initiate the picodiscs formation and the mixture was incubated at 37 °C for 45 min. Purification of the picodiscs was performed on a Superdex 75 10/300 size exclusion column (GE-Healthcare Life Sciences, Piscataway, NJ) equilibrated in 200 mM ammonium acetate (pH 4.8). Finally, picodiscs were concentrated and exchanged into 200 mM ammonium acetate (pH 6.8) and stored at room temperature for a maximum of 1 week. The concentration of SapA in the discs was determined by the UV absorption at 280 nm and the concentration of GM1 was estimated by assuming a 1:1 ratio of GM1:SapA.

Ultracentrifugation and SDS-PAGE

Ultracentrifugation was used to analyze the species present in solutions containing proteins and GL NDs. Briefly, CTB_5 and GM1 ND or CBM and $B2_{NGL}$ ND were incubated in a 200 mM ammonium acetate solution (pH 6.8, 25 °C) and placed in a microconcentrator (EMD Millipore, Billerica, MA) with a MW cutoff of 100 kDa and subjected to ultracentrifugation three times. Each time, 200 mM ammonium acetate buffer was added to the concentrated supernatant solution to maintain the same initial volume. Proteins and protein-ligand complexes with MW

≥100 kDa remained in the supernatant while those with MW <100 kDa passed through the membrane to the filtrate. The supernatant and filtrate were further analyzed by ESI-MS and sodium dodecyl sulfate polyacrylamide (15%) gel electrophoresis (SDS-PAGE). To carry out SDS-PAGE, solutions were diluted with an equal volume of 2× loading buffer (125 mM TrisHCl pH 6.8, 4% (w/v) SDS, 0.01% (w/v) bromophenol blue, 20% glycerol and 200 mM dithiothreitol). The solutions were preheated to ~90 °C for 5 min and then allowed to cool to room temperature prior to loading the samples. Coomassie stain was used to visualize proteins on the gel.

Mass spectrometry

All ESI-MS binding measurements were carried out in positive ion mode (unless otherwise indicated) 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. Nanoflow ESI was performed by inserting a platinum wire into a nanoESI tip, which was 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). The typical voltage applied to the platinum wire was 1.0 kV. The source conditions for the ESI-MS measurements were: source temperature 60 °C, cone voltage 35 V, Trap voltage 5 V, and Transfer voltage 2 V. For each acquisition at least 60 scans (at 2 s $scan^{-1}$) were measured. Data acquisition and processing were performed using Waters MassLynx software (version 4.1).

ESI solutions were prepared using 200 mM aqueous ammonium acetate buffer (pH 6.8, 25°C). For the *direct* ESI-MS measurements, solutions of target protein and GL ND were prepared at the desired concentrations. For the *proxy ligand* ESI-MS assays, solutions containing fixed concentrations of target protein and ligand and varying concentrations of GL ND were

prepared. All solutions were allowed to equilibrate for 15 min at 25 °C prior to ESI-MS analysis, unless otherwise indicated.

ESI-MS affinity measurements

The *direct* ESI-MS assay was used to measure the extent of ligand (oligosaccharide or GL) binding to CTB_5 and CBM and to quantify the interactions. As described in detail elsewhere,[S5] the association constant (K_a) for a 1:1 protein-ligand complex can be determined from the abundance (*Ab*) ratio (*R*) of the ligand-bound (PL) to free protein (P) ions measured from ESI-MS, eq S1:

$$K_a = \frac{R}{[L]_0 - \frac{[P]_0 R}{R+1}} \tag{S1}$$

where *R* is taken to reflect the corresponding equilibrium concentration ratio in the solution, eq S2:

$$R = \frac{\sum Ab(PL)}{\sum Ab(P)} = \frac{[PL]}{[P]} \tag{S2}$$

and $[P]_0$ and $[L]_0$ are the initial concentrations of protein and ligand, respectively.

For a protein with *h* ligand binding sites, the apparent association constant ($K_{a,q}$) for the addition of a q^{th} L to P bound (*q*-1) L can be expressed by eq S3:[S6]

$$K_{a,q} = \frac{R_q / R_{q-1}}{[L]_0 - \frac{[P]_0 \sum_{q=1}^{h} qR_q}{1 + \sum_{q=1}^{h} R_q}} \tag{S3}$$

where R_q is the abundance ratio of ligand-bound (to *q* molecules of L) to free protein measured from ESI-MS, and is taken to be equal to the corresponding concentration ratio at equilibrium, eq S4:

$$R_q = \frac{\sum Ab(\mathrm{PL}_q)}{\sum Ab(\mathrm{P})} = \frac{[\mathrm{PL}_q]}{[\mathrm{P}]} \tag{S4}$$

Also of interest in the present study was the fraction of occupied ligand binding sites (*f*) in P at a given concentration. A general expression for *f*, in terms of abundance or concentration, is given by eq S5:

$$f = \frac{\sum_{q=1}^{h} q \cdot Ab(\mathrm{PL}_q)}{h\left(Ab(\mathrm{P}) + \sum_{q=1}^{h} Ab(\mathrm{PL}_q)\right)} = \frac{\sum_{q=1}^{h} q \cdot [\mathrm{PL}_q]}{h\left([\mathrm{P}] + \sum_{q=1}^{h} [\mathrm{PL}_q]\right)} \tag{S5}$$

***Proxy ligand* ESI-MS assay**

It must be stressed that, in the case of P binding to the GL ligands (L) incorporated into NDs, the PL_q ions detected by ESI-MS are assumed to be associated with NDs in solution and are stripped out of the NDs during the ESI process, *vide infra*,[S7,S8] while the P ions originate from free P in solution. Differences in ionization efficiencies and others effects, such as incomplete extraction of the PL_q complexes from the ND or in-source dissociation of the PL_q ions, could introduce errors to the direct ESI-MS affinity measurements. Consequently, indirect binding measurements were also carried out using the newly developed *proxy ligand* ESI-MS method.

The *proxy ligand* ESI-MS method relies on a proxy ligand (L_{proxy}), which binds to P with known affinity ($K_{a,proxy}$) and competes with the GL ligand (L). The binding of P to L reduces the concentration of free P in solution, resulting in an increase in the concentration of PL_{proxy} complex, relative to P. Consequently, the extent of PL binding in solution can be deduced by monitoring the relative abundance of PL_{proxy} by ESI-MS. For the competitive binding of L and L_{proxy} to a P possessing a single binding site, the relevant equilibrium expressions are given by eqs S6a and S6b:

$$K_{a,proxy} = \frac{[PL_{proxy}]}{[P][L_{proxy}]} = \frac{R_{proxy}}{[L_{proxy}]} \quad \text{(S6a)}$$

$$K_a = \frac{[PL]}{[P][L]} = \frac{R}{[L]} \quad \text{(S6b)}$$

where R_{proxy} corresponds to the abundance ratio of the L_{proxy}-bound P (PL_{proxy}) to free P ions, which is taken to be equal to the corresponding concentration ratio in solution, eq S7a:

$$R_{proxy} = \frac{\sum Ab(PL_{proxy})}{\sum Ab(P)} = \frac{[PL_{proxy}]}{[P]} \quad \text{(S7a)}$$

and R is the concentration ratio of L-bound P (PL) to free P in solution, eq S7b:

$$R = \frac{[PL]}{[P]} \quad \text{(S7b)}$$

The value of R can be found from the experimentally determined R_{proxy} and the following equations of mass balance, eqs S8a-c:

$$[P]_0 = [P] + [PL_{proxy}] + [PL] \quad \text{(S8a)}$$

$$[L_{proxy}]_0 = [L_{proxy}] + [PL_{proxy}] \quad \text{(S8b)}$$

$$[L]_0 = [L] + [PL] \quad \text{(S8c)}$$

Substituting $[PL_{proxy}]$ and [PL] (in eq S8a) with $R_{proxy}[P]$ and $R[P]$, respectively, gives eq S9a:

$$[P] = \frac{[P]_0}{1 + R_{proxy} + R} \quad \text{(S9a)}$$

It follows that $[PL_{proxy}]$ and [PL] can be expressed as eqs S9b and S9c, respectively:

$$[PL_{proxy}] = \frac{R_{proxy}[P]_0}{1 + R_{proxy} + R} \quad \text{(S9b)}$$

$$[PL] = \frac{R[P]_0}{1 + R_{proxy} + R} \quad \text{(S9c)}$$

and $[L_{proxy}]$ and [L] can be expressed as eqs S10a and S10b, respectively:

$$[L_{proxy}] = \frac{R_{proxy}}{K_{a,proxy}} = [L_{proxy}]_0 - \frac{R_{proxy}[P]_0}{1 + R_{proxy} + R} \quad \text{(S10a)}$$

$$[L] = [L]_0 - \frac{R[P]_0}{1 + R_{proxy} + R} \quad \text{(S10b)}$$

Rearranging eq S10a allows R to be expressed in terms of $[P]_0$, $[L]_0$, R_{proxy} and $K_{a,proxy}$, eq S11:

$$R = \frac{R_{proxy}[P]_0}{[L_{proxy}]_0 - \frac{R_{proxy}}{K_{a,proxy}}} - (R_{proxy} + 1) \quad \text{(S11)}$$

and K_a can be calculated from eq S12:

$$K_a = \frac{R}{[L]_0 - \frac{R[P]_0}{1 + R_{proxy} + R}} = \frac{1}{\frac{[L]_0}{R} - \frac{[P]_0}{1 + R_{proxy} + R}}$$

$$= \frac{1}{([L_{proxy}]_0 - \frac{R_{proxy}}{K_{a,proxy}})(\frac{[L]_0}{R_{proxy}[P]_0 - ([L_{proxy}]_0 - \frac{R_{proxy}}{K_{a,proxy}})(R_{proxy} + 1)} - \frac{1}{R_{proxy}})} \quad \text{(S12)}$$

Where necessary, the *proxy ligand* ESI-MS assay was implemented in conjunction with the reference protein method, which was used to quantitatively correct the mass spectra for the occurrence of nonspecific protein-carbohydrate interactions during the ESI process.[S9] This method involves adding a non-interacting reference protein (P_{ref}) to the solution and the extent of nonspecific binding of L to P_{ref} was used to subtract the contribution of nonspecific binding of L to P from the mass spectrum. A complete description of the correction method can be found elsewhere.[S9]

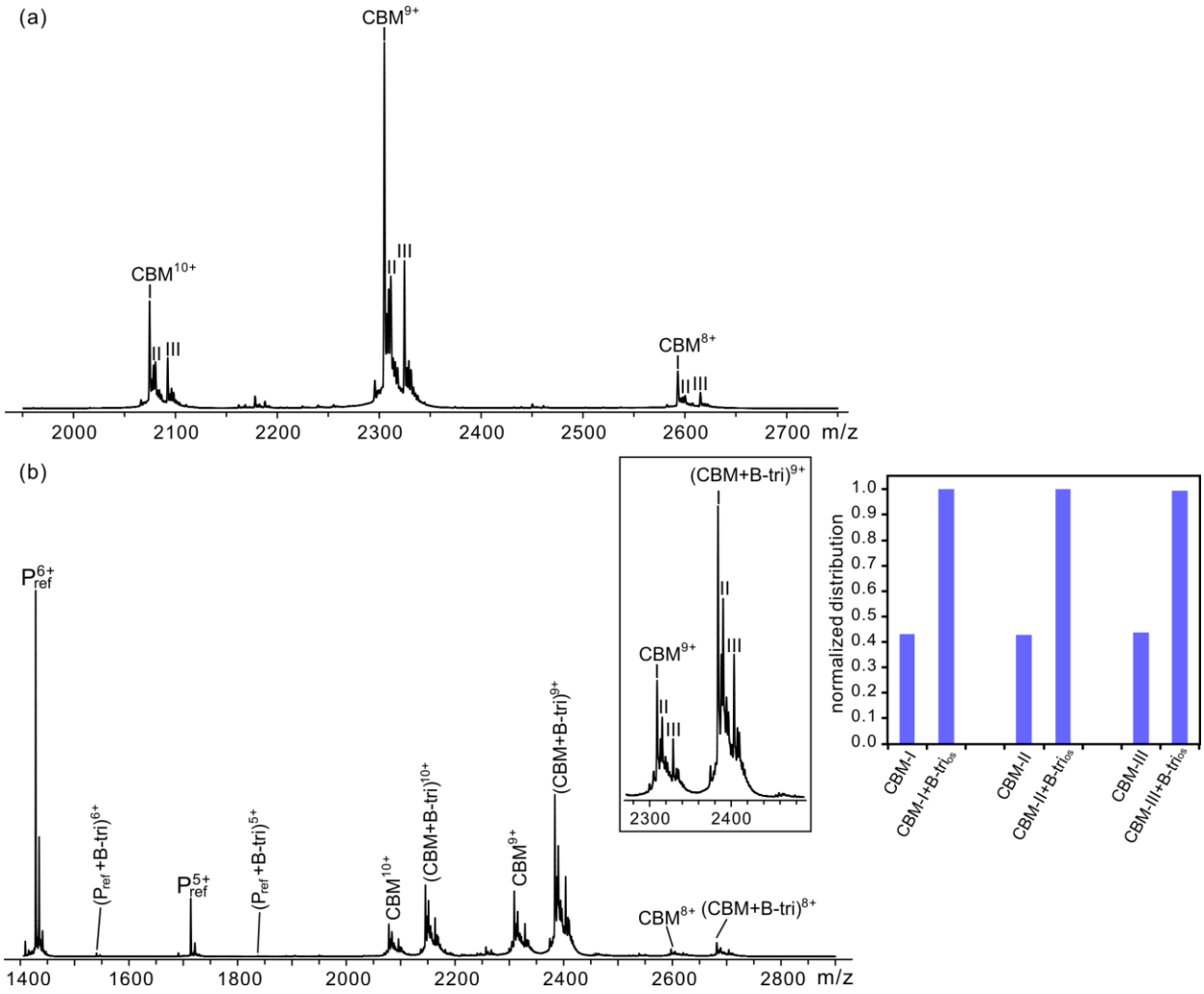
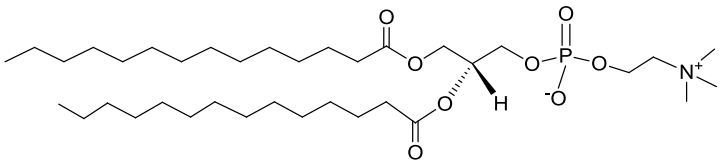


Figure S1. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of (a) CBM (12 μM) alone or (b) CBM (12 μM) with B trisaccharide (B-tri, 40 μM). Inset shows the normalized distributions of free and B-tri-bound CBM measured for the three isoforms (CBM-I, -II and -III).

10	20	30	40	50
GSSHHHHHHS	SGLVPRGSHM	ASTYLSDMDW	SSATHGDIDK	TKTVQKDAPF
60	70	80	90	100
TTGNKGEHTK	ISLLTSDDKV	KYFDKGIGTV	ADSPSVISYD	ISGQGFEKFE
110	120	130	140	150
TYIGIDQSAN	SSRSDHAVVD	RIEIEIDGKV	VYSSSVTNPE	GFRYNTQAQF
160	170	180	190	
ISVTIPQNAK	KISLKSFAGE	HTWGDEVVFA	DAKLIKTVST	

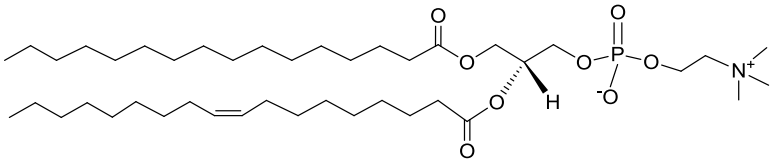
Figure S2. Amino acid sequence of the recombinant fragment of the family 51 carbohydrate binding module (CBM).



DMPC

MW 677.5 Da

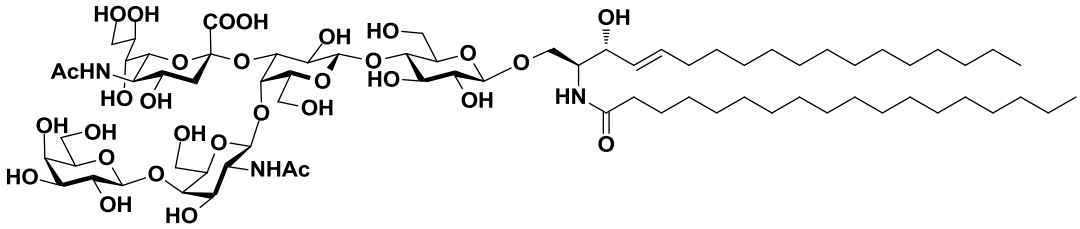
1,2-dimyristoyl-sn-glycero-3-phosphocholine



POPC

MW 760.1 Da

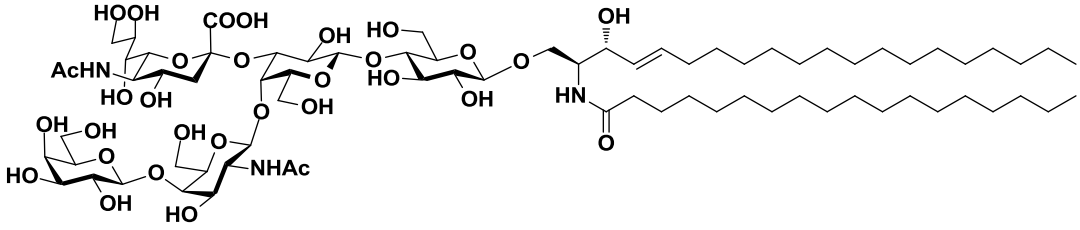
1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine



GM1 (*d*18:1-18:0)

MW 1545.8 Da

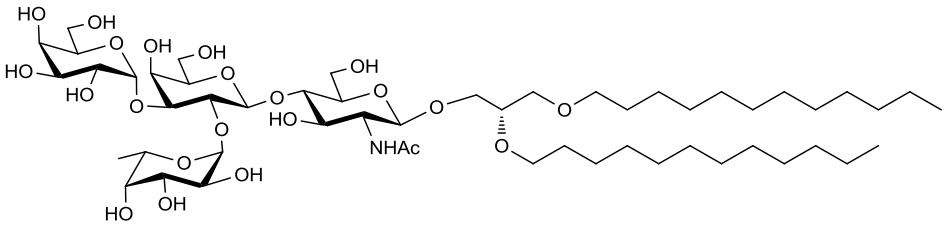
β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-β-D-Glc-ceramide



GM1 (*d*20:1-18:0)

MW 1573.9 Da

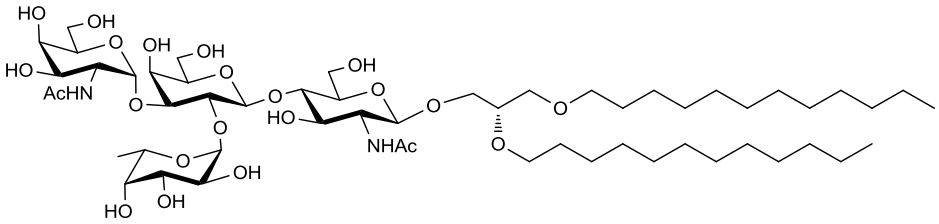
β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-β-D-Glc-ceramide



B type 2 neoglycolipid

MW 1101.7 Da

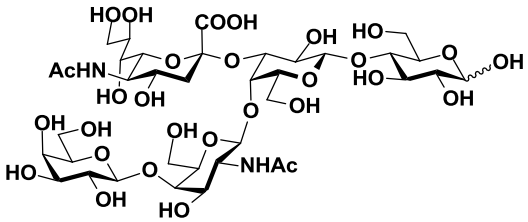
α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-1,2-di-O-dodecyl-sn-glycero



A type 2 neoglycolipid

MW 1142.7 Da

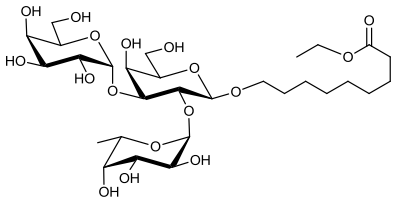
α-D-GalNAc-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-(1→4)-β-D-GlcNAc-1,2-di-O-dodecyl-sn-glycero



$GM1_{os}$

MW 998.3 Da

β-D-Gal-(1→3)-β-D-GalNAc-(1→4)-[α-D-Neu5Ac-(2→3)]-β-D-Gal-(1→4)-D-Glc



B trisaccharide

MW 672.3 Da

α-D-Gal-(1→3)-[α-L-Fuc-(1→2)]-β-D-Gal-O$(CH_2)_8COOCH_2CH_3$

Figure S3. Structures of the phospholipids, glycolipids and oligosaccharides used for the study.

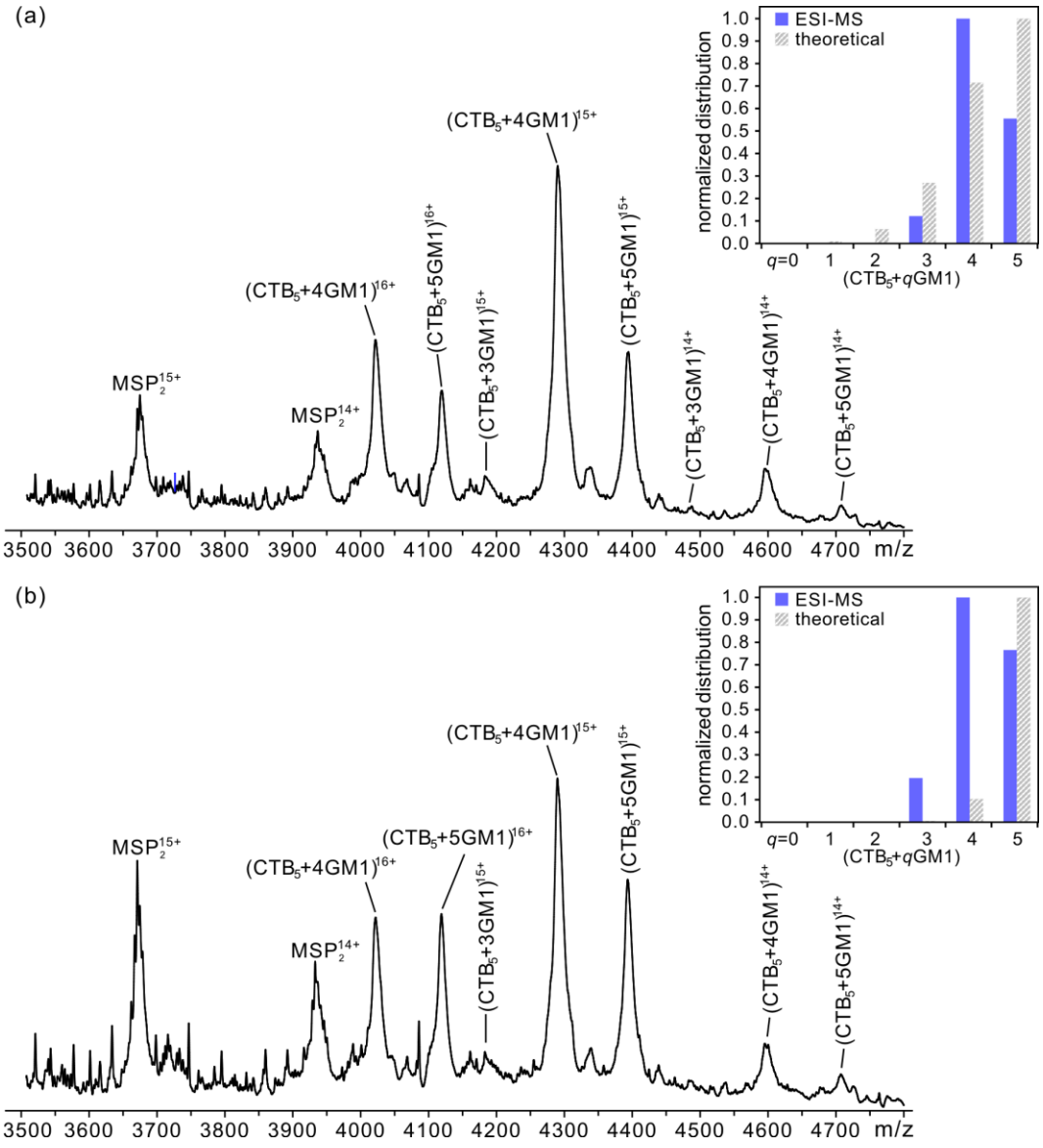


Figure S4. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 (3 μM) with (a) 6.8 μM and (b) 10.2 μM 1% GM1 ND (corresponding to 13.6 and 20.4 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB_5. The theoretical distributions were calculated using association constants reported in reference S10 for the stepwise binding of $GM1_{os}$ to CTB_5.

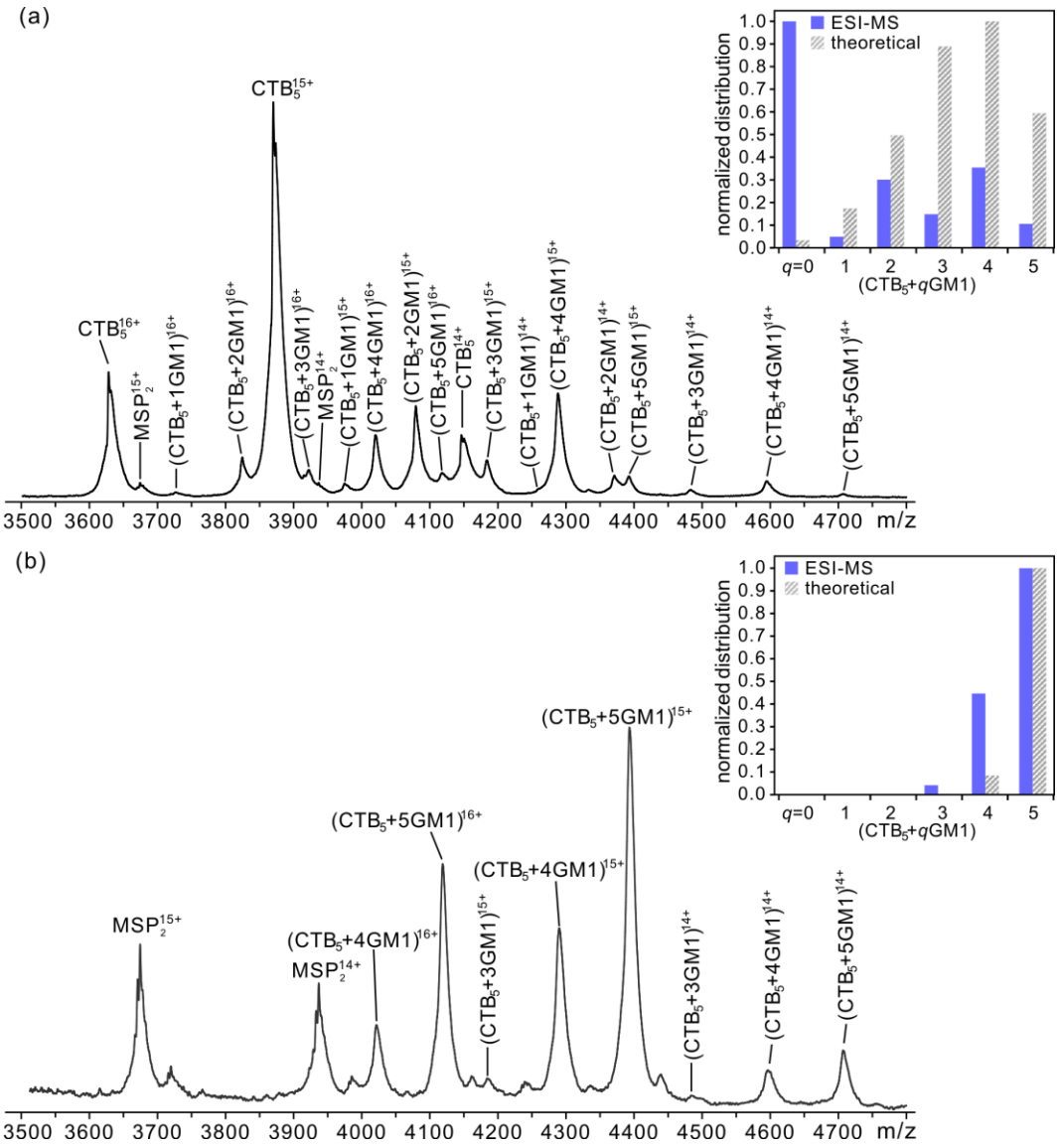


Figure S5. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 (3 μM) with (a) 2.1 μM and (b) 4.3 μM 2.5% GM1 ND (corresponding to 10.5 and 21.5 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB_5. The theoretical distributions were calculated using association constants reported in reference S10 for the stepwise binding of $GM1_{os}$ to CTB_5.

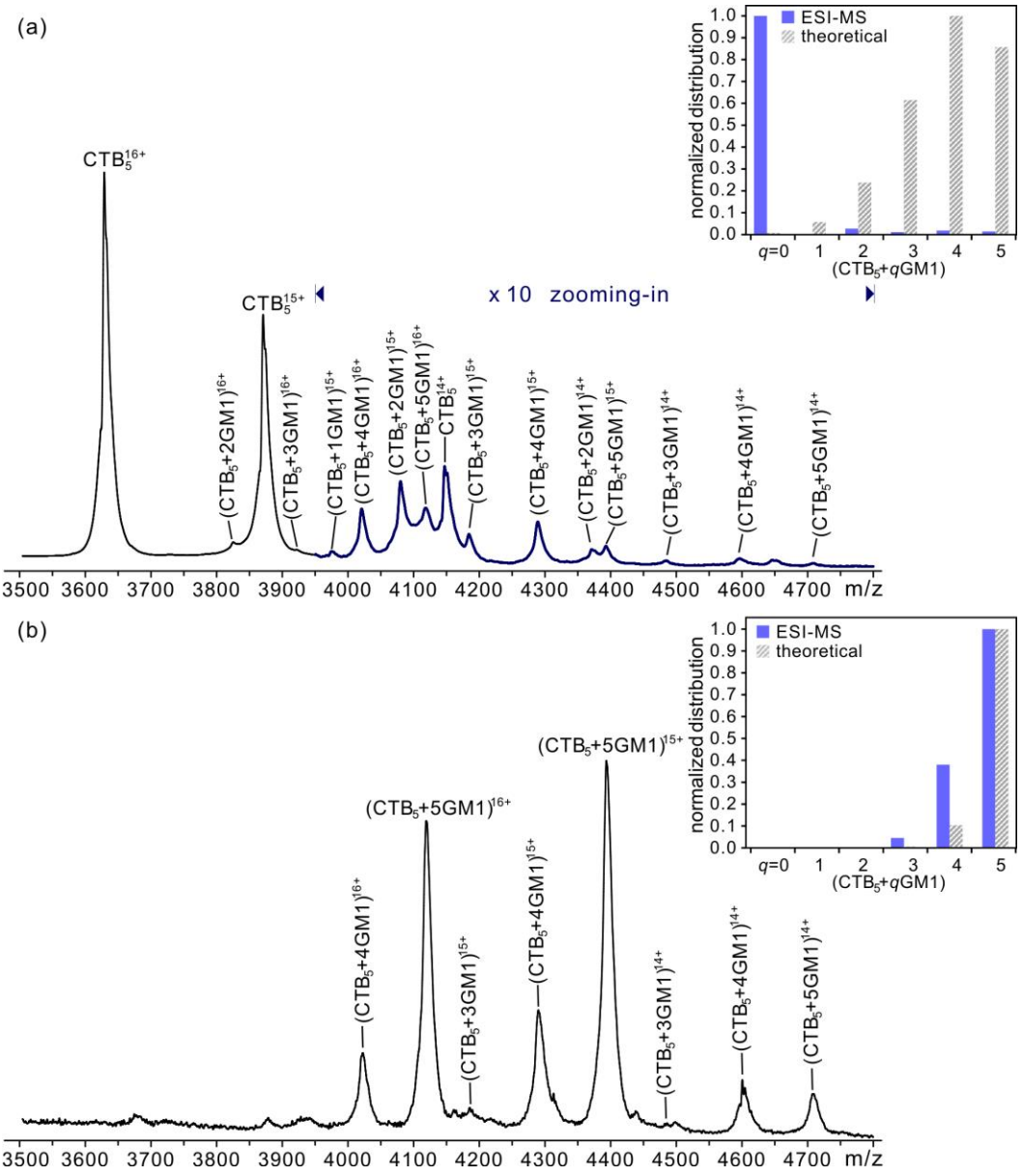


Figure S6. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 (3 μM) with (a) 1.2 μM and (b) 2.0 μM 5% GM1 ND (corresponding to 12 and 20 μM GM1, respectively). Insets show the normalized and theoretical distributions of free and GM1-bound CTB_5. The theoretical distributions were calculated using association constants reported in reference S10 for the stepwise binding of $GM1_{os}$ to CTB_5.

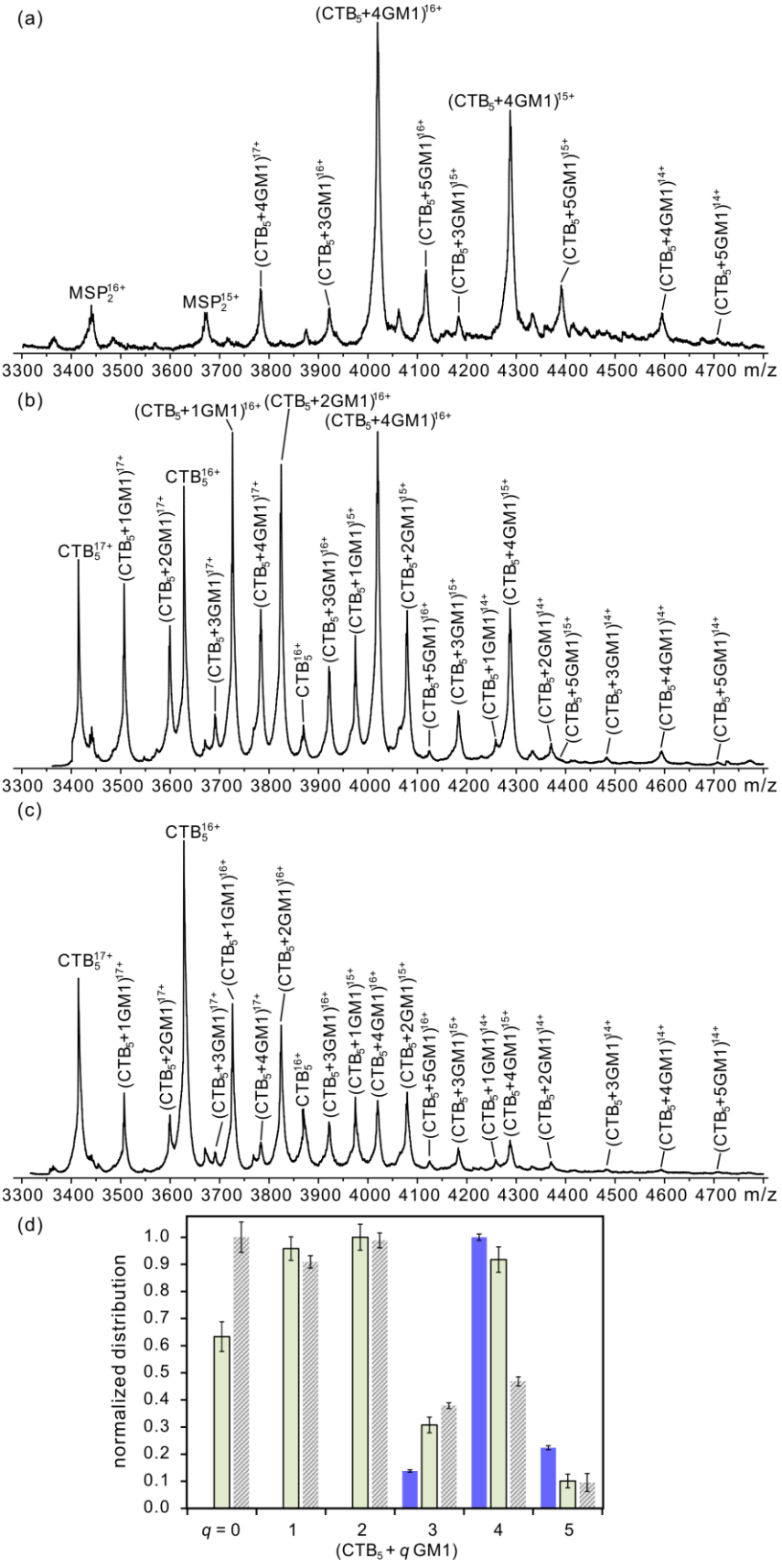

(a)
(CTB5+4GM1)16+
(CTB5+4GM1)15+
(CTB5+4GM1)17+
(CTB5+3GM1)16+
(CTB5+5GM1)16+
(CTB5+3GM1)15+
(CTB5+5GM1)15+
(CTB5+4GM1)14+
(CTB5+5GM1)14+
MSP216+
MSP215+
3300 3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 4700 m/z
(b)
(CTB5+1GM1)16+
(CTB5+2GM1)16+
(CTB5+4GM1)16+
CTB517+
(CTB5+1GM1)17+
CTB516+
(CTB5+2GM1)17+
(CTB5+3GM1)17+
(CTB5+4GM1)17+
(CTB5+3GM1)16+
(CTB5+1GM1)15+
(CTB5+2GM1)15+
(CTB5+5GM1)16+
(CTB5+3GM1)15+
(CTB5+1GM1)14+
(CTB5+4GM1)15+
(CTB5+2GM1)14+
(CTB5+5GM1)15+
(CTB5+3GM1)14+
(CTB5+4GM1)14+
(CTB5+5GM1)14+
3300 3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 4700 m/z
(c)
CTB516+
CTB517+
(CTB5+1GM1)17+
(CTB5+2GM1)17+
(CTB5+3GM1)17+
(CTB5+1GM1)16+
(CTB5+2GM1)16+
(CTB5+4GM1)17+
CTB516+
(CTB5+3GM1)16+
(CTB5+1GM1)15+
(CTB5+4GM1)16+
(CTB5+2GM1)15+
(CTB5+5GM1)16+
(CTB5+3GM1)15+
(CTB5+1GM1)14+
(CTB5+4GM1)15+
(CTB5+2GM1)14+
(CTB5+3GM1)14+
(CTB5+4GM1)14+
(CTB5+5GM1)14+
3300 3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 4700 m/z
(d)
normalized distribution
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
q = 0
1
2
3
4
5
(CTB5 + q GM1)

Figure S7. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 12 μM 0.5% GM1 ND (corresponding to 12 μM GM1) with (a) 3 μM and (c) 6 μM CTB_5. (b) ESI mass spectrum acquired upon addition of another 3 μM CTB_5 to the solution in (a). (d) Normalized distributions of free and GM1-bound CTB_5 measured from mass spectra in (a) ■, (b) ■ and (c) ▨.

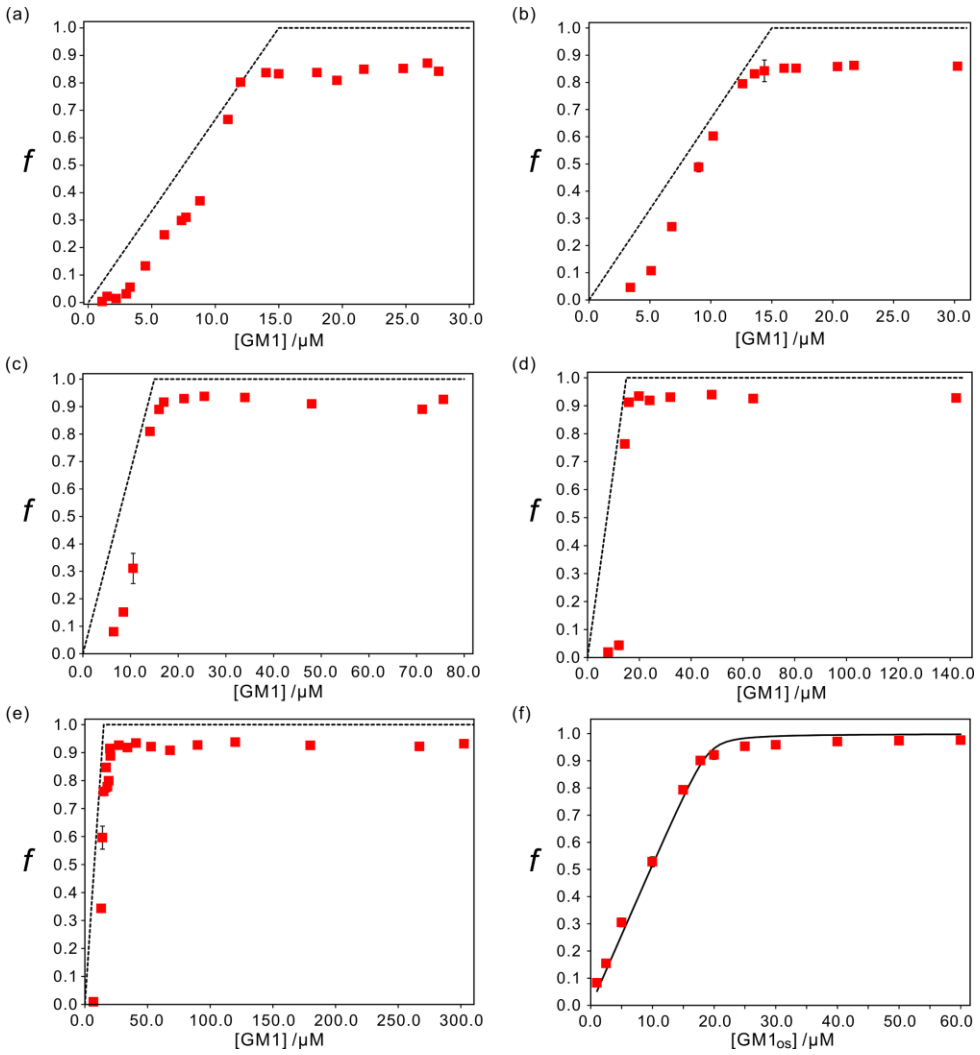


Figure S8. Plot of fraction of occupied ligand binding sites in CTB_5 (*f*) versus GM1 concentration measured by ESI-MS in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of 3 μM CTB_5 and varying concentrations of (a) 0.5%, (b) 1%, (c) 2.5%, (d) 5% and (e) 10% GM1 ND. Dashed lines represent the molar ratio of GM1 to the CTB_5 binding sites. (f) Plot of *f* versus $GM1_{os}$ concentration measured by ESI-MS in positive ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 3.8 μM CTB_5 with $GM1_{os}$ (1.0 μM – 60 μM). The solid curve corresponds to the theoretical plot calculated using affinities for the stepwise binding of $GM1_{os}$ to CTB_5 reported in reference S10. The error bars correspond to one standard deviation.

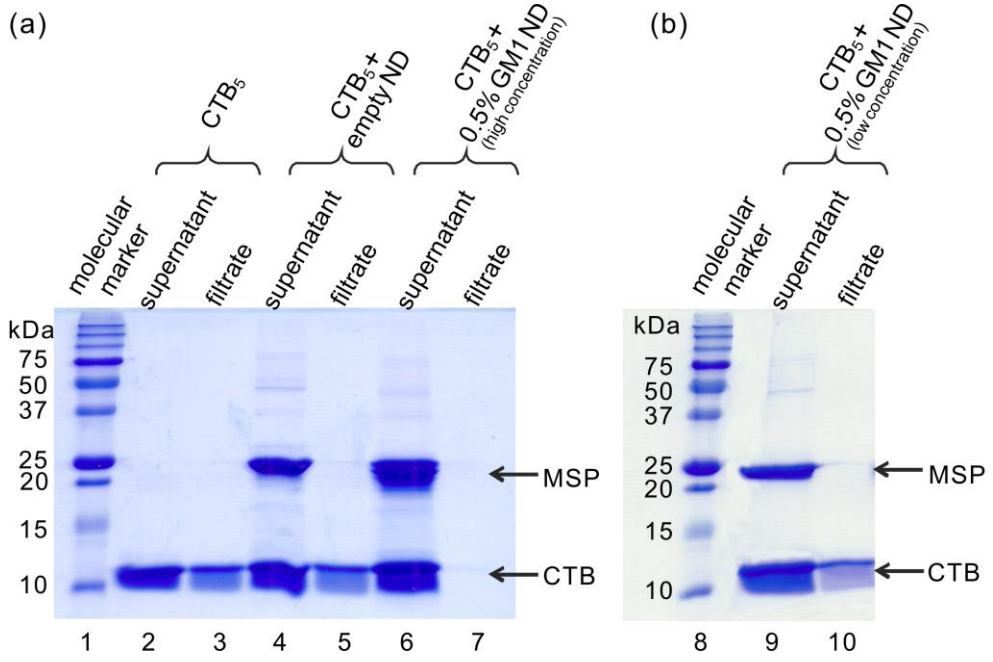


Figure S9. (a) and (b) SDS-PAGE analysis of aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 alone, CTB_5 with empty ND and CTB_5 with 0.5% GM1 NDs subjected to ultracentrifugation using a filter with a MWCO of 100 kDa.; supernatant (MW ≥100 kDa) and filtrate (MW ≤100 kDa). Molecular weight markers (lane 1 and 8); supernatant and filtrate for solution of CTB_5 (5 μM) alone (lanes 2 and 3, respectively); supernatant and filtrate for solution of CTB_5 (5 μM) with ND (11 μM) containing no GM1 (lanes 4 and 5, respectively); supernatant and filtrate for solution of CTB_5 (5 μM) with 0.5% GM1 ND (24 μM) (lanes 6 and 7, respectively); and supernatant and filtrate for solution of CTB_5 (5 μM) with 3 μM 0.5% GM1 ND (lanes 9 and 10, respectively).

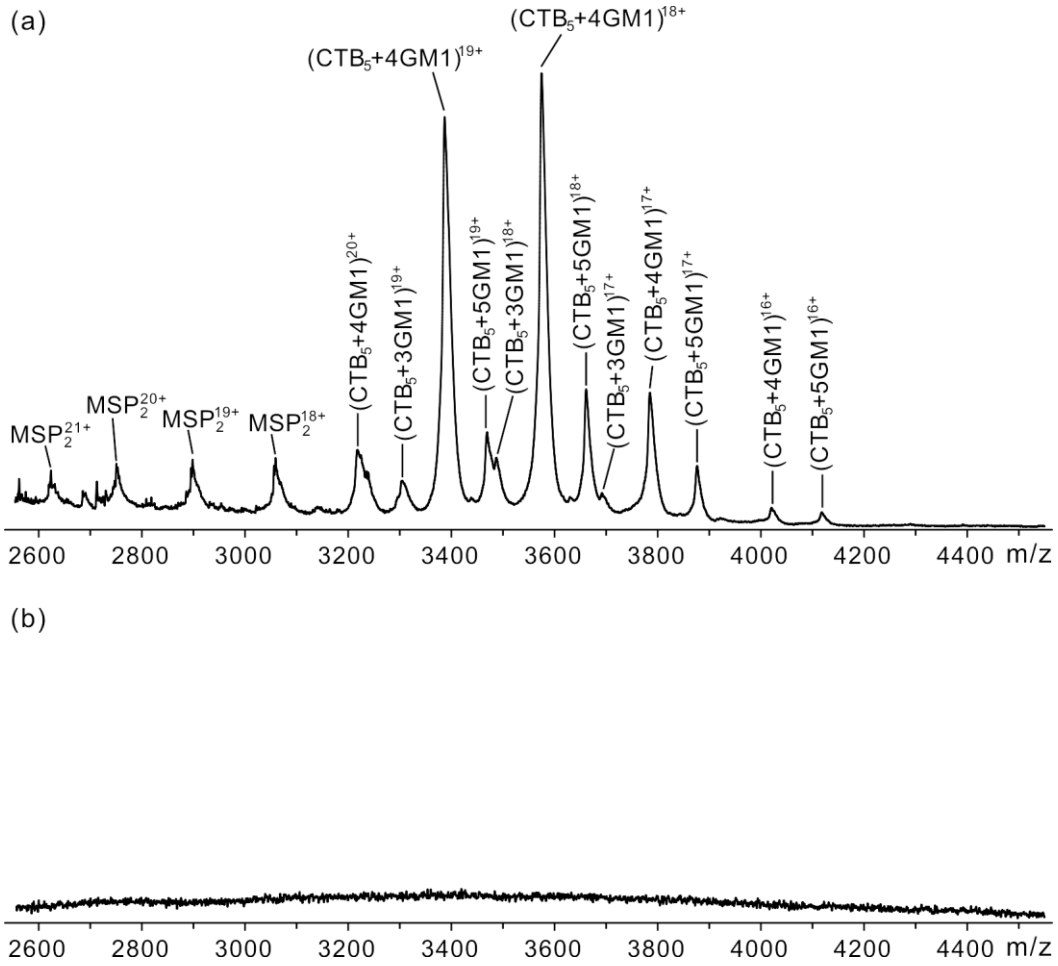


Figure S10. ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 (5 μM) and 0.5% GM1 ND (24 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa.; (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa).

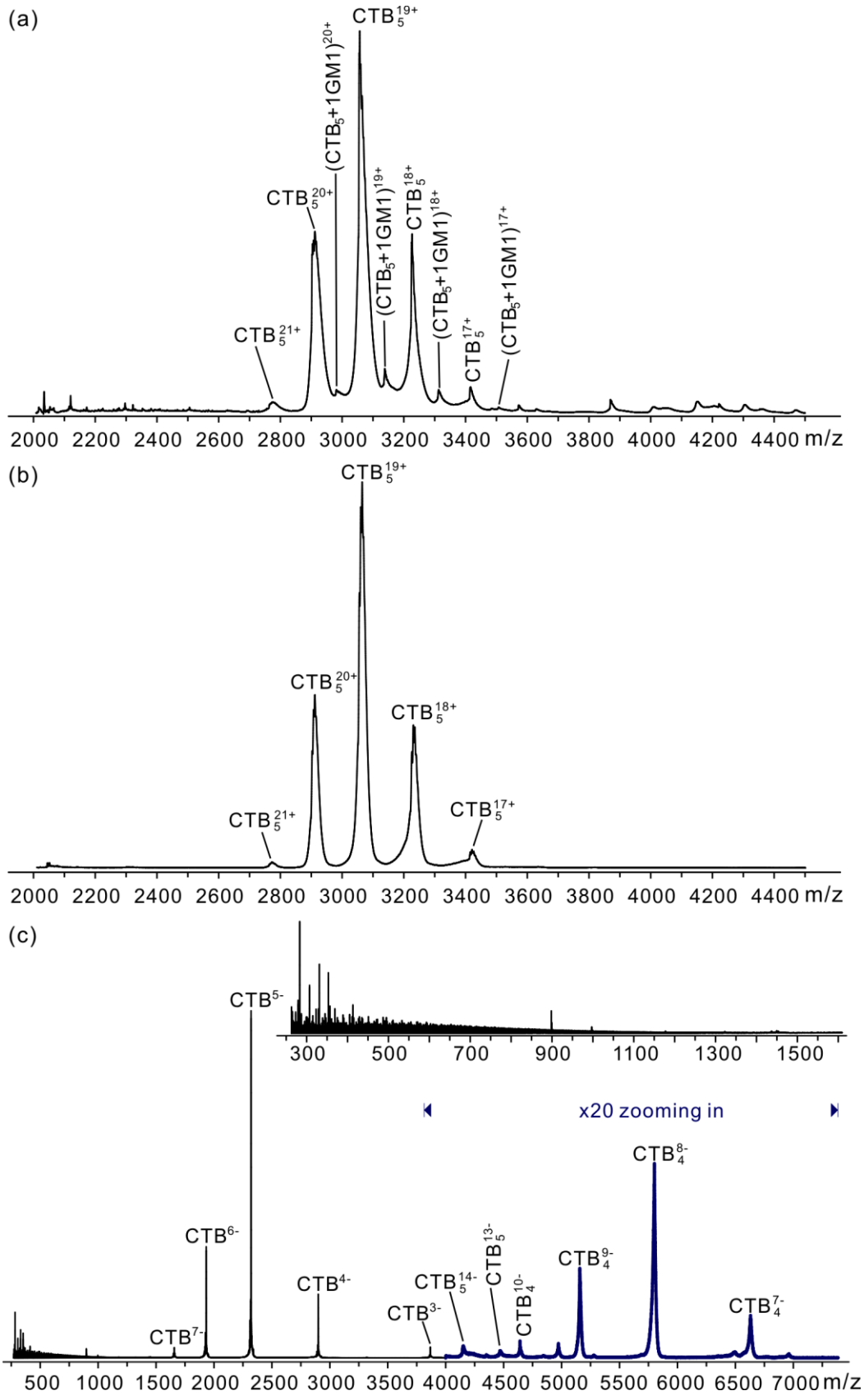

(a)
CTB_5^{19+}
$(CTB_5+1GM1)^{20+}$
CTB_5^{20+}
$(CTB_5+1GM1)^{19+}$
CTB_5^{18+}
$(CTB_5+1GM1)^{18+}$
CTB_5^{17+}
$(CTB_5+1GM1)^{17+}$
CTB_5^{21+}
2000 2200 2400 2600 2800 3000 3200 3400 3600 3800 4000 4200 4400 m/z
(b)
CTB_5^{19+}
CTB_5^{20+}
CTB_5^{18+}
CTB_5^{17+}
CTB_5^{21+}
2000 2200 2400 2600 2800 3000 3200 3400 3600 3800 4000 4200 4400 m/z
(c)
CTB^{5-}
300 500 700 900 1100 1300 1500
x20 zooming in
CTB^{8-}_4
CTB^{6-}
CTB^{13-}_5
CTB^{9-}_4
CTB^{4-}
CTB^{14-}_5
CTB^{10-}_4
CTB^{3-}
CTB^{7-}_4
CTB^{7-}
500 1000 1500 2000 2500 3000 3500 4000 4500 5000 5500 6000 6500 7000 m/z

Figure S11. ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 (5 μM) and 0.5% GM1 ND (3 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa: (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa). (c) CID mass spectrum acquired in negative ion mode for ions of m/z >2,500 produced in (b) using a collision energy of 120 V in the Trap.

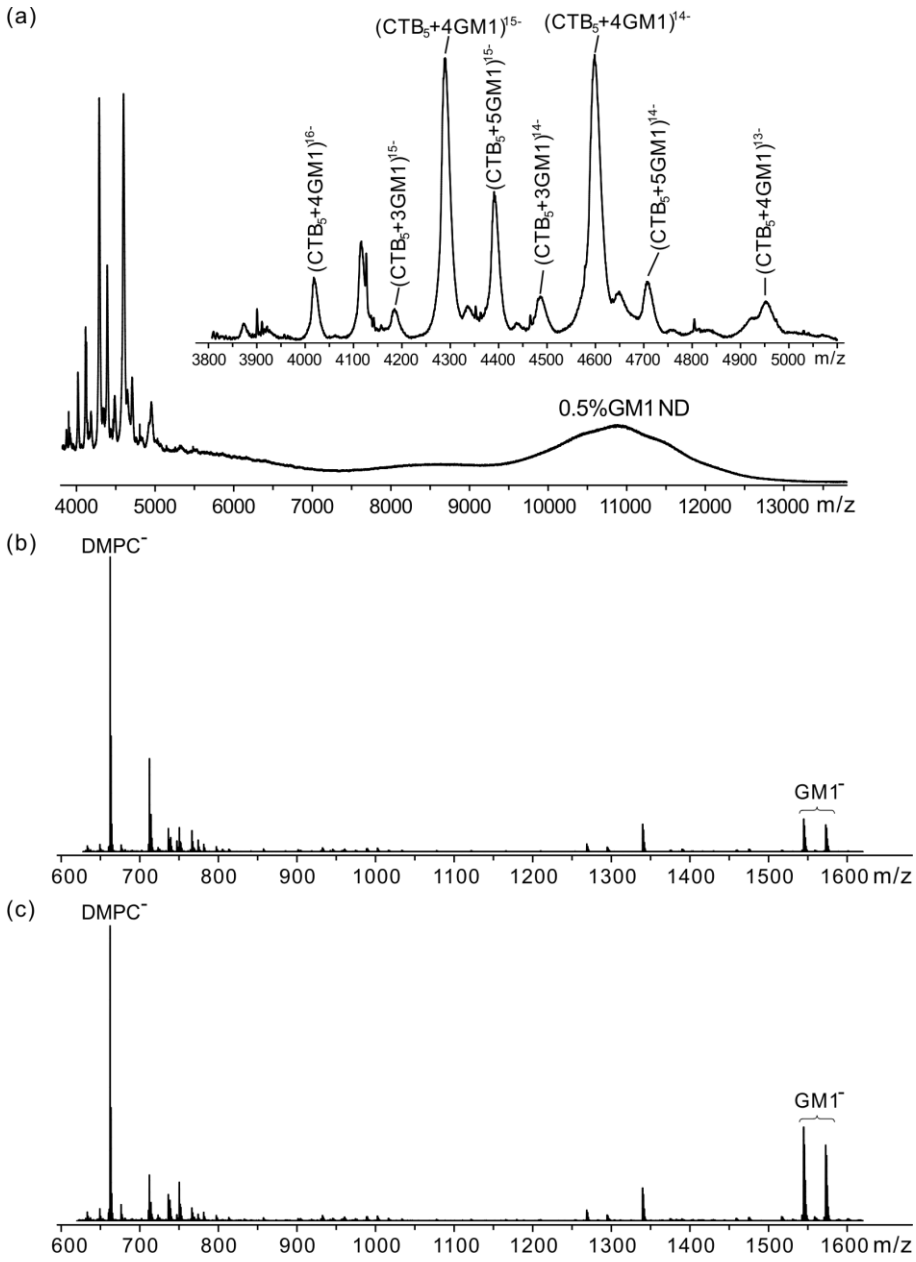


Figure S12. (a) ESI mass spectrum acquired in negative ion mode for aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of CTB_5 (3 μM) and 0.5% GM1 ND (14 μM). (b) CID mass spectrum for ions produced in (a) and centred at m/z 11,000 (which correspond to GM1 ND). (c) CID mass spectrum for ions centred at m/z 11,000 produced from aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 0.5% GM1 ND (14 μM). A collision energy of 200 V in Trap was used for the CID experiments.

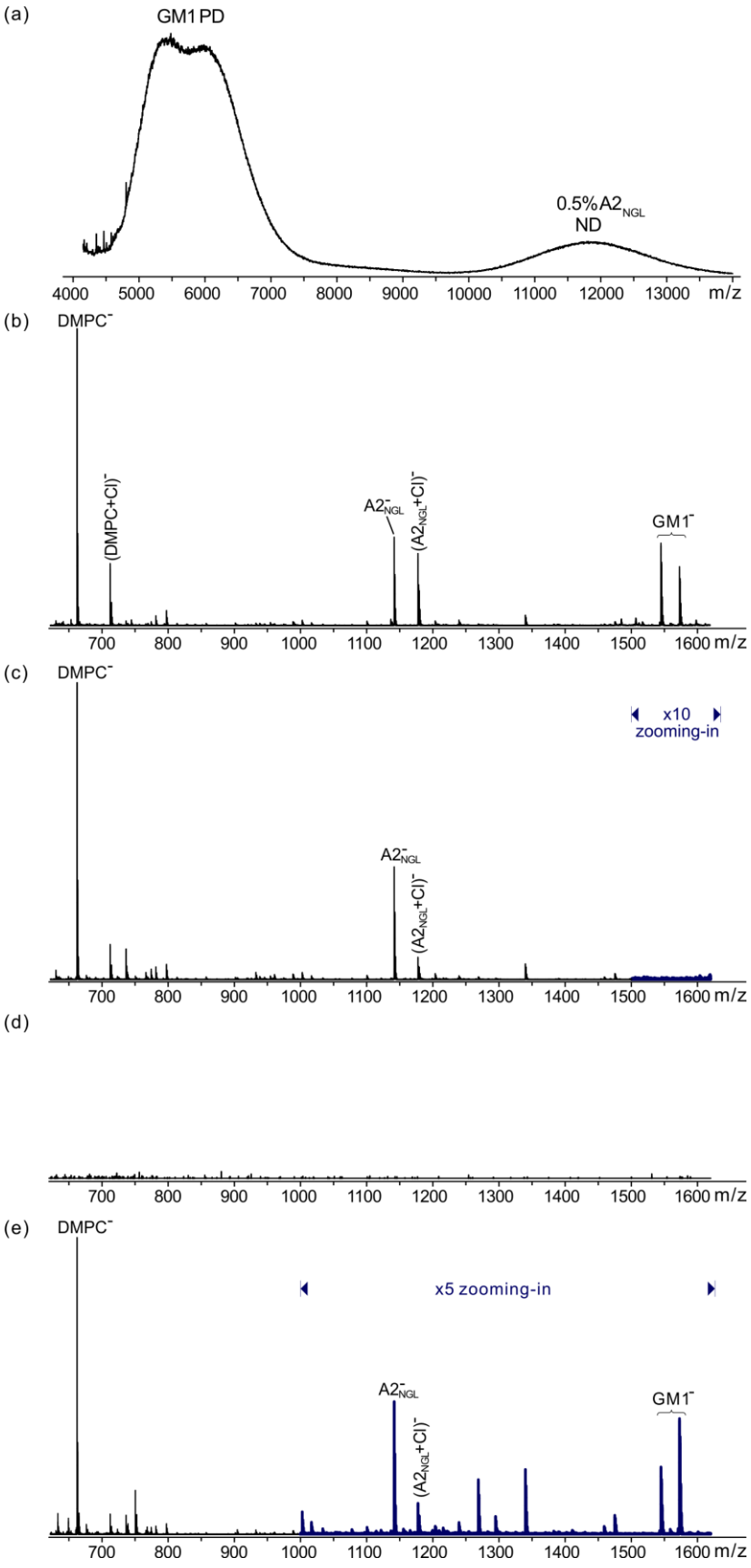

(a)
GM1 PD
0.5%A2NGL
ND
4000 5000 6000 7000 8000 9000 10000 11000 12000 13000 m/z
(b)
DMPC-
(DMPC+Cl)-
A2-NGL
(A2NGL+Cl)-
GM1-
700 800 900 1000 1100 1200 1300 1400 1500 1600 m/z
(c)
DMPC-
x10
zooming-in
A2-NGL
(A2NGL+Cl)-
700 800 900 1000 1100 1200 1300 1400 1500 1600 m/z
(d)
700 800 900 1000 1100 1200 1300 1400 1500 1600 m/z
(e)
DMPC-
x5 zooming-in
A2-NGL
(A2NGL+Cl)-
GM1-
700 800 900 1000 1100 1200 1300 1400 1500 1600 m/z

Figure S13. (a) ESI mass spectrum acquired in negative ion mode for an aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of 0.5% $A2_{NGL}$ ND (12 μM) and picodisc (PD, 54 μM) containing POPC and GM1. (b) – (e) CID mass spectra acquired for ions centred at m/z 11,500 produced from: (b) the same solution as in (a); aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of (c) 0.5% $A2_{NGL}$ ND (12 μM); (d) PD (54 μM) containing POPC and GM1; and (e) 0.5% $A2_{NGL}$ ND (12 μM) and 0.5% GM1 ND (12 μM). For all measurements a collision energy of 200 V in the Trap was used.

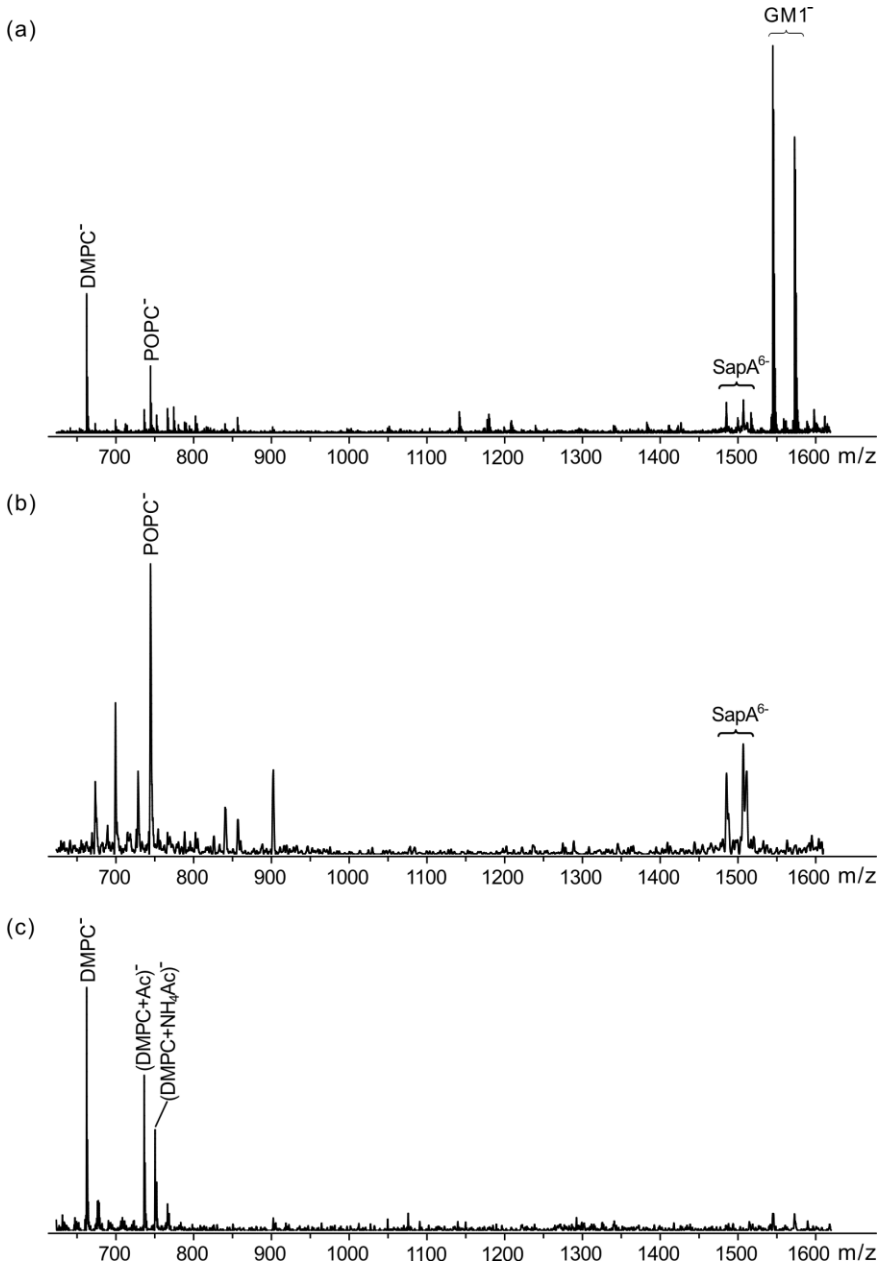


Figure S14. CID mass spectra acquired in negative ion mode for ions centred at m/z 5,500 produced by ESI from an aqueous ammonium acetate solution (200 mM, 25 °C and pH 6.8) of (a) 1% GM1 ND (16 μM) with picodisc (PD, 60 μM) containing POPC; (b) PD (60 μM) containing POPC; and (c) 1% GM1 ND (16 μM). For all measurements, a collisional energy of 120 V in the Trap was used.

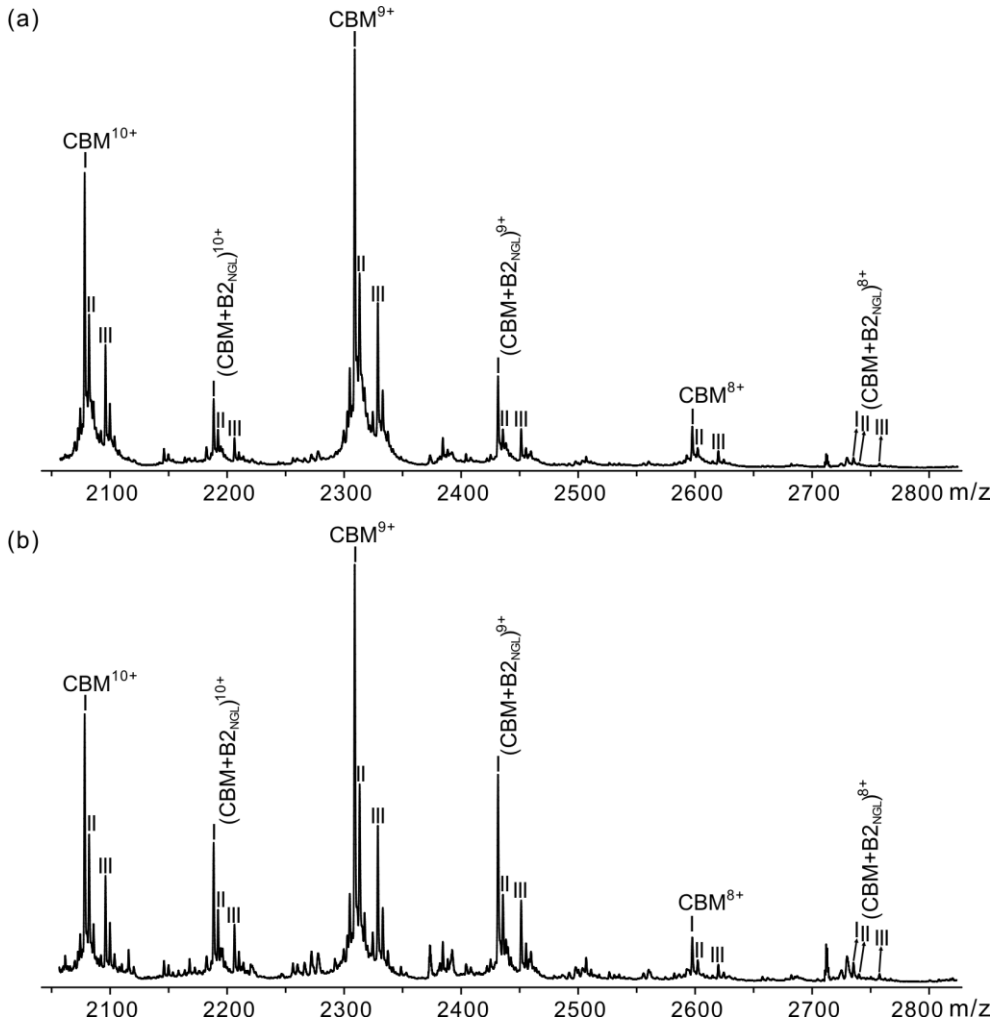


Figure S15. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 µM) and (a) 2.5% $B2_{NGL}$ ND (14 µM) (corresponding to 70 µM $B2_{NGL}$) or (b) 2.5% $B2_{NGL}$ ND (30.8 µM) (corresponding to 154 µM $B2_{NGL}$).

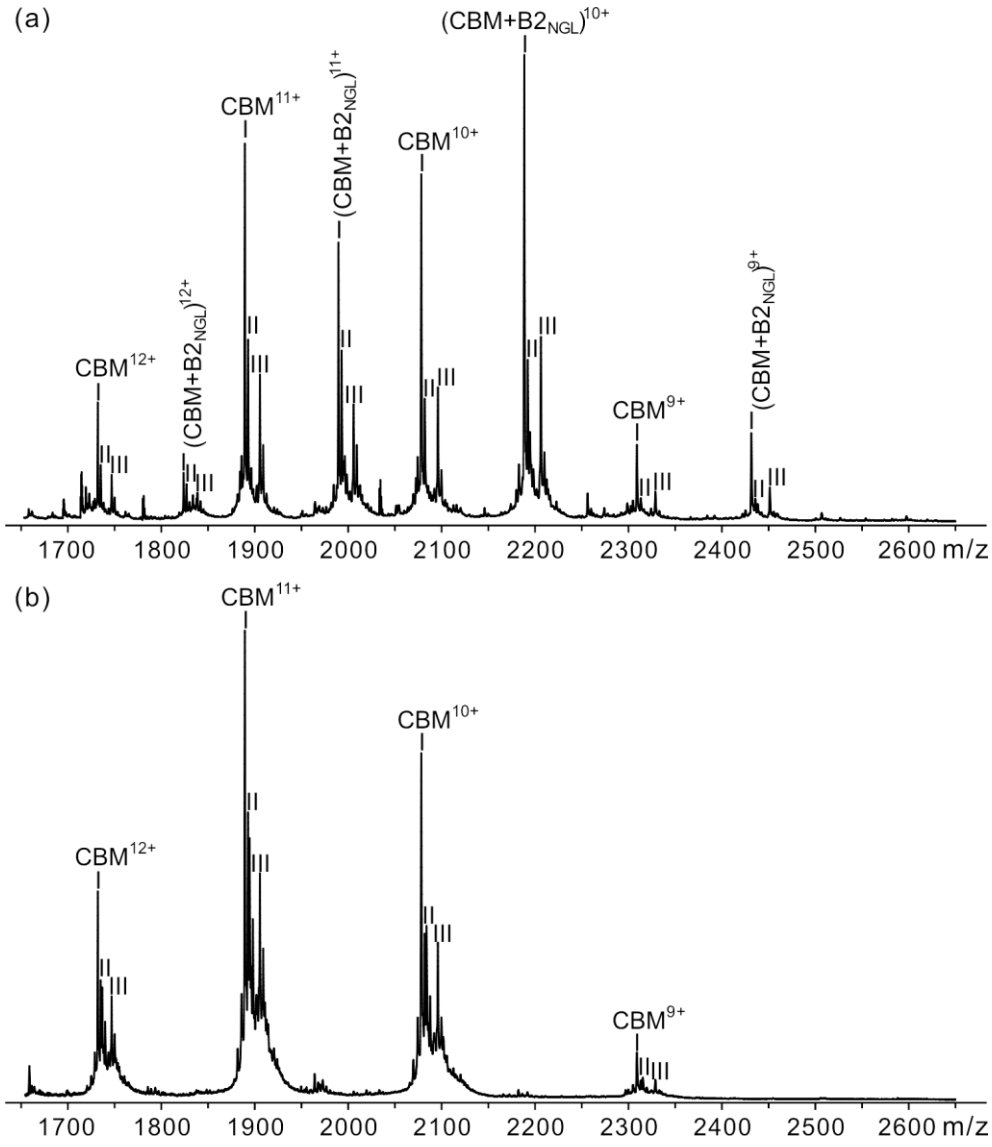


Figure S16. ESI mass spectra acquired in positive ion mode for an aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 μM) and 10% $B2_{NGL}$ ND (25 μM) subjected to ultracentrifugation using a filter with a MWCO of 100 kDa: (a) supernatant solution (MW ≥100 kDa) and (b) filtrate solution (MW ≤100 kDa).

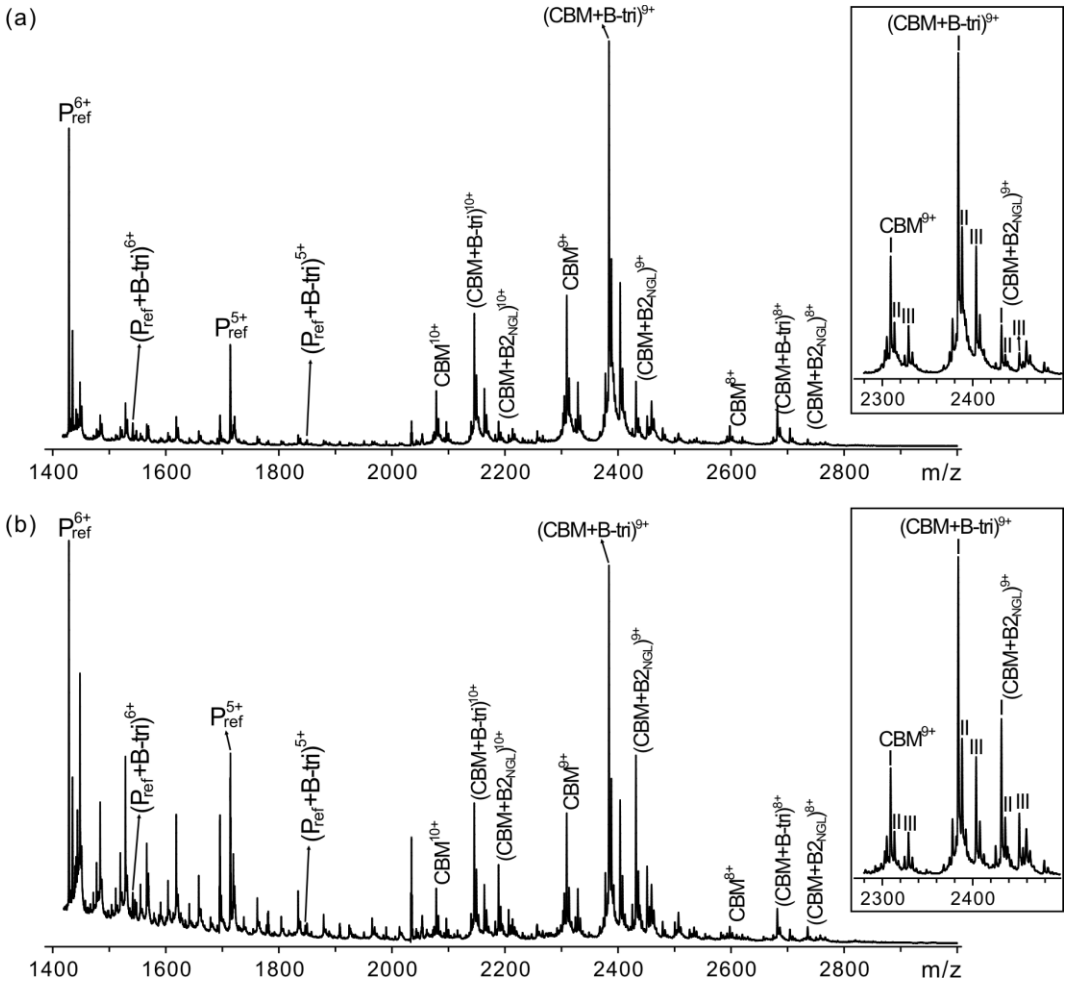


Figure S17. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CBM (12 µM), B-tri (L_{proxy}, 40 µM) with (a) 7 µM and (b) 21 µM 10% $B2_{NGL}$ ND (corresponding to 140 µM and 420 µM $B2_{NGL}$, respectively). 5 µM P_{ref} (Ubq) was added to each solution to correct for the nonspecific ligand binding during ESI process.

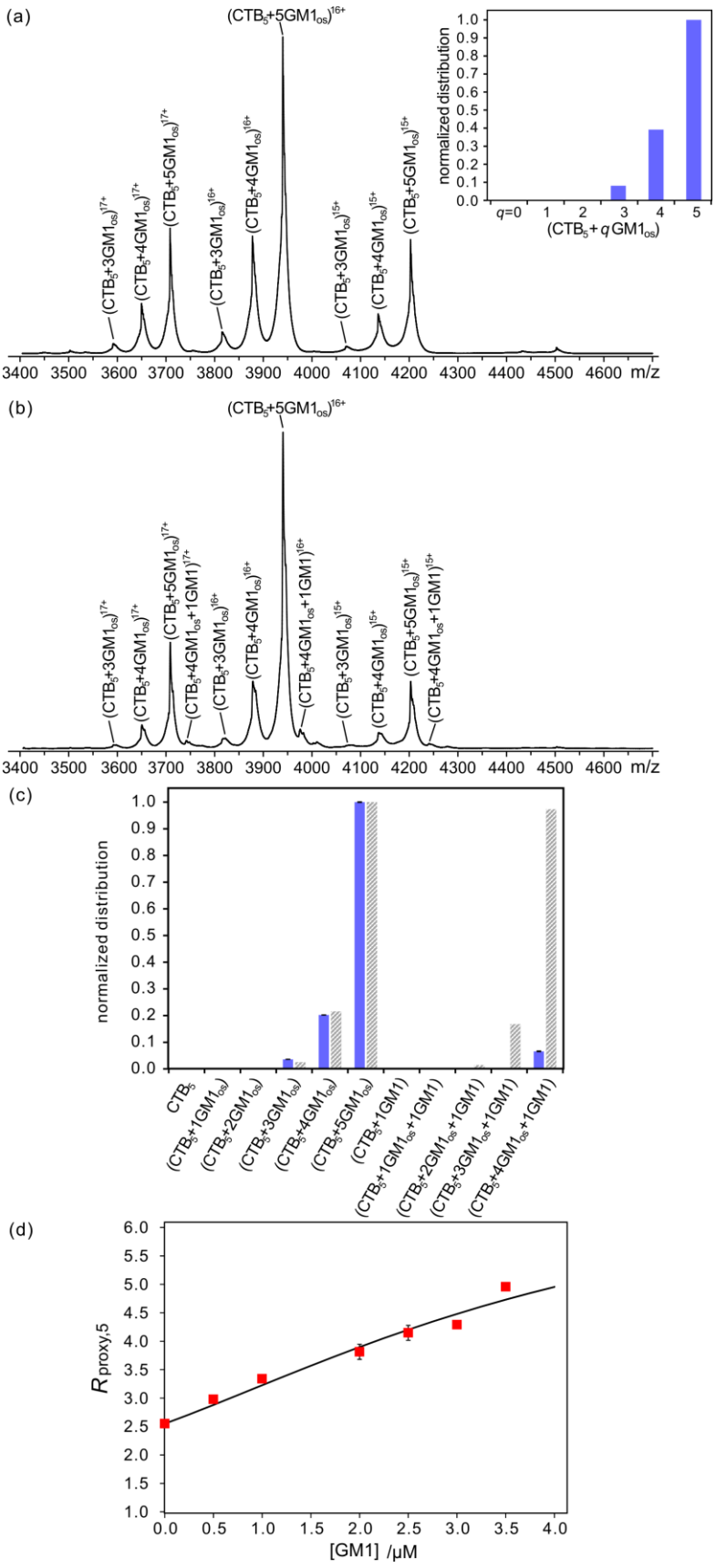
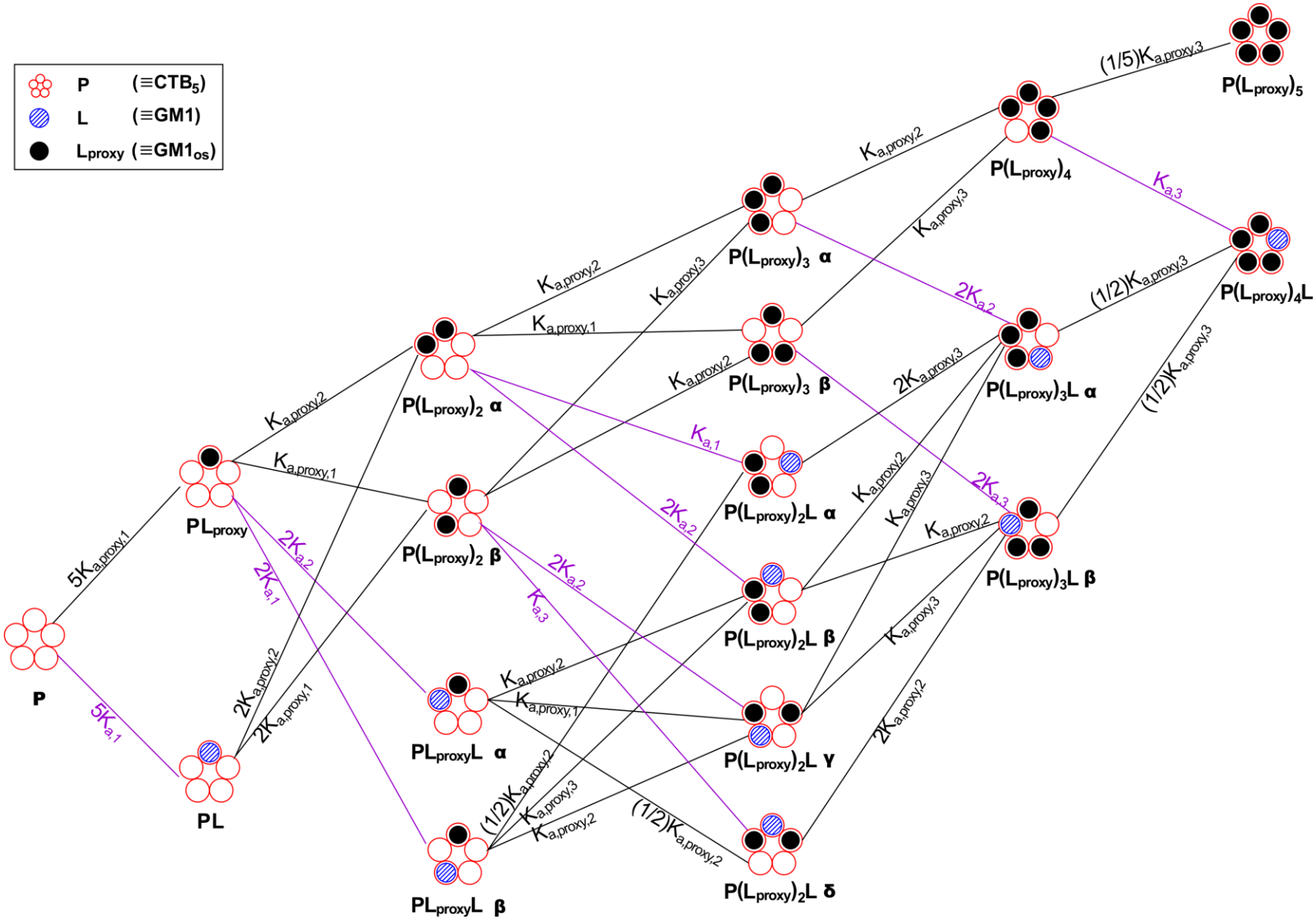

(a)
(CTB5+5GM1os)16+
(CTB5+3GM1os)17+
(CTB5+4GM1os)17+
(CTB5+5GM1os)17+
(CTB5+3GM1os)16+
(CTB5+4GM1os)16+
(CTB5+3GM1os)15+
(CTB5+4GM1os)15+
(CTB5+5GM1os)15+
normalized distribution
q=0 1 2 3 4 5
(CTB5+qGM1os)
3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 m/z
(b)
(CTB5+5GM1os)16+
(CTB5+3GM1os)17+
(CTB5+4GM1os)17+
(CTB5+5GM1os)17+
(CTB5+4GM1os+1GM1)17+
(CTB5+3GM1os)16+
(CTB5+4GM1os)16+
(CTB5+4GM1os+1GM1)16+
(CTB5+3GM1os)15+
(CTB5+4GM1os)15+
(CTB5+5GM1os)15+
(CTB5+4GM1os+1GM1)15+
3400 3500 3600 3700 3800 3900 4000 4100 4200 4300 4400 4500 4600 m/z
(c)
normalized distribution
CTB5
(CTB5+1GM1os)
(CTB5+2GM1os)
(CTB5+3GM1os)
(CTB5+4GM1os)
(CTB5+5GM1os)
(CTB5+1GM1)
(CTB5+1GM1os+1GM1)
(CTB5+2GM1os+1GM1)
(CTB5+3GM1os+1GM1)
(CTB5+4GM1os+1GM1)
(d)
Rproxy,5
[GM1] /μM

Figure S18. ESI mass spectra acquired in positive ion mode for aqueous ammonium acetate solutions (200 mM, 25 °C and pH 6.8) of CTB_5 (4.0 µM) and $GM1_{os}$ (20 µM) with (a) 0 µM and (b) 1.75 µM 1% GM1 ND (corresponding to 3.5 µM GM1). Insets show the normalized distributions of free and $GM1_{os}$-bound CTB_5. (c) (■) Normalized distributions of free and ligand-bound CTB_5 measured from mass spectrum in (b); (▨) theoretical distributions calculated from the association constants reported for the stepwise binding of $GM1_{os}$ (from reference S10) and 1% GM1 ND (determined in the present study) to CTB_5. (d) Plot of $R_{proxy,5}$ (≡ $Ab(CTB_5+5\ GM1_{os})/Ab(CTB_5+4\ GM1_{os})$) versus GM1 concentration. The experimental conditions of are same as in (a) and (b), but with addition of 0 – 1.75 µM 1% GM1 ND. The error bars correspond to one standard deviation.


P (≡CTB5)
L (≡GM1)
Lproxy (≡GM1os)
P
PL
PLproxy
PLproxyL α
PLproxyL β
P(Lproxy)2 α
P(Lproxy)2 β
P(Lproxy)2L α
P(Lproxy)2L β
P(Lproxy)2L γ
P(Lproxy)2L δ
P(Lproxy)3 α
P(Lproxy)3 β
P(Lproxy)3L α
P(Lproxy)3L β
P(Lproxy)4
P(Lproxy)4L
P(Lproxy)5
5Ka,proxy,1
5Ka,1
2Ka,proxy,2
2Ka,proxy,1
2Ka,1
2Ka,2
Ka,3
Ka,proxy,2
Ka,proxy,1
Ka,proxy,3
Ka,1
Ka,2
2Ka,proxy,3
2Ka,3
(1/2)Ka,proxy,2
(1/2)Ka,proxy,3
(1/5)Ka,proxy,3

Scheme S1. Graphical representation of possible protein-ligand interactions involving a monovalent ligand (L), present in a ND, and a monovalent proxy ligand (L_{proxy}) that binding competitively with a protein (P) with five binding sites. Based on the experimental data described in Figures 7b and S16b, only one L is bound to P under the experimental conditions used in this study. The binding model is based on the cooperative model for the stepwise binding of $GM1_{os}$ to CTB_5 proposed by Homans,[S11] wherein ligand affinity is enhanced when a neighbouring binding site is occupied. The α, β, γ and δ labels are used to distinguish the positional isomers. The intrinsic association constants ($K_{a,proxy,1}$, $K_{a,proxy,2}$, and $K_{a,proxy,3}$ and $K_{a,1}$, $K_{a,2}$ and $K_{a,3}$,), along with the statistic coefficients, are given for each interaction.

Application of *proxy ligand* ESI-MS method to quantify CTB_5-GM1 ND interactions

The binding model (Scheme 1), which is an extension of the Homans cooperative binding model established for the stepwise binding of $GM1_{os}$ to CTB_5,[S11] treats both $GM1_{os}$ (L_{proxy}) and GM1 in NDs (L) as monovalent ligands capable of interacting at any of the five binding sites of CTB_5 (P). Based on experimental observations, the model is restricted to the case where, at most, a single GM1 binds but up to five $GM1_{os}$ can bind. Notably, the binding $GM1_{os}$ and the GM1 are described by three association constants ($K_{a,proxy,1}$, $K_{a,proxy,2}$, and $K_{a,proxy,3}$ and $K_{a,1}$, $K_{a,2}$ and $K_{a,3}$, respectively), which reflect the dependence of the affinity on the number of neighbouring subunits that are bound to ligand. According to a recent ESI-MS study,[S10] the binding of $GM1_{os}$ is increased by a factor of 1.7 when one neighbouring binding site is occupied and by a factor of 2.9 when both neighbouring sites are occupied:

$$K_{a,proxy,2} = 1.7K_{a,proxy,1} \quad \text{(S13a)}$$

$$K_{a,proxy,3} = 2.9K_{a,proxy,1} \quad \text{(S13b)}$$

A similar enhancement was assumed for GM1 ND binding:

$$K_{a,2} = 1.7K_{a,1} \quad \text{(S14a)}$$

$$K_{a,3} = 2.9K_{a,1} \quad \text{(S14b)}$$

A summary of all the possible binding interactions, along with the corresponding statistical factors, is given in Scheme 1. Based on this model, the equations of mass balance are given by eqs. S15–S17:

$$[P]_0 = [P] + [PL_{proxy}] + [P(L_{proxy})_2\alpha] + [P(L_{proxy})_2\beta] + [P(L_{proxy})_3\alpha] + [P(L_{proxy})_3\beta] + [P(L_{proxy})_4] + [P(L_{proxy})_5] + [PL] + [PL_{proxy}L\alpha] + [PL_{proxy}L\beta] + [P(L_{proxy})_2L\alpha] + [P(L_{proxy})_2L\beta] + [P(L_{proxy})_2L\gamma] + [P(L_{proxy})_2L\delta] + [P(L_{proxy})_3L\alpha] + [P(L_{proxy})_3L\beta] + [P(L_{proxy})_4L] \quad \text{(S15)}$$

$$[L_{proxy}]_0 = [L_{proxy}] + [PL_{proxy}] + 2\times[P(L_{proxy})_2\alpha] + 2\times[P(L_{proxy})_2\beta] + 3\times[P(L_{proxy})_3\alpha] + 3\times[P(L_{proxy})_3\beta] + 4\times[P(L_{proxy})_4] + 5\times[P(L_{proxy})_5] + [PL_{proxy}L\alpha] + [PL_{proxy}L\beta] + 2\times[P(L_{proxy})_2L\alpha] + 2\times[P(L_{proxy})_2L\beta] + 2\times[P(L_{proxy})_2L\gamma] + 2\times[P(L_{proxy})_2L\delta] + 3\times[P(L_{proxy})_3L\alpha] + 3\times[P(L_{proxy})_3L\beta] + 4\times[P(L_{proxy})_4L] \quad \text{(S16)}$$

$$[L]_0 = [L] + [PL] + [PL_{proxy}L\alpha] + [PL_{proxy}L\beta] + [P(L_{proxy})_2L\alpha] + [P(L_{proxy})_2L\beta] + [P(L_{proxy})_2L\gamma] + [P(L_{proxy})_2L\delta] + [P(L_{proxy})_3L\alpha] + [P(L_{proxy})_3L\beta] + [P(L_{proxy})_4L] \quad \text{(S17)}$$

where $[P]_0$, $[L_{proxy}]_0$ and $[L]_0$ are the initial concentrations of P, L_{proxy} and L in solution, respectively. The relevant equilibrium expressions for the binding interactions are given by eqs S18–S34:

$$\frac{1}{5K_{a,proxy,1}} = \frac{[P][L_{proxy}]}{[PL_{proxy}]} \quad \text{(S18)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[PL_{proxy}][L_{proxy}]}{[P(L_{proxy})_2\alpha]} \quad \text{(S19)}$$

$$\frac{1}{K_{a,proxy,1}} = \frac{[PL_{proxy}][L_{proxy}]}{[P(L_{proxy})_2\beta]} \quad \text{(S20)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2\alpha][L_{proxy}]}{[P(L_{proxy})_3\alpha]} \quad \text{(S21a)}$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_2\beta][L_{proxy}]}{[P(L_{proxy})_3\alpha]} \quad \text{(S21b)}$$

$$\frac{1}{K_{a,proxy,1}} = \frac{[P(L_{proxy})_2\alpha][L_{proxy}]}{[P(L_{proxy})_3\beta]} \quad \text{(S22a)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2\beta][L_{proxy}]}{[P(L_{proxy})_3\beta]} \quad \text{(S22b)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_3\alpha][L_{proxy}]}{[P(L_{proxy})_4]} \quad \text{(S23a)}$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_3\beta][L_{proxy}]}{[P(L_{proxy})_4]} \quad \text{(S23b)}$$

$$\frac{5}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_4][L_{proxy}]}{[P(L_{proxy})_5]} \quad \text{(S24)}$$

$$\frac{1}{5K_{a,1}} = \frac{[P][L]}{[PL]} \quad \text{(S25)}$$

$$\frac{1}{2K_{a,2}} = \frac{[PL_{proxy}][L]}{[PL_{proxy}L\alpha]} \quad \text{(S26a)}$$

$$\frac{1}{2K_{a,proxy,2}} = \frac{[PL][L_{proxy}]}{[PL_{proxy}L\alpha]} \quad \text{(S26b)}$$

$$\frac{1}{2K_{a,1}} = \frac{[PL_{proxy}][L]}{[PL_{proxy}L\beta]} \quad \text{(S27a)}$$

$$\frac{1}{2K_{a,proxy,1}} = \frac{[PL][L_{proxy}]}{[PL_{proxy}L\beta]} \quad \text{(S27b)}$$

$$\frac{1}{K_{a,1}} = \frac{[P(L_{proxy})_2\alpha][L]}{[P(L_{proxy})_2L\alpha]} \quad \text{(S28a)}$$

$$\frac{2}{K_{a,proxy,2}} = \frac{[PL_{proxy}L\beta][L_{proxy}]}{[P(L_{proxy})_2L\alpha]} \quad \text{(S28b)}$$

$$\frac{1}{2K_{a,2}} = \frac{[P(L_{proxy})_2\alpha][L]}{[P(L_{proxy})_2L\beta]} \quad \text{(S29a)}$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[PL_{proxy}L\beta][L_{proxy}]}{[P(L_{proxy})_2L\beta]} \quad \text{(S29b)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[PL_{proxy}L\alpha][L_{proxy}]}{[P(L_{proxy})_2L\beta]} \quad \text{(S29c)}$$

$$\frac{1}{2K_{a,2}} = \frac{[P(L_{proxy})_2\beta][L]}{[P(L_{proxy})_2L\gamma]} \quad \text{(S30a)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[PL_{proxy}L\beta][L_{proxy}]}{[P(L_{proxy})_2L\gamma]} \quad \text{(S30b)}$$

$$\frac{1}{K_{a,proxy,1}} = \frac{[PL_{proxy}L\alpha][L_{proxy}]}{[P(L_{proxy})_2L\gamma]} \quad \text{(S30c)}$$

$$\frac{1}{K_{a,3}} = \frac{[P(L_{proxy})_2\beta][L]}{[P(L_{proxy})_2L\delta]} \quad \text{(S31a)}$$

$$\frac{2}{K_{a,proxy,2}} = \frac{[PL_{proxy}L\alpha][L_{proxy}]}{[P(L_{proxy})_2L\delta]} \quad \text{(S31b)}$$

$$\frac{1}{2K_{a,2}} = \frac{[P(L_{proxy})_3\alpha][L]}{[P(L_{proxy})_3L\alpha]} \quad \text{(S32a)}$$

$$\frac{1}{2K_{a,proxy,3}} = \frac{[P(L_{proxy})_2L\alpha][L_{proxy}]}{[P(L_{proxy})_3L\alpha]} \quad \text{(S32b)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2L\beta][L_{proxy}]}{[P(L_{proxy})_3L\alpha]} \quad \text{(S32c)}$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_2L\gamma][L_{proxy}]}{[P(L_{proxy})_3L\alpha]} \quad \text{(S32d)}$$

$$\frac{1}{2K_{a,3}} = \frac{[P(L_{proxy})_3\beta][L]}{[P(L_{proxy})_3L\beta]} \quad \text{(S33a)}$$

$$\frac{1}{K_{a,proxy,2}} = \frac{[P(L_{proxy})_2 L\beta][L_{proxy}]}{[P(L_{proxy})_3 L\beta]} \quad \text{(S33b)}$$

$$\frac{1}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_2 L\gamma][L_{proxy}]}{[P(L_{proxy})_3 L\beta]} \quad \text{(S33c)}$$

$$\frac{1}{2K_{a,proxy,2}} = \frac{[P(L_{proxy})_2 L\delta][L_{proxy}]}{[PL(L_{proxy})_3 L\beta]} \quad \text{(S33d)}$$

$$\frac{1}{K_{a,3}} = \frac{[P(L_{proxy})_4][L]}{[P(L_{proxy})_4 L]} \quad \text{(S34a)}$$

$$\frac{2}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_3 L\alpha][L_{proxy}]}{[P(L_{proxy})_4 L]} \quad \text{(S34b)}$$

$$\frac{2}{K_{a,proxy,3}} = \frac{[P(L_{proxy})_3 L\beta][L_{proxy}]}{[P(L_{proxy})_4 L]} \quad \text{(S34c)}$$

Given the initial concentrations ($[P]_0$, $[L_{proxy}]_0$ and $[L]_0$) and $K_{a,proxy,1}$, $K_{a,proxy,2}$, and $K_{a,proxy,3}$ values, the equilibrium concentrations of all CTB_5 species can be calculated for a given set of $K_{a,1}$, $K_{a,2}$, and $K_{a,3}$ values. The theoretical distribution can then be compared to the experimental distribution. Because of differences in the ESI-MS response factors for the (CTB_5 + $q$$GM1_{os}$) and ($CTB_5$ + $q$$GM1_{os}$ + GM1) complexes, the $K_{a,1}$, $K_{a,2}$, and $K_{a,3}$ values were found by considering only the abundance ratio $Ab(P(L_{proxy})_5)/Ab(P(L_{proxy})_4)$ ($\equiv R_{proxy,5}$). Optimum $K_{a,1}$, $K_{a,2}$, and $K_{a,3}$ values were found using a least square analysis, where the sum of squares of residuals (*SSR*) between the experimental to theoretical $R_{proxy,5}$ values were minimized, eq S35:

$$SSR = \sum_n \left(R_{proxy,5}(\text{experimental}) - R_{proxy,5}(\text{theoretical}) \right)^2 \quad \text{(S35)}$$

in which n is the total number of data points.

References

(S1) Nath, A.; Atkins, W. M.; Sligar, S. G. *Biochemistry* **2007**, *46*, 2059.

(S2) Bayburt, T. H.; Sligar, S. G. *FEBS Lett.* **2010**, *584*, 1721.

(S3) Popovic, K.; Holyoake, J.; Pomès, R.; Privé, G. G. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 2908.

(S4) Leney, A. C.; Rezaei Darestani, R.; Li, J.; Nikjah, S.; Kitova, E. N.; Zou, C.; Cairo, C. W.; Xiong, Z. J.; Prive, G. G.; Klassen, J. S. *Anal. Chem.* **2015**, in press, DOI: 10.1021/acs.analchem.5b00170.

(S5) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431.

(S6) Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127.

(S7) Zhang, Y. X.; Liu, L.; Daneshfar, R.; Kitova, E. N.; Li, C. S.; Jia, F.; Cairo, C. W.; Klassen, J. S. *Anal. Chem.* **2012**, *84*, 7618.

(S8) Leney, A. C.; Fan, X. X.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2014**, *86*, 5271.

(S9) Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.

(S10) Lin, H.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 104.

(S11) Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1047.