JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

# Glycans Confer Specificity to the Recognition of Ganglioside Receptors by Botulinum Neurotoxin A

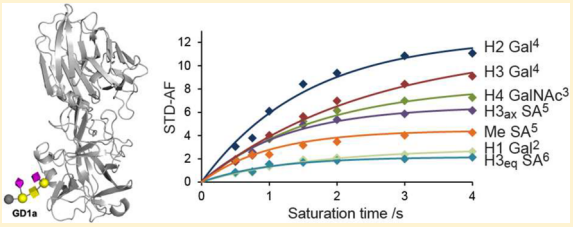
Christoffer Hamark,[†,§] Ronnie P.-A. Berntsson,[‡,§,∥] Geoffrey Masuyer,[‡] Linda M. Henriksson,[‡] Robert Gustafsson,[‡] Pål Stenmark,[*,‡] and Göran Widmalm[*,†]

[†]Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden

[‡]Department of Biochemistry and Biophysics, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden

Supporting Information

**ABSTRACT:** The highly poisonous botulinum neurotoxins, produced by the bacterium *Clostridium botulinum*, act on their hosts by a high-affinity association to two receptors on neuronal cell surfaces as the first step of invasion. The glycan motifs of gangliosides serve as initial coreceptors for these protein complexes, whereby a membrane protein receptor is bound. Herein we set out to characterize the carbohydrate minimal binding epitope of the botulinum neurotoxin serotype A. By means of ligand-based NMR spectroscopy, X-ray crystallography, computer simulations, and isothermal titration calorimetry, a screening of ganglioside analogues together with a detailed characterization of various carbohydrate ligand complexes with the toxin were accomplished. We show that the representation of the glycan epitope to the protein affects the details of binding. Notably, both branches of the oligosaccharide GD1a can associate to botulinum neurotoxin serotype A when expressed as individual trisaccharides. It is, however, the terminal branch of GD1a as well as this trisaccharide motif alone, corresponding to the sialyl-Thomsen−Friedenreich antigen, that represents the active ligand epitope, and these compounds bind to the neurotoxin with a high degree of predisposition but with low affinities. This finding does not correlate with the oligosaccharide moieties having a strong contribution to the total affinity, which was expected to be the case. We here propose that the glycan part of the ganglioside receptors mainly provides abundance and specificity, whereas the interaction with the membrane itself and protein receptor brings about the strong total binding of the toxin to the neuronal membrane.




## INTRODUCTION

*Clostridium botulinum* are pathogenic anaerobic bacteria that produce botulinum neurotoxins (BoNTs). These toxins cause a persistent muscle paralysis, a human disease that is known as botulism. As BoNTs are the most potent toxins known, they are also classified as highly dangerous potential bioterrorism agents.[1,2] However, due to their paralyzing effect, they are also extensively used to treat a multitude of human diseases, as well as in cosmetics.[3,4] Currently, there are seven known serotypes of BoNT, denominated BoNT/A−G.[5−8] Overall, these toxins have the same protein structure and modus operandi, but they differ in their protein sequence as well as in which substrates and receptors they utilize. All BoNTs consist of three domains, viz., the N-terminal proteolytically active light chain (LC), the translocation domain ($H_N$), and the binding domain ($H_C$).

BoNTs are selectively targeted to the neuronal membrane via the so-called double receptor mechanism.[9] The two independent receptors that are targeted are polysialogangliosides (PSG) (Figure 1a) and a membrane protein receptor located in synaptic vesicles.[10−15] PSG are likely used as the initial receptors by BoNT due to their abundance on the presynaptic membrane. Their oligosaccharide, or glycan, part that the BoNTs bind to protrudes out from the membrane and is flexible. A conserved ganglioside binding site (GBS), with an SXWY motif, has been identified in BoNT/A, B, E, F, and G, as well as in the homologous tetanus toxin.[16−23] BoNT/C and D and DC have analogous sites for ganglioside binding at a similar position.[24−28] The second receptor, which BoNT binds to after the initial binding to PSG, varies between the BoNT serotypes. BoNT/B, DC, and G bind to synaptotagmin I and II.[10,21,24,29] BoNT/A and E have been shown to utilize SV2 as a receptor.[30−34] It has also been implicated that BoNT/A can utilize another protein as its protein receptor, namely, FGFR3.[35] The binding domain ($H_C$) of BoNTs is the carbohydrate recognition domain (CRD),[36] and as such, it can be classified as a lectin.[37] The GBS of $H_C$ in serotype A shares many of the common features of this class of proteins, viz., high abundance of aromatic amino acid residues, with the ability to interact with the carbohydrate ligand through CH/$\pi$ stacking.[38]

In this study, we have focused on the PSG binding to BoNT/A. It has previously been demonstrated that PSG are critical for BoNT/A toxicity. In cells that do not synthesize complex PSG, BoNT/A cannot enter the cell and is thus not activated.[39] The structure of BoNT/A-$H_C$ complexed to the glycan part of GT1b, which is the PSG with the highest affinity,[39] has previously been solved.[22] However, there is a series of PSG, of

**Received:** September 14, 2016
**Published:** December 13, 2016





DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

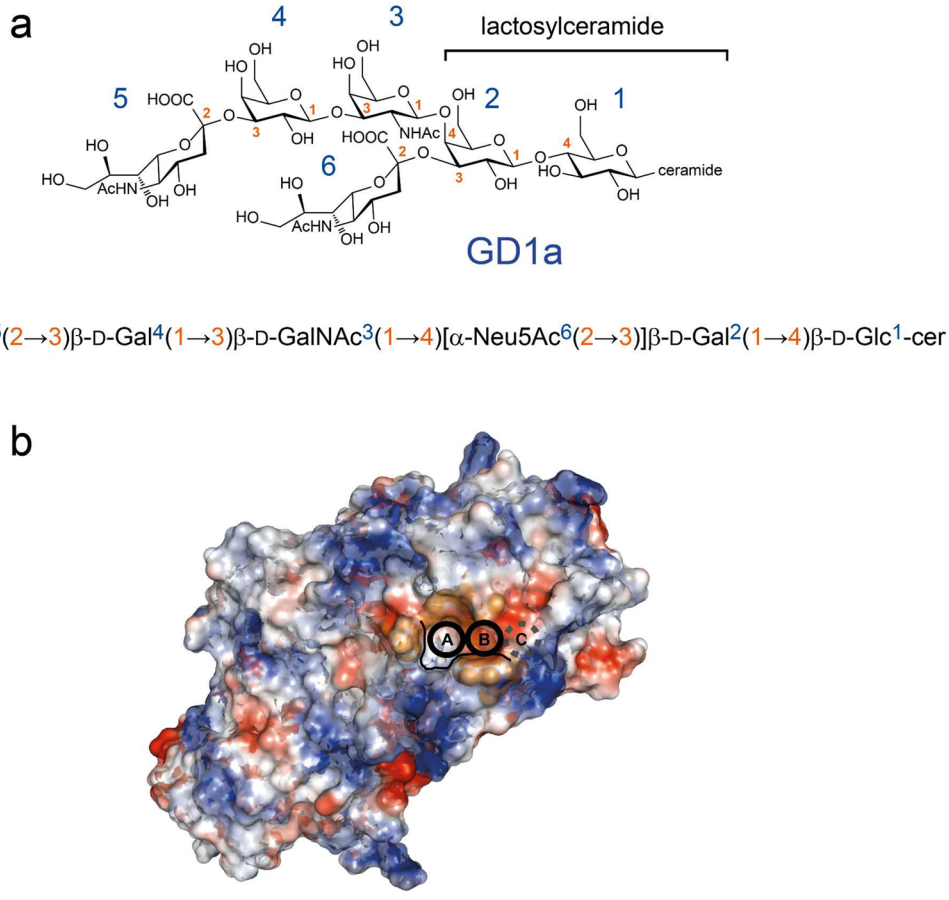



**Figure 1.** GD1a and GBS of BoNT/A-$H_C$. (a) Schematic representation of ganglioside structure exemplified for GD1a. Gangliosides are glycosphingolipids, and the common feature of all gangliosides is the lactosyl ceramide core-structure as well as different degrees of sialylation. The glycosidic linkages are highlighted, and the residue numbering is included. The same systematic numbering is employed for all compounds of this study. (b) Electrostatic surface representation of BoNT/A-$H_C$. The ganglioside binding site (GBS) with its defined subsites (A–C) is highlighted.

different complexity, to which BoNT/A can bind. Via a combination of X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, isothermal titration calorimetry (ITC), and computational methods, the sugar moieties of gangliosides were investigated with the aim of elucidating the structural prerequisites of their binding to BoNT/A. A systematic protocol was devised (Figure 2) where the strengths and complementarities of the engaged techniques were exploited. Ligand-based NMR spectroscopy, herein applied in the form of saturation transfer difference (STD) NMR[40] (for recent reviews, see refs 41–43) together with transferred nuclear Overhauser effect spectroscopy (trNOESY)[44,45] (for recent reviews, see refs 46 and 47), is a viable approach for studying transient receptor–ligand complexes and has successfully been employed in studies of carbohydrate–protein binding in solution.[48–50]

## RESULTS AND DISCUSSION

**Protocol—Ligands and Methodology.** With reference to the published BoNT/A-$H_C$•GT1b complex,[22] the similar oligosaccharide GD1a, also a reported BoNT/A binder and only lacking a nonparticipating sialic acid residue compared with the former,[51] was chosen as a benchmark ligand herein. On the basis of this GD1a template, the hexasaccharide and oligosaccharides thereof were selected to constitute the compound library to be analyzed in association with BoNT/A (Figure 3). Among the ganglioside derivatives in the

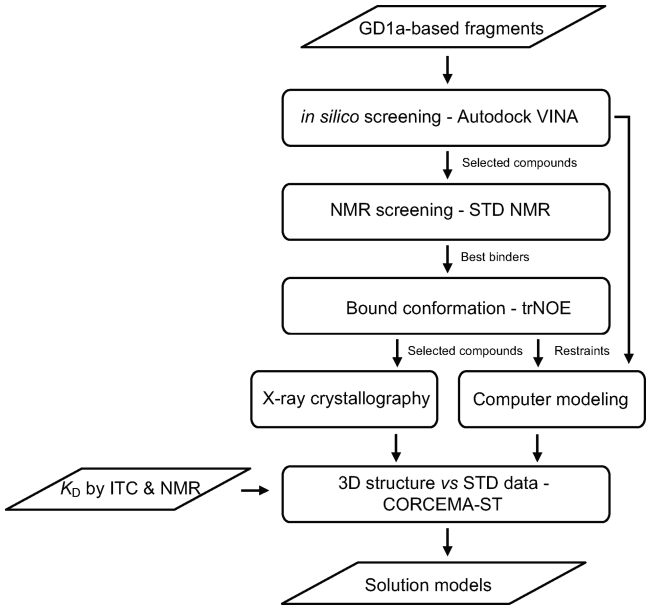



**Figure 2.** Flowchart for the generation of BoNT/A-$H_C$•ganglioside solution models. Parallelograms represent input/output entries, and rectangular shapes denote processing steps.

compound selection, the tumor-associated Thomsen–Friedenreich carbohydrate antigen (T) and its sialylated analogue


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218–230

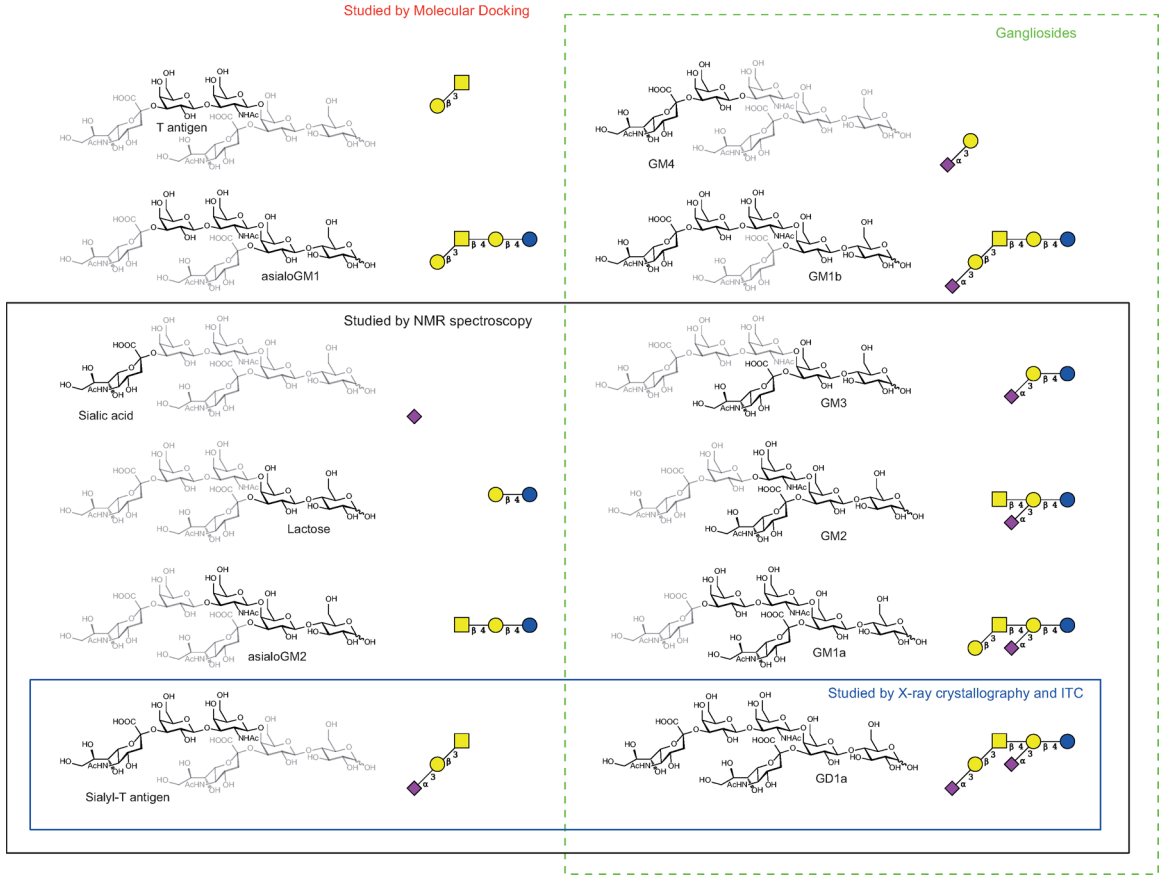



**Figure 3.** Molecular structure and CFG representation of the compounds of this study. They all represent fragments of the glycan part of GD1a, the ganglioside without its sphingolipid part. Different sets and classifications of the ligands are indicated by differently colored frames.

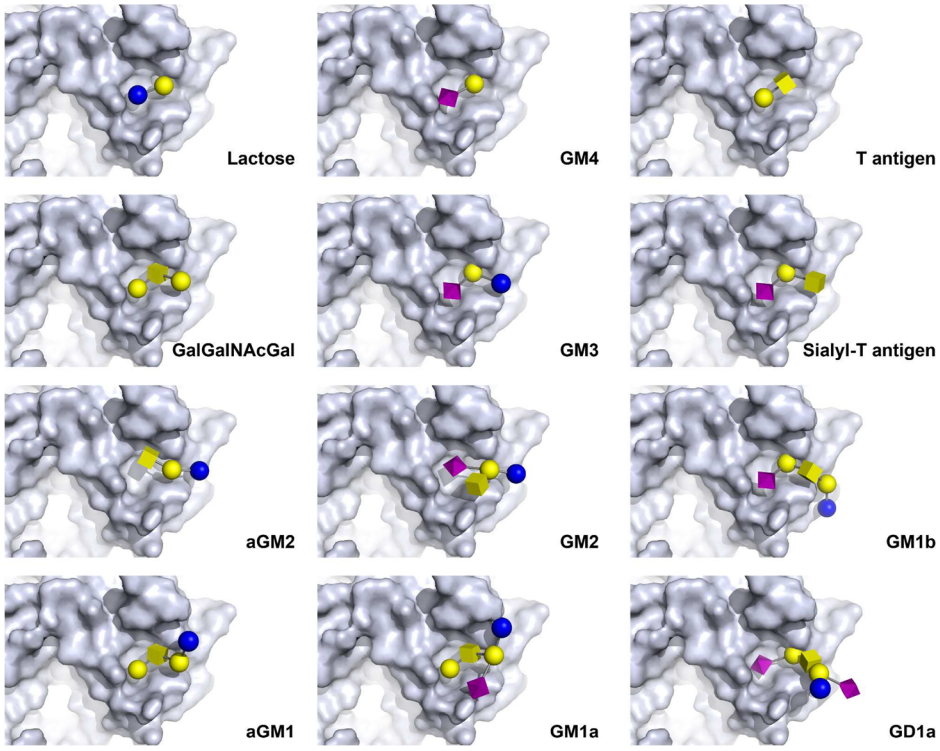



**Figure 4.** Representative output for each studied ligand of the BoNT/A-$H_C$•ligand models from molecular docking simulations with the Autodock VINA software. The ligands are represented as 3D-CFG symbols.[91] The GD1a-containing complex was obtained from a redocking.



DOI: 10.1021/jacs.6b09534
*J. Am. Chem. Soc.* 2017, 139, 218−230

(sialyl-T) should be noted.[52] Employing GD1a and fragments thereof, an *in silico* screening on the full set of carbohydrate compounds was initially carried out (Figure 2) in order to acquire preliminary information on potential binding poses. This was followed by an experimental NMR screening with eight representative compounds. The number of ligands was subsequently reduced further to comprise a small number of compounds that were analyzed in detail using STD and trNOESY experiments, X-ray crystallography/computer modeling, quantitative STD analysis based on equilibrium constants, and generated molecular models resulting in solution state models with respect to their BoNT/A GBS interactions.

**Molecular Docking.** Molecular docking simulations on BoNT/A-$H_C$ were performed with the compound library employing Autodock VINA (ADV),[53] a program previously used for docking of carbohydrate ligands.[54,55] Protein coordinates were extracted from the BoNT/A-$H_C$•GT1b complex (PDB ID: 2VU9),[22] and a restricted search-space was centered at the GBS. Due to the fact that water molecules from the input crystal structure could potentially hinder interactions between the structurally diverse ligands and the protein, simulations were performed without water present. Noninteracting residues of the ligand will thus display a typical so-called vacuum effect, striving to interact artificially with the protein surface instead of pointing out in the bulk.[54] The resulting poses from the molecular docking should therefore be analyzed with care and only with respect to the interacting residues, or minimal binding determinant, within each ligand. Limited scoring success has been reported for ADV with ligands exceeding 20 active torsions, approximately corresponding to a tetrasaccharide.[53] Indeed, an increased spread in both accommodated subsites and conformational space was observed for the larger ligands. Nevertheless, reasonable clusters of poses were obtained among the 10 highest ranked output structures for all studied ligands except for GD1a. The largest docked ligand GD1a, assigned 46 flexible torsions by the program, clearly exceeds the productive limit with respect to the degrees of freedom using ADV. In order to still obtain realistic structures and at the same time validate the docking simulations, the partially desialylated GT1b (i.e., GD1a) from structure 2VU9 was removed from the cocrystal and docked back into the protein, which resulted in a cluster of highly ranked poses with similarities to the original cocrystal structure. The success of such treatment can be attributed to the prearranged fit of both protein and input ligand structure.[56]

The GBS can interact with a limited number of glycan residues, and typically, for some of the serotypes of BoNT, the terminal Sia(2→3)Gal moiety is particularly important for the selective association, as revealed by X-ray crystallography.[10,57] Indeed, the resulting docked poses of the ensemble of protein−ligand combinations revealed that a *galacto*-configured sugar residue (Gal or GalNAc) almost exclusively occupies the same position of the GBS (Figure 4). This subsite (denoted site B in Figure 1b), conserved in most BoNT serotypes, is defined by, e.g., the Trp1266 anchor and His1253 in BoNT/A, both engaged in hydrophobic stacking. Ligands carrying a terminal Sia(2→3)Gal motif occupy site B together with the adjacent subsite restricted by the Tyr1117 anchor (denoted site A in Figure 1b), in a majority of the top-ranked poses from the docking simulation. The results also suggest that in order to be efficiently accommodated in the A−B sites the conformation at the glycosidic $\alpha$-(2→3) linkage has to be *−synclinal* (*−sc*) and *synperiplanar* (*sp*) with respect to the torsion angles $\phi$ (C1′−C2′−O3−C3) and $\psi$ (C2′−O3−C3−H3), respectively. For compounds devoid of a Sia residue, instead bearing a terminal Gal(1→3)GalNAc moiety, this Gal residue is bound in site A, whereas GalNAc occupies site B and the glycosidic torsion is then typically in an *antiperiplanar-*$\phi$ arrangement. Compounds lacking both of these motifs as a terminal end, particularly GM2 and its asialo derivative, aGM2, displayed a nonuniform distribution of the docked binding modes and complementarity to the GBS appeared to be ambiguous.

**STD NMR—Screening and Quantitation.** For the subsequent ligand-based NMR screening, eight compounds were chosen (Figure 3), representing various structural aspects of GD1a and with a varying number of residues, from monosaccharide (Sia alone) to hexasaccharide (GD1a). 1D $^1$H STD NMR spectroscopy was performed at 5 °C in $D_2O$ with these eight ligands in the presence of BoNT/A-$H_C$. Protein resonances were saturated in the aromatic as well as in the aliphatic spectral regions, i.e., outside of the region where $^1$H NMR resonances from the oligosaccharides reside (Figures S1 and S2), typically at 7 and −1 ppm, respectively (Figure S3). Off-resonance spectra were acquired with irradiation at 60 ppm to generate the difference spectra. In order to produce buildup curves, protein irradiation was performed with different saturation times between 0.5 and 4.5 s. For all compounds, the experimental setup was applied also in the absence of protein and it was asserted that no STD signals or only negligible difference artifacts were observed in the spectra under the conditions employed. STD amplification factors (STD-AF), which indirectly give information on the concentrations of protein−ligand complexes in solution, were calculated from the absolute STD effects according to established practice, thus enabling direct comparison of the buildup curves between the ligands.[58] From the slopes of these curves and the magnitude of the STD-AFs (Figure S4), it was determined that sialic acid did not bind to BoNT/A-$H_C$, whereas methyl lactoside, asialo GM2, as well as GM2 displayed significant but close to negligible effects, indicating a very weak, possibly unspecific binding, a finding consistent with the results from the docking simulation. GM3 on the other hand yielded STD effects of a magnitude suggesting weak but specific association to the protein of all three monosaccharide units. As expected, GD1a revealed conspicuous effects but only for the terminal trisaccharide motif (residues 3−5 in Figure 1a). However, the other constituent trisaccharides, corresponding to GM3, within the molecule (residues 1, 2, and 6 in Figure 1a), carrying the branch point, did not demonstrate any substantial effects, which suggests that the terminal residues constitute the active epitope. Consistently, sialyl-T, representing the terminal trisaccharide of GD1a, also binds to the protein, which was established from STD data. GM1a, only missing the terminal Sia unit compared to GD1a, was also identified as a binder from the NMR screening with the terminal disaccharide experiencing significantly stronger STD effects compared with the other residues within the molecule (Figure S3 shows representative STD spectra of these four binders). The results underscore that representation matters for binding; viz., GM3 constitutes part of the larger compound GM2, but the presence of an additional $\beta$-(1→4)-linked GalNAc in the latter hinders binding to BoNT/A. Similarly, GM3 is structurally present as the branched internal trisaccharide of the hexasacharide GD1a and the pentasaccharide GM1a, but the latter two associate to the protein with their terminal ends.


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

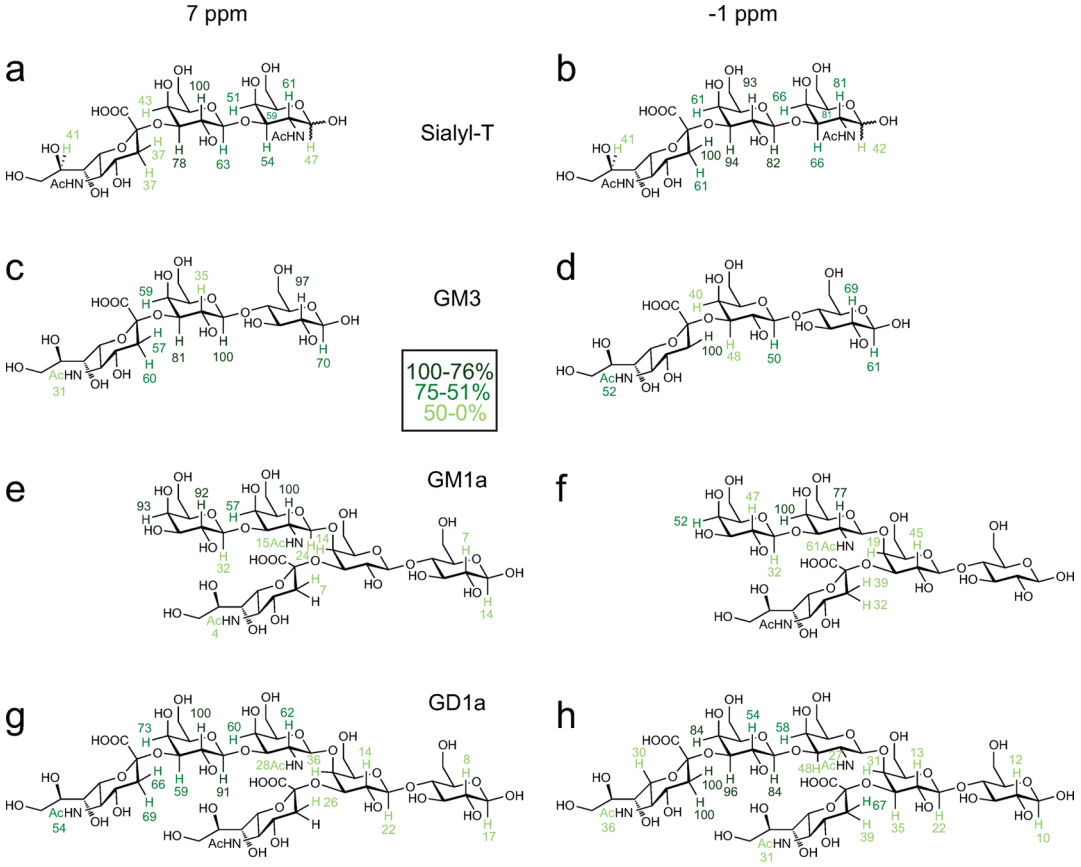



**Figure 5.** Epitope mapping presented as normalized levels of saturation (against the most intense signal arbitrarily assigned to 100%) from STD-$AF_0$ for sialyl-T (a, b), GM3 (c, d), GM1a (e, f), and GD1a (g, h) in complex with BonT/A-$H_C$ with on-resonance irradiation at 7 ppm (a, c, e, g) and −1 ppm (b, d, f, h). Nonexchangeable protons explicitly presented in the figures indicate detected STD effects.

The acquired STD data of the four identified binders, GM3, sialyl-T, GM1a, and GD1a, were subject to quantitative examination. By an initial-slope treatment of the STD-AF buildup curves, yielding STD-$AF_0$, group epitope maps (GEM) could be deduced (Figure 5) unaffected by relaxation bias.[43,59] The epitope maps, composed of relative STD effects from both saturation frequencies, showed that GD1a and sialyl-T bind in a similar fashion but slightly different from GM1a and GM3. Both GD1a and sialyl-T receive the most prominent STD effects for the $Gal^4$ residue. Particularly, proton H2 of $Gal^4$ exhibits the strongest effects when saturating the protein resonances at 7 ppm. Upon saturation at −1 ppm, the H3 protons of Sia5 receive the most conspicuous effect. For GM1a, the strongest effects are observed for H2 and H4 of $GalNAc^3$ when irradiation takes place at 7 and −1 ppm, respectively. In the case of GM3, the corresponding protons were H1 of $Gal^2$ and H3 of $Sia^6$, respectively.

**NOESY and trNOESY NMR Experiments.** A qualitative analysis of relative distances determining conformations at the glycosidic linkages of the free and bound ligands was performed by means of NOESY and transferred NOESY (trNOESY) experiments. 1D and 2D $^1H,^1H$-NOESY experiments were performed on the free and BoNT/A-associated ligands typically employing mixing times of 100−300 ms. In the absence of protein (Figure S5), the NOESY analysis of the ligands revealed some interesting conformational behavior for, in particular, the Sia(2→3)Gal linkages. GD1a contains two such linkages, and they adopt different conformational preferences under the herein applied conditions. By analyzing the interglycosidic NOEs of the $Sia^6(2\rightarrow3)Gal^2$, it was observed that, upon inversion of the resonance $H3_{ax}$ from $Sia^6$, the resonance intensity of H3 in $Gal^2$ was prominent and significantly larger compared with the intraresidual correlation to H5. A distinct correlation to H4 of $Gal^2$ was also present. Inversion of the resonance $H3_{eq}$ yielded NOE correlations of similar magnitude as for H5 of $Sia^6$ and H3 of $Gal^2$. Such an outcome is only possible if an *ap* (*antiperiplanar*) conformation is adopted as the major one at the $\phi$ torsion angle. In the corresponding examination of the $Sia^5(2\rightarrow3)Gal^4$ linkage in GD1a, inversion of the $H3_{ax}$ resonance yielded instead a considerably weaker interresidual correlation to H3 in $Gal^4$, slightly less intense than that to H5 in $Sia^5$. Upon inversion of the $H3_{eq}$ resonance, the interglycosidic correlation to H3 of $Gal^4$ was very weak and correlations to H2 or H4 in $Gal^4$ were not detected. These observations are consistent with a major −*sc* conformer at the $\phi$ torsion of the $Sia^5(2\rightarrow3)Gal^4$ linkage. A previous NMR spectroscopic study of GD1a under comparable conditions (295 K in $D_2O$) indicated that the *ap* conformation is the predominant one at the $\phi$ torsion angle for the internal sialic acid, whereas the −*sc* conformer could also be present in addition to the *ap* conformation for the external sialic acid of the hexasaccharide.[60] The two observed conformations of the Sia(2→3)Gal linkages are both *exo*-anomeric with respect to the $\phi$ torsion angle. The preference for an *ap*-$\phi$ arrangement of the $Sia^6(2\rightarrow3)Gal^2$ linkage has been suggested to stem from an interresidual hydrogen bonding interaction between the $COO^-$ group of $Sia^6$ and the N-acetyl group of GalNAc; the previous studies were, however, performed in a mixture of DMSO-


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

$d_6$:$D_2O$ (98:2) on the intact GD1a ganglioside containing its ceramide residue.[61,62] In analogy with $Sia^6$ in GD1a, the $Sia^6(2\rightarrow3)Gal^2$ linkage of GM1a adopted an *ap-ϕ* arrangement according to NOE data (Figures S6 and S7), whereas the trisaccharides GM3 and sialyl-T (Figures S8−S10) reside in the *−sc-ϕ* conformation at their corresponding linkages. The trisaccharide GM3 represents the internal motif with a branching sialic acid in the GD1a and GM1a structures, and the difference at the $Sia^6(2\rightarrow3)Gal^2$ linkage demonstrates the importance of molecular context and supports the contribution of a stabilizing internal hydrogen bond. For all oligosaccharides, NOE data consistent with *sc-ϕ* conformations were observed for all glycosidic linkages other than those containing a Sia residue; viz., prominent interglycosidic correlations were detected from H1′ to H3 or H4, clearly more intense than any intraresidual ones.

In measuring trNOEs arising from protein•ligand complexation (Figure S11), it is desirable to apply conditions at which contributions from the free ligand NOEs will be negligible.[63,64] In order to determine such conditions, translational diffusion coefficients ($D_t$) were measured for the larger GD1a ligand and the smaller trisaccharide GM3 by $^1H$ NMR pulsed-field-gradient experiments in $D_2O$ at 25 °C.[65] For GD1a, $D_t = 2.25 \times 10^{-6}$ $cm^2 \cdot s^{-1}$, and for GM3, $D_t = 3.12 \times 10^{-6}$ $cm^2 \cdot s^{-1}$; these values were used to calculate the NOE zero-crossing temperature at 500 MHz (11.7 T) being 58 and 18 °C, respectively. Due to the sensitive nature of the studied protein, the elevated temperature was considered unfeasible for subsequent analyses, but also 18 °C raised some stability concerns for longer experimental times. Nevertheless, as a proof of principle, a trNOESY experiment was performed at 18 °C with sialyl-T as a ligand, whereas subsequent NOESY and trNOESY analyses for all selected ligands were acquired at 5 °C. At this temperature, care has to be taken in evaluating the resulting data, since the free ligand NOEs are of the same phase as the trNOEs and can thus contribute to the latter correlations. Due to the different magnitude and rate of buildup between bound ligand trNOEs and free ligand NOEs, qualitative analysis of the bioactive conformation can still be performed, even at conditions where the ligand NOE ≠ 0.

In trNOESY spectra of sialyl-T in complex with BoNT/A-$H_C$ acquired at 18 °C, negative trNOEs were observed, whereas correlations were not detected for the compound free in solution at the same temperature (Figure S12). Performing trNOESY experiments of the same BoNT/A-$H_C$•sialyl-T sample at 5 °C, only small differences in relative resonance intensities were observed (Figures S9 and S10), suggesting that the contributions of the free ligand NOEs are small. Furthermore, for all four ligands studied by NOE experiments, the magnitude of and buildup rate was stronger for the trNOESY compared with the corresponding NOESY correlations, which confirms binding as well as being an indication that trNOESY data can be exploited under these conditions. The trNOESY experiments did not reveal any drastic changes of relative distances as compared to the NOE correlations of the unbound sialyl-T at 5 °C (*vide supra*). The *+sc* conformation for the *ϕ* torsion angle at the glycosidic linkage of $Gal^4(1\rightarrow3)GalNAc^3$ was also persistent in the protein-associated state of the ligand. The terminal trisaccharide of GD1a exhibited the same results as for sialyl-T, indicating a high degree of conformational predisposition for their binding to BoNT/A-$H_C$. The reducing end trisaccharide moiety of GD1a does not change the conformation to any significant extent when bound to the protein. This is important, as it suggests that the BoNTs have evolved to recognize the most prominent glycan conformation. Likewise, the major solution conformation of GM3 as well as GM1a was conserved in the bound state, exemplified for the latter by the similar increase in NOE buildup rates from solution to the protein•ligand complex (Figure S13).

**X-ray Crystallography—Structure of BoNT/A-$H_C$•GD1a and BoNT/A-$H_C$•sialyl-T.** Co-crystallization studies of BoNT/A-$H_C$ with GD1a and with sialyl-T were performed using purified BoNT/A-$H_C$. The crystals grew in space group $C222_1$ (BoNT/A-$H_C$•GD1a) and $P2_1$ (BoNT/A-$H_C$•sialyl-T), and diffracted to 2.0 and 2.6 Å, respectively (Table 1). The overall structure of BoNT/A-$H_C$ was the same as that reported in previous studies.[22,66] The structures were solved using the molecular replacement (MR) technique. BoNT/A-$H_C$•GD1a contained one molecule per asymmetric unit (ASU), with well-defined electron density, not part of the protein, appearing at the GBS after MR. This electron density could readily be modeled as GD1a. BoNT/A-$H_C$•sialyl-T contained two molecules per ASU, with well-defined electron density, not part of the protein, appearing at the GBS of one of the molecules in the ASU. This electron density corresponded to sialyl-T. Due to crystal packing, the GBS in the other protein chain was disrupted, and sialyl-T could thus not bind there.

**Table 1. Data Collection and Refinement Statistics for Crystal Structures of BoNT/A-$H_C$•Ligand Complexes**

| data collection | BoNT/A-$H_C$•GD1a | BoNT/A-$H_C$•sialyl-T |
|---|---|---|
| space group | $C222_1$ | $P2_1$ |
| cell dimensions | | |
| *a*, *b*, *c* (Å) | 73.9, 114.5, 106.4 | 65.5, 104.3, 69.4 |
| *α*, *β*, *γ* (deg) | 90, 90, 90 | 90, 116, 90 |
| resolution (Å) | 40.4−2.0 (2.05−2.0) | 39.1−2.6 (2.67−2.60) |
| $R_{merge}$ | 0.11 (0.55) | 0.12 (0.52) |
| $I/\sigma$ ($I$) | 10.0 (2.6) | 4.9 (1.7) |
| completeness (%) | 98.3 (99.0) | 99.8 (99.7) |
| CC(1/2)* | 0.99 (0.74) | 0.99 (0.69) |
| redundancy | 3.5 | 3.4 |
| **refinement** | **BoNT/A-$H_C$•GD1a** | **BoNT/A-$H_C$•sialyl-T** |
| resolution | 40.4−2.0 | 39.1−2.6 |
| no. unique reflections | 30107 | 25875 |
| $R_{work}/R_{free}$ | 0.19/0.23 | 0.22/0.24 |
| no. atoms | | |
| protein | 3540 | 6751 |
| ganglioside ligand | 77 | 46 |
| water | 326 | 111 |
| *B*-factors | | |
| protein | 24.6 | 54.2 |
| carbohydrate | 65.8 | 83.1 |
| water | 31.0 | 40.8 |
| rms deviations | | |
| bond lengths (Å) | 0.008 | 0.011 |
| bond angles (deg) | 1.37 | 1.21 |

Both GD1a and sialyl-T bind in the defined A−C subsites of the GBS of BoNT/A-$H_C$. Similar to what was observed previously[22] and in the molecular docking simulations, it is clear that $Gal^4$ is the monosaccharide residue that is most firmly bound. This is evident both from its electron density and from having the lowest *b*-factors of the sugar units in both of the bound glycans. The conformations at the $Sia^5(2\rightarrow3)Gal^4$


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

Table 2. Torsion Angles of Glycosidic Linkages in Crystal Structures of BoNT/A-$H_C$•Ligand Complexes

| ligand | linkage | $\phi$ (deg) | annotation[a] | $\psi$ (deg) | annotation[a] |
|---|---|---|---|---|---|
| sialyl-T[b] | $Sia^5(2\rightarrow3)Gal^4$ | 281 | −*sc* | 3 | +*sp* |
| | $Gal^4(1\rightarrow3)GalNAc^3$ | 35 | +*sc* | 338 | −*sp* |
| GD1a[b] | $Sia^5(2\rightarrow3)Gal^4$ | 315 | −*sc* | 18 | +*sp* |
| | $Sia^6(2\rightarrow3)Gal^2$ | 173 | +*ap* | 349 | −*sp* |
| | $Gal^4(1\rightarrow3)GalNAc^3$ | 27 | +*sp* | 325 | −*sc* |
| GT1b[c] | $Sia^5(2\rightarrow3)Gal^4$ | 313 | −*sc* | 5 | +*sp* |
| | $Sia^6(2\rightarrow3)Gal^2$ | 302 | −*sc* | 341 | −*sp* |

[a]According to the Klyne−Prelog system for describing conformations around a single bond. [b]X-ray crystal structures of this study. [c]Stenmark et al.[22]

linkage in GD1a and sialyl-T are similar, being −*sc* and +*sp* at the $\phi$ and $\psi$ torsion angles, respectively. The $Gal^4(1\rightarrow3)GalNAc^3$ linkages in GD1a and in sialyl-T were equally comparable, being ∼30 and ∼330° for $\phi$ and $\psi$, respectively (Table 2). Comparing the herein determined BoNT/A-$H_C$•GD1a structure to the previously described BoNT/A-$H_C$•GT1b complex (PDB code: 2VU9), the glycosidic linkage conformations are highly similar except for the $Sia^6(2\rightarrow3)Gal^2$ linkage (Table 2). It is possible that the additional Sia residue of GT1b ($Sia^7$) affects the conformation of $Sia^6$ and consequently its ability to interact with the protein. GT1b forms one additional hydrogen bond compared to GD1a, viz., from $Sia^6$ to Arg1276 of BoNT/A-$H_C$.[22] This is likely due to the steric hindrance from $Sia^7$ in GT1b forcing $Sia^6$ to adopt a different conformation, since in solution the $Sia^6$ residue of GD1a occupies mainly the same conformation as seen in the BoNT/A-$H_C$•GD1a structure. Sia(2→3)Gal linkages are indeed flexible and can adopt different conformations both in solution and in the bound state depending on the specific ligand as well as the protein.[45,67,68] The bioactive conformation in gangliosides and similar carbohydrate ligands is, however, for this linkage predominantly observed in the −*sc*-$\phi$ state, an exception being the GM1a pentasaccharide binding to Cholera toxin.[69] For glycosidic linkages of the other constituent disaccharides, viz., Gal(1→3)GalNAc, GalNAc(1→4)Gal, and Gal(1→4)Glc, the $\phi$ torsions are exclusively observed in +*sc* conformations in structures present in the RCSB Protein Data Bank.[68]

Interestingly, there are relatively large differences in the Sia5 position between the two solved structures, also in comparison to the previously determined structure of BoNT/A-$H_C$•GT1b. While Gal4 is located in virtually the same position, $Sia^5$ has shifted its position between the GD1a and sialyl-T (Figure 6). This results in different hydrogen-bonding networks for the $Sia^5$ between the BoNT/A-$H_C$•GD1a, BoNT/A-$H_C$•sialyl-T, and BoNT/A-$H_C$•GT1b complexes. While $Sia^5$ in sialyl-T only contributes with two potential hydrogen bonds, to Y1267 and G1279, $Sia^5$ in GD1a has one hydrogen bond from Y1117 and a further three hydrogen bonds, via two bridging water molecules, to Y1267, R1276, and G1279 (Figure 7). In the GT1b complex, $Sia^5$ contributes with three hydrogen bonds, two to Y1117 and one to S1275. It should be noted here that the electron density for the $Sia^5$ in GD1a, and particularly Sia in sialyl-T, is quite weak. Taken together, this strongly indicates that this moiety is flexible and can likely occupy different conformations within the binding site.

Further comparisons of the solved crystal complexes with the STD NMR data revealed close interproton distances between the H2 of $Gal^4$ in both GD1a and sialyl-T to the ring CH protons of His1253; H2 of $Gal^4$ generated the largest observed STD effects in the respective ligand upon protein irradiation at

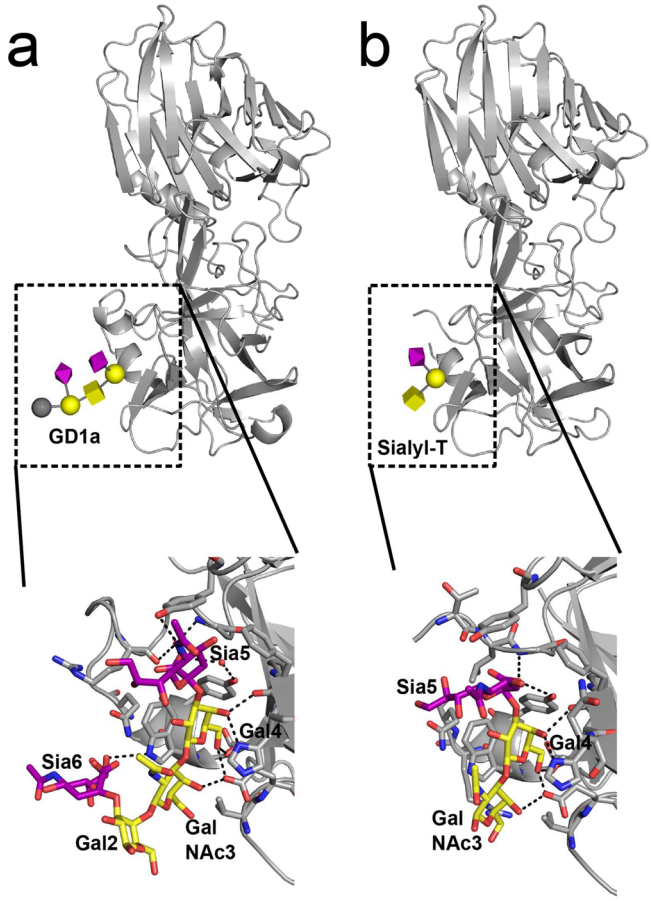



**Figure 6.** Glycan binding to BonT/A-$H_C$ as obtained by X-ray crystallography. (a) Overall view of the BoNT/A-$H_C$•GD1a complex (protein in gray color and sugar in 3D-CFG representation). Glc1 is disordered in the complex and thus colored gray. The lower panel shows a zoomed-in view on the detailed interactions with the GD1a residues colored according to CFG. Possible intermolecular hydrogen bonds are shown as black dotted lines. (b) Corresponding information for the BoNT/A-$H_C$•sialyl-T complex.

7 ppm (Figure 5), targeting mainly aromatic proton resonances. Histidine is assumed to be instantaneously saturated when reached by the selective irradiation and can thereby efficiently mediate saturation transfer to the ligand,[70] thus consistent with these results. The binding modes of the ligands in the GBS also added evidence to the observations that the methyl groups in the Sia and GalNAc residues receive relatively weak STD effects.

**Oligosaccharide Affinity.** The affinities of GD1a and sialyl-T to BoNT/A-$H_C$ were subsequently investigated by NMR spectroscopy and isothermal titration calorimetry (ITC).


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

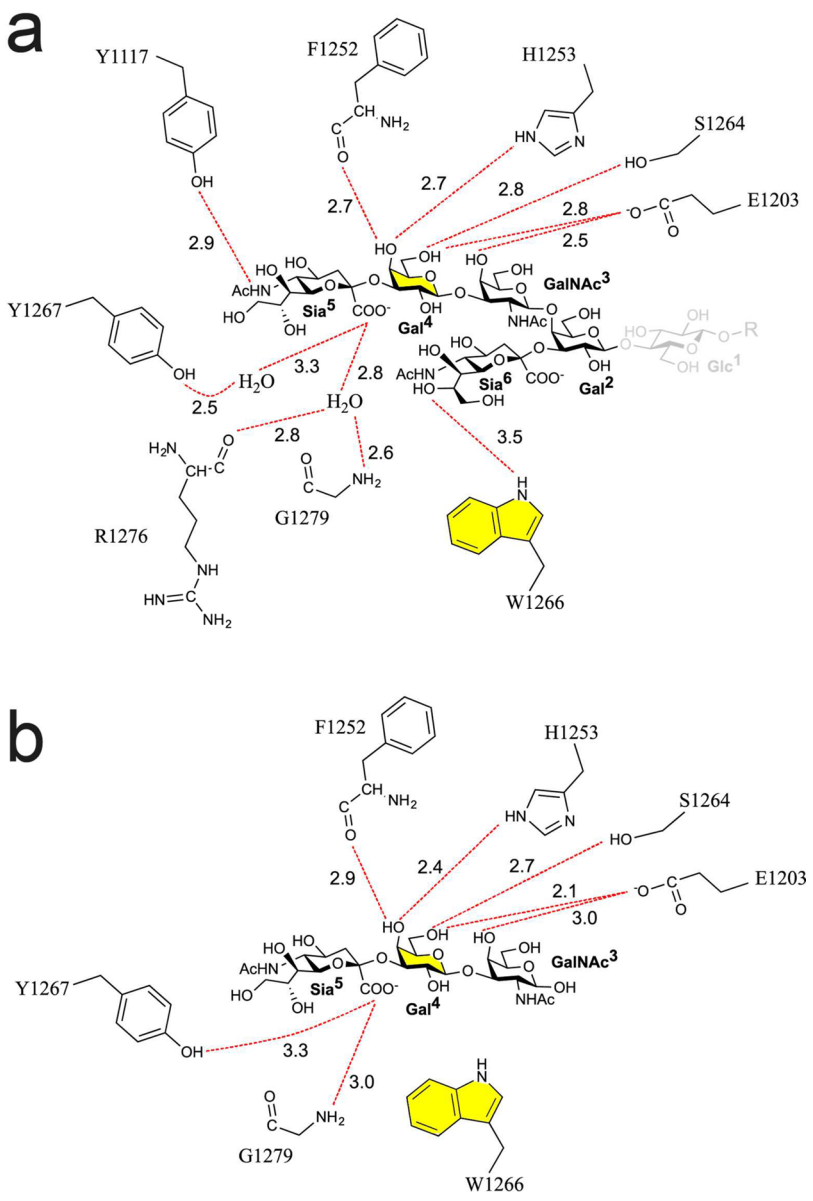



**Figure 7.** Schematic 2D plots of the glycan hydrogen bonding of GD1a (a) and sialyl-T (b) to BonT/A-$H_C$, as obtained by X-ray crystallography. Glc1 is disordered in the GD1a complex and thus colored gray. Possible intermolecular hydrogen bond interactions are shown as red dotted lines, with the hydrogen bond distance annotated for each bond (Å). Yellow coloring in $Gal^4$ and W1266 highlights the stacking interaction between these residues.

Previous studies with ganglioside binding to BoNT/A indicated that the affinity should be in the nanomolar range.[57,71] However, initial NMR experiments failed to show any binding event at these concentrations. The affinity to BoNT/A-$H_C$ was measured for sialyl-T by NMR in a direct way by single ligand titrations exploiting ligand-observed transverse relaxation rates, $1/T_2$, with the CPMG experiment and the resulting $K_D$ of 0.5 mM indicated a much lower affinity (Figure S14). We therefore performed ITC titrations at these higher oligosaccharide concentrations. This yielded a $K_D$ of 1.0 ± 0.1 mM for GD1a and 2.6 mM ± 0.5 for sialyl-T. Due to the low affinity of the interaction, resulting in a lack of plateau levels in the titrations, we can only estimate the dissociation constants to be in the mM range, like for an octasaccharide-tailspike protein interaction,[72] but not the enthalpy and entropy contributions to the binding.

These relatively low affinities are at least 4 orders of magnitude lower than what was previously measured for the gangliosides GD1a (600 nM) and GT1b (200 nM) to BoNT/A, which were determined using a ganglioside-coated 96-well plate assay and surface plasmon resonance (SPR), respectively.[57,71] One major difference between the experiments is that we here only measure the affinity of the glycan part of the PSG to BoNT/A, whereas in the previous experiments the entire PSG were used.[57,71] Furthermore, we only used the binding domain of BoNT/A, as was done in the plate assay,[57] whereas the entire BoNT/A was used in the SPR assay.[71] These results indicate that the membrane itself strongly contributes to the apparent affinity of the toxins for gangliosides. In the case of BoNT/B, BoNT/C, and BoNT/DC, there are clearly exposed hydrophobic residues that are likely to mediate part of this membrane interaction. BoNT/A also has exposed hydrophobic residues in a similar region, although less pronounced. It is


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

likely that electrostatic interactions to the membrane surface play an important role for BoNT/A. We propose that the glycan part of the PSG contributes with specificity but that the affinity of BoNT/A to its receptors comes from its membrane interaction and its protein receptor.

**CORCEMA-ST Analysis of NMR and X-ray Derived Structures.** In order to compare the epitopes deduced by NMR spectroscopy with the structures obtained by X-ray crystallography or molecular docking simulations, full relaxation-matrix calculations with the CORCEMA-ST program were performed.[73] From such an approach, STD buildup curves can be simulated from 3D coordinates and thus related to the experimental data. The program requires a range of different input variables including experimental conditions, kinetic and thermodynamic data of the studied complex, protein chemical shifts (predicted), and various rotational diffusion correlation times, some of which have to be predicted. In the presented simulations, all variables were employed as experimentally deduced or initially predicted except for the protein−ligand association rate constant, $k_{on}$, and $K_D$ (for GM1a), which were iterated (see Materials and Methods). The agreement between the NMR data and those calculated on the basis of the X-ray crystal structure of the BoNT/A-$H_C$•GD1a complex was excellent (Figure 8), as displayed by a normalized root-mean-square deviation (RMSD) value, termed the $R_{NOE}$ factor in the CORCEMA-ST approach, of 0.22.[74,75] For the corresponding comparison with sialyl-T as the ligand, the agreement was less prominent, resulting in an $R_{NOE}$ factor of 0.50. A general interpretation of these results is that the solid phase crystal structure cannot describe the behavior in solution as accurately for sialyl-T as for GD1a. In particular, it can be rationalized that the hexasaccharide GD1a imposes a higher degree of geometrical restraints than the trisaccharide, the latter being devoid of a branched scaffold moiety. Such an explanation implies a higher degree of conformational preorganization of GD1a compared with sialyl-T.

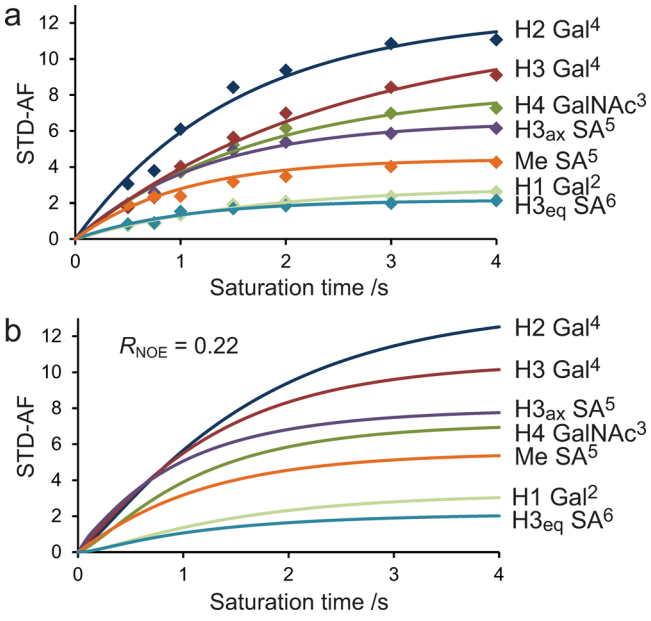



**Figure 8.** STD buildup curves with STD-AF as a function of $t_{sat}$ of the BoNT/A-$H_C$•GD1a complex. (a) Experimental STD-AF with corresponding exponential fits. (b) Simulated STD-AF from CORCEMA-ST calculations of the complex obtained by X-ray crystallography.

GM1a was also subjected to a CORCEMA-ST analysis using structures from molecular docking simulations as input. To evaluate whether the terminal disaccharide of the ligand binds to BoNT/A in the B−C sites, as implied from the trNOESY data or in the A−B sites, as suggested by docking, two docked structures were compared. The complexes were energy-minimized and the subsequent CORCEMA-ST analysis was performed only with respect to the interacting disaccharide moiety. The CORCEMA-ST simulated data of the docked structures in comparison with the experimental STD NMR data (Figure S15) yielded $R_{NOE}$ factors of 0.74 for the A−B binding complex (*−sc-ϕ* conformation of the Sia(2→3)$Gal^2$ linkage) and 0.44 for the B−C binding complex (*ap-ϕ* conformation of the Sia(2→3)$Gal^2$ linkage). Experimental NMR data thus suggest binding of GM1a in the B−C site analogous to the binding mode of the same disaccharide motif in GD1a and sialyl-T.

## CONCLUSIONS

We herein investigated glycan receptor binding to BoNT/A-$H_C$ employing a stepwise protocol. The minimal binding epitope was examined, and it was shown that GM1a, binding with its terminal disaccharide, and sialyl-T, binding with all three of its residues, both form a complex with the protein in an efficient manner. The conformationally predispositioned GD1a ligand represents both of these structural features and consequently shows a higher affinity to BoNT/A. Other motifs being constituents of GD1a showed weaker binding or absence thereof, demonstrating that presentation of the glycan epitope to the protein is highly important. The affinity measured by ITC for GD1a, $K_D$ of 1.0 mM, was, however, 4 orders of magnitude lower than values obtained in previous studies where entire PSG were investigated, showing that the glycan part of gangliosides contributes mainly with specificity. It is rather the other receptor as well as the interaction with the membrane as such that form the basis of the strong affinity of BoNT/A to neuronal tissue. These results are important for the future development of BoNTs as drugs, e.g., in engineering the toxins for interactions to specific neuronal cell subtypes.

Two new X-ray crystal structures were solved, viz., GD1a and sialyl-T in complex with BoNT/A-$H_C$, at 2.0 and 2.6 Å resolutions, respectively. The interacting glycan moieties showed similar binding modes in both structures, and they were comparable also to previously reported complexes with glycans binding to BoNTs of different serotypes. The herein presented BoNT/A-$H_C$•GD1a complex agreed very well with solution NMR data as analyzed by STD-AF buildup curves in the CORCEMA-ST approach, resulting in an excellent fit with an $R_{NOE}$ factor as low as 0.22. The less consistent description of the behavior in solution observed for the BoNT/A-$H_C$•sialyl-T complex is believed to originate from a higher flexibility for this trisaccharide being devoid of the internal branched motif present in GD1a. This study demonstrates a powerful approach for analyzing glycan−lectin association where the proficiency of NMR spectroscopy to probe transient ligand binding was complemented with ITC and high resolution X-ray crystallography.

## MATERIALS AND METHODS

**Nomenclature and Definitions.** Neu*p*5Ac = Sia (or SA in figures), D-Gal*p* = Gal, D-Gal*p*NAc = GalNAc, and D-Glc = Glc, and


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

correspondingly, glycosidic linkages will be denoted, e.g., Sia(2→3)Gal referring to α-Neu*p*5Ac-(2→3)-D-Gal*p*. Desialylated derivatives are termed asialo, which is abbreviated by the letter a, e.g., asialo GM2 = aGM2. NMR definition of torsion angles: for Sia glycosidic linkages related to $\phi$ (C1′−C2′−O3−C3) and $\psi$ (C2′−O3−C3−H3); for other torsions $\phi$ (H1′−C1−O*n*−C*n*) and $\psi$ (C1′−O*n*−C*n*−H*n*), where *n* denotes substitution position.

**Molecular Docking Simulations.** For molecular docking simulations, the crystal structure of the complex between BoNT/A-$H_C$ and the GT1b oligosaccharide (PDB ID: 2UV9) with the ligand removed was used as the protein with Autodock VINA 1.1.2 (ADV).[53] 3D models of the ligands having the β-anomeric configuration at the reducing end were built with CarbBuilder,[76] as implemented in the CASPER program.[77,78] Bond order and partial charges were added, and the potential energy was minimized (steepest descent, conjugate gradients, and truncated Newton) in VEGA ZZ.[79] Gasteiger charges were subsequently added in Autodock Tools (ADT).[80] The protein was prepared in Maestro (Schrödinger 2010) using the *Protein Preparation Wizard*: hydrogens were added, water molecules were removed, and the protonation states were defined; Kollman charges were added to atoms in ADT. A restricted grid (22 Å × 22 Å × 22 Å, 1 Å spacing (point separation)) was centered at W1266, and the simulations were performed with an exhaustiveness value of 32 and different random seeds. The 10 best ligand poses were analyzed and clustered according to their protein subsite occupation and glycosidic torsion angles, in that order. The dockings were performed on an Intel(R) Core(TM) i5 CPU Q 3.2 GHz processor with 8 GB RAM running a Windows 7 operating system in 64 bit.

**Glycans.** The glycan part of gangliosides, removed from their ceramide part, and their derivatives were obtained according to the following: GD1a, GM1a, GM2, aGM2, and GM3 were purchased from Elicityl (Crolles, France); sialyl-T was purchased from Carbosynth (Bershire, U.K.); sialic acid was purchased from Sigma-Aldrich (St Louis, MO, USA). Methyl β-lactoside, which was available from a previous study,[81] was used as the lactose model compound in the NMR studies.

**Protein Expression and Purification.** The construct of BoNT/A-$H_C$ was the same as previously described.[22] Transformed BL21 *E. coli* cells were precultured in LB medium containing kanamycin (50 μg·mL$^{-1}$) at 37 °C overnight. The preinoculum was diluted 1000-fold into aerated 2 L flasks with TB media containing 50 μg·mL$^{-1}$ kanamycin and incubated at 37 °C until they reached an $OD_{600}$ value of ca. 0.5−1.5, whereupon the temperature was lowered to 20 °C and expression induced with 0.5 mM IPTG. After overnight induction, the cells were harvested, pelleted, and frozen at −80 °C. For purification, the cells were thawed and resuspended to an $OD_{600}$ value of ∼100 in either 20 mM phosphate buffer, pH 7.4, 300 mM NaCl, and 10% glycerol or 20 mM Hepes, pH 7.5, 300 mM NaCl, and 10% glycerol. Cell lysis was performed by passing the cell suspension two to three times through an Emulsiflex-C3 (Avestin, Germany) at 20 kPsi. Unlysed cells and cell debris were spun down via ultracentrifugation at 4 °C, 267k × *g* for 60 min. The supernatant was collected; imidazole pH 7.8 was added to a final concentration of 15 mM, and incubated with 0.3−0.5 mL of Ni-NTA resin per 10 mL of supernatant at 4 °C for 60 min while rotating slowly. The material was subsequently packed in a disposable 25 mL column (Bio-Rad, CA, USA), washed with 20 column volumes of wash buffer (50 mM Hepes, pH 7.8, 300 mM NaCl, and 45 mM imidazole pH 7.8, or 20 mM Hepes, pH 7.5, 300 mM NaCl, and 50 mM imidazole, pH 7.8). The protein was eluted by using wash buffer supplemented with 0.5 M imidazole, pH 7.0. Purification was then carried out by size exclusion chromatography, using a Superdex 200 10/300 GL column, pre-equilibrated with 20 mM Bis−Tris, pH 7.0, 150 mM NaCl, or using a Superdex 200 16/60 column, pre-equilibrated with 20 mM Hepes, pH 7.0, 150 mM NaCl. The fractions were pooled and concentrated to 17 mg·mL$^{-1}$. Glycerol was added to a final concentration of 10% (making the final protein concentration 15.3 mg·mL$^{-1}$), and the protein was subsequently flash frozen in liquid nitrogen and stored at −80 °C. For NMR experiments, the protein was thawed, run again over the size exclusion column pre-equilibrated with 20 mM $KP_i$, pH 7.0, 150 mM NaCl (buffer A), to remove the glycerol in the protein sample. The fractions were pooled, concentrated on a Vivaspin 30 kDa MWCO concentrator (Sartorius, Göttingen, Germany), diluted 200-fold in buffer A with $D_2O$ instead of $H_2O$, and concentrated to a 100 μM solution.

**NMR Spectroscopy.** If not otherwise stated, NMR spectroscopy experiments were carried out at 5 °C on a Bruker Advance 500 MHz spectrometer equipped with a 5 mm PFG triple-resonance CryoProbe on samples containing potassium phosphate buffer (20 mM; pH 7) and NaCl (150 mM) in $D_2O$. Both 3 mm (0.18 mL sample volume) and 5 mm (0.55 mL) NMR tubes were used.

$^1$H NMR chemical shifts of the glycans, subjected to the NMR screening, were predicted by the CASPER program[77,78] or obtained from the literature.[60] If needed, additional resonance assignment experiments, e.g., band-selective $^1$H,$^{13}$C-CT-HMBC, 1D $^1$H,$^1$H-TOCSY, were performed on a Bruker 500 MHz (*vide supra*) or a Bruker Avance III 700 MHz spectrometer with 5 mm TCI *Z*-gradient high resolution CryoProbe. The translational diffusion coefficients of GM3 and GD1a were measured using 10−14 pulsed-field-gradient (PFG) $^1$H NMR experiments on a sample containing the oligosaccharides, 15 and 4 mM, respectively. The experiments were performed at 25 °C on a 600 MHz Bruker Avance III NMR spectrometer equipped with a TXI ($^1$H/$^{13}$C/$^{31}$P) probe, where the *Z*-gradient had been calibrated to compensate for gradient inhomogeneities by using a gadolinium doped water sample (1% $H_2O$ in $D_2O$ + 1 mg mL$^{-1}$ $GdCl_3$) and a literature value of $D_t$ = 1.90 × 10$^{-9}$ m$^2$ s$^{-1}$ for the HDO resonance.[82] The diffusion time delay (Δ) was set to 100 ms, and the gradient pulse length (δ) was set to 2 ms for GM3; the corresponding values for GD1a were 300 and 4 ms, respectively. Each experiment was acquired with 32 data points and gradient strengths starting from 5% up to 95% of the maximum (55.79 G·cm$^{-1}$). The decay of the resonances from the sugar bulk region (3.0−4.5 ppm) was used to calculate the diffusion coefficient by fitting a Stejskal−Tanner type equation to the data.[65]

$^1$H,$^1$H-NOESY experiments were performed on samples containing the glycans at concentrations of 4−15 mM employing mixing times between 100 and 500 ms. 1D NOESY experiments with suppression of zero-quantum coherences[83] were carried out with selective excitation of a target proton resonance by using 60−100 ms long r-SNOB shaped pulses.[84] The experiments were performed using 8k data points with a spectral width of 4 kHz, yielding an acquisition time of 1 s, together with a relaxation delay of 2 s. The number of transients used was between 256 and 1k in addition to 16 dummy scans. Phase-sensitive 2D NOESY experiments were used together with excitation sculpting to suppress the residual HDO solvent peak.[85] Spectra were acquired using 3k−12k data points in the direct dimension, 256 increments, with a sweep width of 4−6 kHz in both dimensions, and a relaxation delay of 1.3−2.5 s. The FIDs were acquired using 32−88 scans in addition to 32 dummy scans. $^1$H,$^1$H-trNOESY spectra were all acquired with the corresponding experiments on samples containing the glycans (3−5 mM) and BoNT/A-$H_C$ (50−87 μM), yielding protein−ligand ratios from 1:35 to 1:100.

1D STD NMR spectra were recorded using the standard pulse sequence[40] together with excitation sculpting and a 60 ms 5 kHz spin-lock. A ligand concentration of 3 mM was used in conjunction with BoNT/A-$H_C$ at a concentration of 30 μM, except for GD1a that was used at a 5 mM concentration. Protein saturation was achieved by irradiating on-resonance at either 7 or −1 ppm with a 50 ms train of Gaussian pulses using a power level corresponding to a hard square pulse of 65 Hz. The same pulse was used for off-resonance irradiation at 60 ppm in order to obtain the difference spectra. Irradiation times of 0.5, 0.75, 1, 1.5, 2, 3, and 4 or 4.5 s were employed together with an additional relaxation delay making the recycle time 5.3 s, except in the case of sialyl-T as a ligand where the delay was set to 1 s. Spectra were acquired with 1k−2k scans, 32 dummy scans, 13k data points, and a spectral width of 8 kHz. A line-broadening window function of 2 Hz was applied prior to Fourier transformation, and STD amplification factors (STD-AF)[58] were calculated as the difference between on- and off-resonance signal strength, normalized to the off-resonance signal strength and scaled with the amplification factor (the ligand/protein


DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

ratio). STD buildup curves were constructed in Matlab (the MathWorks, MA, USA) by fitting exponential equations to the STD-AF data; subsequently, STD-$AF_0$ was calculated as the derivative at $t = 0$.

The NMR spectroscopic $K_D$ measurement of sialyl-T in complex with BoNT/A-$H_C$ was performed with $T_2$ CPMG spin−echo experiments as previously described.[49] The BoNT/A-$H_C$ concentration was kept constant, and $T_2$ measurements were performed on sialyl-T at four different concentrations. By plotting the ligand concentration vs the difference between $T_2$ of the ligand in exchange with the protein and $T_2$ of the free ligand, the $K_D$ value was obtained.

**X-ray Crystallography and Structure Determination.** The protein was thawed, ganglioside (GD1a or sialyl-T) was added to a final concentration of 2.5 mM, and it was subsequently crystallized using the vapor diffusion technique. Diffraction quality crystals grew in a solution containing 20% PEG6000, 0.2 M $MgCl_2$, 0.1 M Hepes, pH 7.0 (BoNT/A-$H_C$•GD1a), or 20% PEG3350, 0.2 M potassium thiocyanate, 0.1 M Bis−Tris propane, pH 6.5 (BoNT/A-$H_C$•sialyl-T). The crystals were cryo-protected by the addition of well solution complemented with 20% glycerol and flash frozen in liquid nitrogen. Diffraction data was collected at 0.918 Å wavelength at beamline 14.1, BESSY, Berlin. The crystals diffracted to 2.0 Å (BoNT/A-$H_C$•GD1a) and 2.6 Å (BoNT/A-$H_C$•sialyl-T). Data reduction and processing were carried out using XDS[86] and programs from the CCP4 suite.[87] Relevant statistics are shown in Table 1. The structure was solved via molecular replacement, using a previously solved structure of $H_C$A as the search model (PDB code: 2VU9, A chain only). Refinement was carried out in Refmac5, interspersed with model building in Coot. In the BoNT/A-$H_C$•GD1a structure, all monosaccharide units of GD1a are visible in the electron density map except for $Glc^1$; in BoNT/A-$H_C$•sialyl-T, all monosaccharide units of sialyl-T were visible. However, Sia5 has badly defined electron density in both gangliosides, as well as $Sia^6$ in the GD1a structure. Restrained refinement of these glycan moieties yielded conformations that were highly unlikely or incorrect. Therefore, after the protein refinement was complete, the sialyl moieties, $Sia^5$ and $Sia^6$ in the GD1a structure and $Sia^5$ in the sialyl-T structure, were manually positioned in reasonable chair conformations,[88] and a final round of refinement was performed where only the *b*-factors were refined. However, the $sp^3$ hybridization at atom C2 of $Sia^5$ in the sialyl-T was still suboptimal with respect to the C2−C1 bond, and this part of the structure must therefore be treated cautiously. The crystal structures have been deposited in the RCSB PDB with accession numbers 5TPB (BoNT/A-$H_C$•sialyl-T) and 5TPC (BoNT/A-$H_C$•GD1a).

**Isothermal Titration Calorimetry.** Association of ganglioside oligosaccharides to the binding domain of the A serotypes of BoNT was measured via isothermal titration calorimetry on an ITC200 (GE Healthcare, Little Chalfont, U.K.) at 25 °C and 1000 rpm. A 200 μL solution of $H_C$A (with a concentration of 100 μM) was added to the cell. Binding was measured upon the addition of GD1a or sialyl-T in a stepwise manner, typically 16 injections of 2.5 μL each, at a concentration of 5 mM. The first titration was set to 0.5 μL, and was subsequently deleted in the data analysis. Data analysis was performed using the Origin software provided by the manufacturer. Due to the low affinity of the ligands, $N$ was set to unity during fitting, due to the fact that there is only one GBS. Five titrations were performed for GD1a using two different batches and three titrations for sialyl-T. The error reported for the $K_D$ is the standard deviation.

**CORCEMA-ST Simulations.** Theoretical STD buildup curves were calculated from the crystal structures of BoNT/A-$H_C$ with GD1a or sialyl-T as ligands using CORCEMA-ST;[73] $K_D$ values used were 0.3 or 0.6 mM, respectively. A generalized order parameter $S^2$ of 0.85 and a uniform leakage relaxation of 0.30 $s^{-1}$ were assumed. Ligand correlation times ($\tau_C^{Ligand}$) were calculated from the PFG diffusion measurements, resulting in 1.48 ns for GD1a and 0.55 ns for GM3 at 5 °C. The value obtained for GM3 was used for sialyl-T, with both molecules having the same molecular mass. The methyl group internal correlation time ($\tau_m$) is rapid[89] and was chosen to be 10 ps. The protein correlation time ($\tau_C^{Protein}$) was approximated to 280 ns using Stokes' law. The conformation of the ligand was assumed to be the same in both the free and bound states. The SHIFTX2 software[90] was used to calculate the $^1$H chemical shifts of the protein, and the protons resonating between 6.93 and 7.07 ppm were assumed to be saturated when irradiation was set at 7 ppm, given the 65 Hz irradiation pulse. A binding site cutoff of 8 Å was employed. The ligand on-rate ($k_{on}$) was iterated, and a value outside the diffusion limited range gave the best fit for both sialyl-T and GD1a, namely, $5 \times 10^4$ $s^{-1}$ $M^{-1}$. This value is close to that obtained in a previous study of GT1b in complex with BoNT/A-$H_C$, measured by SPR[71] to $k_{on} \sim 1 \times 10^5$ $s^{-1}$ $M^{-1}$ and is consistent with those found in a recent study for ganglioside oligosaccharides binding to the myelin-associated glycoprotein.[45] STD NMR spectra revealed that ligand binding to BoNT/A was still in the fast exchange regime on the chemical-shift scale.[63] Structures of GM1a in complex with the proteins were taken from the docking simulations representing two different clusters. These structures were energy minimized (heavy atom restraint of 0.6 Å) and H-bond optimized in Maestro (Schrödinger 2010). The same values as those for GD1a and sialyl-T were used except for the $\tau_C^{Ligand}$ which was calculated to be 0.97 ns based on the obtained values for the former ligands and $K_D$, which was iterated to 0.83 mM, giving the best fit. For the comparison with the experimental STD NMR data, $R_{NOE}$ factors were calculated according to eq 1.

$$R_{NOE} = \sqrt{\frac{\sum_{i=1}^{N} (\text{STD-AF}_{sim,i} - \text{STD-AF}_{expt,i})^2}{\sum_{i=1}^{N} (\text{STD-AF}_{expt,i})^2}} \quad (1)$$

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b09534.

NMR spectra of oligosaccharides and protein−ligand preparations as well as analysis data derived therefrom (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

*stenmark@dbb.su.se

*goran.widmalm@su.se

### ORCID

Göran Widmalm: 0000-0001-8303-4481

### Present Address

‖R.P.-A.B.: Department of Medical Biochemistry and Biophysics, Umeå University, S-901 87 Umeå, Sweden.

### Author Contributions

§C.H. and R.P.-A.B. contributed equally.


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Research Council (2010-5200, 2014-5667) to P.S. and (2013-4859) to G.W., The Knut and Alice Wallenberg Foundation to G.W., and the Wenner-Gren Foundations and the Swedish Cancer Society to P.S. and by an EMBO Long Term Fellowship and Marie Curie Actions (EMBOCOFUND2010, GA-2010-267146) to R.P-A.B. We thank the beamline scientists at BESSY, Berlin, and Max-Lab, Lund, for their support in data collection and Biostruct-X for support.

## ■ REFERENCES

(1) Arnon, S. S.; Schechter, R.; Inglesby, T. V.; Henderson, D. A.; Bartlett, J. G.; Ascher, M. S.; Eitzen, E.; Fine, A. D.; Hauer, J.; Layton,



DOI: 10.1021/jacs.6b09534
J. Am. Chem. Soc. 2017, 139, 218−230

M.; Lillibridge, S.; Osterholm, M. T.; O'Toole, T.; Parker, G.; Pearl, T. M.; Russell, P. K.; Swerdlow, D. L.; Tonat, K. *JAMA* **2001**, *285*, 1059−1070.
(2) Gill, D. M. *Microbiol. Rev.* **1982**, *46*, 86−94.
(3) Dolly, J. O.; Lawrence, G. W.; Meng, J.; Wang, J.; Ovsepian, S. V. *Curr. Opin. Pharmacol.* **2009**, *9*, 326−335.
(4) Johnson, E. A. *Annu. Rev. Microbiol.* **1999**, *53*, 551−575.
(5) Barash, J. R.; Arnon, S. S. *J. Infect. Dis.* **2014**, *209*, 183−191.
(6) Montal, M. *Annu. Rev. Biochem.* **2010**, *79*, 591−617.
(7) Schiavo, G.; Matteoli, M.; Montecucco, C. *Physiol. Rev.* **2000**, *80*, 717−766.
(8) Swaminathan, S. *FEBS J.* **2011**, *278*, 4467−4485.
(9) Montecucco, C. *Trends Biochem. Sci.* **1986**, *11*, 314−317.
(10) Berntsson, R. P.-A.; Peng, L.; Dong, M.; Stenmark, P. *Nat. Commun.* **2013**, *4*, 2058.
(11) Chai, Q.; Arndt, J. W.; Dong, M.; Tepp, W. H.; Johnson, E. A.; Chapman, E. R.; Stevens, R. C. *Nature* **2006**, *444*, 1096−1100.
(12) Jin, R.; Rummel, A.; Binz, T.; Brunger, A. T. *Nature* **2006**, *444*, 1092−1095.
(13) Rummel, A. *Curr. Top. Microbiol. Immunol.* **2012**, *364*, 61−90.
(14) Simpson, L. L.; Rapport, M. M. *J. Neurochem.* **1971**, *18*, 1761−1767.
(15) van Heyningen, W. E. *J. Gen. Microbiol.* **1959**, *20*, 310−320.
(16) Fotinou, C.; Emsley, P.; Black, I.; Ando, H.; Ishida, H.; Kiso, M.; Sinha, K. A.; Fairweather, N. F.; Isaacs, N. W. *J. Biol. Chem.* **2001**, *276*, 32274−32281.
(17) Fu, Z.; Chen, C.; Barbieri, J. T.; Kim, J.-J. P.; Baldwin, M. R. *Biochemistry* **2009**, *48*, 5631−5641.
(18) Rummel, A.; Häfner, K.; Mahrhold, S.; Darashchonak, N.; Holt, M.; Jahn, R.; Beermann, S.; Karnath, T.; Bigalke, H.; Binz, T. *J. Neurochem.* **2009**, *110*, 1942−1954.
(19) Rummel, A.; Mahrhold, S.; Bigalke, H.; Binz, T. *Mol. Microbiol.* **2004**, *51*, 631−643.
(20) Schmitt, J.; Karalewitz, A.; Benefield, D. A.; Mushrush, D. J.; Pruitt, R. N.; Spiller, B. W.; Barbieri, J. T.; Lacy, D. B. *Biochemistry* **2010**, *49*, 5200−5205.
(21) Stenmark, P.; Dong, M.; Dupuy, J.; Chapman, E. R.; Stevens, R. C. *J. Mol. Biol.* **2010**, *397*, 1287−1297.
(22) Stenmark, P.; Dupuy, J.; Imamura, A.; Kiso, M.; Stevens, R. C. *PLoS Pathog.* **2008**, *4*, e1000129.
(23) Swaminathan, S.; Eswaramoorthy, S. *Nat. Struct. Biol.* **2000**, *7*, 693−699.
(24) Berntsson, R. P.-A.; Peng, L.; Svensson, L. M.; Dong, M.; Stenmark, P. *Structure* **2013**, *21*, 1602−1611.
(25) Karalewitz, A. P.-A.; Fu, Z.; Baldwin, M. R.; Kim, J.-J. P.; Barbieri, J. T. *J. Biol. Chem.* **2012**, *287*, 40806−40816.
(26) Strotmeier, J.; Gu, S.; Jutzi, S.; Mahrhold, S.; Zhou, J.; Pich, A.; Eichner, T.; Bigalke, H.; Rummel, A.; Jin, R.; Binz, T. *Mol. Microbiol.* **2011**, *81*, 143−156.
(27) Strotmeier, J.; Lee, K.; Völker, A. K.; Mahrhold, S.; Zong, Y.; Zeiser, J.; Zhou, J.; Pich, A.; Bigalke, H.; Binz, T.; Rummel, A.; Jin, R. *Biochem. J.* **2010**, *431*, 207−216.
(28) Zhang, Y.; Buchko, G. W.; Qin, L.; Robinson, H.; Varnum, S. M. *Biochem. Biophys. Res. Commun.* **2010**, *401*, 498−503.
(29) Peng, L.; Berntsson, R. P.-A.; Tepp, W. H.; Pitkin, R. M.; Johnson, E. A.; Stenmark, P.; Dong, M. *J. Cell Sci.* **2012**, *125*, 3233−3242.
(30) Benoit, R. M.; Frey, D.; Hilbert, M.; Kevenaar, J. T.; Wieser, M. M.; Stirnimann, C. U.; McMillan, D.; Ceska, T.; Lebon, F.; Jaussi, R.; Steinmetz, M. O.; Schertler, G. F. X.; Hoogenraad, C. C.; Capitani, G.; Kammerer, R. A. *Nature* **2014**, *505*, 108−111.
(31) Dong, M.; Liu, H.; Tepp, W. H.; Johnson, E. A.; Janz, R.; Chapman, E. R. *Mol. Biol. Cell* **2008**, *19*, 5226−5237.
(32) Dong, M.; Yeh, F.; Tepp, W. H.; Dean, C.; Johnson, E. A.; Janz, R.; Chapman, E. R. *Science* **2006**, *312*, 592−596.
(33) Mahrhold, S.; Rummel, A.; Bigalke, H.; Davletov, B.; Binz, T. *FEBS Lett.* **2006**, *580*, 2011−2014.
(34) Mahrhold, S.; Strotmeier, J.; Garcia-Rodriguez, C.; Lou, J.; Marks, J. D.; Rummel, A.; Binz, T. *Biochem. J.* **2013**, *453*, 37−47.
(35) Jacky, B. P. S.; Garay, P. E.; Dupuy, J.; Nelson, J. B.; Cai, B.; Molina, Y.; Wang, J.; Steward, L. E.; Broide, R. S.; Francis, J.; Aoki, K. R.; Stevens, R. C.; Fernández-Salas, E. *PLoS Pathog.* **2013**, *9*, e1003369.
(36) Drickamer, K.; Taylor, M. E. *Annu. Rev. Cell Biol.* **1993**, *9*, 237−264.
(37) *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009.
(38) Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. *Acc. Chem. Res.* **2013**, *46*, 946−954.
(39) Yowler, B. C.; Kensinger, R. D.; Schengrund, C.-L. *J. Biol. Chem.* **2002**, *277*, 32815−32819.
(40) Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784−1788.
(41) Wagstaff, J. L.; Taylor, S. L.; Howard, M. J. *Mol. BioSyst.* **2013**, *9*, 571−577.
(42) Bhunia, A.; Bhattacharjya, S.; Chatterjee, S. *Drug Discovery Today* **2012**, *17*, 505−513.
(43) Angulo, J.; Nieto, P. M. *Eur. Biophys. J.* **2011**, *40*, 1357−1369.
(44) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1982**, *48*, 402−417.
(45) Bhunia, A.; Schwardt, O.; Gäthje, H.; Gao, G.-P.; Kelm, S.; Benie, A. J.; Hricovini, M.; Peters, T.; Ernst, B. *ChemBioChem* **2008**, *9*, 2941−2945.
(46) Post, C. B. *Curr. Opin. Struct. Biol.* **2003**, *13*, 581−588.
(47) Cala, O.; Guillière, F.; Krimm, I. *Anal. Bioanal. Chem.* **2014**, *406*, 943−956.
(48) Lycknert, K.; Edblad, M.; Imberty, A.; Widmalm, G. *Biochemistry* **2004**, *43*, 9647−9654.
(49) Landström, J.; Bergström, M.; Hamark, C.; Ohlson, S.; Widmalm, G. *Org. Biomol. Chem.* **2012**, *10*, 3019−3032.
(50) Bewley, C. A.; Shahzad-Ul-Hussan, S. *Biopolymers* **2013**, *99*, 796−806.
(51) Kitamura, M.; Iwamori, M.; Nagai, Y. *Biochim. Biophys. Acta, Gen. Subj.* **1980**, *628*, 328−335.
(52) Cazet, A.; Julien, S.; Bobowski, M.; Burchell, J.; Delannoy, P. *Breast Cancer Res.* **2010**, *12*, 204.
(53) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455−461.
(54) Neumann, D.; Kohlbacher, O. In *Proc. Int. Beilstein Symp. Glyco-Bioinformatics*; Hicks, M. G., Kettner, C., Eds.; Beilstein-Institut: Frankfurt/Main, 2009; pp 101−122.
(55) Landström, J.; Persson, K.; Rademacher, C.; Lundborg, M.; Wakarchuk, W.; Peters, T.; Widmalm, G. *Glycoconjugate J.* **2012**, *29*, 491−502.
(56) Nivedha, A. K.; Makeneni, S.; Lachele Foley, B.; Tessier, M. B.; Woods, R. J. *J. Comput. Chem.* **2014**, *35*, 526−539.
(57) Benson, M. A.; Fu, Z.; Kim, J.-J. P.; Baldwin, M. R. *J. Biol. Chem.* **2011**, *286*, 34015−34022.
(58) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108−6117.
(59) Mayer, M.; James, T. L. *J. Am. Chem. Soc.* **2004**, *126*, 4453−4460.
(60) Sabesan, S.; Duus, J. Ø.; Fukunaga, T.; Bock, K.; Ludvigsen, S. *J. Am. Chem. Soc.* **1991**, *113*, 3236−3246.
(61) Koerner, T. A. W., Jr.; Prestegard, J. H.; Demou, P. C.; Yu, R. K. *Biochemistry* **1983**, *22*, 2676−2687.
(62) Scarsdale, J. N.; Prestegard, J. H.; Yu, R. K. *Biochemistry* **1990**, *29*, 9843−9855.
(63) Meyer, B.; Peters, T. *Angew. Chem., Int. Ed.* **2003**, *42*, 864−890.
(64) Landström, J.; Nordmark, E.-L.; Eklund, R.; Weintraub, A.; Seckler, R.; Widmalm, G. *Glycoconjugate J.* **2008**, *25*, 137−143.
(65) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288−292.
(66) Lacy, D. B.; Tepp, W.; Cohen, A. C.; DasGupta, B. R.; Stevens, R. C. *Nat. Struct. Biol.* **1998**, *5*, 898−902.
(67) DeMarco, M. L.; Woods, R. J. *Glycobiology* **2009**, *19*, 344−355.
(68) Lütteke, T.; Frank, M.; von der Lieth, C.-W. *Nucleic Acids Res.* **2005**, *33*, D242−D246.


DOI: 10.1021/jacs.6b09534
*J. Am. Chem. Soc.* 2017, 139, 218−230

(69) Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. J. *J. Mol. Biol.* **1998**, *282*, 1043−1059.
(70) Jayalakshmi, V.; Rama Krishna, N. *J. Magn. Reson.* **2004**, *168*, 36−45.
(71) Yowler, B. C.; Schengrund, C.-L. *Biochemistry* **2004**, *43*, 9725−9731.
(72) Kang, Y.; Gohlke, U.; Engström, O.; Hamark, C.; Scheidt, T.; Heinemann, U.; Widmalm, G.; Santer, M.; Barbirz, S. *J. Am. Chem. Soc.* **2016**, *138*, 9109−9118.
(73) Jayalakshmi, V.; Krishna, N. R. *J. Magn. Reson.* **2002**, *155*, 106−118.
(74) Rama Krishna, N.; Jayalakshmi, V. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *49*, 1−25.
(75) Enríquez-Navas, P. M.; Marradi, M.; Padro, D.; Angulo, J.; Penadés, S. *Chem. - Eur. J.* **2011**, *17*, 1547−1560.
(76) Kuttel, M. M.; Ståhle, J.; Widmalm, G. *J. Comput. Chem.* **2016**, *37*, 2098−2105.
(77) Lundborg, M.; Widmalm, G. *Anal. Chem.* **2011**, *83*, 1514−1517.
(78) Rönnols, J.; Pendrill, R.; Fontana, C.; Hamark, C.; Angles d'Ortoli, T.; Engström, O.; Ståhle, J.; Zaccheus, M. V.; Säwén, E.; Hahn, L. E.; Iqbal, S.; Widmalm, G. *Carbohydr. Res.* **2013**, *380*, 156−166.
(79) Pedretti, A.; Villa, L.; Vistoli, G. *J. Mol. Graphics Modell.* **2002**, *21*, 47−49.
(80) Morris, G.; Goodsell, D.; Halliday, R.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(81) Roslund, M. U.; Säwén, E.; Landström, J.; Rönnols, J.; Jonsson, K. H. M.; Lundborg, M.; Svensson, M. V.; Widmalm, G. *Carbohydr. Res.* **2011**, *346*, 1311−1319.
(82) Mills, R. *J. Phys. Chem.* **1973**, *77*, 685−688.
(83) Cano, K. E.; Thrippleton, M. J.; Keeler, J.; Shaka, A. J. *J. Magn. Reson.* **2004**, *167*, 291−297.
(84) Kupče, E.; Boyd, J.; Campbell, I. D. *J. Magn. Reson., Ser. B* **1995**, *106*, 300−303.
(85) Hwang, T.; Shaka, A. *J. Magn. Reson., Ser. A* **1995**, *112*, 275−279.
(86) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 125−132.
(87) Collaborative Computational Project, Number 4. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *D50*, 760−763.
(88) Grant, O.; Woods, R. *Curr. Opin. Struct. Biol.* **2014**, *28*, 47−55.
(89) Widmalm, G.; Pastor, R. W.; Bull, T. E. *J. Chem. Phys.* **1991**, *94*, 4097−4098.
(90) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. *J. Biomol. NMR* **2011**, *50*, 43−57.
(91) Pendrill, R.; Jonsson, K. H. M.; Widmalm, G. *Pure Appl. Chem.* **2013**, *85*, 1759−1770.


DOI: 10.1021/jacs.6b09534
*J. Am. Chem. Soc.* 2017, 139, 218−230