Modeling the Carbohydrate-Binding Specificity of Pig Edema Toxin†

Maxwell D. Cummings,‡,§ Hong Ling,‡ Glen D. Armstrong,|| James L. Brunton,⊥ and Randy J. Read*,‡,||

Departments of Biochemistry and Medical Microbiology & Immunology, University of Alberta, Edmonton, Alberta T6G 2H7, Canada, and Departments of Medicine and Microbiology, The Toronto Hospital at the University of Toronto and Samuel Lunenfeld Research Institute at Mount Sinai Hospital, Toronto, Ontario M5G 2C4, Canada

Received July 24, 1997; Revised Manuscript Received November 25, 1997

ABSTRACT: The wild-type binding pentamer of Shiga-like toxin IIe (SLT-IIe) binds both the globotriaosylceramide (Gb_3) and globotetraosylceramide (Gb_4) cell surface glycolipids, whereas the double mutant GT3 (Q65E/K67Q) exhibits a marked preference for Gb_3 [Tyrrell, G. J., et al. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 524−528]. We modeled three unique sites (sites 1−3) for binding of the carbohydrate moiety of Gb_3 to GT3 and SLT-IIe, on the basis of the three sites observed for the SLT-I pentamer [Ling, H., et al. (1998) *Biochemistry 37*, 1777−1788]. Examination of the three sites in light of various mutation and binding data strongly suggested that one of the binding sites plays a role in the change of specificity observed for the GT3 mutant. We applied several modeling techniques, and developed a model for binding of the carbohydrate moiety of Gb_4 to this site of the SLT-IIe binding pentamer. This model is consistent with a wide variety of mutation and binding data and clearly shows the importance of the terminal Gal*N*Ac residue of Gb_4, as well as that of the two mutated residues of GT3, to the intermolecular interaction.

The Shiga-like toxins (SLTs),[1] also known as verotoxins (VTs), are a class of bacterial toxins expressed by certain *Escherichia coli* serotypes known to cause several human and animal diseases. Members of this group include Shiga-like toxins I (SLT-I or VT-1) and II (SLT-II or VT-2), pig edema toxin (SLT-IIe, VT-2e, or VT-E), and Shiga-like toxin IIc [SLT-IIc or VT-2c (*1*)]. They have a hexameric AB_5-subunit composition; A represents the catalytically active monomeric subunit that ultimately elicits the toxic effect, and B_5 denotes the homopentameric binding subunit. The binding subunits of these toxins facilitate entry into certain host cells, by virtue of their ability to bind to specific host cell glycolipids.

There is greater than 60% sequence identity between the binding subunits of all Shiga-like toxins (*1*). SLT-I, SLT-II, and SLT-IIc bind globotriaosylceramide [Gb_3, αGal(1−4)*β*Gal(1−4)*β*Glc-Cer, where Cer is ceramide] preferentially, and the prevalence of this glycolipid in kidney tissue contributes to their renal toxicity (*1*). The binding subunit of SLT-IIe, which shows 84% sequence identity with that of SLT-II, prefers to bind globotetraosylceramide [Gb_4, *β*Gal*N*Ac(1−3)αGal(1−4)*β*Gal(1−4)*β*Glc-Cer], although it does bind Gb_3 as well (*1*). Brunton and co-workers explored these differences in binding preference by constructing several mutants of the SLT-I and SLT-IIe binding subunits (*2*). One SLT-IIe double mutant (GT3, Q65E/K67Q) was particularly interesting, in that its binding preference was switched from Gb_4 to Gb_3 (*2*). Later it was shown that this change in binding preference gave rise to a corresponding change in the in vivo activity of the toxin (*3*). GT3, with Gb_3/Gb_4 binding activity similar to that of SLT-I, had a pathology similar to that of SLT-I, and this was distinct from that of wild-type SLT-IIe (*3*).

Structural studies of these lectins, both free and complexed with their carbohydrate receptors, will help in understanding the binding interaction at the atomic level. In turn, such knowledge may aid in the development of vaccines and/or chemotherapeutic agents that block toxin binding. Armstrong and co-workers have recently demonstrated the feasibility of such an approach to the diagnosis and/or treatment of enterohemorrhagic *E. coli* infections (*4*, *5*). Hol and colleagues are compiling a body of structural information

† This research was supported by the Medical Research Council of Canada, and in part by an International Research Scholarship to R.J.R. from the Howard Hughes Medical Institute. R.J.R. is a Scholar of the Alberta Heritage Foundation for Medical Research. M.D.C. was the holder of a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship while on leave from SynPhar Laboratories, Inc., Edmonton. Financial assistance to G.D.A. was from the Canadian Bacterial Diseases Network and to J.L.B. was from the Medical Research Council (MRC Grants PG11123 and MT13071).

* To whom correspondence should be addressed. Telephone: 403-492-4305. Fax: 403-492-7521. E-mail: Randy.Read@ualberta.ca.

‡ Department of Biochemistry, University of Alberta.

§ Present address: Computational Chemistry/Molecular Modeling, SmithKline Beecham Pharmaceuticals, 709 Swedeland Rd., King of Prussia, PA 19406.

|| Department of Medical Microbiology & Immunology, University of Alberta.

⊥ The Toronto Hospital at the University of Toronto and Samuel Lunenfeld Research Institute at Mount Sinai Hospital.

[1] Abbreviations: Gal, galactose; Gal*N*Ac, *N*-acetylgalactosamine; Glc, glucose; galabiose, α Gal(1−4)*β*Gal; lactose, *β*Gal(1−4)*β*Glc; Gb_3, globotriaosylceramide, αGal(1−4)*β*Gal(1−4)*β*Glc-Cer, where Cer is ceramide; Gb_4, globotetraosylceramide, *β*Gal*N*Ac(1−3)αGal(1−4)*β*Gal-(1−4)*β*Glc-Cer (note that the terms Gb_3 and Gb_4 are used throughout to refer to the carbohydrate moieties of these glycolipids, as well as to the complete glycolipids; when the distinction between carbohydrate and glycolipid is not clearly implied by context, it is specified); SLT-I, Shiga-like toxin I or verotoxin 1; SLT-IIe, Shiga-like toxin IIe, verotoxin-2e or pig edema toxin; GT3, Q65E/K67Q double mutant of the binding subunit of SLT-IIe; *ϕ*, H1−C1−O1−C*x*; *ψ*, C1−O1−C*x*−H*x*.

S0006-2960(97)01807-2 CCC: $15.00
Published on Web 01/30/1998

useful as a basis for the development of a similar approach for the treatment of cholera (ref *6* and references therein). Previous reports from this laboratory have described the crystallographic structure of the binding subunit of SLT-I (*7*), and its structural relationship to the binding subunit of the cholera toxin family (*8*). Nyholm et al. (*24*) modeled the binding of the Gb_3 to one site of SLT-I and, more recently, that of several Gb_3 and Gb_4 analogues to two distinct sites on SLT-I and several SLT-II variants (*9*). Workers in this laboratory have reported the structure of SLT-I with a Gb_3 analogue bound at three distinct sites (*12*), and the refinement of the structure of GT3 complexed with the same Gb_3 analogue is also underway (H. Ling and R. J. Read, unpublished results).

Since Gb_4 differs from Gb_3 only by the addition of a terminal *N*-acetylgalactosamine moiety, it is likely that Gb_3 and Gb_4 bind to SLT-IIe with similar binding modes. Related similarities in the binding of various carbohydrates have been observed in structural studies of the heat-labile enterotoxin of *E. coli* (*10*). If this is the case, then the loss of Gb_4 affinity observed for the SLT-IIe double mutant GT3 may be due to the gain or loss of one or a few specific interactions with the mutated residues (*2*, *11*) or proximal side chains that interact with either or both of the mutated residues. Here we report the results of modeling studies that provide an explanation for the difference in carbohydrate binding preferences of SLT-IIe and the double mutant GT3.

MATERIALS AND METHODS

Protein Structure Preparation. The structure of the binding pentamer of SLT-I in complex with the Gb_3 analogue Pk-MCO (the carbohydrate moiety of which is identical to that of Gb_3) has been determined (*12*), and the structure of a similar complex with the SLT-IIe double mutant GT3 is currently being refined in this laboratory. The SLT-I complex has four pentamers in the asymmetric unit; the GT3 complex has one. The refinement of the Pk-MCO/GT3 complex had just begun when this modeling study was performed; the experimentally observed GT3 protein structure was used as a starting point for this work (see below), but none of the (more recently) observed Pk-MCO binding sites on GT3 were used. Instead, we modeled the Gb_3 binding sites of GT3 by comparing GT3 with the Pk-MCO/SLT-I complex (*12*) (Gb_3 binding sites were inferred directly from the observed Pk-MCO binding sites). However, our modeling results would not have been significantly different if we had used the Pk-MCO positions from the complex with GT3. We arbitrarily chose pentamer 1 (VBA1-VBE1) of the SLT-I structure for our studies. For both SLT-I and GT3, we used a trimer consisting of monomers A, D, and E of the binding pentamer for all of our comparisons and minimizations, to reduce the size of the system being studied. The binding sites of monomer E were studied, with the assumption that the binding sites of this monomer would, in the presence of the neighboring subunits A and D, behave in a manner similar to those of this monomer in the pentamer.

Protein structures were initially prepared according to the general method employed in our previous docking studies (*13*, *14*). Monomers B and C, water molecules, and bound carbohydrates were deleted. Hydrogens were added to the protein in the standard way with INSIGHTII (Biosym/MSI, San Diego) at neutral pH. Polar hydrogen positions were then optimized with NETWORK (*15*), which maximizes hydrogen-bonding networks. This was followed by a round each of steepest descents and conjugate gradient minimization (200 steps maximum in each round, heavy atoms fixed) with the CVFF potential function of DISCOVER.

To construct the initial wild-type version of pig edema toxin from the crystal structure of the double mutant GT3, we simply converted the two mutant residues to those of wild-type SLT-IIe (E65Q/Q67K) using the standard procedure in INSIGHTII (Biosym/MSI). For the initial model, we selected the rotamers [using the rotamer library of Ponder and Richards (*16*), as implemented in INSIGHTII] that were most similar to the conformations observed for the two mutant residues in the crystal structure. These side chains were adjusted during manual docking, and were also unconstrained during the subsequent energy minimizations of the complexes.

The three Gb_3 binding sites observed in the complex with SLT-I were reproduced in our GT3/SLT-IIe models by superimposing the Pk-MCO/SLT-I complex (see below). Coordinates were superimposed and compared using unpublished programs written by Trevor Hart. Figures 2 and 4−7 were prepared with the MOLSCRIPT program (*17*).

Calculation of Carbohydrate Conformations. Starting models of the carbohydrate portions of Gb_3 and Gb_4 were generated by a combination of simulation methods, similar to one of the procedures suggested by Tvaroška and Pérez (*23*). Rigid grid searches were carried out with the hard-spheres program GEGOP (*18*) to find low-energy glycoside conformations for each of the two relevant disaccharides, lactose [βGal(1−4)βGlc] and galabiose [αGal(1−4)βGal], of Gb_3. In these searches, ϕ and ψ were searched in 10° steps, and the glycosidic bond angle and the C5−C6 torsion were optimized at each step. We then selected several low-energy conformations of each of the two disaccharides for further minimization with GEGOP, in which the glycosidic linkage, the C5−C6 torsion, and all hydroxyl groups were allowed to vary. Following this minimization, we constructed the possible Gb_3 conformers from the final disaccharide conformations and repeated the final minimization. The resultant Gb_3 structures were then imported into the DISCOVER (Biosym/MSI) version of AMBER (*19*), with the additional carbohydrate parameters developed by Homans (*20*). Charges and atom types were assigned as described by Homans (*20*). All heavy atoms were fixed, and hydrogen positions were minimized by the steepest descents method for a maximum of 200 steps. Following this, all atoms were optimized for a maximum of 500 steps of conjugate gradients minimization. For the AMBER minimizations, a distance-dependent dielectric ($\epsilon = 4r$) was used, as described in the program documentation (Biosym/MSI).

We generated three low-energy conformers of the terminal disaccharide of Gb_4 [βGal*N*Ac(1−3)α Gal], as described above for the Gb_3 disaccharides. Three Gb_4 conformers were then generated by combining the Gb_3 conformer of interest (see below) with each of the three low-energy conformations of βGal*N*Ac(1−3)αGal. These three Gb_4 structures were then minimized first with GEGOP and then with DISCOVER/AMBER, as described above for the Gb_3 conformers.

Carbohydrate/Protein Complexes. The starting point for modeling the interaction of the Gb_3 and Gb_4 carbohydrates

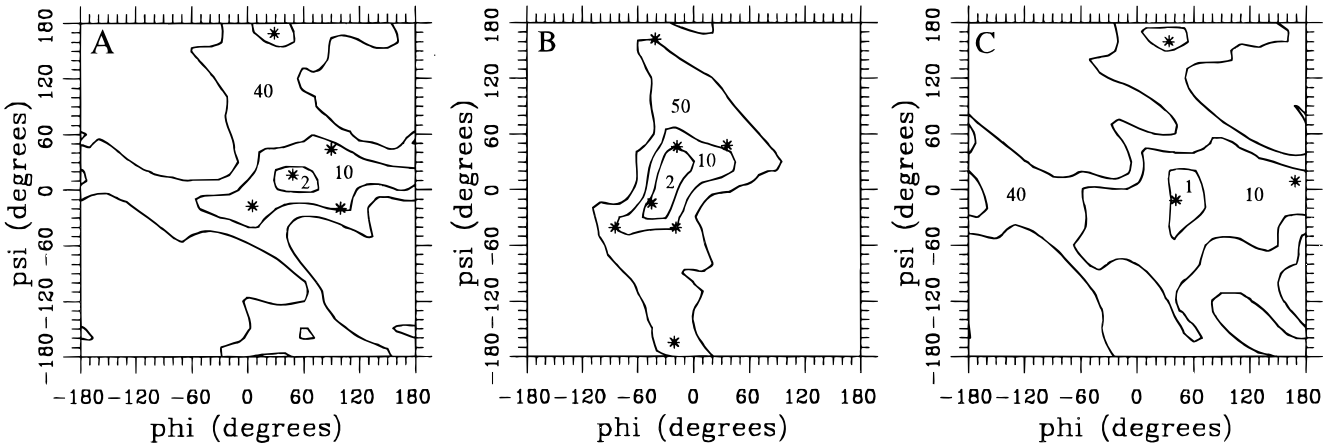

FIGURE 1: Conformational analysis of the glycosidic linkages of the constituent disaccharides of Gb_3 and Gb_4. Plots were obtained from the results of the original rigid grid searches with GEGOP, as described in the text. The contours contain the conformational space occupied by conformers that fall below the energy values (in kilocalories per mole) shown within the contours. Asterisks denote the conformers chosen for further refinement after the initial round of GEGOP analysis (see the text). (A) For lactose, the two final ϕ/ψ pairs used in Gb_3 construction were 52/4 and 30/168. (B) For galabiose, the two final ϕ/ψ pairs used in Gb_3 construction were −40/−16 and −23/−157. (C) For βGal*N*Ac(1−3)αGal, the three final ϕ/ψ pairs used in Gb_4 construction were 54/−6, 170/−1, and 36/159.

with wild-type SLT-IIe was the three distinct Gb_3 binding sites modeled for monomer E of the binding subunit of SLT-IIe by superimposing the SLT-I complex onto our SLT-IIe model. Gb_4 was initially modeled in these sites by superimposing the galabiose heavy atoms of Gb_4 onto those of Gb_3. Selected side chains (see below), the exocyclic moieties of the *N*-acetylgalactosamine ring, and the glycosidic linkage of the βGal*N*Ac(1−3)αGal disaccharide of Gb_4 were manually adjusted to optimize the protein−carbohydrate interaction and alleviate any bad contacts. These manually built complexes were then minimized with the DISCOVER version of AMBER, as described above for the carbohydrate structures. In some cases (see below), the cycle of manual model optimization followed by constrained energy minimization was repeated once or twice.

RESULTS AND DISCUSSION

Calculated Gb_3 Carbohydrate Conformations. The ϕ/ψ plots of lactose and galabiose generated by the rigid grid searches with GEGOP are shown in Figure 1. For these simple disaccharides, we expect this relatively crude hard-spheres method to give a reasonable estimate of the low-energy conformers, although the energy barriers separating conformers will be overestimated (*21*−*23*). We chose several conformers of each of the two disaccharides (Figure 1) for further minimization with GEGOP (see Materials and Methods). After this more rigorous flexible minimization, we were left with just two distinct ϕ/ψ conformers for each of the two disaccharides (Figure 1, Table 1). Four possible Gb_3 conformers were constructed by combining these disaccharide conformations. Table 1 shows the ϕ and ψ angles and energy values for the Gb_3 conformers after refinement with GEGOP and then with DISCOVER/AMBER. We note that conformer 1, the lowest-energy conformation, is similar to the carbohydrate moiety of the low-energy Gb_3 conformer modeled by Nyholm et al. (*24*) in their docking study of Gb_3 and SLT-I and also resembles the lowest-energy conformer described by Poppe et al. (*21*) in their combined NMR and computational analysis of Gb_3 conformation.

Comparison of Calculated and Bound Gb_3 Conformations. At the time that this modeling study was carried out, three distinct Gb_3 binding sites per monomer had been observed for the binding subunit of SLT-I (see ref *12*), whereas only one Gb_3 site per monomer had been observed for the GT3 complex (Figure 2). Therefore, these four binding sites were used for our analysis of bound Gb_3 conformers (Figure 2, Table 2). More recently, a second site, similar to site 1 of the SLT-I complex shown in Figure 2, has been observed for the GT3 complex (H. Ling and R. J. Read, unpublished results). The ranges and average ϕ and ψ values of the Gb_3 conformers bound at each of the four sites analyzed are shown in Table 2. Our values differ in detail, but not in essence, from those presented in the preceding structural paper (*12*), reflecting more extensive crystallographic refinement. For the SLT-I structure, the glycoside conformations of the lactose moieties show wide variation over the 20 copies of each of the three unique binding sites (Table 2). This is a reflection of the inherent flexibility of this glycosidic linkage (Figure 1; ref *21* and references therein), the lack of contact between the protein and the glucose residue, and, in the case of a few of the sites, crystal contacts. The average values of these parameters are, however, similar for the three SLT-I sites, as well as for the single GT3 site analyzed (Table 2). The galabiose moieties are much more restricted,

Table 1: Energies and Glycoside Conformations of Low-Energy Gb_3 Conformers

stage of refinement[a]	energy (kcal mol^{-1})	ϕ^b (deg)	ψ^b (deg)	ϕ (deg)	ψ (deg)
		galabiose		lactose	
conformer 1					
GEGOP	−4.7	−39.6	−16.3	52.9	3.6
DISCOVER	−8.9	−46.6	−9.7	49.1	1.5
conformer 2					
GEGOP	1.6	−39.7	−16.1	30.3	168.4
DISCOVER	−7.3	−46.6	−9.7	52.2	−164.2
conformer 3					
GEGOP	11.6	−23.2	−157.1	56.6	1.1
DISCOVER	−5.0	−20.3	−161.5	49.4	1.6
conformer 4					
GEGOP	18.3	−23.1	−156.9	31.9	168.1
DISCOVER	−3.5	−19.9	−163.8	52.0	−165.5

[a] Energies and dihedral values refer to the final two stages of refinement of the modeled Gb3 conformations (see the text). [b] ϕ, H1−C1−O1−C*x*; ψ, C1−O1−C*x*−H*x*.

FIGURE 2: Crystallographic Gb_3 binding sites of SLT-I and GT3. The three unique Gb_3 binding sites of SLT-I (left) are shown, as well as the numbering used throughout the text. The single Gb_3 binding site of GT3 available at the time of this study is also shown (right). The figure clearly shows the relatedness of the Gb_3 binding site 2 for GT3 and SLT-I.

Table 2: Glycoside Conformations of Gb_3 Carbohydrates Bound to GT3 and SLT-I[a]

	ϕ^b (deg)	ψ^b (deg)	ϕ (deg)	ψ (deg)
conformer	galabiose		lactose	
GT3	−42.8	−17.1	38.3	−23.4
	(−46.6, −40.4)	(−17.9, −15.9)	(36.9, 39.4)	(−25.7, −19.0)
SLT-I site 1	−46.7	−13.6	57.6	−8.5
	(−67.9, −13.4[c])	(−34.4, 1.4)	(15.7, 139.8)	(−87.6, 68.4)
SLT-I site 2	−50.3	−11.1	45.2	1.5
	(−58.2, −37.3)	(−20.5, 0.8)	(30.1, 70.3)	(−21.2, 36.2)
SLT-I site 3	−43.4	−2.6	32.2	−8.8
	(−48.9, −36.7)	(−17.4, 11.6)	(−92.8, 78.0)	(−60.9, 63.1)

[a] Hydrogens were added in INSIGHTII in the standard way at neutral pH. [b] For SLT-I, values shown are the average of the 20 copies found in the asymmetric unit; for GT3, values shown are the average of the 5 copies found in the asymmetric unit. Values in parentheses define the range of values observed for that parameter; ϕ and ψ are defined in Table 1. [c] Only one value was above −36.5°; omission of this point (−13.4) gave a mean of −48.4.

reflecting the greater contact that this disaccharide makes with the protein. Of the four low-energy Gb_3 conformers that we modeled, the one most similar to the bound conformers is conformer 1, the lowest-energy conformer (Tables 1 and 2). This result validates our carbohydrate modeling methodology, which was also applied to modeling Gb_4 (see below).

Gb_3 Binding Sites on SLT-I and GT3. Several residues of the binding subunits of SLT-I and SLT-IIe have been studied by site-directed mutagenesis, and some of these residues have been inferred to be involved in receptor binding (*2*, *26*, *27*). These mutations are discussed in more detail in the preceding paper (*12*); here we focus on substitutions that affect the specificity of binding to Gb_3 or Gb_4. If alteration of a side chain has a marked effect on ligand binding, but is shown not to affect significantly other important properties of the protein, then the simplest interpretation is that the side chain in question interacts with the ligand. Although such reasoning is commonly invoked, in the absence of direct experimental verification of these interactions, such inferences are always somewhat tenuous. This point is of some concern to us, since the distance constraints offered by specific intermolecular interactions are potentially of great utility in computer-based docking simulations.

Residues Gln65 and Lys67 of SLT-IIe have been shown to be important for Gb_4 binding (*2*). The SLT-IIe double mutant GT3 (Q65E/K67Q) had a markedly reduced affinity for Gb_4, but Gb_3 binding was moderately increased (*2*). A reasonable conclusion is that the Gal*N*Ac moiety of Gb_4 makes specific favorable contacts with Gln65 and Lys67, and that these contacts are not possible for the double mutant GT3 (*2*). Our goal in this study was to provide an explanation for the effect of the GT3 mutations on Gb_3/Gb_4 binding.

To generate the three SLT-IIe Gb_3 binding sites from those observed for SLT-I (Figure 2), we superimposed the Cα atoms of selected residues of monomer E of SLT-I onto the corresponding atoms of SLT-IIe (Figure 3). The final rms difference for the superposed atoms was 0.39 Å.

Figure 4 shows the central monomer (monomer E) of the SLT-IIe binding subunit trimer used in our modeling experiments, with the residues discussed above highlighted. Also shown are the three Gb_3 "binding sites" of monomer E of SLT-IIe, modeled as described above. Visual inspection of this model shows clearly that the two residues mutated in GT3 are associated with site 1. Assuming a similar binding mode for the analogous portions of Gb_3 and Gb_4, both ligands would have the greatest interaction with these residues when

```
                  10        20        30        40        50        60
SLT-I:TPDCVTGKVEYTKYNDDDTFTVKVGDKELFTNRWNLQSLLLSAQITGMTVTIKTNACHNGGGFSEVIFR
      ||  ||  |  ||| | ||||||   |  ||||||| || ||| |||||||  | |  | || || |
GT3: ADCAKGKIEFSKYNEDNTFTVKVSGREYWTNRWNLQPLLQSAQLTGMTVTIISNTCSSGSGFAEVQFN
                  10        20        30        40        50        60    Q K
```

FIGURE 3: Alignment used for superposition of SLT-I onto GT3. Asterisks denote the two mutated residues of GT3, which we replaced in our model of SLT-IIe. The Cα's of the identical residues of SLT-I, denoted by vertical lines between the two sequences, were superimposed onto the corresponding atoms of GT3/SLT-IIe.

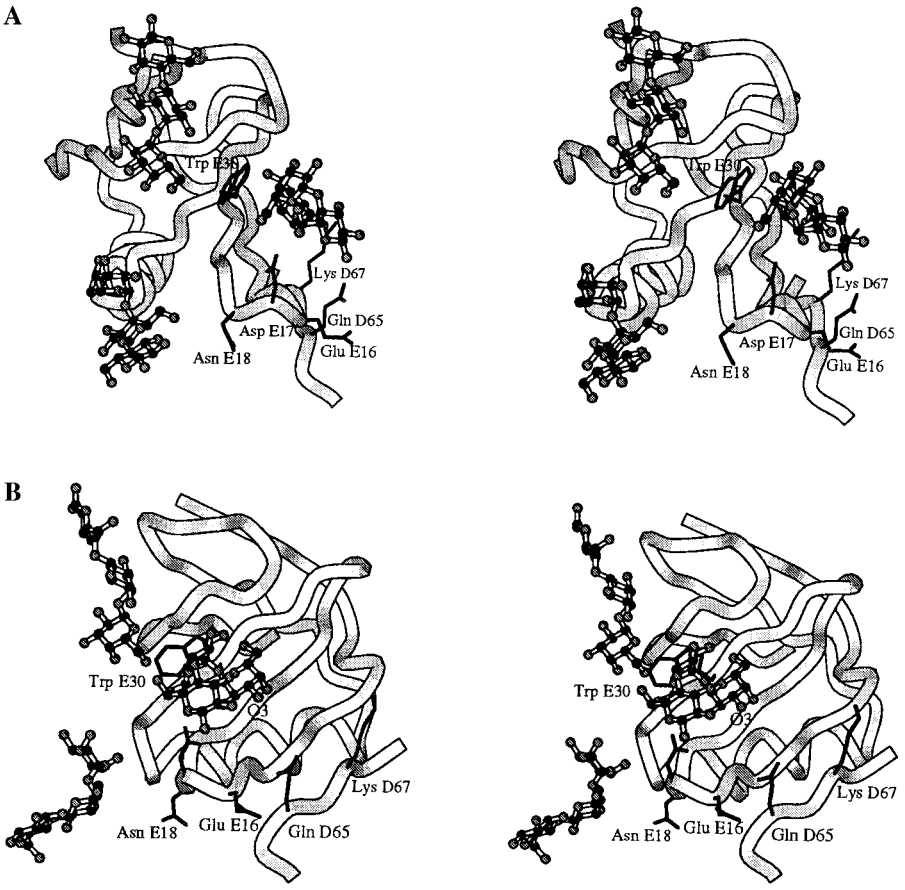

FIGURE 4: Initial model of SLT-IIe with three unique Gb_3 binding sites. The backbone of monomer E, and residues D63−D69, is included, with the side chains of several key residues (see the text) shown. The bottom figure is a 60° rotation of the top figure. O3 of the terminal Gal residue of Gb_3, the Gal*N*Ac linkage point for Gb_4, is labeled.

bound in site 1 (Figure 4). Since a direct interaction with the mutated residues offers the simplest explanation of the change in affinity, we decided to explore Gb_4 binding in site 1 of our modeled wild-type SLT-IIe.

Gb_4 Conformations. We assumed that the Gb_3 moiety of Gb_4 would bind at site 1 (discussed above) of our modeled wild-type pig edema toxin (SLT-IIe), in a mode similar to that of Gb_3 bound at the analogous site in the SLT-I complex. Thus, to generate reasonable conformers of Gb_4 for modeling the complexes, we had only to concern ourselves with the terminal βGal*N*Ac(1−3)αGal glycosidic linkage.

Figure 1 shows the original ϕ/ψ plot for this disaccharide generated with the GEGOP rigid grid search (as described for lactose and galabiose), and the three conformers chosen for further minimization with GEGOP. Table 3 lists the ϕ and ψ angles and energies of the three Gb_4 carbohydrate conformers constructed by combining Gb_3 conformer 1 (Table 1) and the three refined βGal*N*Ac(1−3)αGal conformers (Figure 1), following the final two rounds of energy minimization with GEGOP and DISCOVER/AMBER (as described above for the Gb_3 conformers).

The final conformation of the Gb_3 moiety of the three calculated Gb_4 conformers (essentially Gb_3 conformer 1) is in good agreement with the results of Poppe et al. (*21*, *25*). However, none of the three βGal*N*Ac(1−3)αGal conformations can be considered identical to those of the low-energy Gb_4 conformers described by Poppe et al. (*25*). In this respect, conformer 1 is most similar to the most popular low-energy Gb_4 conformer(s) of Poppe et al. (*25*); ϕ and ψ differ by approximately 20 and 40°, respectively. The free Gb_4 conformer modeled by Nyholm et al. (*9*) was also very similar to conformer 1 of this study.

Gb_4 Binding at Site 1 of SLT-IIe. Our modeling of the Gb_4/SLT-IIe complex commenced with the superposition of the galabiose moiety of our three calculated Gb_4 conformers onto the corresponding fragment of Gb_3 in site 1 of SLT-IIe. Figure 5 shows that the Gal*N*Acβ(1−3)Gal moiety of conformer 2 projects away from the protein.[2] Without major conformational changes to both the protein and carbohydrate, close contact between the Gal*N*Ac ring of Gb_4 and either residue GlnD65 or LysD67 seems unlikely for this Gb_4 conformation. Conformers 1 and 3 bound at this site seem

Table 3: Energies and Glycoside Conformations of Low-Energy Gb_4 Conformers[a]

stage of refinement[a]	energy (kcal mol^{-1})	ϕ^b (deg)	ψ^b (deg)	ϕ (deg)	ψ (deg)	ϕ (deg)	ψ (deg)
		GalNAcβ(1−3)Gal		galabiose		lactose	
conformer 1							
GEGOP	−6.8	51.8	2.7	−39.5	−16.2	53.5	3.4
DISCOVER	−25.3	53.3	−8.0	−43.2	−2.2	49.8	2.2
conformer 2							
GEGOP	−6.9	168.3	0.0	−39.8	−17.7	53.1	−1.0
DISCOVER	−27.6	175.6	−3.6	−46.6	−9.8	49.1	0.5
conformer 3							
GEGOP	−0.6	36.8	158.3	−39.6	−16.2	53.2	3.0
DISCOVER	−23.2	36.3	159.0	−47.0	−11.6	49.2	1.1

[a] Energies and dihedral values refer to the final two stages of refinement of the modeled Gb_4 conformations (see the text). [b] See Table 1.

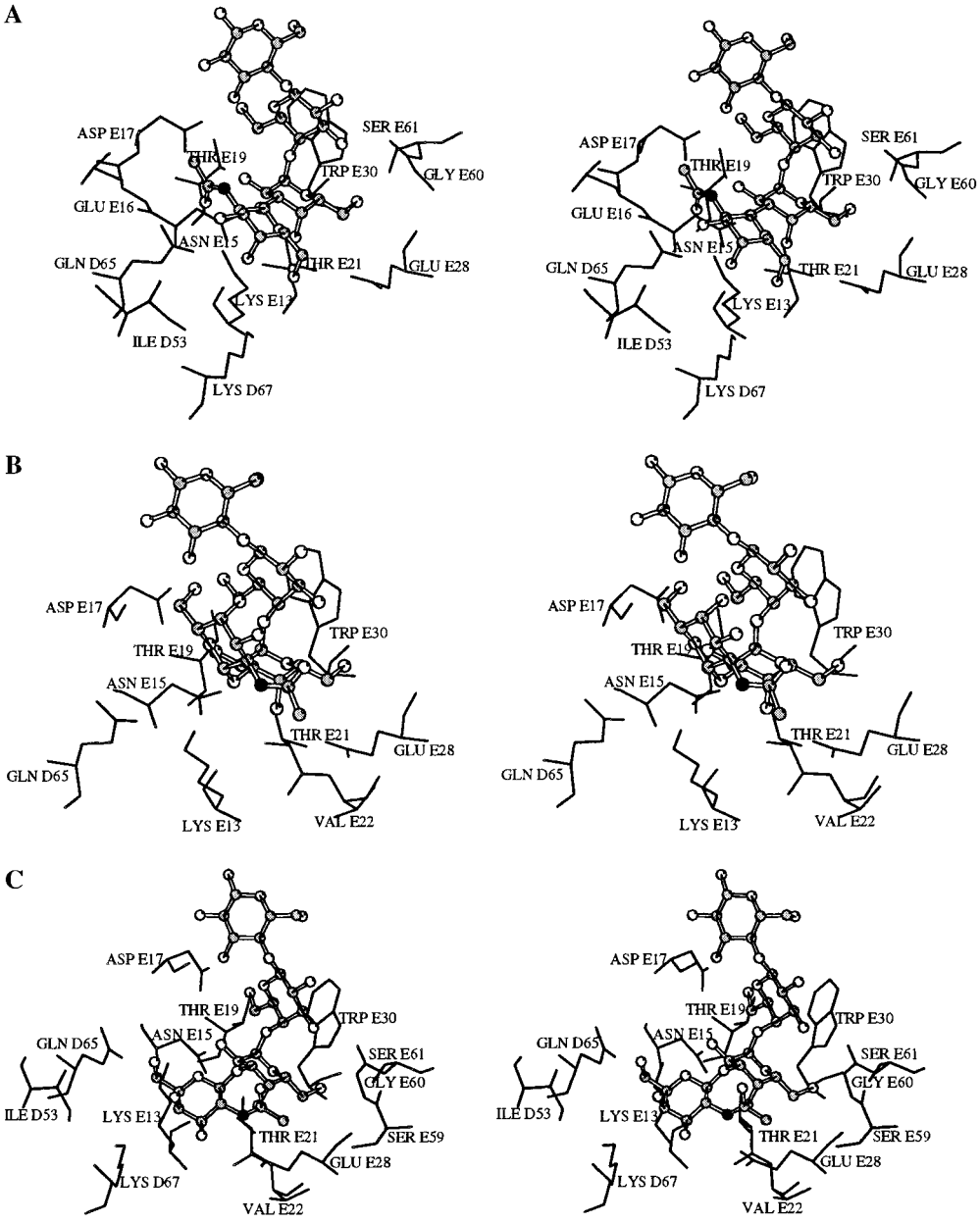

FIGURE 5: Minimized complexes of each of the three Gb_4 conformers at site 1 of SLT-IIe. Most of the residues with one or more atoms within 5 Å of a Gb_4 atom are shown: (A−C) Gb_4 conformers 1−3, respectively.

more promising, although both clash with the side chain of LysE13[3] (not shown, but see Figure 5).

We subjected each of the three complexes to energy minimization with DISCOVER/AMBER, using an initial 100 steps of highly constrained steepest descents minimization, followed by a more relaxed cycle of 300 steps of conjugate gradients minimization. Table 4 lists the constraints used in each of the two minimization protocols. Essentially, we allowed flexibility for side chains and backbone regions in close contact with Gb_3/Gb_4 "bound" at this site. This

Table 4: Constraints for Minimization of Gb_4/SLT-IIe Site 1 Complexes[a]

initial steepest descents		final conjugate gradients	
residues[b]	constraint[c]	residues	constraint
Gb_4 1	none	Gb_4 1	none
−	−	Gb_4 2−4	none
SLT-IIe D53	bb	SLT-IIe D53	bb
SLT-IIe D65	bb	SLT-IIe D65	bb
SLT-IIe D67	bb	SLT-IIe D67	bb
SLT-IIe E13	bb	SLT-IIe E11−E24	bb
SLT-IIe E15	bb	−	−
−	−	SLT-IIe E27−E31	bb
−	−	SLT-IIe E58	bb
−	−	SLT-IIe E59−E61	none

[a] All atoms were constrained to their initial positions, except those listed here. Initial Gb_4/SLT-IIe complexes were minimized first with a more highly constrained round of steepest descents minimization (initial steepest descents) with DISCOVER/AMBER, followed by a longer and less constrained round of conjugate gradients minimization (final conjugate gradients) with the same potential function (see the text). [b] Gb_4 is numbered progressively, with the Gal*N*Ac ring being residue 1. During steepest descents minimization, the βGal*N*Ac(1−3)αGal glycosidic linkage was unconstrained. Each monomer of the SLT-IIe binding pentamer consists of 68 residues. The trimers used in our experiments were numbered A2−A69, D2−D69, and E2−E69, according to SLT-I numbering (see Figure 3). [c] None, unconstrained; bb, Cα, N, C, and O atoms constrained.

included several residues previously implicated in receptor binding (discussed above). Our goal with the first round of minimization was to fix the Gb_3 moiety of the Gb_4s, as well as most of the protein, and allow the terminal Gal*N*Ac ring and selected parts of the protein to adjust to accommodate this docking. The less constrained second round of minimization should then allow further finer adjustment of slightly more of the protein, as well as the complete Gb_4 molecule. Figure 5 shows site 1 of each of the minimized complexes, and Table 5 lists the energies of the important favorable residue−residue interactions for each minimized complex.

Given our assumption [first discussed in general terms by DeGrandis et al. (*11*) and later in more specific terms by Tyrrell et al. (*2*)] that the difference in binding preference between SLT-IIe and GT3 is due to interactions between the Gal*N*Ac moiety of Gb_4 and GlnD65 and LysD67 (for site 1 of monomer E), the complex of SLT-IIe with Gb_4 conformer 2 seems unlikely. This is somewhat surprising, since our calculations describe this as the most stable Gb_4 conformation, at least of the three conformers we considered (Table 3). In this complex, Gal*N*Ac projects away from the protein surface (Figure 5), and most of this residue does not contact the protein. Only 242 Å^2 of nonpolar surface (all carbons) is buried in this complex, compared to 291 and 347 Å^2 for the conformer 1 and 3 complexes, respectively. The complex with Gb_4 conformer 2 also involves the fewest (6, vs 9 and 7 for conformers 1 and 3, respectively) intermolecular hydrogen bonds of the three complexes. Table 5 shows that Gal*N*Ac is not a major contributor of binding energy for this complex. Of particular relevance is the observation that contacts between Gal*N*Ac and both GlnD65 and LysD67 are not important for this complex (Figure 5, Table 5). Thus, we see that conformer 2 does not provide a satisfactory model of Gb_4 bound to (site 1 of) SLT-IIe.

The minimized complexes of Gb_4 conformers 1 and 3 with SLT-IIe inspire more confidence. The complex of conformer 1 is favored over those of 2 and 3 by approximately 10 kcal mol^{-1} (Table 5). In both of these models (complexes with conformers 1 and 3), the Gal*N*Ac moiety projects toward the protein, and makes extensive contact with it (Figure 5). For conformer 3, contact between Gal*N*Ac and GlnE65 makes a relatively small favorable contribution to the stability

Table 5: Important Residue−Residue Interactions for the Three Initial Complexes of Gb_4 Bound at SLT-IIe Site 1[a]

	Gb4 residue				
SLT-IIe residue[b]	GalNAc1	Gal2	Gal3	Glc4	total
Gb_4 conformer 1					
Ile:D53	−0.5				−0.6
GlnD65	−3.4				−3.6
LysD67	−1.6				−1.3
LysE13	−8.0				−8.1
AsnE15	−1.5	−2.0			−3.8
GluE16	−1.0	−1.2			
AspE17			−2.8	−0.8	−3.1
ThrE19		−0.6	−1.2		−1.8
ThrE21		−2.7			−3.2
GluE28		−3.4			−3.6
TyrE29		−0.5			−0.5
TrpE30		−2.4	−4.8		−7.4
GlyE60			−1.3		−1.8
total	−17.8	−13.5	−11.5	−1.0	−43.9
Gb_4 conformer 2					
LysE13	−2.7	−2.7			
AsnE15	−2.3	−2.9			
AspE17	−2.7	−0.7			−2.6
ThrE19	−1.1	−1.6			
ThrE21	−2.1	−2.6			
GluE28	−0.5	−4.9			−6.0
TyrE29	−1.1	−1.2			
TrpE30	−2.4	−4.9			−7.6
GlyE60	−1.3	−2.1			
total	−5.1	−15.3	−11.7	−1.1	−33.1
Gb_4 conformer 3					
GlnD65	−0.8	−1.0			
LysE13	−1.1	−0.8			
AsnE15	−1.0	−1.0			−2.1
AspE17	−1.9	−0.8			−2.2
ThrE19	−0.9	−1.2			
ThrE21	−2.2	−2.5			
GluE28	−4.6	−3.5			−8.6
TyrE29	−1.6	−1.7			
TrpE30	−1.7	−4.7			−6.7
GlyE60	−1.0	−1.4			−2.6
total	−8.4	−13.4	−11.3	−1.2	−34.2

[a] Initial complexes following the first cycle of steepest descents and conjugate gradients minimization using the DISCOVER implementation of the AMBER potential function (see the text). All residue−residue interactions with a total energy of −0.5 kcal mol^{-1} or less are listed, in units of kilocalories per mole. However, totals are the net total for that residue. [b] All of our simulations involved a trimer of subunits A, D, and E of the pentameric binding subunit, and we focused on the Gb_3/Gb_4 binding sites of monomer E. The three subunits of the trimer were numbered A2−A69, D2−D69, and E2−E69, according to SLT-I numbering (see Figure 3).

[2] The reader should note that the postminimization complexes are shown in Figure 5. However, minimization did not significantly affect the major differences between the three complexes. Projection of the Gal*N*Ac residue away from the protein in this complex is a result of the conformation of the βGal*N*Ac(1−3)αGal glycosidic linkage in this Gb_4 conformer.

[3] As for footnote 2, the postminimization complexes are shown in Figure 5, and the clashes with LysE13 have largely been eliminated. It is clear, however, that LysE13 is near to the Gal*N*Ac ring in the complexes of conformers 1 and 3, but not in that of conformer 2.

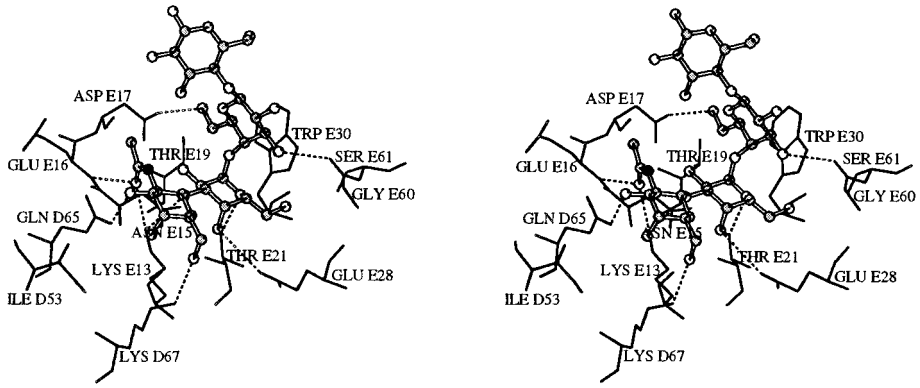

FIGURE 6: Final model of Gb_4 bound at site 1 of SLT-IIe. The possible intermolecular hydrogen bonds descibed in Table 7 are shown as dotted lines.

of the complex (Table 5), and LysD67 does not contact Gb_4 at all (Figure 5, Table 5). Compared to conformer 2, Gal*N*Ac makes a larger contribution to the stability of this complex (Table 5). The complex with conformer 1, however, is by far the most attractive of the three models. The contribution of Gal*N*Ac to the stability of this complex is greater than 3 times that for conformer 2 (Table 5), and more than twice that for conformer 3 (Table 5). Given the binding results observed with GT3 (discussed above), the energetic consequence of the contacts between Gb_4 and both GlnD65 and LysD67 (Table 5) also contributes to the attractiveness of this model.

All of our simulations were performed in vacuo, without any explicit consideration of solvent effects. A simple way of partially compensating for this deficiency is to apply a desolvation correction based on the area of nonpolar surface buried upon complex formation (ref *28* and references therein; *29*−*31*). If we ascribe an energetic reward of 20 cal $Å^{-2}$ of nonpolar surface (all carbons), a compromise between several similar reported values (ref *28* and references therein; *29*−*31*), the complexes of conformers 1−3, would be corrected by −5.8, −4.8, and −6.9 kcal mol^{-1}, respectively.

Following inspection of the minimized complexes of conformers 1 and 3 (Figure 5), we concluded that further optimization of the interactions would require manual intervention. We restricted ourselves to manipulation of the glycosidic linkage of βGal*N*Ac(1−3)αGal, the exocyclic *N*-acetyl group of the Gal*N*Ac ring, and the side chains of LysD67 and LysE13 of the protein. The dihedrals of the side chains were adjusted in steps of 120°. Figure 1 indicates that the βGal*N*Ac(1−3)αGal glycosidic linkage has considerable torsional flexibility in the regions occupied by both conformers 1 and 3 (Table 3).

For both of these minimized complexes, the biggest problem was the prohibitive interaction with LysE13[3] (Figure 5). Minor manual adjustment of ϕ and ψ of βGal*N*Ac(1−3)αGal and the LysE13 and LysD67 side chains appears to turn this problem into an advantage for the complex of conformer 1 (Figure 6, Table 6). Such manipulations proved fruitless for the complex involving Gb_4 conformer 3 (results not shown). Minor adjustments of these and other torsions did not yield a complex involving direct interactions between Gal*N*Ac and the side chains of GlnD65 and LysD67, and at this point, we focused on the complex involving conformer 1.

Table 6: Important Residue−Residue Interactions for the Final Model of the Complex of Gb_4 Bound at Site 1 of SLT-IIe[a]

SLT-IIe residue	Gb4 residue				
	Gal*N*Ac1	Gal2	Gal3	Glc4	total
IleD53	−0.6				−0.6
GlnD65	−3.4				−3.6
LysD67	−3.5				−3.2
LysE13	−8.0				−8.1
AsnE15	−1.6	−2.0			−3.8
GluE16	−1.0				−1.2
AspE17			−2.8	−0.8	−3.2
ThrE19		−0.6	−1.2		−1.8
ThrE21		−2.7			−3.2
GluE28		−3.3			−3.5
TyrE29		−0.5			−0.5
TrpE30		−2.4	−4.8		−7.4
GlyE60			−1.3		−1.8
total	−20.0	−13.4	−11.5	−1.0	−45.9

[a] The final modeled complex following initial minimization, manual adjustment, further minimization, manual adjustment, and a final round of minimization. All residue−residue interactions with a total energy of −0.5 kcal mol^{-1} or less are listed, in units of kilocalories per mole. However, totals are the net total for that residue (see the text and also Table 5).

Two rounds of both minor manual adjustment (see above) and 100 steps of conjugate gradients minimization, with constraints as in step 2 of our original two-step minimization scheme (Table 4), yielded a good model of the complex (Figure 6). Table 6 lists the intermolecular residue−residue interactions that make important contributions to the stability of the modeled complex. Our final model is slightly more stable than our initial complex with conformer 1, and this improvement is largely due to interactions involving the Gal*N*Ac moiety (Table 6). Hydrogen bonds between the Gal*N*Ac ring of Gb_4 and both GlnD65 and LysD67 stabilize the complex (Figure 6, Tables 6 and 7).

Plausibility of the Gb₄/SLT-IIe Model. We can determine the plausibility of our final model of the Gb_4/SLT-IIe complex in several ways. First, we can examine the final conformation of Gb_4 itself, and compare it to our calculated low-energy conformer(s). Second, we can examine the final protein conformation, to determine if construction of our model has resulted in the generation of any unreasonable structural features. Third, we can compare the Gb_3 moiety of our modeled Gb_4/SLT-IIe complex to our initial model of Gb_3 bound to site 3 of SLT-IIe. Fourth, and finally, we can examine the consistency of our model with a variety of experimental mutation and binding data.

Table 7: Possible Intermolecular Hydrogen Bonds of the Final Model of the Complex of Gb_4 Bound at Site 1 of SLT-IIe[a]

donor	acceptor	H···A distance[b] (Å)	D···H···A angle (deg)
Gal*N*Ac HO3	GlnD65 OE1	2.36 (3.18)	141.9
LysD67 HZ1	Gal*N*Ac O6	1.89 (2.87)	161.1
LysE13 HZ1	Gal*N*Ac O7	2.17 (2.80)	119.1
LysE13 HZ3	Gal*N*Ac O1	2.29 (3.11)	136.8
GluE16 HN	Gal*N*Ac O7	2.81 (3.73)	152.6
ThrE21 HG1	Gal2 O4	2.28 (3.23)	168.5
ThrE21 HG1	Gal2 O5	2.42 (2.93)	113.1
Gal2 HO4	GluE28 OE2	2.69 (3.67)	177.8
Gal3 HO6	AspE17 OD2	1.90 (2.78)	150.3
Gal3 HO3	GlyE60 O	2.51 (3.16)	123.6

[a] Note that, since the structure being analyzed is a model, we used a relatively permissive hydrogen bond filter of 90° for the D···H···A angle and 3.0 Å for the H···A distance. [b] The corresponding D···A distances are shown in parentheses.

The final glycoside dihedrals (ϕ/ψ) of our model of bound Gb_4 are 32.2/−46.5, −41.9/−8.2, and 53.7/0.4 for Gal*N*Acβ-(1−3)Gal, galabiose, and lactose, respectively. Similar values were reported for free Gb_4 (*9*, *21*) and for a model of the Gb_4/SLT-IIe complex (*9*). In this study, these values were determined following minimization of the hydrogen atoms in the absence of protein, to allow for a meaningful comparison with our earlier result (100 steps of conjugate gradients minimization with DISCOVER/AMBER, all heavy atoms fixed). The energy of this conformer was −24.5 kcal mol^{-1}, almost identical to that of the original conformer 1 (Table 3), despite the changes in ϕ and ψ, −21 and −38°, respectively, of the βGal*N*Ac(1−3)αGal linkage. This confirms our earlier observation (see above) that this disaccharide shows considerable torsional freedom in the immediate vicinity of this energy minimum (Figure 1). Of particular interest is the observation that our final Gb_4 conformation is much closer than our initial calculated Gb_4 conformer to the majority of low-energy conformers described in the NMR/computation study of Poppe et al. (*25*). The galabiose and lactose dihedrals of our final Gb_4 model (see above) remain essentially unaltered from their initial calculated values (Table 3). Overall, therefore, the final modeled conformation of Gb_4 bound to SLT-IIe is not much different from our initial calculated low-energy Gb_4 conformation. Indeed, our manual adjustments led to what is probably a more stable conformation for free Gb_4 in solution. It is intriguing to observe that the carbohydrate moieties of both Gb_3 and Gb_4 bind to these toxins in conformations very similar to those favored by the free carbohydrates in solution.

It is also important to consider whether our model of the Gb_4 carbohydrate bound to SLT-IIe is consistent with a Gb_4 glycolipid molecule that is a membrane component. The terminal glucose residue of our model is completely exposed to solvent on the membrane binding face of the pentamer. Therefore, the glucose residue is "pointing" in a functionally sensible direction, and is unconstrained by any contacts with protein. This is similar to the experimentally observed Gb_3 glycolipid analogues bound to SLT-I (*12*) and GT3 (H. Ling and R. J. Read, unpublished results) and to the modeling results of Nyholm et al. (*9*, *24*). Furthermore, the flexibility of the βGal(1−4)βGlc disaccharide is well-documented (*25*). This, in conjunction with the lack of glucose−protein contacts (noted above), indicates that our model does not exclude any of the glucose−ceramide conformers described by Strömberg et al. (*32*) accessible for the Gb_4 glycolipid in a membrane (they used the HSEA-derived low-energy conformer of Gb_4, a conformation similar to that of our model). Our model of the carbohydrate portion of Gb_4 bound to site 1 of SLT-IIe is therefore consistent with the membrane receptor function of the Gb_4 glycolipid.

We used the protein analysis program PROCHECK (*33*) to analyze the initial and final model of the SLT-IIe binding subunit trimer used throughout our simulations [results not shown, but see the preceding paper (*12*)]. With the exception of side chain conformations in the region of the modeled Gb_4 binding site, the final and initial SLT-IIe models are virtually identical. Several peptide bonds in this region showed slight increases (1−3°) in the deviation of ω from ideality, and a few others showed improvements of similar magnitudes. Side chain rms differences were all less than 1 Å (for this comparison, we ignored the two modified residues). Our manual manipulations of LysD67 and LysE13 (see above) were restricted to the two terminal dihedrals, which are not considered by PROCHECK during side chain conformation analysis. However, as noted above, these angles were adjusted in 120° steps, and the final conformations seem reasonable (Figure 6). All secondary structural features of the initial model were preserved in the final model. Therefore, construction of our model of Gb_4 bound to SLT-IIe wrought minimal changes in the protein conformation, and did not result in the generation of any unreasonable features in the conformation of the protein.

Our starting model of Gb_3 bound at site 1 of SLT-IIe was generated by superimposing the SLT-I structure onto our modeled SLT-IIe structure and then superimposing the galabiose moieties of our calculated Gb_4 conformers onto the corresponding atoms of the three Gb_3s bound to SLT-IIe (see above). For conformer 1 at site 1 of SLT-IIe, the rms difference for the Gb_3 heavy atoms of the two molecules was 2.0 Å. For the galabiose heavy atoms, this difference was 1.6 Å. For our final model of Gb_4 bound at this site, these differences were 1.8 and 1.1 Å, respectively. The modeling described above has led to a structure with several favorable contacts between Gal*N*Ac and the protein (see above), while at the same time yielding a conformation of the Gb_3 moiety which is closer to that observed for Gb_3 bound to SLT-I.

Table 6 presents a breakdown of the calculated energetic contributions of residues of the protein and of the trisaccharide to the stability of the complex. That 45% of the net favorable binding energy derives from interactions involving the Gal*N*Ac residue of Gb_4 (Table 6) is compelling when we consider that this is the moiety that interacts with the two mutated residues of GT3. The possibility that site 1 is strictly a Gb_4 binding site in SLT-IIe is consistent with, but not established by, this result. However, the recent experimental observation of Gb_3 binding at this site of GT3 (H. Ling and R. J. Read, unpublished results) would seem to contradict this conclusion.

The importance of Gln65 and Lys67 in the specific interaction with the Gal*N*Ac moiety of our modeled complex is unequivocally convincing (Tables 6 and 7, Figure 6). The reduced Gb_4 affinity of GT3 may be explained by the absence or disturbance of one or both of these interactions in the double mutant (discussed in detail below). Such interactions were first postulated by DeGrandis et al. (*11*), and later

discussed by Tyrrell et al. (*2*) in their original description of this mutant. Our model shows that such interactions are indeed possible for a Gb_4 molecule bound at this site of SLT-IIe and that this explanation of the altered binding activity of GT3 is reasonable. In this light, it is interesting to compare our model of the Gb_4/SLT-IIe complex with that proposed by Nyholm et al. (*9*), focusing particularly on the Gal*N*Ac moiety of Gb_4. Both models focus on roughly the same area of the SLT-IIe surface (site 1 in our numbering scheme); however, close inspection reveals significant differences between the two. The present model proposes four of five possible intermolecular hydrogen bonds involving *N*-acetylgalactosamine and four different SLT-IIe side chains (Figure 6, Table 7), whereas that of Nyholm et al. (*9*) includes only two such bonds, both involving Gln64 (Gln65 in our numbering scheme, Table 1 in ref *9*). Involvement of Lys67 in Gb_4 binding by SLT-IIe was a central focus of our study (cf. GT3 double mutant), and this residue is a key contributor to the stability of our final model of the complex (Tables 6 and 7). Somewhat surprisingly, this residue does not seem significant in the alternative model [cf. Table 1 and Figure 4 of Nyholm et al. (*9*)]. Solution studies suggest that the pendent *N*-acetyl group of free Gb_4 adopts a conformation similar to that of our final model (*21*); Nyholm et al. (Figure 4 of ref *9*) propose an apparently somewhat less favorable conformation for this group. The differences between the two models arise, at least in part, from the different models of the Gb_3/SLT-IIe complex used as starting points for the two modeling studies [discussed in the preceding paper (*12*)].

One less than satisfactory aspect of our model is that we do not see any clearly prohibitive interactions between GT3 and Gb_4 (not shown). Although GlnD65 and LysD67 make two, or possibly three (see below), hydrogen bonds with the Gal*N*Ac residue of Gb_4, their absence does not obviously preclude Gb_4 binding at this site. Neither substitution seems to lead to an unfavorable steric or electrostatic interaction (not shown). It may be that the interactions of these two residues with Gal*N*Ac contribute a critical amount of binding energy, and that this critical level is still exceeded when either one of the residues is substituted [either substitution alone does not seem to affect Gb_4 binding (*2*)]. The Q65E substitution would only lead to partial loss of the contribution of this residue, and the remaining hydrogen bond with HO3 of the Gal*N*Ac residue would probably be slightly more favorable with a charged Glu than with Gln. To complete this scenario, then, substitution of both residues would result in loss of a critical amount of binding energy so that Gb_4 binding decreases dramatically (*2*). Of course, it is also possible that binding of Gb_3 and/or Gb_4 at this site invokes some subtle conformational change(s) that would more clearly explain the altered Gb_3 and Gb_4 affinities of GT3. There seems to be no clear evidence, however, that strongly supports either of these speculations.

Further inspection of the model (Figure 6) suggests that the acetyl group of Gb_4 is a relatively minor contributor to complex stability. This is consistent with the slight decrease in affinity observed for de-acetylated Gb_4 (*11*). The importance of LysE13 to the intermolecular interaction (Tables 6 and 7, Figure 6) suggests that mutagenesis of this residue might provide further information regarding Gb_4 binding to SLT-IIe. Also, in single, double, and triple mutants, the I53K substitution has implicated Ile53 in receptor binding (*2*, *26*). However, this issue is clouded by the observation that this substitution (I53K) may have major effects on the conformation of the B-subunit, as well as on the association of the A- and B-subunits (*2*). Our model places Gal*N*Ac near IleD53, thus posing a plausible explanation of this result, and also suggesting that more conservative substitution of this residue would be informative.

CONCLUSION

It is clear that there are compelling reasons to accept the model of Gb_4 binding at site 1 of SLT-IIe that we have described. The conformation of the bound Gb_4 model is similar to our calculated minimum energy conformation and is very similar to the reported solution structure of this molecule (*21*). The structure of the final protein model is also reasonable; only relatively minor side chain adjustments of the modeled SLT-IIe binding site were required to optimize the interaction with Gb_4. The relative energetic importance of the contacts between bound Gb_4 and the two residues that were mutated in the GT3 double mutant presents a reasonable explanation of the $Gb_4 \rightarrow Gb_3$ switch in binding preference exhibited by the double mutant, and strongly supports our model. Several other residues known to be important to Gb_3 and/or Gb_4 binding are also at or near the modeled binding site, and we look forward to the experimental determination of the structure of the complex.

ACKNOWLEDGMENT

The authors thank Cliff Clark for supplying the GT3 protein and Ami Boodhoo for preparation of the crystals. M.D.C. thanks Trevor Hart for help with computer programs and Bart Hazes for help preparing the figures. J.L.B. thanks Rumanna Agha for expert technical assistance.

REFERENCES

1. Lingwood, C. A. (1993) in *Advances in lipid research*, (Bell, R. M., Hannun, Y. A., and Merrill, A. H., Jr., Eds.) pp 189−212, Academic Press, San Diego.
2. Tyrrell, G. J., Ramotar, K., Toye, B., Boyd, B., Lingwood, C. A., and Brunton, J. L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 524−528.
3. Boyd, B., Tyrrell, G. J., Maloney, M., Gyles, C., Brunton, J. L., and Lingwood, C. A. (1993) *J. Exp. Med. 177*, 1745−1753.
4. Armstrong, G. D., Fodor, E., and Vanmaele, R. (1991) *J. Infect. Dis. 164*, 1160−1167.
5. Armstrong, G. D., Rowe, P. C., Goodyer, P., Orrbine, E., Klassen, T. P., Wells, G., Mackenzie, A., Lior, H., Blanchard, C., Auclair, F., Thompson, B., Rafter, D. J., and McLaine, P. N. (1995) *J. Infect. Dis. 171*, 1042−1045.
6. Merritt, E. A., Sarfaty, S., van den Akker, F., L'hoir, C., Martial, J. A., and Hol, W. G. J. (1994) *Protein Sci. 3*, 166−175.
7. Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L., and Read, R. J. (1992) *Nature 355*, 748−750.
8. Sixma, T. K., Stein, P. E., Hol, W. G. J., and Read, R. J. (1993) *Biochemistry 32*, 191−198.
9. Nyholm, P.-G., Magnusson, G., Zheng, Z., Norel, R., Binnington-Boyd, B., and Lingwood, C. A. (1996) *Chem. Biol. 3*, 263−275.
10. Merritt, E. A., Sixma, T. K., Kalk, K. H., van Zanten, B. A. M., and Hol, W. G. J. (1994) *Mol. Microbiol. 13*, 745−753.
11. DeGrandis, S., Law, H., Brunton, J., Gyles, C., and Lingwood, C. A. (1989) *J. Biol. Chem. 264*, 12520−12525.

12. Ling, H., Boodhoo, A., Hazes, B., Armstrong, G. D., Brunton, J. L., and Read, R. J. (1998) *Biochemistry 37*, 1777−1788.
13. Hart, T. N., and Read, R. J. (1992) *Proteins: Struct., Funct., Genet. 13*, 206−222.
14. Cummings, M. D., Hart, T. N., and Read, R. J. (1995) *Protein Sci. 4*, 885−899.
15. Bass, M. B., Hopkins, D. F., Jaquysh, W. A. N., and Ornstein, R. L. (1992) *Proteins: Struct., Funct., Genet. 12*, 266−277.
16. Ponder, J. W., and Richards, F. M. (1987.) *J. Mol. Biol. 193*, 775−791.
17. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
18. Stuike-Prill, R., and Meyer, B. (1990) *Eur. J. Biochem. 194*, 903−918.
19. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765−784.
20. Homans, S. W. (1990) *Biochemistry 29*, 9110−9118.
21. Poppe, L., von der Lieth, C.-W., and Dabrowski, J. (1990) *J. Am. Chem. Soc. 112*, 7762−7771.
22. French, A. D., and Brady, J. W. (1990) in *Computer modeling of carbohydrate molecules* (French, A. D., and Brady, J. W., Eds.) pp 1−19, American Chemical Society, Washington, DC.
23. Tvaroška, I., and Pérez, S. (1986) *Carbohydr. Res. 149*, 389−410.
24. Nyholm, P.-G., Brunton, J. L., and Lingwood, C. A. (1995) *Int. J. Biol. Macromol. 17*, 199−204.
25. Poppe, L., Dabrowski, J., von der Lieth, C.-W., Koike, K., and Ogawa, T. (1990) *Eur. J. Biochem. 189*, 313−325.
26. Jackson, M. P., Wadolkowski, E. A., Weinstein, D. L., Holmes, R. K., and O'Brien, A. D. (1990) *J. Bacteriol. 172*, 653−658.
27. Clark, C., Bast, D., Sharp, A., St. Hilaire, P., Agha, R., Stein, P. E., Toone, E. J., Read, R. J., and Brunton, J. L. (1996) *Mol. Microbiol. 19*, 891−899.
28. Cummings, M. D., Hart, T. N., and Read, R. J. (1995) *Protein Sci. 4*, 2087−2099.
29. Eisenberg, D., Wesson, M., and Yamashita, M. (1989) *Chem. Scr. 29A*, 217−221.
30. Eisenberg, D., and McLachlan, A. D. (1986) *Nature 319*, 199−203.
31. Chothia, C. (1974) *Nature 248*, 338−339.
32. Strömberg, N., Nyholm, P.-G., Pascher, I., and Normark, S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9340−9344.
33. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

BI971807F