A mutant Shiga-like toxin IIe bound to its receptor Gb_3: structure of a group II Shiga-like toxin with altered binding specificity

Hong Ling[1], Navraj S Pannu[2], Amechand Boodhoo[3], Glen D Armstrong[3], Clifford G Clark[4], James L Brunton[5,6] and Randy J Read[1,2,3]*

Background: Shiga-like toxins (SLTs) are produced by the pathogenic strains of *Escherichia coli* that cause hemorrhagic colitis and hemolytic uremic syndrome. These diseases in humans are generally associated with group II family members (SLT-II and SLT-IIc), whereas SLT-IIe (pig edema toxin) is central to edema disease of swine. The pentameric B-subunit component of the majority of family members binds to the cell-surface glycolipid globotriaosyl ceramide (Gb_3), but globotetraosyl ceramide (Gb_4) is the preferred receptor for SLT-IIe. A double-mutant of the SLT-IIe B subunit that reverses two sequence differences from SLT-II (GT3; Gln65→Glu, Lys67→Gln, SLT-I numbering) has been shown to bind more strongly to Gb_3 than to Gb_4.

Results: To understand the molecular basis of receptor binding and specificity, we have determined the structure of the GT3 mutant B pentamer, both in complex with a Gb_3 analogue (2.0 Å resolution; $R = 0.155$, $R_{free} = 0.194$) and in its native form (2.35 Å resolution; $R = 0.187$, $R_{free} = 0.232$).

Conclusions: These are the first structures of a member of the medically important group II Shiga-like toxins to be reported. The structures confirm the previous observation of multiple binding sites on each SLT monomer, although binding site 3 is not occupied in the GT3 structure. Analysis of the binding properties of mutants suggests that site 3 is a secondary Gb_4-binding site. The basis for understanding the tissue specificities and pathogenic mechanisms of members of the SLT family.

Addresses: [1]Department of Biochemistry, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada, [2]Department of Haematology, University of Cambridge, Wellcome Trust Centre for the Study of Molecular Mechanisms in Disease, Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Hills Road, Cambridge, CB2 2XY, UK, [3]Department of Medical Microbiology & Immunology, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada, [4]National Laboratory for Enteric Pathogens, Bureau of Microbiology, Laboratory Centre for Disease Control, Ottawa, K1A 0L2, Canada, [5]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, M5G 1X5, Canada and [6]Departments of Medicine and Laboratory Medicine, University of Toronto, Toronto, Ontario, M5G 2C4, Canada.

*Corresponding author.
E-mail: rjr27@cam.ac.uk

Key words: bacterial toxins, glycolipid, mutagenesis, protein–carbohydrate recognition, receptor binding.

Received: **19 August 1999**

Accepted: **22 December 1999**

Published: **22 February 2000**

Structure 2000, **8**:253–264

0969-2126/00/$ – see front matter

Introduction

In 1983 it was recognized that Shiga-like toxin-producing *Escherichia coli* are associated with hemorrhagic colitis (HC) and the hemolytic uremic syndrome (HUS) [1–3]. Since then, an increasing number of outbreaks owing to the consumption of contaminated food have been reported. In North America, where this serious and sometimes fatal infection is commonly called 'hamburger disease', *E. coli* O157:H7 is the most common pathogen. In the United States alone, approximately 20,000 people become ill and 250 people die from infection by the pathogenic strains each year [4–6]. Shiga-like toxins (SLTs or verotoxins) are the major virulence factors of the pathogenic *E. coli* strains that cause disease in humans and animals (Table 1). Antibiotics have not proved useful and might even increase the risk of complications of infection, as killing the bacteria may accelerate the release of toxins [7,8]. Other therapies aimed at toxin neutralization are therefore needed.

SLTs are AB_5 toxins composed of one enzymatic (A) subunit and five copies of a cell-binding subunit (the B pentamer). The A subunit (32 kDa) of the holotoxin is the toxic component that acts within the target host cell. The B pentamer (7.5 kDa × 5) is responsible for toxin attachment to globoseries glycolipids on the cell surface [9,10]. This attachment is required for internalization of the toxin and retrograde routing through the Golgi apparatus to the endoplasmic reticulum (ER). It is thought that the A subunit enters the cytosol at this level (reviewed in [11]), enzymatically inactivates the ribosomes [12] and triggers cell death. The glycolipid Gb_3 (globotriaosyl ceramide; Figure 1) functions as a receptor for SLTs and is present on the surface of target cells, such as epithelial cells in the intestine and endothelial cells in the kidney [13–16]. Cells without Gb_3 on their surface are resistant to the toxins [9,17,18]. Gb_3 expression correlates with the tissue specificity of toxin damage and, in turn, the disease symptoms in patients [19].

Table 1

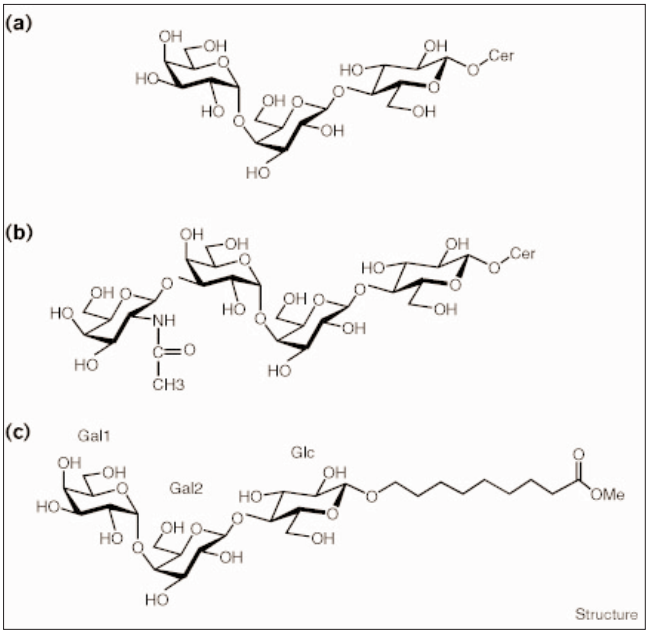
The Shiga toxin family*.

Toxin	Alternative names	Binding specificity	Disease caused	Crystal structures	No. of sugars bound
ST	Shiga toxin	Gb_3	Dysentery	Holotoxin† (– sugar)	–‡
SLT-I	Shiga-like toxin I, verotoxin 1	Gb_3	Human HC and HUS	B Pentamer§ (+ sugar) B Pentamer# (– sugar)	Three per monomer –
SLT-II	Shiga-like toxin II, verotoxin 2	Gb_3	Human HC and HUS	None	–
SLT-IIc	Shiga-like toxin IIc, verotoxin 2c	Gb_3	Human HC and HUS	None	–
SLT-IIe	Shiga-like toxin IIe, pig edema toxin, verotoxin 2v	Gb_4 Gb_3¶	Edema disease of swine	None	–
GT3	Q65E/K67Q mutant of SLT-IIe	Gb_3	SLT-I-like disease in pigs	B Pentamer¥ (+ sugar)	Two per monomer

*Reviewed in [22]. †[41]. ‡The Shiga toxin B subunit is identical to the SLT-I B subunit and should have the same three distinct receptor-binding sites as SLT-I B. §[37]. #[40]. ¶Binding affinity for Gb_3 is significantly lower than for Gb_4. ¥This study.

The SLTs are similar to Shiga toxin from *Shigella dysenteriae* in both structure and function [20,21]. Members of the Shiga toxin family are divided into two groups, originally distinguished by their immunological properties (reviewed in [22]). Group I includes Shiga toxin (ST) and SLT-I, which differ by only one amino acid in the A subunit. Group II includes SLT-II, SLT-IIc and SLT-IIe, which are very similar to each other in sequence but significantly different from SLT-I (Figure 2). All family members except SLT-IIe are associated with HC and HUS in humans [1,23,24]. SLT-IIe, or pig edema toxin, is associated with edema disease of swine, which is characterized by microvascular damage causing edema in the subcutaneous tissue of the eye lids, the central nervous system, and the spiral colon [25]. This disease is an important cause of death associated with weaning in swine herds in parts of Europe.

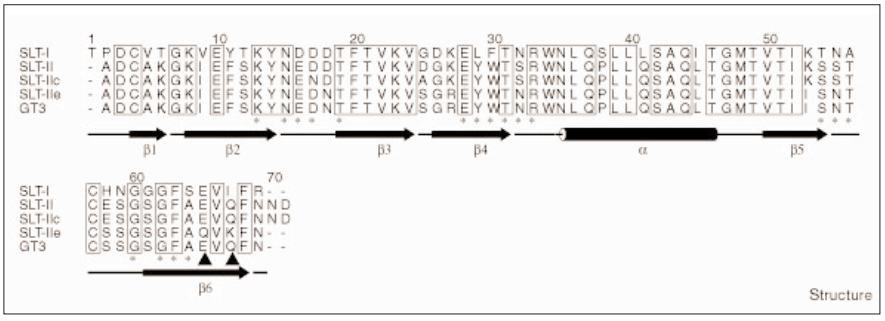
Figure 1



Cell surface receptors and the Gb_3 analogue. **(a)** Gb_3 **(b)** Gb_4 and **(c)** the Gb_3 analogue 8-methoxycarbonyloctyl trisaccharide (Pk-MCO). The Pk trisaccharide terminus is the same as the carbohydrate portion of the glycolipid Gb_3 (Gal(α1–4)Gal(β1–4)Glc(β1–8) ceramide) that is recognized by the B pentamer of SLTs. In this paper, we refer to this trisaccharide as Pk, and to the three sugar residues αGal, βGal and βGlc as Gal1, Gal2 and Glc, respectively.

E. coli strains producing SLT-II have been associated with more severe disease in humans than those strains which produce SLT-I [18,26–28]. As there is no detectable difference in enzymatic activity of the A subunits, this observation has been ascribed to differences in receptor binding. Indeed, A–B hybrid toxins of SLT-I and SLT-II demonstrate that cytotoxicity is modulated by the B subunits [29]. The binding affinity of SLT-I for Gb_3-bearing target cells ($K_d = 4.6 \times 10^{-8}$ M) is higher than that of group II toxins ($K_d = 3.7 \times 10^{-7}$ M) [29,26]. It has been proposed that SLT-II's lower receptor-binding affinity enables it to stay longer in the circulation and to reach the kidneys more easily than SLT-I [27,26]. Furthermore, SLT-II has been shown to be much more cytotoxic than SLT-I towards kidney microvascular endothelial cells *in vivo* [26,28]. The B subunits are an ideal target for rational drug design, because binding to the cell surface is a crucial step in cytotoxicity. A new drug (Synsorb-Pk), derived from the glycolipid analogue Pk-MCO (Figure 1), is being tested against HUS in children infected with O157:H7 *E. coli* [30,31]. It is proposed that Synsorb-Pk should adsorb toxins in the intestinal lumen, preventing them from entering the tissues and circulatory system. Synsorb-Pk is presently in phase III clinical trials. Phase II trials showed a trend towards protection against HUS; however, this did not achieve statistical significance because the study did not enrol sufficient patients with proven *E. coli* O157:H7 infection [32]. We hope to find

Figure 2



Alignment of the amino acid sequences of B subunits of SLTs. GT3 is a mutated form of SLT-IIe. Invariant residues are shown boxed. Asterisks denote residues involved in sugar binding and filled triangles denote the two residues of GT3 that were mutated to be the same as those in SLT-II. Secondary structure elements are indicated by broad arrows (β strands) and a cylinder (α helix). The B subunit of Shiga toxin from *Shigella dysenteriae* is identical to that of SLT-I. (This figure was generated with the program ALSCRIPT [71].)

higher affinity inhibitors through close examination of atomic-level interactions between the B pentamers and their receptors.

All of the SLTs associated with diseases in humans bind to Gb_3. Although the SLT-IIe B subunit shares 62% amino acid sequence identity with SLT-I B and 87% identity with SLT-II B, it binds the glycolipid Gb_4 (globotetraosyl ceramide; Figure 1) in preference to Gb_3 [33]. Cells bearing only Gb_3 are still susceptible to the action of SLT-IIe, but not as susceptible as those bearing Gb_4 [34]. To explore which amino acids determine binding specificity, Tyrrell *et al.* [35] constructed a series of SLT-IIe mutants by substituting residues from the SLT-II B sequence. One double-mutant of the SLT-IIe B subunit (Gln65→Glu, Lys67→Gln, designated GT3; residues numbered according to the alignment with SLT-I B in Figure 2) changes its binding preference from Gb_4 to Gb_3 [35]. The change in preference occurs through the combination of a dramatic reduction in binding to Gb_4 and an increase in binding to Gb_3. Moreover, GT3 causes an SLT-I-like disease in pigs [36], thus providing strong evidence that the binding specificity dominates the cytotoxic tissue specificity *in vivo*.

Previously, we studied the binding of Gb3 to a group I SLT by determining the structure of SLT-I B in complex with a Gb3 analogue [37]. Three distinct Gb3 receptor binding sites are present in each SLT-I B monomer. The study of group II B subunits has been hindered by the difficulty of expressing large amounts in a biologically active form [38]. In this study, we report the first group II SLT structures: the GT3 mutant B pentamer in complex with the Gb3 analogue Pk-MCO (Figure 1) at 2.0 Å resolution and the native GT3 structure at 2.35 Å resolution. The GT3 B subunit differs from the SLT-II B subunit in only seven of 68 common positions and is thus an excellent model for other group II toxins. Having the structure of this SLT-II group member allows comparisons with the SLT-I structure and sheds light on the relative receptor-binding activities of the group I and II B subunits.

Results and discussion

Quality of the structures

The molecular replacement solution of the GT3–Pk-MCO complex was determined first and revealed one GT3 B pentamer per asymmetric unit. Crystallographic refinement of the complex with all data from 31.0–2.0 Å produced a residual R factor of 0.155 (R_{free} = 0.194; see Table 2). The final refined model of the structure includes all 340 amino acid residues of the B pentamer, seven carbohydrate moieties of Pk-MCO molecules and 160 water molecules. Excellent electron density is observed for the entire GT3 B pentamer, whereas the quality of density for carbohydrates varies in different binding sites (Table 3; Figure 3), as discussed below. An analysis of protein stereochemistry with the program PROCHECK [39] indicates that, for all tested properties, the model geometry is equal to or better than that expected for a 2.0 Å structure (data not shown). A Ramachandran plot indicates that over 95% of non-glycine residues are in the most favoured regions and the remainder are in additional allowed regions. The root mean square deviation (rmsd) among the five B monomers is low (0.15 Å for all Cα atoms and 0.23 Å for all non-hydrogen protein atoms), although a very loose noncrystallographic symmetry (NCS) restraint (w_{ncs} = 10 kcal mol^{-1} $Å^{-2}$) was used in the final refinement.

The molecular replacement solution of native GT3 revealed four GT3 B pentamers per asymmetric unit. All data from 21.0–2.35 Å were used in the crystallographic refinement, producing a working R factor of 0.187 and an R_{free} of 0.232. The model contains coordinates for all 1360 amino acids of the four B pentamers, and 359 water molecules. Over 94% of the non-glycine residues are in the most favored region of the Ramachandran plot as defined by PROCHECK [39]. As with the structure of the complex, the rmsds among the 20 B monomers are low (0.17 Å for all Cα and 0.59 Å for all non-hydrogen protein atoms). Most of the deviations arise from the differing conformations of Trp34 in the monomers, discussed below.

Table 2

Crystallographic data.

Structure	GT3–Pk-MCO complex	GT3 native
Space group	$P2_12_12_1$	$P2_1$
Unit cell		
a, b, c (Å)	62.3, 78.9, 78.8	113.5, 54.5, 116.9
β (°)		109.1
No. unique reflections	25,704 (31.0–2.0 Å)	34,187 (21.0–2.35 Å)
No. measurements	308,496	62,524
R_{merge}*		
Overall (%)	10.2 (31.0–2.00 Å)	8.4 (21.0–2.35 Å)
Highest resolution shell (%)	22.2 (2.15–2.00 Å)	33.0 (2.49–2.35 Å)
Completeness of data		
Overall (%)	95.3 (∞–2.00 Å)	60.4 (∞–2.35 Å)
Highest resolution shell (%)	87.2 (2.15–2.00 Å)	12.3 (2.39–2.35 Å)
Model		
Protein (non-hydrogen atoms)	2665	10,660
Carbohydrate (non-hydrogen atoms)	228	N/A
Solvent (water molecules)	160	359
Average B factor (Å^2)		
Protein	19.5	28.9
Carbohydrate	41.7	N/A
Solvent	33.8	31.4
R factor† (working data)	0.155 (31.0–2.00 Å)	0.187 (21.0–2.35 Å)
R_{free}‡	0.194 (2541 reflections)	0.232 (1055 reflections)
Rmsd bond lengths (Å)	0.015	0.012
Rmsd bond angles (°)	1.6	1.7

*$R_{merge} = \Sigma |I_i - \langle I_i \rangle| / \Sigma |I|$, where I is the intensity of the reflections. †R factor $= \Sigma||F_o|-|F_c|| / \Sigma|F_o|$, where F_o is the observed structure-factor amplitude from diffraction data and F_c is the calculated structure-factor amplitude from the molecular model. ‡R_{free} is calculated as for the R factor, using only an unrefined subset of the diffraction data [61].

Structure of GT3

The GT3 B subunit has a typical oligomer-binding (OB) fold that consists of a six-stranded antiparallel β barrel capped by an α helix [40]. Each GT3 B subunit monomer contains 68 residues (numbered 2–69, Figure 2), one residue fewer than the SLT-I B monomer [35]. A disulphide bond formed between Cys4 and Cys57 of GT3 is also present in SLT-I B. The two mutant residues of GT3 are located in strand β6 near the C terminus, at the interface between neighbouring monomers (Figure 4a).

Table 3

Quality of the electron density and extent of the carbohydrate model for each binding site in the GT3–Pk-MCO complex.

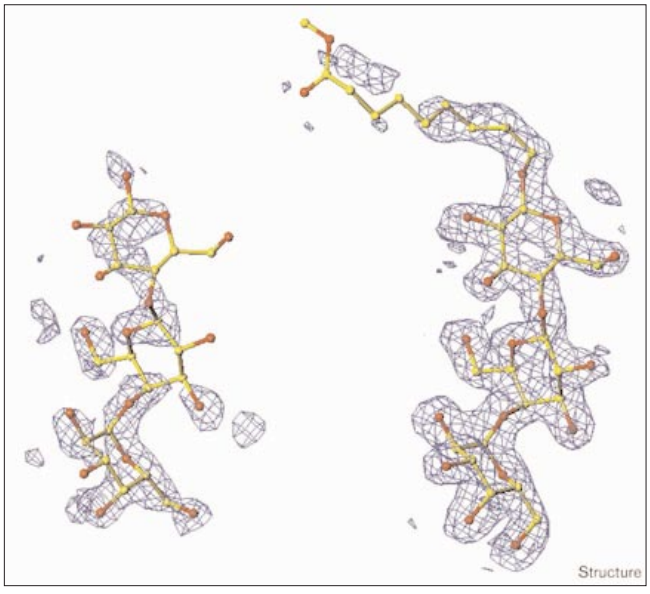
Monomer	Binding site	Density quality	Sugar molecule modelled
B1	Site 1	Poor, discontinuous	Gal1–Gal2–Glc
	Site 2	Excellent	Gal1–Gal2–Glc–tail*
B2	Site 1	Poor, discontinuous	Gal1–Gal2
	Site 2	Excellent	Gal1–Gal2–Glc
B3	Site 1	No density	None
	Site 2	Excellent	Gal1–Gal2–Glc–tail*
B4	Site 1	Very little density	None
	Site 2	No density	None
B5	Site 1	Poor, discontinuous	Gal1–Gal2
	Site 2	Excellent	Gal1–Gal2–Glc–tail*

*Only part of the MCO tail (3–4 carbon atom chain) was built and refined.

Carbohydrate binding has little effect on the conformation of the GT3 B subunit monomer: the average rmsd for comparisons of Cα atoms in monomers of complexed and uncomplexed GT3 is approximately 0.2 Å. Furthermore, the average rmsd for similar comparisons of GT3 monomers (complexed or uncomplexed) with SLT-I monomers (complexed or uncomplexed) is approximately 0.5 Å. The small difference between the GT3 and SLT-I monomers is not surprising, as the molecules share 63% sequence identity.

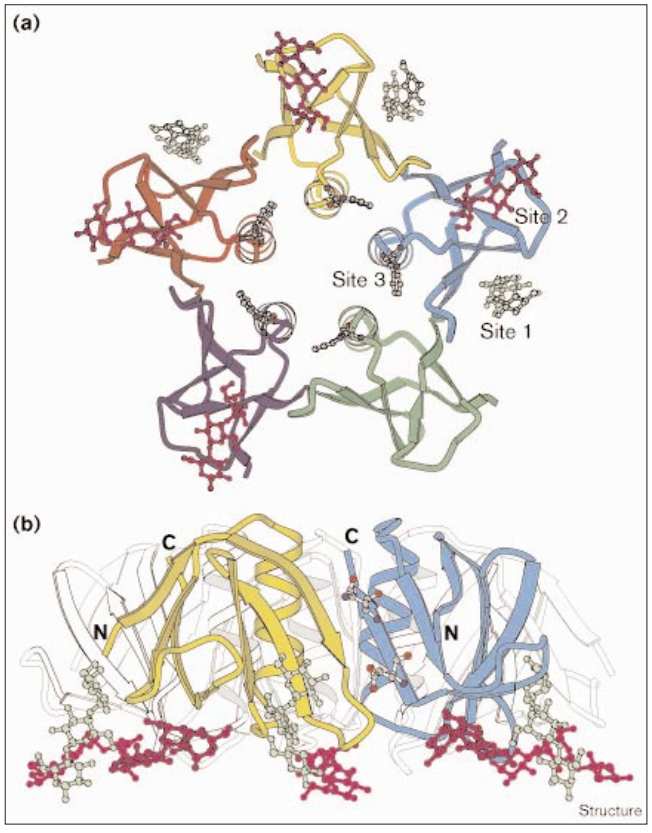
Larger differences emerge when quaternary structures are compared. In the native SLT-I B pentamer [40] there is a screw component to the fivefold symmetry with a translation of 1.3 Å parallel to the fivefold axis — the structure, therefore, resembles a lock–washer. In contrast, no significant screw component is seen in the GT3–Pk-MCO complex, the SLT-I–Pk-MCO complex [37] or the ST holotoxin [41], and the rotation angles between neighbouring monomers are characteristic of good fivefold symmetry. The native GT3 pentamers, in general, show greater deviations from perfect fivefold symmetry, although they do not have a lock–washer structure. The analysis of deviations from perfect fivefold symmetry, summarized in Table 4,

Figure 3



Typical electron-density maps for Pk-MCO. Pk-MCO is shown in ball-and-stick representation with carbon atoms in yellow and oxygen atoms in red. Electron density is shown for Pk-MCO bound to **(a)** site 1 (monomer B1) and **(b)** site 2 (monomer B1). Electron-density maps are contoured at 1.0 × root mean square electron density, prepared with the program O [59].

Figure 4



Two orthogonal views of the GT3 B pentamer bound to Pk-MCO. Each monomer is shown in a different colour. **(a)** View along the fivefold axis. The sugar-receptor binding surface is towards the viewer, corresponding to the bottom surface in (b). There is no sugar binding at site 3 in this structure. Trp34 sidechains in site 3 are shown in ball-and-stick representation. **(b)** Side view. The N and C termini of the two monomers in colour are labelled. The top face of the B pentamer is the interface associated with the A subunit of the holotoxin; the bottom face of the B pentamer is the receptor-binding surface. Sugars bound at site 1 of each monomer are shown in green and sugars bound at site 2 are in purple. The mutated residues Glu65 and Gln67 are shown in ball-and-stick representation for one subunit. (These figures were generated with the program MOLSCRIPT [72].)

provides more evidence for the hypothesis that interactions with a ligand or an A subunit enforce a closer to perfect fivefold symmetry on the B pentamer. It has been speculated that reduced stability of the pentamer interactions in the absence of the A subunit could be relevant to toxin assembly (WGJ Hol, personal communication).

Most amino acid sidechains in the structures exhibit a single, well-defined conformation. The sidechains of residue Glu10 of each monomer show two alternative conformations in the complex structure at 2.0 Å resolution. Both the SigmaA-weighted $2F_o–F_c$ map [42,43] and fivefold-averaged electron density accommodate the two alternative sidechain models. Although there are indications of similar static disorder in the native structure, the resolution was not considered sufficient to model two conformations.

An interesting example of multiple conformations is seen for the Trp34 sidechain. Collectively, the crystal structures of SLT B subunits indicate that this sidechain is intrinsically flexible. When exposed to solvent, as in the ST holotoxin structure [41] or in four of the five monomers of the GT3–Pk-MCO complex, this tryptophan sidechain lacks clear electron density. Disordered Trp34 sidechains are also observed in other SLT structures where the indole rings lack specific interactions (HL and RJR, unpublished observations). Trp34 is only well-ordered when involved in interactions that select one of its possible conformations; and a variety of conformers is seen. For example, in the native SLT-I B crystal structure, the screw translation along the fivefold axis allows four of the indole rings to interact with one another and to maintain a common well-defined conformation. In the SLT-I–Pk-MCO complex the Trp34 indole ring stacks against the β-galactose ring of Gb_3 in site 3, and there is clear density for what would otherwise be an unfavourable sidechain conformation [37]. In the native GT3 structure, many of the Trp34 sidechains from neighbouring pentamers stack against each other. In addition, because of the deviations from perfect fivefold symmetry, a number of other Trp34 sidechains stack in the manner seen in the native SLT-I B structure.

Table 4

Deviations of GT3 B subunit pentamers from perfect fivefold noncrystallographic symmetry.

Pentamer	Rotation angles* (°)		Screw translations* (Å)	
	range†	rmsd‡	range†	rmsd‡
GT3–Pk-MCO	71.23–72.77	0.70	−0.73–0.92	0.60
GT3 native 1	69.68–74.70	2.08	−0.63–1.27	0.66
GT3 native 2	69.60–74.84	2.07	−0.24–0.25	0.17
GT3 native 3	71.37–72.70	0.58	−1.05–1.90	1.07
GT3 native 4	70.70–72.93	0.77	−0.84–1.69	0.91

*Rotation angles and screw translations refer to the noncrystallographic symmetry (NCS) operation required to superimpose one monomer on its neighbour in a pentamer. †The range of an angle or translation refers to the extreme values seen for that pentamer. ‡The rmsd statistics refer to deviations from the values expected for a pentamer with perfect fivefold symmetry: 72° for angles and zero for translations.

Pk-MCO trisaccharide binding

The OB-fold has been found in several proteins that bind oligosaccharides or oligonucleotides (reviewed in [44]). All these proteins have a single binding site in a structurally conserved location. In our previous study of the SLT-I–Pk-MCO complex, however, three distinct binding sites were found in each SLT-I B subunit (i.e., 15 binding sites per B pentamer [37]). Site 2 (Figure 4b) is topologically equivalent to the binding sites found in other OB-fold proteins, whereas site 1 corresponds to a previously predicted binding site [40]. In this study, two binding sites have been found in the GT3–Pk-MCO complex: sites 1 and 2 (Table 3; Figures 4 and 5). The conformations of the bound trisaccharides are similar to those in the SLT-I–Pk-MCO complex. The presence of all trisaccharides on the flat surface opposite to the A subunit confirms this region as the membrane-binding surface. As discussed previously for the SLT-I B structure [37], the different orientations of the Gb_3 trisaccharide in the distinct binding sites of the B subunits would allow the sugar–lipid linkers to adopt favourable conformations when the toxin binds to the host cell membrane.

Electron density is visible for sugar residues at the majority of binding sites. The quality of the density varies between sites 1 and 2 as well as between equivalent sites related by NCS (Table 3; Figure 3, in Table 3 and throughout the rest of the paper we will refer to the B subunits from the GT3–Pk-MCO complex as B1–B5). This variation in quality suggests that the isolated Pk-MCO binds more tightly to site 2 than to site 1. A similar variation among sites was observed in the structure of the SLT-I–Pk-MCO complex. Furthermore, a nuclear magnetic resonance (NMR) study of trisaccharide binding to the SLT-I B subunit showed that, at a 1:1 ratio of carbohydrate to protein, binding was seen only at site 2 [45]. This indicates that site 2 has the highest affinity for the isolated carbohydrate. It is important, however, to note that presentation of the trisaccharide by a glycolipid embedded in a membrane could well change the relative affinities. For the binding sites related by NCS, the differences appear to be mainly due to packing environments that are not crystallographically equivalent. With the exception of site 2 of B4, crystal-packing contacts do not block sugar access to any binding sites.

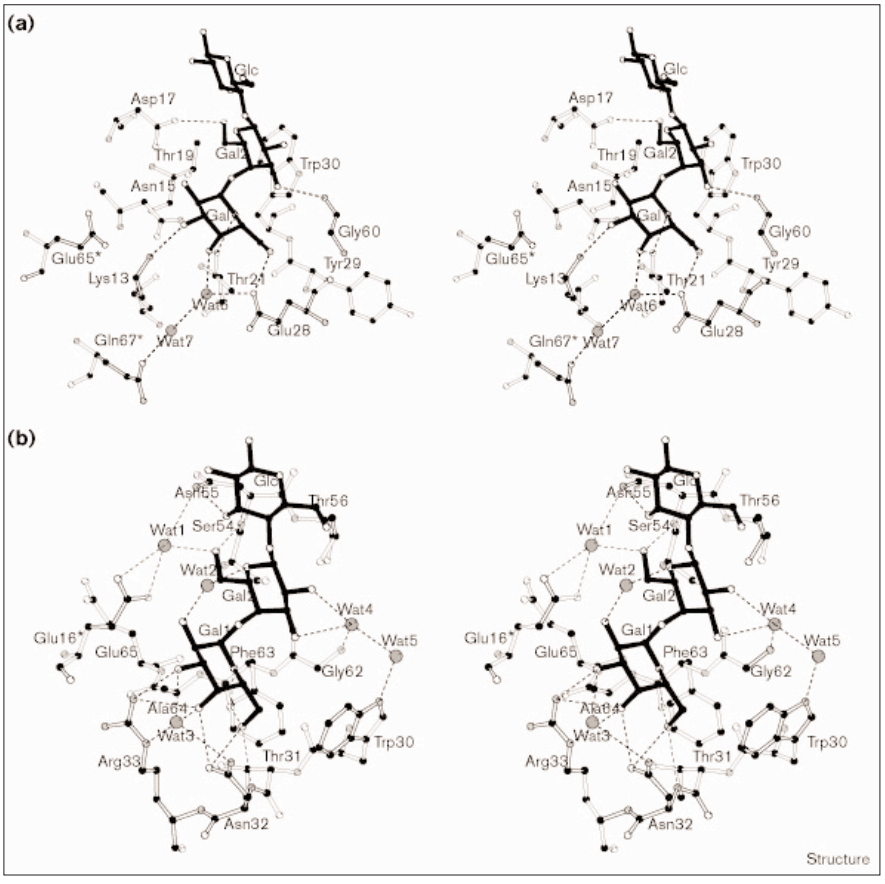
There are no direct interactions between the trisaccharides in the different binding sites of the pentamer. The binding sites are constructed from residues of two adjacent monomers and thus require the pentameric assembly of the B pentamer (Figure 5). The trisaccharides sit in shallow clefts on the relatively flat sugar-binding surface. Site 1 is located in a cleft near the interface with an adjacent B monomer, whereas site 2 is in an elongated shallow cleft along the flat surface (Figure 4a). The residues involved in sugar binding are either conserved in the Shiga toxin family or are conservatively substituted (Figure 2).

Binding at site 1

Electron density for the Pk-MCO trisaccharide at site 1 is visible, with varying quality, in four of the five B monomers of the GT3 pentamer (Table 3). In some cases, electron density reveals only the two terminal galactose residues, which make the primary interactions with the protein. Clear density for a trisaccharide is seen in monomer B1, and the interactions described below are based primarily on this binding site.

The interactions in site 1 (see Figure 5a) are very similar to those already described for the closely related binding site 1 in SLT-I [37]. Hydrophobic interactions are dominated by the aromatic ring of Trp30, which stacks against the B-face of Gal2 as also observed for Phe30 in the structure of SLT-I–Pk-MCO [37]. One additional interaction is observed: a hydrogen bond from Lys13 to O3 of Gal1 (Figure 5a). The substantial increase in resolution, however, allows us to observe the role of solvent, as in the water-mediated hydrogen bond from Glu28 to Gal1 (Figure 5a). With few exceptions, the site 1 contact residues are conserved in the SLTs. Trp30 is a conservative substitution for Phe30 in SLT-I, and only the mainchain of Tyr29 (leucine in SLT-I) contacts the sugar (Figure 5a). In SLT-IIc, however, the charged Asp17 residue (Figure 5a) is replaced by a neutral asparagine (Figure 2), resulting in a weaker hydrogen bond. This would correlate with the reduced cytoxicity of this toxin variant and mutants of SLT-II [46]. Site-directed mutagenesis has shown that the conserved residue Gly60 is essential for cytotoxicity of SLT-I and SLT-II [47]. Gly60 is a binding residue in site 1 of both SLT-I B and GT3. It adopts a mainchain conformation ($\phi \approx 90°$, $\psi \approx 0°$) that is inaccessible to non-glycine residues; substitution of this

Figure 5



Stereo diagrams of the Pk-MCO-binding sites in a GT3 B subunit. Protein segments are drawn with open bonds and the oligosaccharide with filled bonds. Dashed lines indicate intermolecular hydrogen bonds; asterisks indicate amino acid residues from an adjacent B subunit of the pentamer. Water molecules are depicted as large filled circles. **(a)** The Pk-MCO trisaccharide at site 1. **(b)** The Pk-MCO trisaccharide at site 2. (The figure was generated using the program MOLSCRIPT [72].)

residue would alter the conformation of the β5–β6 loop, which is involved not only in site 1 but also in site 2.

Binding at site 2

Four B monomers have clear electron density for Pk-MCO in site 2 (Table 3), defining the entire Pk-MCO trisaccharide and part of its hydrophobic tail (Figure 3b). Binding at site 2 of B4 is blocked by crystal packing. The trisaccharide is oriented approximately parallel to the protein surface (Figure 4), allowing all three sugar residues to interact directly with the protein. This extensive interaction probably explains the strong density observed at site 2, which is similar to that observed for the SLT-I–Pk-MCO complex. Hydrogen bonds with Gal1 dominate the binding interactions (Figure 5b). Five water molecules mediate hydrogen bonding between the protein and sugar molecule at site 2. The majority of the residues in site 2 are conserved and play a similar role to those in the SLT-I–Pk-MCO complex. One exception is the substitution of Glu16 for aspartic acid in SLT-I B, where the water-mediated hydrogen bond seen for Glu16 (Figure 5b) is replaced by a direct interaction for Asp16. This substitution might be responsible in part for the reduced Gb_3 affinity of the group II toxins.

Site-directed mutagenesis has shown that Arg33 plays an important role in the cytotoxicity of SLT-I and SLT-II [47]. The results can be explained by the extensive involvement of its sidechain in hydrogen bonding to the terminal galactose of Gb_3 trisaccharides in site 2 of both SLT-I [37] and GT3. By analogy, Arg33 should play a similar role in SLT-II. Mutations of Arg33 could influence binding activity indirectly, however, as its sidechain is also extensively involved in interactions that stabilize the tertiary structure of the toxin pentamer.

Lack of Pk-MCO binding at site 3

In the SLT-I–Pk-MCO complex, the trisaccharide chain at site 3 is nearly perpendicular to the protein surface and therefore has fewer contacts with the protein than the two other binding sites [37]. Nonetheless, good electron density for the trisaccharide is clearly observed in all 20 equivalent binding sites over the four pentamers in the asymmetric unit of the SLT-I–Pk-MCO complex. This binding is also observed in a Gly62→Thr mutant of SLT-I B in a different crystal form (HL and RJR, unpublished observations), which provides further evidence that binding in site 3 of SLT-I is not a crystallization artifact.

The lack of binding in site 3 of GT3 was unexpected, as changes in this region are minimal. All but one of the residues surrounding the site are conserved from SLT-I, with the sole exception being the substitution of Asn18 for Asp18; as discussed below, residue 18 has been implicated in discriminating between Gb_3 and Gb_4. Apart from the exposed Trp34 sidechains, none of the residues in the binding site differ significantly in conformation. Unless more distant differences have subtle effects on solvent structure, the substitution of an amide for a carboxylate in residue 18 must be responsible for the reduction in affinity for the Gb_3 trisaccharide. In the SLT-I–Pk-MCO complex, Asp18 accepts a bifurcated hydrogen bond from O4 of Gal1. Asn18, on the other hand, must either donate or accept a single hydrogen bond. The lack of observed binding, however, does not necessarily imply that Gb_3 fails to bind in site 3, but may only indicate that the concentration of Pk-MCO was too low. Even in the other, presumably stronger, binding sites, a 50 mM concentration of Pk-MCO is not sufficient to compete with crystal-packing interactions and provide universal occupation. It should also be noted, that in earlier attempts to cocrystallize the related SLT-I B subunit and Pk-MCO, no carbohydrate binding was observed until the carbohydrate concentration was raised above 20 mM (HL and RJR, unpublished observations); this implies that the level of saturation at 50 mM will be low at all sites and that the observation of binding will be very sensitive to minor reductions in affinity.

Correlation with binding data for carbohydrate analogues

Nyholm *et al.* [48] have tested the binding of various deoxy analogues of Gb_3 to several members of the Shiga toxin family. As discussed previously, the results for SLT-I are consistent with the observed binding interactions, with the exception being that no direct interactions with the 2′-OH of Gal2 were observed to explain the importance of this moiety in recognition [37]. We had speculated that there might be water-mediated interactions, which are now seen to exist in site 2 of the GT3 mutant complex (Figure 5b).

In general, as expected from the similarity of observed binding interactions, the same hydroxyl groups are important for binding to all of the toxins studied [48]. SLT-I and SLT-IIe differ significantly only for Gal1–3′-deoxy, which eliminates binding in SLT-IIe but has little effect in SLT-I. This hydroxyl group makes a direct hydrogen bond to SLT-I only in site 2, whereas in GT3 (and presumably also in SLT-IIe and SLT-II) it interacts with Lys13 in site 1. Detailed analysis is complicated, however, by the possibility that the relative affinities of the multiple binding sites could differ among family members. For example, SLT-I binding might also be less affected by Gal1–3′-deoxy because site 3, which does not interact with this hydroxyl group [37], plays a more important role in Gb_3 binding than it does in SLT-IIe.

In the same study [48], computer models for the binding of the Gb_3 trisaccharide in sites 1 and 2 were presented. As discussed previously, the general location of the predicted binding sites agrees with the experimental results, but the details of the interactions are quite different [37]. Subsequent NMR experiments have confirmed the experimentally observed orientation of soluble trisaccharide in binding site 2 [45]. The results of fluorescence-resonance energy-transfer experiments have been suggested to be more consistent with the computer model for site 2 than the experimental binding results [49]. However, the differences in predicted distances are small; given experimental errors in the calibration of distances and, more importantly, uncertainty in the position of the fluorescent probe at the end of a flexible linker to the glucose residue, the energy-transfer results are also consistent with the experimentally observed ligand positions.

Mutations and binding specificity

Our structure reveals that the two mutated residues (Gln65→Glu and Lys67→Gln) are both associated with site 1 (Figure 5a) and that one (Gln65→Glu) is also in the vicinity of site 2 (Figure 5b). Neither single mutation can change the binding preference of SLT-IIe from Gb_4 to Gb_3 [35], indicating that each of the two residues makes a significant contribution to the preference for Gb_4 over Gb_3. The location of the mutations and the strength of their effect also suggests that Gb_4 binds most strongly to site 1 of SLT-IIe. Possible binding interactions for Gb_4 at this site of SLT-IIe have been proposed in our previous modelling paper [50], showing that direct interactions could be made between the βGal*N*Ac moiety and Gln65 and Lys67. (Subsequent refinement of the GT3 complex structure has not shown any significant differences in the interactions with Pk-MCO, which formed the scaffold for the Gb_4 model.) The two mutations might also affect the orientation of the sidechains of Asn15 and Lys13. GT3 has the same binding residues in site 1 as SLT-II, and therefore Gb_3 binding to SLT-II in this site can be inferred to be the same (Figure 5a).

The two GT3 mutations reverse the only sequence differences in binding site 1 between SLT-IIe and SLT-II, with a dramatic reduction of Gb_4 binding and a slight increase in Gb_3 binding. In contrast to SLT-II, however, which does not bind Gb_4, GT3 still retains some Gb_4 binding [35], presumably in one or both of sites 2 and 3. There is room for the extra βGal*N*Ac residue of Gb_4, as seen by placing low-energy conformations of Gb_4 [50] into these sites (results not shown). It is interesting to consider whether the remaining sequence differences shed light on possible Gb_4 binding in sites 2 and 3. In total, the B subunit of SLT-IIe differs from that of SLT-II in 11 positions (Figure 2), including a C-terminal deletion of two residues. Nine sequence differences remain in GT3; of these, seven are situated far from any of the binding sites and would not be

expected to contribute to any difference in binding specificity. Thus, this leaves only two binding site residues that differ between GT3 and SLT-II — at positions 18 and 32. The Gb_4 superpositions show that Asn32 (serine in SLT-II) could interact with the βGal*N*Ac residue of a Gb_4 molecule bound to site 2. This residue, however, is also an asparagine in SLT-I, which binds exclusively to Gb_3, and therefore the presence of Asn32 cannot be pivotal to specificity. The sole remaining residue is Asn18, which is an aspartate in both SLT-I and SLT-II. Interestingly, a mutation of Asp18 to asparagine in SLT-I confers Gb_4-binding activity [35]. As discussed above, the interactions of an asparagine with Gb_3 in site 3 would not be quite as favourable as the interaction of an aspartate. On the other hand, the Gb_4 superpositions show that an asparagine in this position could interact with the βGal*N*Ac residue of Gb_4 bound to site 3. It seems likely, therefore, that site 3 is a secondary Gb_4-binding site in SLT-IIe.

Implications for Gb_3 binding to SLT-II

SLT-II binds to Gb_3 less avidly than SLT-I, which may be relevant to SLT-II's association with human disease [27,26]. As discussed above, the substitution of Asp16 in SLT-I with a glutamate in both SLT-II and GT3 may be responsible in part for the reduced binding affinity. In addition, SLT-II differs from SLT-I in ways that are not shared with GT3 (or SLT-IIe). Figure 2 shows that SLT-II is unique in the substitution of two serine residues for asparagines, at positions 32 and 55, that are involved in binding at site 2. These substitutions will change the nature and possibly the strength of interactions in this binding site.

Biological implications

In foodborne infections caused by *Escherichia coli* O157:H7 the Shiga-like toxins (SLTs) are thought to be responsible for life-threatening vascular complications, such as the hemolytic uremic syndrome (HUS). The SLTs associated with disease in humans can be divided into group I and II members. Previous studies have focused on the group I toxin SLT-I, partly because of difficulties in producing sufficient quantities of the receptor-binding B subunit of group II toxins. The group II toxins SLT-II and SLT-IIc, however, appear to be associated with more severe cases of disease. A variety of studies have shown that cytotoxicity, tissue specificity and pathology are determined by the B subunit and its properties in binding to the glycolipid receptor Gb_3. One important determinant may be the lower receptor-binding affinity exhibited by the group II toxins, which has been speculated to allow the more widespread dissemination of these toxins in the body.

The group II toxins include SLT-IIe, which is associated with edema disease of swine and differs from other SLTs by binding to Gb_4 in preference to Gb_3. A double-mutant of the SLT-IIe B subunit, designated GT3, reverses two sequence differences from SLT-II and switches the binding preference to Gb_3. We report here a structural study of the binding of a Gb_3 analogue to GT3. The study was carried out not only to provide a molecular explanation for the change of specificity, but also to obtain the first structural insights into the group II SLTs.

The results from this structure and our previous study on SLT-I [37] indicate the existence of multiple receptor- binding sites in SLTs. The structures provide a molecular basis for the differences in binding affinity *in vitro*, which are essential to understanding the tissue specificities of the two SLT groups *in vivo* and the pathogenic mechanisms of the disease. Finally, the receptor-binding sites revealed in these structures are excellent targets for drug design and vaccine development.

Materials and methods

GT3 B subunit overproducing clone

The B subunit encoding the sequence of pGT3 [35] was amplified by the polymerase chain reaction using the primers 5′-ATAGAATTCGT-GAATGAAAGG-3′ for + strand synthesis and 5′-TGCAAGCTTAAA-CAAAAGACGCGC-3′ for – strand synthesis [51]. The former primer consists of an *Eco*RI recognition sequence followed by residues 1156–1167 of the *vte* sequence [52] (Genbank accession number M36727); while the latter primer consists of a *Hind*III recognition sequence followed by residues 1473–1459 of the – strand of the *vte* sequence. The resulting amplicon was cleaved with *Eco*RI and *Hind*III, gel purified, and ligated to the *Eco*RI and *Hind*III sites of the expression vector pKK223-3 [53]. The construct was transformed into the host strain *E. coli* JM101, Δ *lac pro sup* E *thi* (F′ *traD36 lacZ* Δ *M15 proAB* $lacI^q$) and the sequence of the B subunit cistron was confirmed. The plasmid was designated pGT3B.

Expression and purification of the GT3 B subunit

Strain JM101 (pGT3B) was grown in Luria broth supplemented with carbenicillin 50 μg ml^{-1} to mid-exponential phase ($A_{600} = 0.6$). B subunit production was then induced with isopropyl β-D-thiogalactoside (2 mM) and incubation was continued for 4 h with aeration. Periplasmic extracts were prepared in phosphate-buffered saline (PBS) containing 0.1 mg ml^{-1} polymyxin B, as described previously [54]. The extract was concentrated by ultrafiltration using an Amicon YM10 membrane, and dialyzed against 50 mM Tris pH 8.0. Chromatography on Q sepharose (Pharmacia) in 50 mM Tris pH 8.0 yielded virtually pure B subunit in the flow-through fractions. These fractions were pooled and dialyzed against 25 mM ethanolamine HCl pH 9.4 and the material was applied to a chromatofocusing column of polybuffer exchanger 94 (Pharmacia). GT3 B subunit was eluted (at pH 8.5) using polybuffer 96 pH 7.0. The ampholytes were removed by gel filtration on Sephadex G-50 and the purified B subunit was lyophilized from solution in 50 mM NH_4CO_3.

Crystallization and data collection

The GT3–Pk-MCO complex was crystallized by the hanging-drop vapour diffusion technique. The hanging drops were formed by mixing 3 μl of the GT3–Pk-MCO complex (5 mg ml^{-1} GT3 + 50 mM Pk-MCO) with 3 μl well solution (1 M NaCl, 10 mM Tris-HCl pH 8.4). Crystals grew to a typical size of $0.3 \times 0.4 \times 0.6$ mm^3 and diffracted to 2.0 Å resolution. Data were collected at room temperature using a Siemens multiwire area detector and were processed with XENGEN 2.0 [55].

GT3 native crystals were grown with a well solution containing 8% polyethylene glycol (PEG) 8000, 0.1 M NaCl and 0.1 M imidazole

(pH 7.4) using the hanging-drop vapour diffusion method. The hanging drops were formed with equal volumes of well solution and protein solution (10 mg ml^{-1}). Data to 2.35 Å resolution were collected at room temperature using a Siemens multiwire area detector and were processed with XENGEN 2.0 [55]. Unlike the data for the complex, these data were weak and relatively incomplete (Table 2).

Structure determination of the GT3–Pk-MCO complex

The GT3–Pk-MCO structure was solved by molecular replacement using the SLT-I B pentamer from the SLT-I–Pk-MCO complex [37] as the search model. The orientation of the search model was determined with the AMoRe package [56]. Ten equal weight peaks related by fivefold NCS and twofold crystallographic symmetry had the highest correlation coefficient in the rotation search ($r = 0.31$, 8.0–4.0 Å), twice as high as the highest noise peak. The search model in the averaged orientation derived from the top ten rotation peaks was used in a translation search with the program BRUTE [57]. The translation search gave a correlation coefficient of 0.60 for the best solution.

Refinement calculations were carried out with X-PLOR [58] and manual rebuilding with O [59]. Electron-density maps used for model rebuilding were improved by fivefold NCS averaging and solvent flattening with the DEMON package [60]. At all stages, the introduction of new degrees of freedom was validated with R_{free} [61]. Following initial rigid-body refinement, non-conserved residues were substituted and the model was adjusted manually. Further rigid-body refinement was followed by positional refinement by energy minimization, maintaining strict fivefold NCS. The model was then subjected to simulated annealing to remove the remaining bias from the initial model. Further refinement used NCS restraints instead of constraints, initially with a relatively tight NCS weight of 300 kcal mol^{-1} Å^{-2}. In the course of the refinement the Gb_3 trisaccharides and water molecules were incorporated into the model and the strength of the NCS restraint was gradually reduced to 10 kcal mol^{-1} Å^{-2}. The relaxation of the NCS restraint improved both the R factor and R_{free}. At this stage, individual temperature factors (B factors) were refined with a restraint of $\sigma = 1.5$ Å^2 for NCS-related atoms and tight restraints of 0.25 Å^2 for mainchain atoms and 0.5 Å^2 for sidechain atoms. In the final stage, a bulk-solvent correction [62] was applied in refinement against the MLF maximum-likelihood refinement target [63]. The final R factor is 0.155 for all data in the range 31.0–2.0 Å, and R_{free} is 0.194 (Table 2). The presence of fivefold NCS in the crystal will reduce the difference between R and R_{free} because correlations between NCS-related reflections will cause overfitting to propagate to a certain extent into the cross-validation data. Nonetheless, numerical experiments indicate that R_{free} is still a useful indicator of the validity of a refinement process in the presence of NCS [64].

Structure determination of native GT3

The unit-cell volume was judged sufficient to contain four pentamers, with a Matthews' coefficient [65] of 2.25 Å^3 Da^{-1}. A rotation function using the complexed GT3 model described above was computed in AMoRe [56] from the CCP4 suite [66] using data from 5–10 Å resolution. This yielded ten equivalent peaks at about four times the rmsd above the mean (with the next highest peaks at $2 \times$ rms), corresponding to two unique solutions to the orientation problem. A solution to the translation function, however, could not be found for either rotation function solution. The pentamer models were improved by expanding the data to space group P1 and carrying out rigid-body refinement of the constituent monomers in CNS [67]. A correct translation solution was then readily obtained for both pentamers, using AMoRe [56] and BRUTE [57].

Manual inspection in XtalView [68] of the packing of the two pentamers indicated space for two more pentamers in the asymmetric unit. To find them, a 'lock–washer' GT3 pentamer model was created by superimposing a monomer of the refined GT3–Pk-MCO complex on the refined structure of the SLT-I B pentamer [40]. Again, ten significant rotation function peaks were obtained, and pentamers corresponding to the two unique orientations were subjected to rigid-body refinement in CNS. Following this procedure, the translation for one of the pentamers was clear from the output of AMoRe and BRUTE. The correct translation solution for the fourth and final pentamer, however, could only be chosen among the top several possibilities by inspecting the crystal packing using XtalView. The final molecular replacement solution with the four pentamers resulted in an R factor of 0.442 and an R_{free} of 0.453 with data from 4–10 Å resolution. In retrospect, it is not clear why the lock–washer model was required in molecular replacement. Although pentamers 3 and 4 have larger screw translations than pentamers 1 and 2 (Table 4), the systematic screw component of the initial lock–washer model disappeared during subsequent refinement.

Cross-validation data were selected in thin shells to reduce the propagation of overfitting in this structure with high NCS. Rigid-body refinement, first defining the four pentamers and then the 20 monomers as rigid groups, was performed in CNS. Strict NCS refinement was then performed followed by restrained NCS refinement in CNS using the maximum-likelihood function MLI [63] and torsion-angle dynamics [69]. Intensity data were used in refinement, although all R factors reported are in terms of structure-factor amplitudes. Refinement cycles were combined with manual rebuilding using the program O [59] and 20-fold NCS-averaged maps created by the DEMON package [60]. The weight of the NCS term in the CNS energy function was gradually reduced from 300 to 50 kcal mol^{-1} Å^{-2} during subsequent refinement cycles. Water molecules were added and removed using a combination of ARP [70] with DEMON and CNS. The final R factor is 0.187 for all data, and the R_{free} is 0.232.

Accession numbers

The atomic coordinates and structure factors have been deposited with the Protein Data Bank with accession codes 1QOH and 2BOS for the native and complexed forms, respectively.

Acknowledgements

Norma Duke collected the native GT3 diffraction data. We thank Bart Hazes for useful discussions, and Penelope Stein and Catherine McPhalen for careful reading of the manuscript. NSP thanks Robin Carrell for the allocation of resources necessary for the completion of this project. NSP is supported by a Wellcome Trust Prize Studentship and by the Natural Sciences and Engineering Research Council of Canada. RJR was supported by the Medical Research Council of Canada (MRC), the Alberta Heritage Foundation for Medical Research and an International Research Scholarship from the Howard Hughes Medical Institute. RJR is now supported by the Wellcome Trust (UK). JLB is supported by the MRC, grant number MT13071.

References

1. Riley, L.W., *et al.*, & Cohen, M.L. (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**, 681-685.
2. Karmali, M.A., Steele, B.T., Petric, M. & Lim, C. (1983). Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* **1**, 619-620.
3. O'Brien, A.D., Lively, T.A., Chang, T.W. & Gorbach, S.L. (1983). Purification of *Shigella dysenteriae* 1 (Shiga)-like toxin from *Escherichia coli* O157:H7 strain associated with haemorrhagic colitis. *Lancet* **2**, 573.
4. Armstrong, G.L., Hollingsworth, J. & Morris, J.G., Jr. (1996). Emerging foodborne pathogens: *Escherichia coli* O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiol. Rev.* **18**, 29-51.
5. Mahon, B.E., Griffin, P.M., Mead, P.S. & Tauxe, R.V. (1997). Hemolytic uremic syndrome surveillance to monitor trends in infection with *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli*. *Emerg. Infect. Dis.* **3**, 409-412.
6. Qadri, S.M. & Kayali, S. (1998). Enterohemorrhagic *Escherichia coli*. A dangerous food-borne pathogen. *Postgrad. Med.* **103**, 179-180, 185-187.
7. Cimolai, N., Carter, J.E., Morrison, B.J. & Anderson, J.D. (1990). Risk factors for the progression of *Escherichia coli* O157:H7 enteritis to hemolytic-uremic syndrome. *J. Pediatr.* **116**, 589-592.
8. Proulx, F., Turgeon, J.P., Delage, G., Lafleur, L. & Chicoine, L. (1992). Randomized, controlled trial of antibiotic therapy for *Escherichia coli*

O157:H7 enteritis. *J. Pediatr.* **121**, 299-303.
9. Weinstein, D.L., Jackson, M.P., Perera, L.P., Holmes, R.K. & O'Brien, A.D. (1989). *In vivo* formation of hybrid toxins comprising Shiga toxin and the Shiga-like toxins and the role of the B subunit in localization and cytotoxic activity. *Infect. Immun.* **57**, 3743-3750.
10. Lingwood, C.A. (1993). Verotoxins and their glycolipid receptors. *Adv. Lipid Res.* **25**, 189-211.
11. Hazes, B. & Read, R.J. (1997). Accumulating evidence suggests that several AB-toxins subvert the endoplasmic reticulum-associated degradation pathway to enter cells. *Biochemistry* **36**, 11051-11054.
12. Endo, Y., Tsurugi, K., Yutsudo, T., Takeda, Y., Ogasawara, K. & Igarashi, K. (1988). Site of action of a Verotoxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes: RNA N-glycosidase activity of the toxins. *Eur. J. Biochem.* **171**, 45-50.
13. Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G.T. (1986). Pathogenesis of *Shigella* diarrhea. XI. Isolation of a Shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J. Exp. Med.* **163**, 1391-1404.
14. Lindberg, A.A., Brown, J.E., Strömberg, N., Westling-Ryd, M., Schultz, J.E. & Karlsson, K.A. (1987). Identification of the carbohydrate receptor for Shiga toxin produced by *Shigella dysenteriae* type 1. *J. Biol. Chem.* **262**, 1779-1785.
15. Lingwood, C.A., *et al.,* & Karmali, M. (1987). Glycolipid binding of purified and recombinant *Escherichia coli*-produced verotoxin *in vitro*. *J. Biol. Chem.* **262**, 8834-8839.
16. Waddell, T., Head, S., Petric, M., Cohen, A. & Lingwood, C. (1988). Globotriosyl ceramide is specifically recognized by *Escherichia coli* verocytotoxin 2. *Biochem. Biophys. Res. Commun.* **152**, 674-679.
17. Cohen, A., Hannigan, G.E., Williams, B.R. & Lingwood, C.A. (1987). Roles of globotriosyl- and galabiosylceramide in Verotoxin binding and high affinity interferon receptor. *J. Biol. Chem.* **262**, 17088-17091.
18. Tesh, V.L., Samuel, J.E., Perera, L.P., Sharefkin, J.B. & O'Brien, A.D. (1991). Evaluation of the role of Shiga and Shiga-like toxins in mediating direct damage to human vascular endothelial cells. *J. Infect. Dis.* **164**, 344-352.
19. Obrig, T.G., Louise, C.B., Lingwood, C.A., Boyd, B., Barley-Maloney, L. & Daniel, T.O. (1993). Endothelial heterogeneity in Shiga toxin receptors and responses. *J. Biol. Chem.* **268**, 15484-15488.
20. Konowalchuk, J., Speirs, J.I. & Stavric, S. (1977). Vero response to a cytotoxin of *Escherichia coli*. *Infect. Immun.* **18**, 775-779.
21. O'Brien, A.D. & LaVeck, G.D.(1983). Purification and characterization of a *Shigella dysenteriae* 1-like toxin produced by *Escherichia coli*. *Infect. Immun.* **40**, 675-683.
22. Brunton, J.L. (1990). The Shiga toxin family: molecular nature and possible role in disease. In *The Bacteria, Vol XI*, *Molecular Basis of Bacterial Pathogenesis.* (Iglewski, B. & Clark, V., eds), pp. 377-398, Academic Press, New York, NY.
23. O'Brien, A.D., Newland, J.W., Miller, S.F., Holmes, R.K., Smith, H.W. & Formal, S.B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694-696.
24. Karmali, M.A., Petric, M., Lim, C., Fleming, P.C., Arbus, G.S. & Lior, H. (1985). The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J. Infect. Dis.* **151**, 775-782.
25. MacLeod, D.L., Gyles, C.L. & Wilcock, B.P. (1991). Reproduction of edema disease of swine with purified Shiga-like toxin-II variant. *Vet. Pathol.* **28**, 66-73.
26. Tesh, V.L., *et al.*, & Samuel, J.E. (1993). Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice. *Infect. Immun.* **61**, 3392-3402.
27. Wadolkowski, E.A., Sung, L.M., Burris, J.A., Samuel, J.E. & O'Brien, A.D. (1990). Acute renal tubular necrosis and death of mice orally infected with *Escherichia coli* strains that produce Shiga-like toxin II. *Infect. Immun.* **58**, 3959-3965.
28. Louise, C.B. & Obrig, T.G. (1995). Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells. *J. Infect. Dis.* **172**, 1397-1401.
29. Head, S.C., Karmali, M.A. & Lingwood, C.A. (1991). Preparation of VT1 and VT2 hybrid toxins from their purified dissociated subunits. Evidence for B subunit modulation of A subunit function. *J. Biol. Chem.* **266**, 3617-3621.
30. Armstrong, G.D., Fodor, E. & Vanmaele, R. (1991). Investigation of Shiga-like toxin binding to chemically synthesized oligosaccharide sequences. *J. Infect. Dis.* **164**, 1160-1167.
31. Armstrong, G.D., *et al.*, & McLaine, P.N. (1995). A phase I study of chemically synthesized verotoxin (Shiga-like toxin) Pk-trisaccharide receptors attached to chromosorb for preventing hemolytic-uremic syndrome. *J. Infect. Dis.* **171**, 1042-1045.
32. Armstrong, G.D., McLaine, P.D. & Roew, P.C. (1998). Clinical trials of Synsorb-Pk in preventing hemolytic-uremic syndrome. In Escherichia coli *O157:H7 and Other Shiga toxin-Producing* E. coli *Strains.* (Kaper, J.B. & O'Brien, A.D., eds), pp. 374-384, ASM Press, Washington, DC.
33. De Grandis, S., Law, H., Brunton, J., Gyles, C. & Lingwood, C.A. (1989). Globotetraosylceramide is recognized by the pig edema disease toxin. *J. Biol. Chem.* **264**, 12520-12525.
34. Keusch, G.T., Jacewicz, M., Acheson, D.W.K., Donohue-Rolfe, A., Kane, A.V. & McCluer, R.H. (1995). Globotriaosylceramide, Gb3, is an alternative functional receptor for Shiga-like toxin 2e. *Infect. Immun.* **63**, 1138-1141.
35. Tyrrell, G.J., Ramotar, K., Toye, B., Boyd, B., Lingwood, C.A. & Brunton, J.L. (1992). Alteration of the carbohydrate binding specificity of verotoxins from Galα1-4Gal to GalNAcβ1-3Galα1-4Gal and vice versa by site-directed mutagenesis of the binding subunit. *Proc. Natl Acad. Sci. USA* **89**, 524-528.
36. Boyd, B., Tyrrell, G., Maloney M., Gyles, C., Brunton, J. & Lingwood, C. (1993). Alteration of the glycolipid binding specificity of the pig edema toxin from globotetraosyl to globotriaosyl ceramide alters *in vivo* tissue targeting and results in a verotoxin 1-like disease in pigs. *J. Exp. Med.* **177**, 1745-1753.
37. Ling, H., *et al.*, & Read, R.J. (1998). Structure of the Shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb_3. *Biochemistry* **37**, 1777-1788.
38. Acheson, D.W., *et al.*, & Keusch, G.T. (1995). Expression and purification of Shiga-like toxin II B subunits. *Infect. Immun.* **63**, 301-308.
39. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl. Crystallogr.* **26**, 283-291.
40. Stein, P.E., Boodhoo, A., Tyrrell, G.J., Brunton, J.L. & Read, R.J. (1992). Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*. *Nature* **355**, 748-750.
41. Fraser, M.E, Chernaia, M.M, Kozlov, Y.V. & James, M.N.G. (1994). Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution. *Nat. Struct. Biol.* **1**, 59-64.
42. Read, R.J. (1986). Improved Fourier coefficents for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140-149.
43. Read, R.J. (1997). Model phases: probabilities and bias. *Methods Enzymol.* **227**, 110-128.
44. Murzin, A.G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences *EMBO J.* **12**, 861-867.
45. Shimizu, H., Field, R.A., Homans, S.W. & Donohue-Rolfe, A. (1998). Solution structure of the complex between the B-subunit homopentamer of verotoxin VT-1 from *Escherichia coli* and the trisaccharide moiety of globotriaosylceramide. *Biochemistry* **37**, 11078-11082.
46. Lindgren, S.W., Samuel, J.E., Schmitt, C.K. & O'Brien, A.D. (1994). The specific activities of Shiga-like toxin type II (SLT-II) and SLT-II-related toxins of enterohemorrhagic *Escherichia coli* differ when measured by Vero cell cytotoxicity but not by mouse lethality. *Infect. Immun.* **62**, 623-631.
47. Perera, L.P., Samuel, J.E., Holmes, R.K. & O'Brien, A.D. (1991). Identification of three amino acid residues in the B subunit of Shiga toxin and Shiga-like toxin type II that are essential for holotoxin activity. *J. Bacteriol.* **173**, 1151-1160.
48. Nyholm, P.-G., Magnusson, G., Zheng, Z., Norel, R., Binnington-Boyd, B. & Lingwood, C.A. (1996). Two distinct binding sites for globotriaosyl ceramide on verotoxins: identification by molecular modelling and confirmation using deoxy analogues and a new glycolipid receptor for all verotoxins. *Chem. Biol.* **3**, 263-275.
49. Picking, W.D., McCann, J.A., Nutikka, A. & Lingwood, C.A. (1999). Localization of the binding site for modified Gb_3 on verotoxin 1 using fluorescence analysis. *Biochemistry* **38**, 7177-7184.
50. Cummings, M.D., Ling, H., Armstrong, G.D., Brunton, J.L. & Read, R.J. (1998). Modeling the carbohydrate-binding specificity of pig edema toxin. *Biochemistry* **37**, 1789-1799.
51. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual,* 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
52. Gyles, C.L., De Grandis, S.A., MacKenzie, C. & Brunton, J.L. (1988). Cloning and nucleotide sequence analysis of the genes determining verocytotoxin production in a porcine edema disease isolate of *Escherichia coli*. *Microb. Pathog.* **5**, 419-426.
53. Amann, E., Brosius, J. & Ptashne, M. (1983). Vectors bearing a hybrid *trp–lac* promoter useful for regulated expression of cloned genes in

Escherichia coli. *Gene* **25**, 167-178.
54. Ramotar, K., Boyd, B., Tyrrell, G., Gariepy, J., Lingwood, C. & Brunton, J. (1990). Characterization of Shiga-like toxin I B subunit purified from overproducing clones of the SLT-I B cistron. *Biochem. J.* **272**, 805-811.
55. Howard, A.J., Gilliland, G.L., Finzel, B.C., Poulos, T.L., Ohlendorf, D.H. & Salemme, F.R. (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Appl. Crystallogr.* **20**, 383-387.
56. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157-163.
57. Fujinaga, M. & Read, R.J. (1987). Experiences with a new translation function program. *J. Appl. Crystallogr.* **20**, 517-521.
58. Brünger, A.T. (1992). *X-PLOR, a System for X-ray Crystallography and NMR, Version 3.1*. Yale University Press, New Haven, CT.
59. Jones, T.A., Zou, J.Y., Cowan, S.W & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of error in these models. *Acta Crystallogr. A* **47**, 110-119.
60. Vellieux, F.M.D.A.P., Hunt, J.F., Roy, S. & Read, R.J. (1995). DEMON/ANGEL: a suite of programs to carry out density modification. *J. Appl. Crystallogr.* **28**, 347-351.
61. Brünger, A.T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.
62. Jiang, J.S. & Brünger, A.T. (1994). Protein hydration observed by X-ray diffraction. Solvation properties of penicillopepsin and neuraminidase crystal structures. *J. Mol. Biol.* **243**, 100-115.
63. Pannu, N.S. & Read, R.J. (1996). Improved structure refinement through maximum likelihood. *Acta Crystallogr. A* **52**, 659-668.
64. Kleywegt, G.J. & Brünger, A.T. (1996). Checking your imagination: applications of the free R value. *Structure* **4**, 897-904.
65. Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
66. *Collaborative Computational Project, No. 4* (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763.
67. Brünger, A.T., *et al*., & Warren, G.L. (1998). Crystallography and NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905-921.
68. McRee, D.E. (1999). XtalView/Xfit – a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156-165.
69. Rice, L.M. & Brünger, A.T. (1994). Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277-290.
70. Lamzin, V.S. & Wilson, K.S. (1993). Automated refinement of protein models. *Acta Crystallogr. D* **49**, 129-147.
71. Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37-40.
72. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.

Because *Structure with Folding & Design* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed (accessed from http://biomednet.com/cbiology/str). For further information, see the explanation on the contents page.