Structural analysis of the catalytic and binding sites of *Clostridium botulinum* neurotoxin B

Subramanyam Swaminathan and Subramaniam Eswaramoorthy

***Clostridium botulinum* neurotoxins are among the most potent toxins to humans. The crystal structures of intact *C. botulinum* neurotoxin type B (BoNT/B) and its complex with sialyllactose, determined at 1.8 and 2.6 Å resolution, respectively, provide insight into its catalytic and binding sites. The position of the belt region in BoNT/B is different from that in BoNT/A; this observation presents interesting possibilities for designing specific inhibitors that could be used to block the activity of this neurotoxin. The structures of BoNT/B and its complex with sialyllactose provide a detailed description of the active site and a model for interactions between the toxin and its cell surface receptor. The latter may provide valuable information for recombinant vaccine development.**

Botulinum and tetanus neurotoxins are solely responsible for the neuroparalytic syndromes of botulism and tetanus. These toxins are among the most poisonous known, with LD_{50} values in humans in the range of 0.1–1 ng kg^{-1} (ref. 1). *Clostridium botulinum* cells produce seven serotypes of botulinum neurotoxin (BoNT, EC 3.4.24.69), called A–G, while tetanus neurotoxin (TeNT, EC 3.4.24.68) is produced by *Clostridium tetani*. These toxins are all produced as single inactive polypetide chains of ~150 kDa that are cleaved by tissue proteinases into two chains: a heavy (H) chain of ~100 kDa and a light (L) chain of ~50 kDa linked by a single disulfide bond. Most of the BoNT serotypes are released in the active, two-chain configuration by *C. botulinum* after proteolysis, although BoNT/E is released as a single chain by a strain lacking the processing protease; BoNT/E is cleaved by host proteases only after uptake into host organisms[2]. These toxins are now used in therapeutics and the Food and Drug Administration in the USA has approved BoNT/A as a therapeutic agent in patients with strabismus, blepharospasm and other facial nerve disorders[3]. Use of BoNT/B as an alternative therapeutic agent is also being investigated.

BoNTs block the release of acetylcholine at the neuromuscular junction to cause flaccid paralysis, while TeNT blocks the release of neurotransmitters such as glycine and γ-aminobutyric acid (GABA) in the inhibitory interneurons of the spinal cord, resulting in spastic paralysis. In spite of different clinical symptoms, their etiological agents are similarly toxic to neuronal cells and these toxins have a similar structural organization[4]. The presence of a conserved sequence in the middle of the L chain led to the discovery that these toxins are metalloendopeptidases acting specifically on protein components of the same neuroexocytosis apparatus present in cytosol. Accordingly, these toxins have a possible common evolutionary origin. An ancestral clostridial gene coding for a nonspecific extracellular zinc protease may have evolved to allow selective cleavage of proteins in the multisubunit complex involved in neuroexocytosis, a fundamental physiological process of higher eukaryotes.

Both BoNTs and TeNT are characterized by similar functional and domain structures. They fall under the general category of AB toxins, which consist of two protomers: protomer A (activating) and protomer B (binding) held together by a disulfide bond after proteolytic cleavage[5]. In the case of neurotoxins, the B protomer is
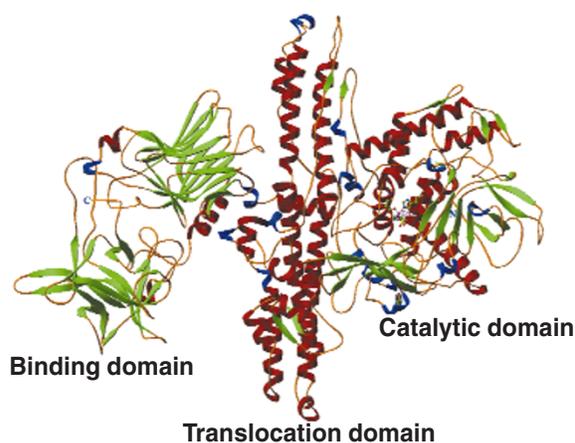

Fig. 1 A RIBBONS[49] representation of BoNT/B. Helices in blue represent 3_{10}-helices. The three functional domains are labeled as binding, translocation and catalytic domains. Zinc and the coordinating residues are shown in ball and stick representation.

the H chain, and the A protomer is the L chain. The H chain can be further cleaved into two chains by papain; the C-terminal chain, H_C, binds to the target cell surface and the N-terminal domain, H_N, is involved in translocating the toxin across the membrane. This three-domain organization is similar to that of diphtheria toxin (DT) and *Pseudomonas aeriginosa* exotoxin A (ETA). However, the effects of these toxins on target cells are different; while in most cases toxins cause cell death, only exocytosis is blocked by botulinum neurotoxins[6]. As DT and ETA consist of three functional units, they have a different structural organization from Cholera toxin (CT), Pertussis toxin (PT), Shiga toxin (ST)[7–10] and other toxins that consist of two functional units, one of which is a pentamer. Even though the general intoxication may be similar, the details are different, with each class having a specific mode.

The inhibition of exocytosis by neurotoxins in the cytosol has been identified as a zinc-dependent specific proteolysis of components in the neuroexocytosis apparatus. Each neurotoxin attacks a specific target component: BoNTs B, D, F and G specifically cleave the vesicle-associated membrane protein (VAMP; also called synaptobrevin); BoNTs A and E cleave a synaptosomal associated

Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA.

Correspondence should be addressed to S.S. *email: swami@bnl.gov*

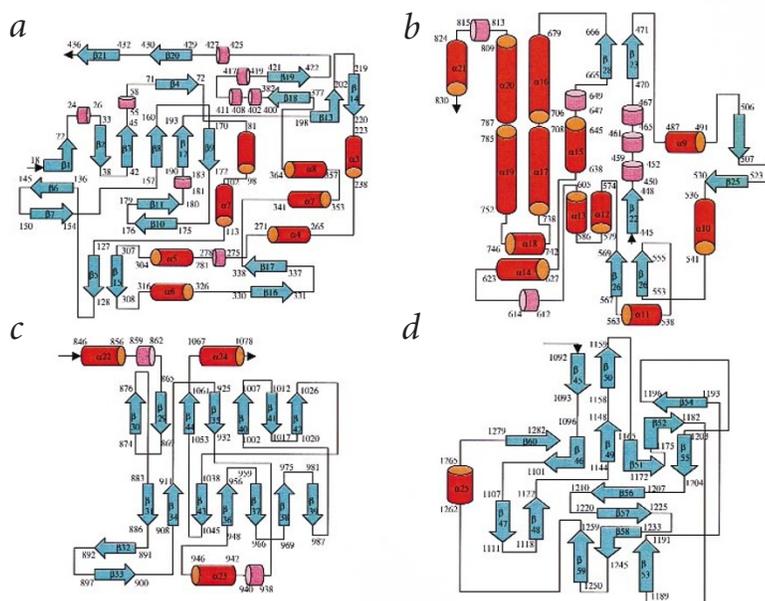# articles

a

b

c

d

Fig. 2 The topography of BoNT/B. β-Strands and α-helices are represented by light blue arrows and red cylinders, respectively. Small purple discs represent 3_{10}-helices. For convenience the topography is given for individual domains. *a*, Catalytic domain; *b*, translocation domain; *c*, the N-terminal half of the binding domain; *d*, the C-terminal half of the binding domain. The starting and ending sequence numbers for each secondary structural element are also given.

tal data were used because these crystals diffracted to better than 1.8 Å resolution. The model is complete except for three residues in the proteolytically cleaved region between the L and H chains.

General description of the structure

The BoNT/B molecule consists of three structural domains corresponding to the three functional domains, catalytic, translocation and binding. These three domains are arranged in a linear fashion with the translocation domain in the middle (Fig. 1). The catalytic domain has a compact globular structure consisting of a mixture of α-helices and β-sheets. The translocation domain is mainly α-helical and the binding domain comprises two subdomains (Fig. 2). Even though the primary sequence starts with the catalytic domain, we discuss below the domain structures starting from the C-terminal end, corresponding to the sequential mode of action of the toxin.

Binding domain

The binding domain consists of two subdomains (Fig. 2*c*,*d*). The N-terminal subdomain consists of two 7-stranded antiparallel β-sheets forming a 14-stranded β-barrel in a jelly roll motif. Helix α22 connects this subdomain to the translocation domain, and α24 connects it to the C-terminal subdomain, which contains a β-trefoil motif. The entire binding domain is tilted away from the central translocation domain and makes minimal inter-

protein of 25 kDa (SNAP-25) by specific hydrolysis; and BoNT/C cleaves syntaxin[11–16]. Söllner *et al.*[17] have shown that together these proteins form part of a complex responsible for mediating vesicle docking and fusion. The importance of Zn^{2+} in the catalysis was established by the fact that removal of Zn^{2+} by chelation blocks the toxic affect. However, in the case of TeNT, Zn^{2+} could be replaced by Ni^{2+} or Co^{2+} without affecting its activity[18].

Even though a large body of chemical, biological, and pathogenic information on these neurotoxins is available, vital information in the form of representative three-dimensional structure is available only for BoNT/A at 3.3 Å resolution[19]. Here we present a high resolution crystal structure of BoNT/B and map its catalytic site. Since the first step in the toxicity of these neurotoxins is binding to gangliosides, we have also determined the crystal structure of BoNT/B in complex with sialyllactose, a partial mimic of a ganglioside, to model and analyze the interactions between the sugar molecule and the binding domain. The differences between the structures of BoNT/A and BoNT/B highlight the different possibilities for designing specific inhibitors. Also, the structure of BoNT/B in complex with sialyllactose presents a model for the interaction between the receptor molecule and the BoNT/B binding domain, revealing prospects for developing recombinant vaccines against these neurotoxins.

The crystal structure of BoNT/B was determined by a combination of multiwavelength anomalous dispersion (MAD) and multiple isomorphous replacement (MIR) methods. Diffraction data were collected from two native crystals. The initial structure determination was carried out with the native I crystal that diffracted to 2.4 Å. In the final stages of model building and refinement, the native II crys-
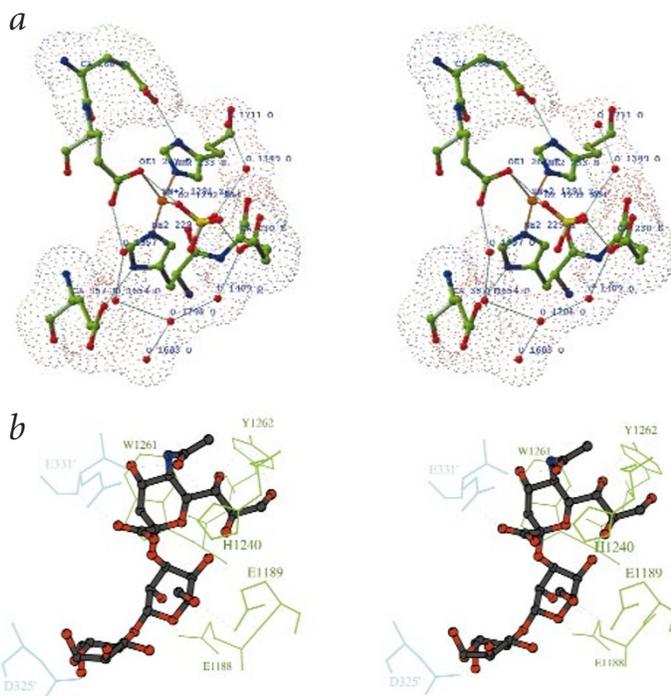
a

b

Fig. 3 The active and binding sites of BoNT/B. *a*, Stereo view of the active site environment. Zinc and bonds to its coordinating residues are shown in orange. Residues in the first shell around the zinc (the immediate surroundings) are shown as ball and stick models, while residues in the second shell and water molecules are shown as ball and stick models with their van der Waals surfaces as dotted surfaces. Hydrogen bonds are shown as thin blue lines. *b*, Stereo view of sialyllactose and the residues interacting with it. Residues from a symmetry related molecule are shown in blue. Potential hydrogen bond contacts are shown as dashed lines.

Fig. 4 Comparison of BoNT/A and BoNT/B. Stereo view of the Cα trace of BoNT/A (magenta) superimposed on that of BoNT/B (green). Significant differences in the orientations of the C-terminal halves of each binding domain with respect to the translocation domain are clearly visible.
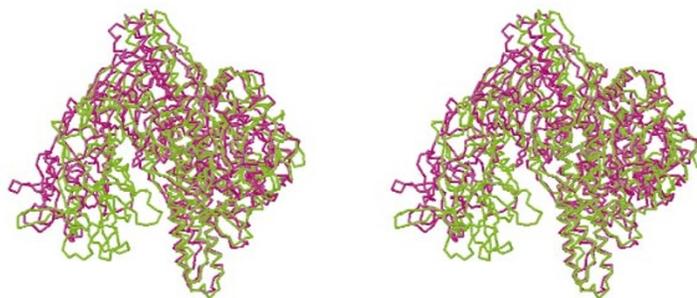
actions with it. The structure of the binding domain is very similar to that of the C-fragment of tetanus toxin and the binding domain of BoNT/A[19,20]. Even though the overall sequence homology is poor for all clostridial neurotoxins in the C-terminal half of the binding domain, it was suggested that they would adopt the same fold, with the differences in sequences accounted for by the extended loop regions[20]. This is shown to be correct, at least for BoNT/A, BoNT/B and TeNT.

Translocation domain

Once toxins are bound to membranes, a temperature- and energy-dependent process internalizes them. After internalization, the toxin can no longer be neutralized by antibodies. To attack their targets in the cytosol, neurotoxins must cross the hydrophobic barrier of the vesicle membrane. This is necessary for all bacterial toxins with intracellular targets and is the least understood step in the process. It has been proposed that acidification of the vesicle lumen by a proton pumping ATPase in the vesicle membrane leads to conformational changes in the toxin that expose a hydrophobic area of the toxin molecule. This area could then create an ion channel in the membrane through which the L chain could be inserted into the cytosol[1,5,21].

The translocation domain consists of two long α-helical regions, each ~105 Å long, that form coiled coils. Two kinks in the coiled coil split the helices into four, each ~50 Å long. The core of the translocation domain consists of a four-helical bundle (α12, α13, α17 and α19) at one end and a three-helical bundle (α16, α20 and α21) at the other (Fig. 2b). The three-helical bundle resembles the translocation region observed in other toxins, for example colicin Ia[22]. The putative membrane spanning region, as predicted by the TMAP program, consists of residues 637–659 (ref. 23). However, this region does not adopt a helical conformation in our structure and extends from the top to the middle of the molecule; the only α-helix in this region is from residues 638–645 in the middle part of the molecule. At acidic pH (6.0), this region has a very flexible conformation that is neither extended nor heli-

cal. A similar conformation is also observed in the structure of BoNT/A determined at neutral pH[19]. Since structures of BoNT/B at lower pH values (between 4 and 6) are not yet known, no conclusions can be made about whether a pH-dependent conformational change takes place. The most interesting part of this domain is a long loop formed by residues 481–532 that wraps around the catalytic domain, as though to keep the catalytic domain in position. This long loop is called the belt region and, as discussed below, plays an important role in shielding the active site.

Catalytic domain

The L chain is released into the cytosol after translocation and attacks a specific target. Segments of the L chains of BoNT and TeNT exhibit sequence homology in the middle of the chain; one segment contains a HEXXH sequence (X is any amino acid) identified as a zinc binding motif in other zinc endopeptidases. Physicochemical measurements have shown that clostridial neurotoxins contain one zinc atom per toxin molecule (except for BoNT/C, which contains two zinc atoms[24]) bound to the L chain. Chemical modification and mutagenesis studies with neurotoxins have suggested that two His residues and one Glu residue provide ligands to zinc, similar to thermolysin[11,18,25]. A fourth ligand is reportedly provided by a water molecule that is responsible for the hydrolysis of a peptide bond in the substrates.

The catalytic domain is a compact globule consisting of a mixture of α-helices and β-sheets and strands (Fig. 2a). The active site zinc is bound deep inside a large open cavity that has a high negative electrostatic potential. This zinc is coordinated by His 229, His 233 and Glu 267. After initial refinement with the complete model and a zinc atom, an F_o - F_c map showed a strong positive peak (10 σ) centered at ~3 Å from the zinc position. The shape of the peak clearly indicated a tetrahedral arrangement of atoms similar to a sulfate or a phosphate ion. Accordingly, this residual density was modeled as a sulfate ion and included in the refinement. The average thermal parameter (23.8 Å²) for the sulfate ion was comparable to those for protein atoms and a subsequent difference Fourier showed no residual peak. The protein was supplied as a precipitate in 60% ammonium sulfate, which was removed by dialysis before crystallization. We hypothesize that the water molecule bound to the zinc ion was displaced by the sulfate ion when the protein was precipitated and remains tightly bound even after dialysis. However, since phosphate buffer was used throughout the protein purification, we cannot completely rule out the possibility of a phosphate
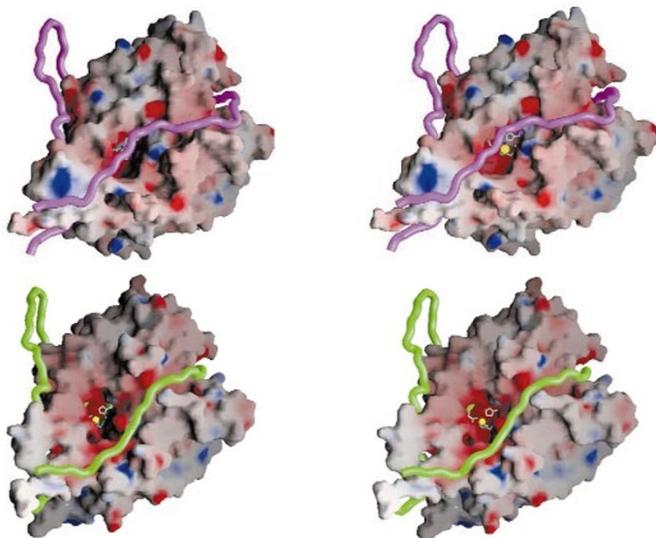

Fig. 5 Comparison of the zinc site in BoNT/A and BoNT/B. Stereo views of the electrostatic potential surfaces of L chains of BoNT/A (top) and BoNT/B (bottom) shown in the same orientation. Zinc and the coordinating residues are shown in yellow and light blue, respectively. For clarity, the sulfate ion in BoNT/B is not shown. The belt regions are shown as worm models in magenta (BoNT/A) and green (BoNT/B). This figure shows the difference in the relative positions of the belt region with respect to the L chain domain.
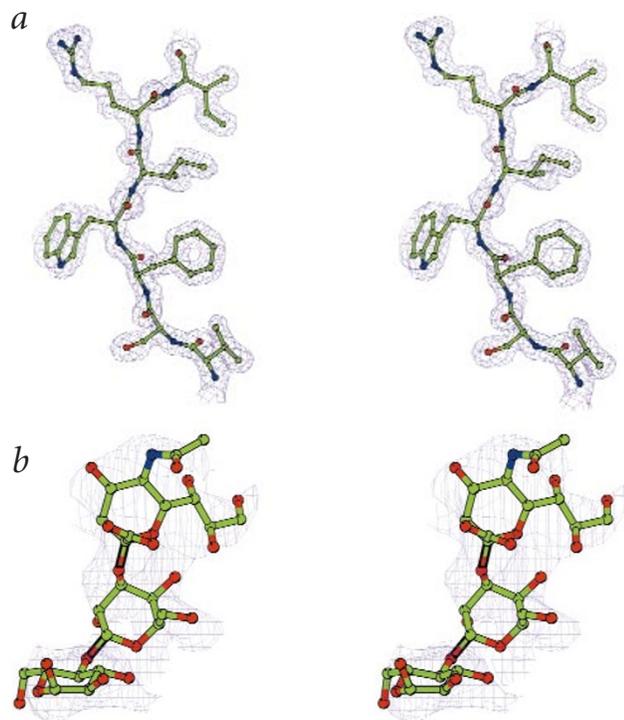
articles

a

b

Fig. 6 Sigma weighted 2F$_o$ - F$_c$ maps. *a*, A representative section of the 2F$_o$ - F$_c$ map of intact BoNT/B structure contoured at 1.2 σ. All reflections up to 1.8 Å resolution were included in the map calculations. *b*, 2F$_o$ - F$_c$ map showing density around sialyllactose contoured at 1.0 σ in the crystal structure of BoNT/B in complex with sialyllactose.

Comparison of BoNT/A and BoNT/B

When the individual functional domains of BoNT/A and BoNT/B are compared they look very similar but with significant differences in certain regions. The root mean square (r.m.s.) deviations between the binding, translocation, and the catalytic domains of BoNT/A and BoNT/B are 1.43, 1.56 and 1.43 Å, respectively, for ~80–85% of Cα atoms of matched residues when calculated with the program LSQMAN[30]. However, when the entire molecules are considered, only 49% of the residues match, with an r.m.s. deviation of 2.06 Å, indicating that the association of the three domains may be slightly different in the two molecules. A stereo view of the Cα trace of BoNT/A superimposed on that of BoNT/B is shown in Fig. 4. While the catalytic and translocation domains (except for the belt region) superimpose fairly well, the binding domains, especially the C-terminal halves, do not. The orientations of the entire binding domains with respect to the rest of the molecule seem different. The binding domain of BoNT/A tilts away from the translocation domain by ~25° more than that of BoNT/B.

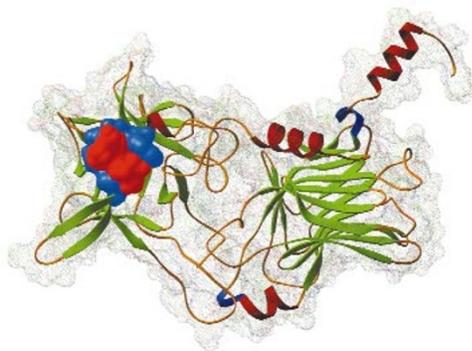## Belt region and the active site cavity

From the present study, it is obvious that the belt region of the translocation domain plays a role in protecting the active site by occluding it from the solvent or exposed region, at least in BoNT/A[19]. In BoNT/A, the cavity is partially covered by the belt, which shields the zinc ion from the environment. In BoNT/B, however, the belt region (residues 481–532) does not shield the zinc ion, thus making it completely accessible to inhibitor or substrate molecules (Fig. 5). A sequence comparison of the residues in this region between BoNT/A and BoNT/B shows that there is a deletion of 14 residues near position 530 in BoNT/A and a deletion of 19 residues near position 492 in BoNT/B. The belt region seems to be shorter in BoNT/B and BoNT/E compared to A, which may explain the differences in the observed conformations. This observation presents interesting possibilities for the design of specific inhibitors for various botulinum neurotoxins. It also suggests that the belt region in BoNT/E may be more similar to that of BoNT/B than BoNT/A. Interestingly, 17 of the 22 residues within 8 Å of the zinc are identical in BoNT/A and BoNT/B. As the radius increases, however, differences become more apparent, with residues from the belt region first appearing at ~15 Å from the zinc. This similarity suggests that the difference in substrate specificity between the two molecules may not be due to the residues in the immediate

ion in this position from crystallographic studies alone. The presence of a sulfate or phosphate ion at the catalytic site has also been reported in the crystal structures of diphtheria repressor, carbonic anhydrase IV, and *Streptomyces griseus* aminopeptidase[26–28]. The coordination distance between zinc and the oxygen atom of the sulfate ion is 2.26 Å and is within the observed values (1.91–2.42 Å; Metalloprotein Database and Browser (MDB) site, http://metallo.scripps.edu) in other structures.

The zinc coordination along with the residues and water molecules in the second shell are shown in Fig. 3a. Zinc is coordinated by Nε2 of His 229, Nε2 of His 233, Oε1 of Glu 267, and O2 of the sulfate ion. The coordination distances are 2.15, 2.15, 2.24 and 2.26 Å, respectively. O2 of the sulfate makes a hydrogen bond contact with Oε1 of Glu 267 (2.83 Å) while O3 of the sulfate is hydrogen bonded to Oε2 of Glu 230 (2.56 Å). We believe that O2 should correspond to the position of the nucleophile. This arrangement is very similar to that in thermolysin except that the sulfate ion replaces a water molecule[29]. However, there is very little similarity between the secondary structures of the two proteins at this site. The active site cavity of BoNT/B has dimensions of 15 × 24 × 25 Å³ with a very wide mouth and an opening opposite it that is partly shielded by Phe 271. Modeling studies based on our polyalanine model of the catalytic domain with 7-N-phenylcarbamoylamino-4-chloro-3-propyloxyisocoumarin have shown that the cavity is large enough to accommodate this molecule (M. Adler, pers. comm.).

Fig. 7 Sequence comparison of the C-terminal region of BoNT/A, BoNT/B, BoNT/E and TeNT. This sequence alignment was obtained with the DIALIGN program[50]. Residues conserved in all four toxins are shown in red. The three conserved Trp residues are indicated by an arrow. His 1292, the site of photo-affinity labeling in TeNT is also indicated by an arrow. Residues aligned in space but not in sequence are shown in green.

```
BoNTA  1200   GVEKILSALE  IPDVGNLSQV  VVMKSkndqg  itnkCKMNLQ  DNNG------
BoNTB  1185   ----------  ------FKKE  EEKLFLAPIS  DSDEFYNTIQ  IKEYDEQPTY
BoNTE  1167   TTNKEKTIKI  SSSGNRFNQV  VVMNSVGn--  ----CTMNFK  NNNG------
TeNT   1202   -NNNEHIVGY  PKDGNAFNNL  DRTLRVGYNA  PGIPLYKKME  AVKLRDLKTY

                                              ↓
BoNTA  1244   ----------  ---NDTGFIG  FHQFnniak-  -------LVA  SNWYNRQIER
BoNTB  1219   SCQLlfkkde  estDEIGLIG  IHRFYESGIV  FEEYKDYFCI  SKWYLKEVKR
BoNTE  1205   ----------  ---NNIGLLG  Fka-------  -----DTVVA  STWYYTHMRD
TeNT   1251   SVQLklyd--  ---DKNASLG  LVGTHNGQIG  NDPNRDILIA  SNWYFNHLKD

                 ↓            ↓
BoNTA  1273   SSRT--LGCS  WEFIPVDDGW  GERpl
BoNTB  1269   KpynlkLGCN  WQFIPKDEGW  TE---
BoNTE  1230   HTNS--NGCF  WNFISEEHGW  QEK--
TeNT   1296   Ki----LGCD  WYFVPTDEGW  Tnd--
```

Fig. 8 A RIBBONS representation of the binding domain of the BoNT/B binding domain inlaid on its van der Waals surface. The sialyllactose and the interacting residues are shown as solid van der Waals surfaces, while the rest of the surface is shown as a dotted surface. The trisaccharide molecule is shown in red and residues interacting with it in blue.

vicinity of the zinc site but may be caused by a long range effect, especially near the mouth of the cavity.

Evidence for ganglioside binding

The toxicity of neurotoxins is initiated by cell binding[21]. Neurotoxins bind first to the large negatively charged surface of the presynaptic membrane, which consists of polysialogangliosides and other acidic lipids[21]. Binding studies have revealed that neurotoxins bind to disialogangliosides and trisialogangliosides (for example, GD1a, GT1b and GD1b), especially the 1b series, with a low affinity[31,32]. However, to produce the high levels of toxicity achieved by such minute concentrations of toxin (subpicomolar) the binding affinity to receptors must be very high. Therefore, a double receptor model for binding has been proposed[31]. In this model, the toxin binds to the negatively charged surface of presynaptic membranes through low affinity, high concentration of polysialogangliosides and then moves laterally to bind to an as yet unknown specific protein receptor[31]. Since the final binding constant is the product of these two binding constants, a very high affinity is achieved. In support of this model, a 58 kDa protein from rat brain synaptosomes has been shown to bind to BoNT/B only in the presence of GT1b or GD1a[33].

To study the binding site of BoNT/B and its interaction with gangliosides, we have determined the structure of BoNT/B in complex with sialyllactose, a partial mimic of one branch of the sugar moiety of GT1b, at 2.6 Å resolution. Sialyllactose was chosen for the soaking study because it is soluble in water, unlike GT1b, and has a terminal sialic acid that was predicted to bind to BoNT/B. Crystals of BoNT/B were soaked in a mother liquor containing 25 mM of sialyllactose for over a week (see Methods). The $F_o - F_c$ map calculated without including the trisaccharide molecule showed clear density near Trp 1261. The trisaccharide molecule could be modeled in this difference density map (Fig. 6b).

In tetanus toxin, mutations in the C-terminal half of the binding domain affect ganglioside binding[34]. Moreover, Shapiro et al.[35] showed that the C-terminal 34 residues (1281–1314) are sufficient for ganglioside binding in TeNT, and photoaffinity labeling occurred predominantly at His 1292. Quenching of BoNT/A tryptophan fluorescence upon ganglioside binding is consistent with the hypothesis that the receptor binding sites in BoNTs and TeNT contain Trp residues[36], which are conserved in the C-terminal regions of a majority of these toxins (Fig. 7). Halpern and Loftus[34] concluded from their mutation study with TeNT that residues 1235–1294 may be particularly critical for direct ganglioside binding of TeNT. They also concluded that residues 1305–1309 may be important for binding but may not bind to gangliosides directly. In BoNT/B, Trp 1261 is the only Trp that is exposed to solvent. Interestingly, the corresponding residues Trp 1265 (BoNT/A) and Trp 1288 (TeNT) are also exposed to solvent. In BoNT/B, this residue is located just before the N-terminus of α25. The observa-

tion that there is fluorescent quenching of tryptophan on ganglioside binding suggests that there may be tryptophan residues near the binding site. The fact that Trp 1261 is exposed to water and is within the residues (1235–1294) found to be critical for direct ganglioside binding suggests that this Trp 1261 is near the binding site. Thus, it was proposed that the ganglioside binding site may be in this region. It was suggested that the two Lys residues close to His 1292 in TeNT might bind to the negatively charged carboxylate groups of the sialic acids of GT1b[35]. By analogy to TeNT, the three Lys residues and one Arg close to Glu 1265 of BoNT/B should bind to the negatively charged carboxylate groups of the sialic acids of GT1b.

In the structure of the BoNT/B–sialyllactose complex, there is a cleft between Trp 1261 and His 1240, and the sialic acid sits between these two with its sugar moiety making several hydrogen bonds with the protein (Fig. 3b). Sialic acid and the galactose make hydrogen bonding contacts with Glu 1188, Glu 1189, His 1240 and Tyr 1262 (Table 1). Sialic acid also makes hydrophobic contacts with Trp 1261. However, the sialic acid does not make any contact with Lys 1267, Arg 1268, Lys 1269 or Glu 1265, as proposed for TeNT[35]. These residues, though close to Trp 1261 in the primary sequence, are spatially far from the binding site. The closest contact the sialic acid makes with a Lys residue is with Lys 1264 (6.4 Å). Only the terminal sialic acid and the adjacent galactose moiety make contacts with the protein molecule while the glucose moiety protrudes outside the neurotoxin. When the C-terminal subdomains of BoNT/A, TeNT and BoNT/B are superimposed after alignment of the LSQMAN sequence (data not shown), His 1252 (BoNT/A) and His 1270 (TeNT) superimpose on His 1240 of BoNT/B, thus creating the same kind of pocket for sialic acid. Similarly, Glu 1189 (BoNT/B), Glu 1202 (BoNT/A) and Asp 1221 (TeNT) also superimpose. However, these residues do not align sequentially (shown in green in Fig. 7). Because sialyllactose mimics one branch of the sugar moiety of GT1b that contains a terminal sialic acid with a 2→3 linkage to the galactose residue, we propose that our structure provides a model for interaction of gangliosides with neurotoxins. His 1292, the site of photoaffinity labeling in TeNT[35] lies on the other side of Trp 1288 relative to His 1240, creating the small cleft between these residues. However, preliminary studies with other branched sugar molecules show no density in this region in the difference density map that could accommodate the sialic acid on the second branch of GT1b. This

Table 1 Potential interactions between BoNT/B and sialyllactose[1,2]

Sia 2000	O10–ND1	His 1240	2.9 Å
Sia 2000	O10–O	Glu 331′	2.6 Å
Sia 2000	O9–OH	Tyr 1262	3.4 Å
Sia 2000	O8–O4	Gal 2001	2.5 Å
Sia 2000	O7–OH	Tyr 1262	3.0 Å
Sia 2000	O1B–N	Glu 331′	2.8 Å
Gal 2001	O6–OE2	Glu 1189	2.9 Å
Gal 2001	O2–OE1	Glu 1188	3.0 Å
Glc 2002	O2–OD2	Asp 325′	3.1 Å

[1]3′-N-Acetylneuraminyl-lactose(α-Neu5Ace[2→3]-β-D-Gal-[1→4]-D-Glc).
[2]Primed numbers indicate residues from symmetry related molecules.

articles

Table 2 Crystal data, phasing, and refinement statistics

Cell parameters

Native I	Space group P2$_1$; a = 78.164, b = 125.580, c = 94.256 Å, β = 112.46°
Native II	Space group P2$_1$; a = 76.080, b = 123.110, c = 95.860 Å, β = 113.00°
BoNT/B–sugar	Space group P2$_1$; a = 76.298, b = 123.651, c = 95.560 Å, β = 113.38°

Phasing statistics

Derivative (all MAD)	Heavy atom sites	Resolution (Å)	Number of reflections	Overall FOM centric / acentric	After solvent flattening
Tungsten	2	2.5	53,356	0.3774 / 0.1881	0.8757
Platinum	5	3.0	33,000	0.5770 / 0.3800	0.8885
Mercury	4	2.8	40,634	0.5018 / 0.2844	0.8844
Osmium	5	2.2	74,774	0.3090 / 0.2120	0.8741

Refinement statistics

	Native II	BoNT/B–sialyllactose
Resolution range (Å)	50.0–1.8	50.0–2.6
Number of reflections	144,901	39,278
R-factor	0.195	0.21
R-factor (all data)	0.199	0.214
R$_{free}$	0.227	0.27
Number of residues	1,287	1,286
Number of protein atoms	10,587	10,438
Number of heterogen atoms	6	49
Number of water molecules	870	98
Average B factor (Å^2)	31.8	19.3
R.m.s. deviations		
Bond lengths (Å)	0.006	0.008
Bond angles (°)	1.226	1.313

is the first time crystallographic data on ganglioside binding to BoNT/B has been presented. A ribbon drawing of the binding domain with the bound sugar molecule shows the sialyllactose and the interacting protein residues forming a nice lock and key arrangement (Fig. 8).

Implications of sugar binding

The structure of the BoNT/B–sialyllactose complex provides a model for the interaction between the sugar moiety of ganglioside GT1b and the binding domain. This structure defines the residues interacting with the sugar moiety that may be very similar for all BoNTs. Since binding of ganglioside to the neuronal cells is the first step in the toxic activity of neurotoxins, the binding site is a possible target for the development of a vaccine of recombinant, inactivated BoNT.

Conclusions

The three-dimensional structure of BoNT/B reveals at high resolution its active site geometry. It also shows that the conformation of the belt region is different from that of BoNT/A. Accordingly, in spite of a similar folding pattern, strategies for designing inhibitors may be different for the two serotypes. The BoNT/B–sialyllactose complex defines the residues that interact with the sialic acid. Since cell binding initiates the toxic activity of BoNT/B, mutating residues that interact with the sugar moiety may render the toxin inactive; this could help in developing recombinant vaccines.

Note added in proof: A cocrystal structure of the catalytic domain of BoNT/B and a synaptobrevin fragment is reported by Hanson and Stevens[51] elsewhere in this issue. We have recently determined the structure of intact BoNT/B with the sulfate ion replaced by a water molecule, and the zinc coordinating distances agree with this report.

Methods

Protein preparation and crystallization. BoNT/B was purchased from the Food Research Institute, (Madison, Wisconsin, USA). Preparation of protein and crystallization methods are as described[37]. Crystals were grown from 10% (w/v) PEG 4000, 0.1 M 2-(N-morpholino)-ethanesulphonic acid (MES) at pH 6.0. Good diffracting quality crystals were obtained by seeding with cat whiskers.

Cryostabilization and heavy metal derivatization. Cryostabilization was found to be better when 5% (v/v) glycerol was present in the original crystallization droplet. Crystals were transferred to mother liquor containing sodium cacodylate buffer instead of MES at pH 6.0 and 10% (v/v) glycerol and then picked up on Hampton loops and immediately transferred to a vial containing liquid nitrogen before being mounted on the goniostat.

Data collection and phase determination. The initial native data set was collected from a crystal of size 0.25 × 0.25 × 0.1 mm^3 at beam line X12C of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory, using a Brandeis B1 CCD based detector. The data set extended to 2.4 Å resolution and was 93.2% complete. Data collection and processing were conducted with the programs MARMAD[38] and DENZO[39]. Derivative data were collected from crystals soaked in sodium tungstate, PIP (di-μ-iodobis(ethylenediamine) diplatinum(II) nitrate), mersalyl acid and potassium hexachloro osmate(IV). For every derivatized crystal, data were collected at three wavelengths, absorption edge, inflection point and remote point, to be phased by a modified MAD procedure[40]. Each derivative data set was considered as a MAD data set. The heavy atom positions were located by Patterson and cross Fourier methods with the PHASES program[41]. Experimental phases for the first derivative were calculated with the program SHARP[42]. For successive derivatives, the previous set of phases was used as external phases in SHARP. In this way, the experimental phases were improved successively. The final derivative data set was with hexachloro osmate, which was the best derivative for this crystal. However, finding the right conditions and duration was very tricky. After many unsuccessful attempts, the heavy atom reagent was added to the cryoprotectant solution and the crystal was transferred and soaked for 20–30 min before being picked up in a loop and plunged into liquid nitrogen for data collection. This procedure of heavy atom soaking may be applicable to other proteins as well. Data collection and refinement statistics are given in Table 2. In our experience we found that the procedure of successively using the previous phase sets was very successful since no single derivative was sufficient to trace the chain. The experimental phases were improved by solvent flipping by SHARP/SOLOMON[43]. Later, we collected native data from the native II crystal, which extended to 1.8 Å resolution. But when this was compared with the original native data set the R$_{merge}$ was very high (~40%). However, this was used for refinement as described below.

Model building and refinement. For initial model building, phases from the last derivative set (osmium) were used with the remote data set as the native data. The map was of excellent quality, showing the secondary structural elements and side chains in most cases very clearly. The initial model was built with a 3.0 Å resolution map with the molecular graphics program O[44]. It was immediately evident that the molecule consists of three well-defined domains. The Cα chain trace was completed except for about 160 residues and most of the side

chains could be identified and fitted in the electron density. A slow cool annealing refinement was carried out with the use of CNS[45] with the initial native data set. At this stage the R-factor was 0.30. From this point on the native II data set was used for model building and refinement as follows. Since the native II crystal was of a different crystal form, the molecule was rotated and positioned in the unit cell with AMoRe[46]. This gave a good correlation coefficient and R-factor (0.50 and 0.41, respectively) and the model looked reasonably good. However, to find the missing residues and to fine-tune the model we used the ARP/wARP program[47]. We used the initial phases from the molecular replacement and conducted an automatic model building procedure that gave 1,123 main chain peptides in 21 chains with a connectivity index of 0.96. The auto side chain fitting procedure could fit most of the residues. To our knowledge, this is the largest structure so far for which the ARP/wARP program has been used successfully. The model was further refined by CNS and refinement and model building were done alternately to fit the complete model using Sigma weighted $2F_o - F_c$ and $F_o - F_c$ maps. The progress and validity of the refinement protocol was checked by monitoring the R_{free} for 5% (7,276 reflections) of the total reflections (144,901). The final R-factor and R_{free} are 0.195 and 0.227, respectively, for reflections in the resolution range 50–1.8 Å. The final model consists of 1,287 residues, 870 water molecules, one zinc and one sulfate ion. Model geometry was analyzed with PROCHECK[48] and 89% of the residues are in the most favored region in the Ramachandran plot. A representative section of the $2F_o - F_c$ map is shown in Fig. 6a.

Structure determination of the BoNT/B–sialyllactose complex. Crystals of BoNT/B were soaked in the mother liquor containing 25mM silalyllactose for over a week. Data were collected at beam line X25 of the NSLS to 2.6 Å resolution. Since the R_{merge} between this data set and the native data set was 0.13, the BoNT/B model obtained from the native crystals was used to refine the structure. After rigid body and simulated annealing refinement, the R-factor dropped to 0.22. A difference Fourier at this stage revealed density for the trisaccharide. Sialyllactose was built into the density and the model was further refined to a final R-factor of 0.21.

Coordinates. The atomic coordinates for the intact BoNT/B structure (accession code 1EPW) and the BoNT/B–sialyllactose structure (accession code 1F31) have been deposited in the Protein Data Bank.

1. Schiavo, G., Rossetto, O. & Montecucco, C. Clostridial neurotoxins as tools to investigate the molecular events of neurotransmitter release. *Semin. Cell Biol.* **5**, 221–229 (1994).
2. Sathyamurthy, V. & Dasgupta, B.R. Separation, purification, partial characterization and comparison of the heavy and light chains of botulinum neurotoxin types A, B and E. *J. Biol. Chem.* **260**, 10461–10466 (1985).
3. Johnson, E.A. & Goodnough, M.C. Preparation and properties of botulinum toxin type A for medical use. In *Handbook of dystonia* (eds Tsui, J.K.C. & Calne, D.B.) 347–365 (Marcel Dekkar, Inc., New York; 1995).
4. Simpson, L.L. Molecular pharmacology of botulinum toxin and tetanus toxin. *Annu. Rev. Pharmacol. Toxicol.* **26**, 427–453 (1986).
5. Montecucco, C., Papini, E. & Schiavo, G. Bacterial protein toxins penetrate cells via a four-step mechanism. *FEBS Lett.* **346**, 92–98 (1994).
6. Schantz, E.J. & Johnson, E.A. Properties and use of botulinum toxin and other microbial neurotoxins in medicine. *Microbiol. Rev.* **56**, 80–99 (1992).
7. Choe, S. *et al.* The crystal structure of diphtheria toxin. *Nature* **357**, 216–222 (1992).
8. Allured, V.S., Collier, R.J., Carroll, S.F. & McKay, D.B. Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0 Å resolution. *Proc. Natl. Acad. Sci. USA* **83**, 1320–1324 (1986).
9. Sixma, T.K. *et al.* Refined structure of *Escherichia coli* heat-labile enterotoxin, a close relative of cholera toxin. *J. Mol. Biol.* **230**, 890–918 (1993).
10. Stein, P.E. *et al.* The crystal structure of pertussis toxin. *Structure* **2**, 45–57 (1994).
11. Schiavo, G., Shone, C.C., Rossetto, O., Alexander, F.C.G. & Montecucco, C. Botulinum neurotoxin serotype F is a zinc endopeptidase specific for VAMP/synaptobrevin. *J. Biol. Chem.* **268**, 11516–11519 (1993).
12. Schiavo, G. *et al.* Tetanus and botulinum-B neurotoxins block neurotransmitter release by a proteolytic cleavage of synaptobrevin. *Nature* **359**, 832–835 (1992).
13. Blasi, J. *et al.* Botulinum neurotoxin C blocks neurotransmitter release by means of cleaving HPC-1/syntaxin. *EMBO J.* **12**, 4821–4828 (1993).
14. Schiavo, G. *et al.* Botulinum neurotoxin serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds. *FEBS Lett.* **335**, 99–103 (1993).
15. Schiavo, G. *et al.* Identification of the nerve terminal targets of botulinum neurotoxin serotypes A, D and E. *J. Biol. Chem.* **268**, 23784–23787 (1993).
16. Schiavo, G. *et al.* Botulinum G neurotoxin cleaves VAMP/synaptobrevin at a single Ala-Ala peptide bond. *J. Biol. Chem.* **269**, 20213–20216 (1994).
17. Söllner, T. *et al.* SNAP receptors implicated in vesicle targeting and fusion. *Nature* **362**, 318–324 (1993).
18. Schiavo, G. *et al.* Tetanus toxin is a zinc protein and its inhibition of neurotransmitter release and protease activity depend on zinc. *EMBO J.* **11**, 3577–3583 (1992).
19. Lacy, D.B., Tepp, W., Cohen, A.C., DasGupta, B.R. & Stevens, R.C. Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nature Struct. Biol.* **5**, 898–902 (1998).
20. Umland, T.C. *et al.* Structure of the receptor binding fragment H_c of tetanus neurotoxin. *Nature Struct. Biol.* **4**, 788–792 (1997).
21. Menestrina, G., Schiavo, G. & Montecucco, C. Molecular mechanisms of action of bacterial protein toxins. *Mol. Aspects Med.* **15**, 79–193 (1994).
22. Weiner, M., Freymann, D., Ghosh, P. & Stroud, R.M. Crystal structure of colicin Ia. *Nature* **385**, 461–464 (1997).
23. Persson, B. & Argos, P. Topology prediction of membrane proteins. *Protein Sci.* **5**, 363–371 (1996).
24. Schiavo, G., Shone, C.C., Bennett, M.K., Scheller, R.H. & Montecucco, C. Botulinum neurotoxin type C cleaves a single Lys-Ala bond within the carboxyl-terminal region of syntaxins. *J. Biol. Chem.* **270**, 10566–10570 (1995).
25. Schiavo, G., Rossetto, O., Santucci, A., Dasgupta, B.R. & Montecucco, C. Botulinum neurotoxins are zinc proteins. *J. Biol. Chem.* **267**, 23479–27483 (1992).
26. Pohl, E., Holmes, R.K. & Hol, W.G.J. Motion of the DNA-binding domain with respect to the core of the diphtheria toxin repressor (DtxR) revealed in the crystal structures of apo- and holo-DtxR. *J. Biol. Chem.* **273**, 22420–22427 (1998).
27. Stams, T. *et al.* Crystal structure of the secretory form of membrane-associated human carbonic anhydrase IV at 2.8-Å resolution. *Proc. Natl. Acad. Sci. USA* **93**, 13589–13594 (1996).
28. Greenblatt, H.M. *et al. Streptomyces griseus* aminopeptidase: X-ray crystallographic structure at 1.75 Å resolution. *J. Mol. Biol.* **265**, 620–636 (1997).
29. Holmen, M.A. & Matthews, B.W. Structure of thermolysin refined at 1.6 Å resolution. *J. Mol. Biol.* **160**, 623–639 (1982).
30. Kleywegt, G.J. & Jones, T.A. Detecting folding motifs and similarities in protein structures. *Methods Enzymol.* **277**, 525–545 (1997).
31. Montecucco, C. How do tetanus and botulinum toxins bind to neuronal membranes? *Trends Biochem. Sci.* **11**, 314–317 (1986).
32. Kitamura, M., Takamiya, K., Aizawa, S., Furukawa, K. & Furukawa, K. Gangliosides are the binding substances in neural cells for tetanus and botulinum toxins in mice. *Biochem. Biophys. Acta* **1441**, 1–3 (1999).
33. Nishiki, T-i. *et al.* The high-affinity binding of *Clostridium botulinum* type B neurotoxin to synaptotagmin II associated gangliosides GT_{1b}/GD_{1a}. *FEBS Lett.* **378**, 253–257 (1996).
34. Halpern, J.L. & Loftus, A. Characterization of the receptor-binding domain of tetanus toxin. *J. Biol. Chem.* **268**, 11188–11192 (1993).
35. Shapiro, R.S. *et al.* Identification of a ganglioside recognition domain of tetanus toxin using a novel ganglioside photoaffinity ligand. *J. Biol. Chem.* **272**, 30380–30386 (1997).
36. Kamata, Y., Yoshimoto, M. & Kozaki, S. Interaction between botulinum neurotoxin type A and ganglioside: ganglioside inactivates the neurotoxin and quenches its tryptophan fluorescence. *Toxicon* **35**, 1337–1340 (1997).
37. Swaminathan, S. & Eswaramoorthy, S. Crystallization and preliminary X-ray analysis of *Clostridium botulinum* neurotoxin type B. *Acta Crystallogr.* **in the press** (2000).
38. Skinner, J.M. & Sweet, R.M. Integrated software for a macromolecular crystallography synchrotron beamline. *Acta Crystallogr. D* **54**, 718–725 (1998).
39. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
40. Ramakrishnan, V. & Biou, V. Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. *Methods Enzymol.* **276**, 538–557 (1997).
41. Furey, W. & Swaminathan, S. PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules. *Methods Enzymol.* **276**, 590–620 (1997).
42. De La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* **276**, 472–493 (1997).
43. Collaborative Computational Project Number 4. CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
44. Jones, T.A., Zou, J., Cowtan, S. & Kjeldgaard, M. Improved methods in building protein models in electron density map and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
45. Brunger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
46. Navaza, J. & Saludjian, P. AMoRe: an automated molecular replacement program package. *Methods Enzymol.* **276**, 581–594 (1997).
47. Perrakis, A., Morris, R. & Lamzin, V.S. Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463 (1999).
48. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
49. Carson, M. Ribbons 2.0. *J. Appl. Crystallogr.* **24**, 958–961 (1991).
50. Morgenstern, B., Werner, T. & Dress, A.W.M. Multiple DNA and protein sequence alignment based on segment-to-segment comparison. *Proc. Natl. Acad. Sci. USA* **93**, 12098–12103 (1996).
51. Hanson, M.A. & Stevens, R.C. Cocrystal structure of synaptobrevin-II bound to botulinum neurotoxin type B at 2.0 Å resolution. *Nature Struct. Biol.* **7**, 687–692 (2000).