Structural Analysis of Botulinum Neurotoxin Serotype F Light Chain: Implications on Substrate Binding and Inhibitor Design†

Rakhi Agarwal,‡ Thomas Binz,§ and Subramanyam Swaminathan*,‡

Biology Department, Brookhaven National Laboratory, Upton, New York 11973, and Department of Biochemistry, Medizinische Hochschule Hannover, 30623 Hannover, Germany

Received May 27, 2005; Revised Manuscript Received July 11, 2005

ABSTRACT: The seven serologically distinct *Clostridium botulinum* neurotoxins (BoNTs A−G) are zinc endopeptidases which block the neurotransmitter release by cleaving one of the three proteins of the soluble *N*-ethylmaleimide-sensitive-factor attachment protein receptor complex (SNARE complex) essential for the fusion of vesicles containing neurotransmitters with target membranes. These metallopeptidases exhibit unique specificity for the substrates and peptide bonds they cleave. Development of countermeasures and therapeutics for BoNTs is a priority because of their extreme toxicity and potential misuse as biowarfare agents. Though they share sequence homology and structural similarity, the structural information on each one of them is required to understand the mechanism of action of all of them because of their specificity. Unraveling the mechanism will help in the ultimate goal of developing inhibitors as antibotulinum drugs for the toxins. Here, we report the high-resolution structure of active BoNT/F catalytic domain in two crystal forms. The structure was exploited for modeling the substrate binding and identifying the S1′ subsite and the putative exosites which are different from BoNT/A or BoNT/B. The orientation of docking of the substrate at the active site is consistent with the experimental BoNT/A-LC:SNAP-25 peptide model and our proposed model for BoNT/E-LC:SNAP-25.

Clostridium botulinum produces seven antigenically distinct neurotoxins (BoNT/A−G)[1] recognized as the most potent biological toxins (*1*). BoNTs cause the disease botulism, a neuromuscular disorder characterized by flaccid paralysis that is due to the blockage of neurotransmitter release (*2*). Botulism is also caused by other members of the genus such as *Clostridium baratii*. There are a few case reports on botulism due to BoNT/F secreted by the *C. baratii* in adults and infants (*3*−*5*). BoNTs are released as inactive molecules of 150 kDa and are cleaved by exogenous or endogenous proteases into active dichain molecules with an N-terminal light chain (LC, 50 kDa) and a C-terminal heavy chain (HC, 100 kDa) held together by a disulfide bond (*6*). The C-terminal domain of the HC mediates binding of the toxin to the specific neuronal receptors, and the N-terminal domain enables the catalytically active LC to translocate to the cytosol, where it recognizes and cleaves one of three SNARE proteins. Though LCs of BoNTs share significant sequence homology (30−60%) and structural similarity, they selectively cleave specific SNARE proteins. BoNT/A, -C, and -E cleave SNAP-25, and BoNT/B, -D, -F, and -G and tetanus toxin (TeNT) cleave VAMP (also known as synaptobrevin) (*7*). BoNT/C is unique since it also cleaves syntaxin. VAMP is a tail anchored membrane protein of exocytotic vesicles. VAMP comprises a family of different isoforms, and neurotoxin sensitive isoforms are present in many non-neuronal tissues (*8*). However, clostridial neurotoxins do not act on non-neuronal cells due to the absence of cell surface neurotoxin receptors to internalize them (*7*, *9*). The LCs are unique zinc endopeptidases as to the requirement of unusually large segments of their substrates for optimal activity. This is reflected in the presence of additional recognition and binding sites, called exosites, on BoNTs besides the active site. All three SNARE proteins contain conserved nine residue segments, termed SSR. There are four (S1−S4) such SSRs in SNAP-25 and two each in VAMP (V1 and V2) and Syntaxin (X1 and X2) (*10*). At least one of the SSRs is required for the substrate cleavage in addition to the cleavage site (*11*, *12*). While the cleavage site is at the C-terminal side of SSRs in most of them, BoNT/F and -D are unique in that they cleave peptide bonds located between V1 and V2. BoNT/F cleaves Gln58−Lys59 while BoNT/D cleaves the adjacent peptide bond, Lys59−Leu60. Since mutation of both V1 and V2 residues affects BoNT/F activity, it may require both V1 and V2 for activity unlike BoNT/D which needs only V1 (*12*).

† Research supported by the U.S. Army Medical Research Acquisition Activity (Award No. DAMD17-02-2-0011) under DOE Prime Contract No. DE-AC02-98CH10886 with Brookhaven National Laboratory. T.B. was supported by Grant RGY0027/2001 from Human Frontier Science Program.

* Author to whom correspondence should be addressed. E-mail: swami@bnl.gov. Tel: 1-(631)-344-3187. Fax: 1-(631)-344-3407.

‡ Brookhaven National Laboratory.

§ Medizinische Hochschule Hannover.

[1] Abbreviations: BoNT, botulinum neurotoxin; TeNT, tetanus neurotoxin; SNARE, soluble *N*-ethylmaleimide-sensitive-factor attachment protein receptor; SSR, SNARE Secondary Recognition; LC, light chain; SNAP-25, 25 kDa synaptosome-associated protein; VAMP, vesicle-associated membrane protein; USAMRIID, United States Army Medical Research Institute of Infectious Diseases; PCR, polymerase chain reaction; GST, glutathione sulfotransferase; PMSF, phenylmethylsulfonyl fluoride; HEPES, *N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]; rmsd, root-mean-square deviation.

10.1021/bi0510072 CCC: $30.25
Published on Web 08/10/2005

Due to the unique substrate requirement and stringent cleavage selectivity, the development of a multivalent inhibitor against all seven serotypes seems unlikely. Thus, the current need is to understand the catalytic mechanism of each one of them, to define the variable substrate specificity and selective cleavages by them. Knowledge of enzyme exosites along with the residues of substrate involved in enzyme−substrate complex formation is a prerequisite for drug design for botulism. The docking of the substrate at the active site needs to be studied more precisely to understand the molecular mechanism of catalysis brought about by these toxins in order to develop inhibitors. To achieve this goal, high-resolution crystal structures of catalytic domains of all serotypes and their complexes with the substrates or substrate analogues would be a priority. So far only two complex structures are available (*13*, *14*).

Here we report the expression, purification, and crystal structure analysis of BoNT/F-LC in two crystal forms. This is the first structural report for any functional fragment of BoNT/F. We also compare BoNT/F-LC with the available crystal structures of BoNT/A, -B, and -E and TeNT. We have identified the S1′ subsite and the putative exosites in the toxin. The suggested orientation of docked substrate at the active site is consistent with thermolysin-inhibitor, BoNT/A-LC:SNAP-25 peptide, and our proposed model for the BoNT/E-LC:SNAP-25 complex structures (*13*, *15*, *16*).

MATERIALS AND METHODS

Cloning, Expression, and Purification of BoNT/F-LC. The pFB4 vector possessing the full length of BoNT/F-LC (Met1−Lys439; identical to strain NCTC 10821; GenBank accession number X81714) with 6xHis tag was used as template. This has four extra residues at the C-terminus. The LC encoding segment was PCR amplified using the forward primer (5′ ATG ACC ATG GGA ATG CCA GTT GTA A 3′) bearing a *Nco*I restriction site at the 5′ end and the reverse primer (5′ GAT GCT CGA GCC CGG GAG TTG GCG G 3′) bearing a *Xho*I site at the 3′ end. The PCR products were digested with the *Nco*I and *Xho*I restriction enzymes and ligated to the pET-28b vector between similar sites. The nucleotide sequence was confirmed for the entire BoNT/F-LC gene in both strands by the Big Dye termination Cycle Sequencing (Applied Biosystems). The plasmid pET-28b-LC was transformed to BL21 (DE3) cells for expression of the protein.

The expression and purification methods are similar to those for BoNT/E-LC (*17*). The 2xYT medium (1.6% bactotryptone, 1.0% bacto-yeast extract, and 0.5% NaCl) containing 100 μg/mL kanamycin was inoculated with the freezer stock of BL21(DE3) cells containing the pET-F-LC vector. Cells were grown in a shaking incubator at 37 °C until A_{600} reached 0.6−0.8. At this point, 1 mM IPTG was added and the cell growth was continued for an additional 14−16 h at 20 °C. Cells were then harvested for the protein preparation, and 30 mL of lysis buffer containing 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 5 mM benzamidine, 0.5 mM PMSF, and 1 μg/mL pepstatin A was added supplemented with two tablets of protease inhibitor cocktail (Roche) and 0.5 mg/mL lysozyme (Sigma), 2 mL of Bugbuster (Novagen), and 6 mM iodoacetamide. The suspension was incubated at room temperature for 20−30 min, and then 2 μL of benzonase was added and the suspension was incubated for an additional 10 min. The suspension was centrifuged at 12 000 rpm for 30 min and the supernatant saved. The supernatant was allowed to mix with Ni-NTA agarose (Qiagen) prewashed with the phosphate buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl) for 1 h, poured into a column, and washed with 100 mL of phosphate buffer. The elutions were done with increasing concentrations of imidazole (10−250 mM) in phosphate buffer. The protein fractions eluted in the 250 mM imidazole concentration were pooled together and purified further in a size exclusion column (Superdex-75). The protein was concentrated to 10 mg/mL using Centricon for crystallization purpose. The protein concentration was measured by absorbance at 278 nm using a Perkin-Elmer spectrophotometer.

Expression and Purification of GST-Tagged VAMP. The pGEX-2T vector encoding VAMP2 (aa 1−115; human) with an N-terminal GST tag was kindly provided by J. Schmidt, USAMRIID. The GST-VAMP fusion protein was expressed in BL21 cells and purified using GST-resin as per user's manual (Novagen).

Enzymatic Activity of BoNT/F-LC. The enzymatic activity of BoNT/F-LC was assayed on its substrate VAMP. The assay was performed at 37 °C in 20 μL total volume (20 mM Hepes buffer, pH 7.4, 2 mM DTT, and 10 μM Zn acetate) containing 10 nM enzyme and 5 μM GST-VAMP. The hydrolytic curve (data not shown) for the light chain (2 nM) was determined by analyzing the cleavage of 5 μM substrate (GST-VAMP) at various time intervals (0, 1, 2, 3, 5, 10, 15, 30, 45, 60, 90, 120, and 180 min). The reaction was stopped by adding the 3x SDS−PAGE sample buffer containing 1 mM EDTA. The evaluation of cleavage was based on the appearance/disappearance and intensity of uncleaved substrate and product in 4−20% gradient Tris-glycine SDS−PAGE gels. Quantification of the digestion was performed by densitometry analysis of the cleaved and uncleaved fragments in Coomassie stained gel.

Screening of Mercurial Compounds as Inhibitors. The inhibitor assay was performed as above with two concentrations (500 and 1000 μM) of five mercurial compounds individually added to the assay buffer and incubated for 40 min at room temperature before adding the substrate. The reaction was stopped after 30 min by adding the 3x SDS−PAGE sample buffer containing 1 mM EDTA. The efficiency of cleavage was calculated as percentage of cleavage by the untreated enzyme as detailed above. The most promising compound was assayed again at finer grid points.

Crystallization and Data Collection. The initial crystallization screening was carried out by the sitting drop vapor diffusion method using the high-throughput crystal screen (Hampton Research) and TECAN crystallization robot. Microcrystals were obtained using a precipitant containing 0.1 M $CdCl_2$, 0.1 M Na-acetate pH 4.6, and 30% PEG 400 at room temperature in 2−3 days. Diffraction quality crystals were obtained by microseeding and varying the concentration of $CdCl_2$. Two morphologically distinct crystal forms, hexagonal (form I) and rod (form II) shaped, appeared under the same crystallization conditions. Preliminary analysis showed that both forms are in the space group *C*2. Form I crystals contain two molecules per asymmetric unit while form II crystals have one molecule per asymmetric unit. A

Table 1: Crystal Data and Refinement Statistics

form	Crystal Data Details
I	$a = 173.40$ Å, $b = 53.24$ Å, $c = 113.87$ Å, $\beta = 119.2°$; space group $C2$
II	$a = 63.41$ Å, $b = 79.46$ Å, $c = 89.66$ Å, $\beta = 110.0°$; space group $C2$

	form I	form II
wavelength (Å)	1.1	1.1
resolution (Å)	50−1.8	50−2.0
no. of reflections	69805	26560
I/σ	7.3	11.1
R-merge (outer shell)	0.09 (0.42)	0.08 (0.21)
redundancy	3.0	6.5
completeness (%)	82.9 (47.4)	93.7 (62.7)
Refinement Statistics		
resolution used (Å)	50−1.8	50−2.0
molecules/asymmetric unit	2	1
no. of reflections	64843	25768
completeness (%)	76.7	90.9
R-factor	0.24	0.23
R-free	0.27	0.28
no. of protein atoms	6338	3264
no. of heteroatoms	9	3
no. of water molecules	241	167
Ramachandran plot most favored region (additionally allowed) (%)	88.4 (10.4)	90 (9.2)
PDB id code	2A97	2A8A

Table 2: Interatomic Distances between Various Conserved Residues

	distance (Å)		distance (Å)
Zn−H227 NE2	2.00	Nu water−R365 NH2	7.58
Zn−H231 NE2	2.22	Nu water−H231 NE2	3.21
Zn−E266 OE1	2.20	E353 OE1−H227 ND1	3.44
Zn−E250 OE2	2.96	E353 OE2−R365 NH1	4.01
Zn−Nu water	2.13	E265 OE1−H231 ND1	2.82
Nu water−E228 OE1	3.74	E265 OE1−H234 ND1	2.63
Nu water−E228 OE2	4.42	Y368-OH−Zn	4.37
Nu water−Y368-OH	4.64	Y368-OH−E266 OE1	2.87
Nu water−E266 OE1	3.49	Y368-OH−E266 OE2	3.71
Nu water−E266 OE2	3.02	R365 NH2−Zn	6.62

self-rotation function calculation showed the presence of a noncrystallographic 2-fold symmetry in form I crystals (*18*).

Crystals were directly mounted on a nylon loop and flash frozen immediately by plunging into liquid nitrogen. X-ray diffraction data were collected at the SGX-CAT beam line of the APS for form I crystals and at the X29 beamline of the NSLS for form II crystals. All data sets were processed with DENZO and scaled and merged using SCALPACK (*19*). The data collection statistics are given in Table 1.

Structure Determination. Crystal structures of forms I and II were determined using the molecular replacement method with BoNT/B-LC (1F82) as search model (*18*). ARP/wARP was used to build the model (*20*). About 80% of the model was built automatically and the rest manually using "O" (*21*). Refinement was carried out using CNS (*22*). Fifteen C-terminal residues could not be located in the electron density maps in both crystal forms and were presumed to be disordered. In addition, form I crystals lacked electron density for the loop regions Asp205−Thr214 and Glu249−Ile262 in both monomers while form II crystals lacked only for Asp205−Thr214. These loops have a tendency to be disordered in all LCs (*14*, *23*, *24*). Refinement details are given in Table 1.

Modeling of Substrate on BoNT/F-LC. Two crystal structures of botulinum neurotoxin catalytic domain−substrate complex are available (*13*, *14*). The two LCs have different substrates, and the mode of binding is also different. To model BoNT/F-LC with its substrate, both models were tried. First BoNT/B-LC with VAMP peptide was aligned with BoNT/F-LC by least-squares fit. In this process the disordered loops were not included. Similar alignment was done with BoNT/A (mutant):SNAP-25 peptide. In the latter case, the SNAP-25 peptide was converted into VAMP peptide via mutation in "O".

RESULTS AND DISCUSSION

Cloning, Expression, Purification, and Catalytic Activity of BoNT/F-LC. The BoNT/F catalytic domain gene from pFB4 plasmid has been recloned into a pET-28b vector with C-terminal 6xHis tag (pET-F-LC). Using pET-F-LC BoNT/F-LC has been overexpressed yielding >25 mg/L cell culture. The solubility of the protein was extremely good, and we could concentrate it to >10 mg/mL in HEPES buffer. The protein is stable when stored in the same buffer between −80 and −20 °C.

VAMP2 (aa 1−115) with the N-terminal GST tag was used for assaying enzymatic activity. The GST-tag (25 kDa) does not interfere with the catalytic activity (*12*). Incubation of GST-VAMP (∼38 kDa) with the BoNT/F-LC led to cleavage of the Gln58−Lys59 peptide bond and produced ∼32 and ∼6 kDa fragments. LC, 10 nM, was able to completely cleave 5 μM VAMP in 30 min of incubation at 37 °C (data not shown).

Description of BoNT/F-LC Structures. BoNT/F-LC has the fold typical of the available botulinum toxin catalytic domains. The protein has one zinc ion per molecule. Form I and form II crystal structures are very similar. The rmsd between the two monomers in form I is 0.40 Å, and that between form II molecule and either one of form I monomers is 0.43 Å. The loop region Asp205−Thr214 is disordered and not modeled in both structures. Loop Glu249−Ile262 disordered and hence not modeled in form I is ordered in the form II crystal. The C-terminal residues (Val422−Lys439) are not modeled due to poor electron density in both structures. The absence of the C-terminal residues and the disordered region (Glu249−Ile262) is not due to autolysis as shown by SDS−PAGE analysis on crystals of both forms (data not shown) (*24*, *25*). In form I crystals the two molecules (A and B) in the asymmetric unit are associated via a noncrystallographic 2-fold symmetry. One β-strand from each molecule is held together by a cadmium ion, which is interacting with OD1 and OD2 of Asp383 of both monomers with distances of 2.27, 2.28, 2.72, and 2.52 Å. Other interactions are primarily through loop regions. Residues 7−13 of molecule A interact with residues 394−396 of molecule B and vice versa. The role of cadmium ions is probably to stabilize the packing and thus help in crystallization. Though the molecule exists as a dimer in the crystal structure, the buried surface area of 1680 Å^2 (about 5% of the total area) on dimerization indicates that it is not a biological dimer. This notion is supported by results of size exclusion chromatography experiments where the peak

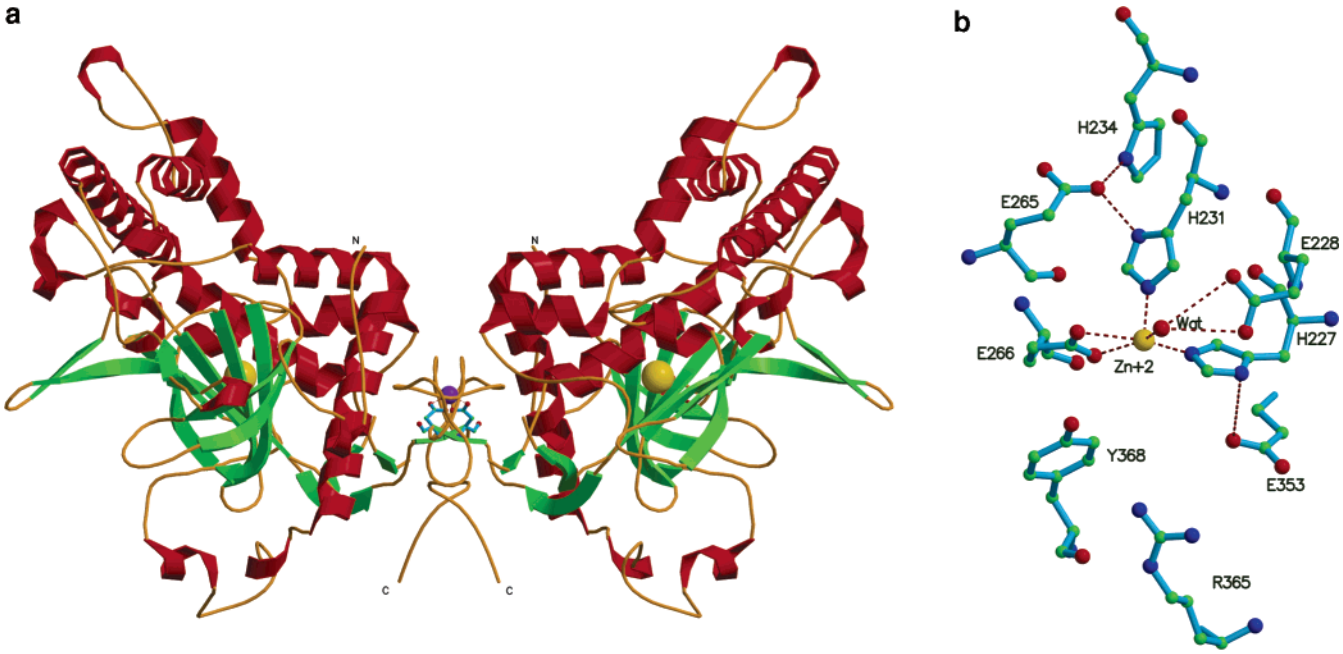

FIGURE 1: (a) Ribbon representation of the dimer formed via crystallographic 2-fold in form II crystals. The dimerization is similar in form I, but via a noncrystallographic 2-fold. The dimer formation is aided by a Cd ion (shown in magenta) coordinated by the carboxylate side chain of Asp383, shown in ball-and-stick model. The active-site zinc is shown in yellow. The dimerization may be an artifact of crystal packing. This figure was produced with MOLSCRIPT (*41*). (b) The active site of BoNT/F-LC. The zinc ion is coordinated by His227, His231, Glu266, and a nucleophilic water (wat) which is hydrogen bonded to Glu228. The interatomic distances are given in Table 2.

due to the protein appeared at 52 kDa when eluted at a concentration of 1−2 mg/mL. In both the molecules the active sites are facing outside the dimeric interface and exposed to solvent (Figure 1a). The dimeric association is different in all the clostridial neurotoxin light chain structures reported so far (*23*, *24*, *26*). Whether or not the dimerization has any biological significance is not yet clear. In form II crystals similar dimerization is achieved via the crystallographic 2-fold, also with a cadmium ion bridging the molecules. Since all monomers and their interatomic distances in both crystal forms agree within experimental errors, the discussion hereafter pertains to form II crystals as the model is more complete (Table 1).

The BoNT/F-LC structure is similar to BoNT/A, -B, and -E and TeNT LCs with respect to the core conformation involving the helices and strands though large variations exist in their loop conformations. Of the full-length catalytic domain structures available so far (BoNT/F and -E and TeNT), the C-terminal region is visible only in BoNT/E-LC (*23*, *26*). In BoNT/F-LC, the active site is well formed and has interactions typical of zinc-binding motifs as in other clostridial neurotoxins. The active site zinc is coordinated by His227 NE2, His231 NE2, Glu266 OE1, and a nucleophilic water molecule (Figure 1b). The latter is kept in its position via Glu228.

Counterparts of the remaining depicted active site residues are present in analogous arrangement in all clostridial neurotoxins crystallized to date. This argues for the functional similarity of these residues. The function of several active sites residues has been studied by mutational analysis (*16*, *27*−*30*). Thus, Glu265 appears to be important for maintaining the orientation and protonation state of His231. Glu353 likely fulfills the same function on His227. Arg365 probably helps in stabilization of the P1′ carbonyl oxygen of the scissile bond of the substrate to give proper orientation. Tyr368 may play a major role in transition state stabilization as mutation of the equivalent residue in BoNT/E (Tyr350) to alanine inactivated the enzyme (*16*). However, moderate differences in phenotypes have been observed for mutations of analogous residues in different clostridial neurotoxins.

Putative Substrate Binding Sites for BoNT/F-LC. Two crystal structures of BoNT LCs with their substrates are available (*13*, *14*). The two toxins have different substrates and could hence represent the mode of binding of the particular substrate to the catalytic domain. In both models, part of the substrate runs close to the β strand formed by residues in the region 160−170 (BoNT/A). In all LC structures and the holotoxins, this region forms a β strand. In our modeling the transition state by comparing BoNT/B with a sulfate ion with thermolysin, we had pointed out that residues in this strand might interact with the substrate (*31*). A similar conclusion has been drawn also from mutational studies in BoNT/E-LC (*16*). However, while the substrate runs antiparallel to this strand to form a β sheet in the BoNT/A:SNAP-25 complex, in the BoNT/B:VAMP complex, it runs close to the β strand but is in the same direction as the strand. This difference could be because of the difference in substrate type and/or the serotype. This raises an interesting possibility that all LCs which have VAMP as substrate will bind similarly to BoNT/B and those with SNAP-25 as substrate similarly to BoNT/A. But a reversal of direction is also a possibility for the same substrate when it binds to a different serotype. To model the substrate binding to BoNT/F-LC we took both possibilities into account. Accordingly, the VAMP peptide was modeled similarly to SNAP-25 in the BoNT/A:SNAP-25 peptide complex (Type 1). It was also modeled by tracing the chain in the reverse direction (Type 2). Both models were analyzed with respect to biochemical and mutational studies (*12*, *32*).

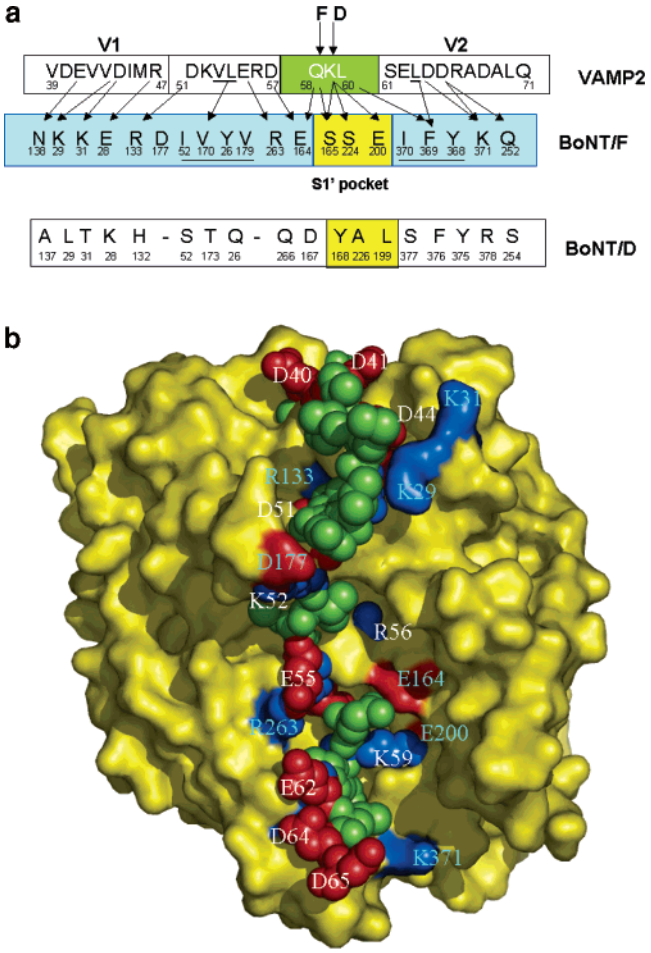

FIGURE 2: (a) Proposed VAMP−BoNT/F interactions. The scissile peptide bonds for BoNT/F and -D are shown with arrow marks. The residues involved in S1′ subsite are shown in yellow. The three residues forming the S1′ subsite will form hydrogen bonds with the P1′ residue, Lys59. The SSR motifs, V1 and V2, are also shown. The putative interactions between VAMP and BoNT/F are shown with arrow marks. Most of the interactions are between residues with complementary charges. Residues involved in the putative hydrophobic interactions are underlined. Strong interactions of V1 and V2 residues with the enzyme explain the need of both motifs for optimal catalytic activity of BoNT/F. BoNT/D residues corresponding to interacting residues of BoNT/F are also shown. Among the interacting residues only Tyr368 and Phe369 (BoNT/F) are conserved in BoNT/D and -F. The putative S1′ subsite for BoNT/D is highlighted in yellow. To allow for the hydrophobic interactions between P1′ and the S1′ subsite, VAMP will have to slide up by one residue. (b) BoNT/F-LC:VAMP complex model. BoNT/F-LC is shown as molecular surface in yellow with charged residues interacting with VAMP in red (negative) and blue (positive). VAMP (39−65) is shown in sphere model. Charged residues are shown in red and blue, while the rest are shown in green. The complementary charges between VAMP and BoNT/F-LC are evident. Charged residues are labeled in white for VAMP and in cyan for BoNT/F-LC. Figures 2b, 3, 4, and 5 were prepared with PyMOL (*42*).

The cleavage site in VAMP2 by BoNT/F-LC lies between the two SSR sites, V1 and V2 (Figure 2a) (*12*). The cleavage of VAMP2 is susceptible to mutation of residues in both SSRs, though more so for V1 residues than V2. When Asp40, Glu41, and Asp44 of V1 were mutated, the activity was almost completely lost. When Asp64, Asp65, and Asp68 of V2 were mutated, the cleavage decreased by 40−60% (*12*). In type 1 alignment, Asp44 has interactions with two positively charged residues (Lys29 and Lys31 of the LC) and could form a salt bridge to stabilize binding. According to the model, i.e., basing on a similar progression of the substrate along the binding channel as found in BoNT/A: SNAP-25, while Glu41 could make a potential hydrogen bond with the side chain of Asn138 of BoNT/F, Asp40 does not appear to form interactions with the LC. In addition, Arg47 of VAMP is near Glu28 of BoNT/F. Asp51, Lys52, Glu55, and Arg56 of VAMP interact with Arg133, Asp177, Arg263, and Glu164 of BoNT/F, respectively. Hydrophobic residues Val53 and Leu54 sit in a hydrophobic pocket formed by Ile52, Val170, Tyr26, and Val179 of BoNT/F. On the C-terminal side of the cleavage site, Asp64 and Asp65 of V2 interact with Lys371 and Gln252 of BoNT/F (Figure 2b). All these interactions indicate strong binding of the substrate to BoNT/F and largely support the results from mutagenesis studies. Together, these interactions could be considered the exosite interactions of BoNT/F. These LC residues are not conserved in BoNT/D (Figure 2a), and this could be the reason for different effects on the activity due to mutation of SSR residues even though the two cleave adjacent peptide bonds (*12*, *33*, *34*). On the other hand, in type 2 alignment, residues with the same kind of charge clash between the substrate and LC. Accordingly, we conclude that VAMP binds BoNT/F in a direction similar to SNAP-25 to BoNT/A. It remains to be shown whether the different direction of substrate binding in BoNT/B is unique or applies to other LCs as well. In discussing the need for the presence of V1 for TeNT activity, it was assumed that the VAMP will bind similarly to BoNT/B (*26*). However, the reverse substrate orientation can at present not be excluded. At any rate, the different specificity of TeNT and BoNTs is due to variation in the charge distribution on the surface of the LCs which allows different segments of the same substrate to bind presenting the specific cleavage site at the active site. Figure 3 provides a comparison of the charge distribution on BoNT/F-, BoNT/B-, and TeNT-LCs, all cleaving the same substrate.

S1′ Pocket of BoNT/F-LC. The requirements on P2, P1, P1′, and P2′ residues with respect to cleavage by clostridial neurotoxins have been studied extensively (*32*, *35*−*37*). While variation of P1 is irrelevant to cleavage (except for BoNT/C (*37*)), P1′ is crucial. The S1′ subsite has been identified in BoNT/E (*16*). The P1′ residue, Ile181, in BoNT/E occupies a hydrophobic site formed by Thr159, Phe191, Gln203, Leu207, Thr208, and Tyr356. A similar modeling with BoNT/F-LC and thermolysin proposes the S1′ subsite of BoNT/F. In this case, the P1′ residue, Lys59, interacts with Ser165, Glu200, and Ser224 (Figure 4). In BoNT/A these residues correspond to Phe162, Thr192, and Thr219 and have the potential to interact with P1′ (Arg198 of SNAP25) with some rearrangement of the side chain orientations. A proper rotamer position of Lys59 of VAMP will allow its side chain to interact with the side chains of these three residues. Changing Lys59 of VAMP to either Arg or Ala rendered VAMP cleavage resistant (*32*). Though Arg is positively charged as Lys, the side chain is too big to be accommodated in this subsite and is detrimental to the activity. As for Ala, the lack of interactions with the subsite residues will produce the same effect. In our substrate complex model, P1 is exposed to the solvent and has minimal interactions with the LC. This may be the case for other complexes and might explain why the type of amino acid at P1 is not crucial for cleavage. For BoNT/F, P2 and P2′ of the substrate are crucial for cleavage. P2 is Asp57, and its

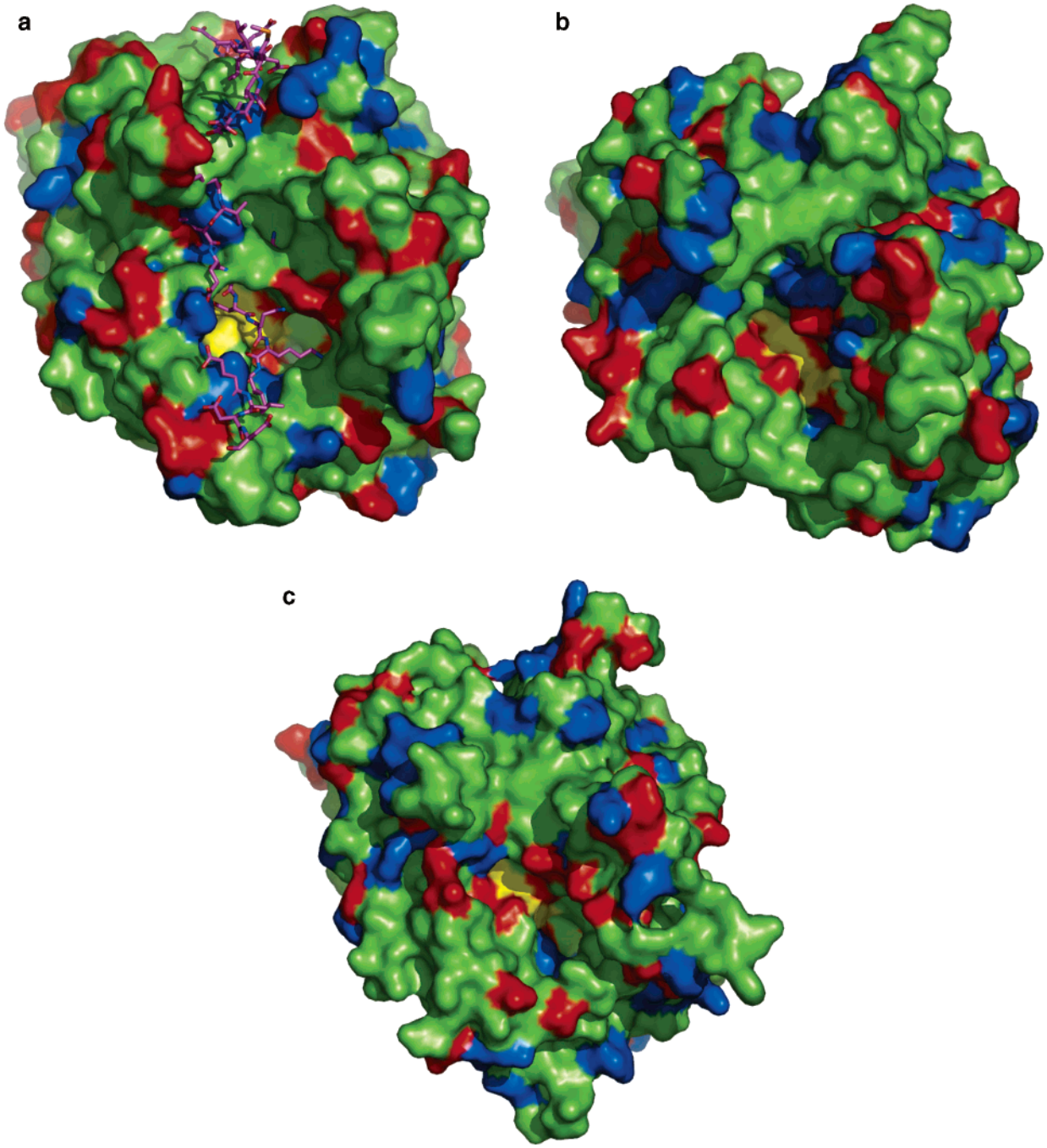

FIGURE 3: Comparison of the surface charges on (a) BoNT/F-LC, (b) TeNT-LC (PDB id 1YVG), and (c) BoNT/B-LC (PDB id 1F83). The three molecules are in the same orientation, and VAMP modeled on BoNT/F-LC is shown in stick model. Residues with positive and negative charges are shown in red and blue, respectively, and the active site region is shown in yellow. The variation in charge distribution explains the specificity for different cleavage sites. (Though BoNT/B and TeNT cleave the same peptide bond of VAMP, the optimal length of VAMP for activity is different. While TeNT requires both V1 and V2 SSR sites, BoNT/B requires only V2.)

side chain interacts with Arg263 of the LC probably via a salt bridge. Changing Asp57 to either Asn or Ala or Glu affected the activity. In the case of Glu the side chain is one bond length larger and might have steric clash. P2′ (Leu60) occupies a pocket formed by bulky hydrophobic residues Tyr368, Phe369, and Ile370. Change of the P2′ residue to Ala compromises the hydrophobic interactions and consequently affects the rate of substrate cleavage.

Mercury Compounds as Inhibitors of Botulinum Neurotoxins. The role of mercurial compounds on the activity of BoNTs has been investigated with respect to the interchain disulfide bridge connecting the heavy and light chains (*38*). However, the same compound had different effects on various serotypes. Ahmed et al. also have shown that similar mercurial compounds completely abolish the activity of BoNT/A-LC at 10−50 μM concentrations (*39*), and this is attributed to the modification of the thiol group of Cys164 which is in the vicinity of the active site. We chose to investigate the effect of mercury compounds on BoNT/F since there are two adjacent cysteines (Cys166 and Cys167) in the β strand that would form a β sheet with VAMP according to our enzyme−substrate complex model described in earlier sections. Cys167 of BoNT/F and Cys164 of BoNT/A are conserved. In addition, Cys364 is close to Arg365 and

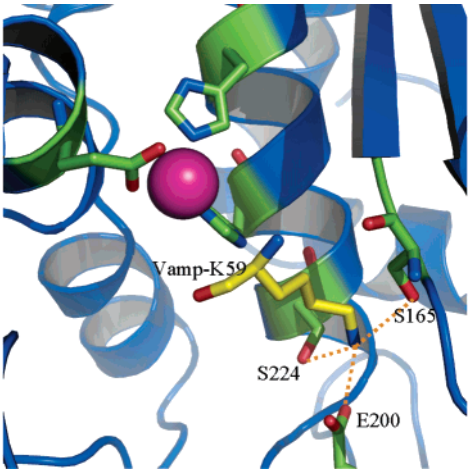

FIGURE 4: Close-up view of the active site and S1′ subsite. The active site zinc is shown in magenta, and the residues involved in the active and S1′ subsite are shown in stick model. The P1′ residue Lys59 is shown in yellow. The three potential hydrogen bonds are shown as dashed lines.

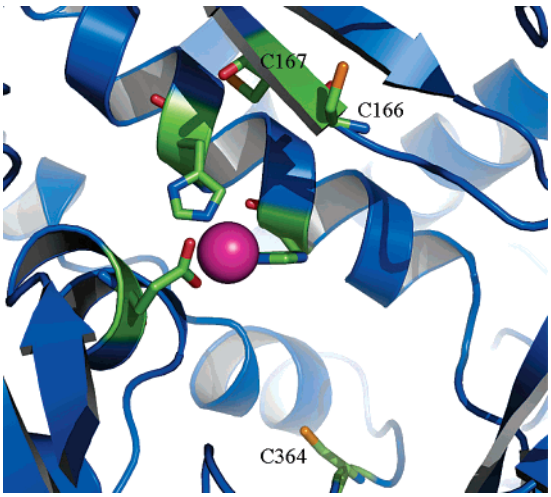

FIGURE 5: Cysteines near the active site. The three cysteines (166, 167, and 364) within 11 Å from the active site zinc are shown. Cys364 is close to Arg365 and Tyr368 which stabilize the transition state. The fourth cysteine (Cys429) was not modeled in the structure.

Tyr368 (Figure 5). The three mercury compounds we tried on BoNT/F-LC had varying effects in inhibiting the catalytic activity. As has been pointed out (*38*), this could be because of the modification of thiol groups or histidines near the active site or both. Future crystallographic investigations may help in understanding the role of mercury compounds in inhibiting the catalytic activity.

The Rationale for Inhibitor Development. BoNT/A, -E, and -C proteolytically cleave SNAP-25 while BoNT/B, -D, -F, and -G cleave VAMP. Small molecule inhibitors such as zinc chelators are relatively inefficient in blocking the activity of botulinum neurotoxins. This may be because either they do not bind tightly at the active site or they could be easily displaced by the substrate. In addition, small molecule zinc chelators could have undesirable side effects since they might inhibit other zinc proteins, with similar zinc binding motifs, essential for metabolism (*40*). In view of this, it would be advantageous to design inhibitors which would bind tightly covering a larger area of the enzyme. Mutations of residues in the light chain farther away from the active site either abolish or reduce the catalytic activity. The substrates are large peptides, and the enzyme−substrate models indicate strong binding of substrate segments in more than one site. This fact could be taken advantage of in developing peptidomimics. The competitive removal of the small substrate peptides by the larger substrate can be overcome by using inhibitors that will bind in more than one site. Another approach would be to block the S1 and S1′ subsites and exosites with two or three small molecules connected by linkers. Such molecules have been tried successfully in other cases before.

A further approach is to manipulate the binding of the substrate to have better affinity than the actual substrate by modifying the charges in the substrate residues thereby increasing the number of interaction points. This is only possible when we have the complex structures available for each of the serotypes with their substrates. However, the idea will be to modify the substrate in such a way that it has the minimal length requirement but with better affinity than the actual substrate and also not being cleaved by the enzyme.

Scientists have attempted to crystallize native enzyme with substrates without success and that prompted the complex formation with a mutant (*13*, *23*). It points to the fact that it is difficult to obtain a complex of the enzyme with cleaved product(s). If there are many interactions between the substrate and the enzyme, it is not clear why cleavage of one bond can let the product move away from the enzyme. Either the cleaved products could change their conformation and lose interactions with the enzyme or uncleaved substrates compete with cleaved products to displace them. In view of this, if an effective inhibitor is to be developed, it needs to have stronger interactions with the enzyme than substrate but not necessarily similar to what is involved in substrate recognition. An understanding of the molecular mechanism of the changes leading to the separation of the cleaved substrate from the enzyme would be helpful in inhibitor design.

ACKNOWLEDGMENT

We thank SGX-CAT (APS) and NSLS for providing beam time for X-ray data collection and J. Romeo for technical assistance. We also thank Dr. Kumaran for his help in data collection at the APS.

REFERENCES

1. Schiavo, G., Rossetto, O., and Montecucco, C. (1994) Clostridial neurotoxins as tools to investigate the molecular events of neurotransmitter release, *Semin. Cell Biol. 5*, 221−229.
2. Montecucco, C., Papini, E., and Schiavo, G. (1994) Bacterial protein toxins penetrate cells via a four-step mechanism, *FEBS Lett. 346*, 92−98.
3. Harvey, S. M., Sturgeon, J., and Dassey, D. E. (2002) Botulism due to *Clostridium baratii* type F toxin, *J. Clin. Micrbiol. 40*, 2260−2262.
4. McCroskey, L. M., Hatheway, C. L., Woodruff, B. A., Greenberg, J. A., and Jurgenson, P. (1991) Type F botulism due to neurotoxigenic *Clostridium baratii* from an unknown source in an adult, *J. Clin. Microbiol. 29*, 2618−2620.
5. Paisley, J. W., Lauer, B. A., and Arnon, S. S. (1995) A second case of infant botulism type F caused by *Clostridium baratii*, *Pediatr. Infect. Dis. J. 14*, 912−914.
6. Humeau, Y., Dousseau, F., Grant, N. J., and Poulain, B. (2000) How botulism and tetanus neurotoxins block neurotransmitter release, *Biochimie 82*, 427−446.
7. Schiavo, G., Matteoli, M., and Montecucco, C. (2000) Neurotoxins affecting neuroexocytosis, *Physiol. Rev. 80*, 717−766.

8. Bennett, M. K., and Scheller, R. H. (1994) Molecular correlates of synaptic vesicle docking and fusion, *Curr. Opin. Neurobiol. 4*, 324−329.
9. Rossetto, O., Gorza, L., Schiavo, G., Schiavo, N., Scheller, R. H., and Montecucco, C. (1996) VAMP/synaptobrevin isoforms 1 and 2 are widely and differentially expressed in nonneuronal tissues, *J. Cell Biol. 132*, 167−179.
10. Rossetto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, F., Lozzi, L., and Shone, C. C. (1994) SNARE motif and neurotoxins, *Nature 372*, 415−416.
11. Pellizzari, R., Rossetto, O., Lozzi, L., Giovedí, S., Johnson, E., Shone, C. C., and Montecucco, C. (1996) Structural determinants of the specificity for synaptic vesicle-associated membrane protein/synaptobrevin of tetanus and botulinum type B and G neurotoxins, *J. Biol. Chem. 271*, 20353−20358.
12. Pellizzari, R., Mason, S., Shone, C. C., and Montecucco, C. (1997) The interaction of synaptic vesicle-associated membrane protein/synaptobrevin with botulinum neurotoxins D and F, *FEBS Lett. 409*, 339−342.
13. Breidenbach, M. A., and Brunger, A. (2004) Substrate recognition strategy for botulinum neurotoxin serotype A, *Nature 432*, 925−929.
14. Hanson, M. A., and Stevens, R. C. (2000) Cocrystal structure of synaptobrevin-ll bound to botulinum neurotoxin type B at 2.0 Å resolution, *Nat. Struct. Biol. 7*, 687−692.
15. Matthews, B. W. (1988) Structural basis of the action of thermolysin and related zinc peptides, *Acc. Chem. Res. 21*, 333−340.
16. Agarwal, R., Binz, T., and Swaminathan, S. (2005) Analysis of active site residues of botulinum neurotoxin E by mutational, functional and structural studies: Glu335Gln is an apoenzyme, *Biochemistry 44*, 8291−8302.
17. Agarwal, R., Eswaramoorthy, S., Kumaran, D., Dunn, J. J., and Swaminathan, S. (2004) Cloning, high level expression, purification and crystallization of the full length Clostridium botulinum neurotoxin type E light chain, *Protein Expression Purif. 34*, 95−102.
18. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron density map, *Acta Crystallogr. D57*, 1451−1456.
19. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
20. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
21. Jones, T. A., Zou, J., Cowtan, S., and Kjeldgaard, M. (1991) Improved methods in building protein models in electron density map and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
22. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszwewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Somonsom, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
23. Agarwal, R., Eswaramoorthy, S., Kumaran, D., Binz, T., and Swaminathan, S. (2004) Structural analysis of botulinum neurotoxin type E catalytic domain and its mutant Glu212→Gln reveals the pivotal role of the Glu212 carboxylate in the catalytic pathway, *Biochemistry 43*, 6637−6644.
24. Segelke, B., Knapp, M., Kadkhodayan, S., Balhorn, R., and Rupp, B. (2004) Crystal structure of *Clostridium botulinum* neurotoxin protease in a product-bound state: Evidence for noncanonical zinc protease activity, *Proc. Natl. Acad. Sci. U.S.A. 101*, 6888−6893.
25. Ahmed, S. A., Byrne, M. P., Jensen, M., Hines, H. B., Brueggemann, E., and Smith, L. A. (2001) Enzymatic autocatalysis of botulinum A neurotoxin light chain, *J. Protein Chem. 20*, 221−231.
26. Rao, K. N., Kumaran, D., Binz, T., and Swaminathan, S. (2005) Structural studies on the catalytic domain of clostridial tetanus toxin, *Toxicon 45*, 929−939.
27. Binz, T., Bade, S., Rummel, A., Kollewe, A., and Alves, J. (2002) Arg^{362} and Tyr^{365} of the botulinum neurotoxin type A light chain are involved in transition state stabilization, *Biochemistry 41*, 1717−1723.
28. Rossetto, O., Caccin, P., Rigoni, M., Tonello, F., Bortoletto, N., Stevens, R. C., and Montecucco., C. (2001) Active-site mutagenesis of tetanus neurotoxin implicates TYR-375 and GLU-271 in metalloproteolytic activity, *Toxicon 39*, 115−1159.
29. Rigoni, M., Caccin, P., Johnson, E. A., Montecucco, C., and Rossetto, O. (2001) Site-directed mutagenesis identifies active-site residues of the light chain of botulinum neurotoxin type A, *Biochem. Biophys. Res. Commun. 288*, 1231−1237.
30. Li, L., Binz, T., Niemann, H., and Singh, B. R. (2000) Probing the mechanistic role of glutamate residues in the zinc-binding motif of type A botulinum neurotoxin light chain, *Biochemistry 39*, 2399−2405.
31. Swaminathan, S., Eswaramoorthy, S., and Kumaran, D. (2004) Structure and enzymatic activity of botulinum neurotoxins, *Movement Disord. 19* (Suppl. 8), S17−S22.
32. Schmidt, J. J., and Stafford, R. G. (2005) Botulinum neurotoxin serotype F: identification of substrate recognition requirements and development of inhibitors with low nanomolar affinity, *Biochemistry 44*, 4067−4073.
33. Lacy, D. B., and Stevens, R. C. (1999) Sequence homology and structural analysis of clostridial neurotoxins, *J. Mol. Biol. 291*, 1091−1104.
34. Binz, T., Kurazono, H., Popoff, M. R., Eklund, M. W., Sakaguchi, G., Kozaki, S., Krieglstein, K., Henschen, A., Gill, D. M., and Niemann, H. (1990) Nucleotide sequence of the gene encoding Clostridium botulinum neurotoxin type D, *Nucleic Acids Res. 18*, 5556.
35. Shone, V. C., and Roberts, A. K. (1994) Peptide substrate specificity and properties of the zinc-endopeptidase activity of botulinum type B neurotoxin, *Eur. J. Biochem. 225*, 263−270.
36. Schmidt, J. J., and Bostian, K. A. (1997) Endoproteinase activity of typeA botulinum neurotoxin:substrate requirements and activation by serumalbumin, *J. Protein Chem. 16*, 19−26.
37. Vaidyanathan, V. V., Yoshino, K., Jahnz, M., Dorries, C., Bade, S., Nauenburg, S., Niemann, H., and Binz, T. (1999) Proteolysis of SNAP-25 isoforms by botulinum neurotoxin types A, C, and E: Domains and amino acid residues controlling the formation of enzyme-substrate complexes and cleavage, *J. Neurochem. 72*, 327−337.
38. Simpson, L. L., Maksymowych, A. B., Park, J.-B., and Bora, R. S. (2004) The role of the interchain disulfide bond in governing the pharmacological actions of botulinum toxin, *J. Pharmacol. Exp. Ther. 308*, 857−864.
39. Ahmed, S. A., and Smith, L. A. (2000) Light chain of botulinum A neurotoxin expressed as an inclusion body from a synthetic gene is catalytically and functionally active, *J. Protein Chem. 19*, 475−487.
40. Fuchs, O., Babusiak, M., Vyoral, D., and Petrak, J. (2003) Role of zinc in eukaryotic cells, zinc transporters and zinc-containing proteins, *Sb. Lek. 104*, 157−170.
41. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, *J. Appl. Crystallogr. 24*, 946−950.
42. DeLano, W. L. (2002) in *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA.

BI0510072