2.3 Å Crystal Structure of Tetanus Neurotoxin Light Chain[†,‡]

Mark A. Breidenbach[§] and Axel T. Brunger*[,||]

Howard Hughes Medical Institute, Departments of Molecular and Cellular Physiology and of Neurology and Neurological Sciences, and Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, California 94305

Received February 13, 2005; Revised Manuscript Received March 25, 2005

ABSTRACT: TeNT is the causative agent of the neuroparalytic disease tetanus. A key component of TeNT is its light chain, a Zn^{2+} endopeptidase that targets SNAREs. Recent structural studies of closely related BoNT endopeptidases indicate that substrate-binding exosites remote from a conserved active site are the primary determinants of substrate specificity. Here we report the 2.3 Å X-ray crystal structure of TeNT-LC, determined by combined molecular replacement and MAD phasing. As expected, the overall structure of TeNT-LC is similar to the other known CNT light chain structures, including a conserved thermolysin-like core inserted between structurally distinct amino- and carboxy-terminal regions. Differences between TeNT-LC and the other CNT light chains are mainly limited to surface features such as unique electrostatic potential profiles. An analysis of surface residue conservation reveals a pattern of relatively high variability matching the path of substrate binding around BoNT/A, possibly serving to accommodate the variations in different SNARE targets of the CNT group.

TeNT[1] is produced by the anaerobic bacterium *Clostridium tetani* and is considered a significant health threat to both humans and animals (*1*−*4*). Neuronal intoxication by any member of the CNT group (including TeNT and BoNT serotypes A−G) substantially reduces the probability of quantal neurotransmitter release at presynaptic nerve termini by inhibition of Ca^{2+}-dependent, K^+-evoked synaptic vesicle exocytosis (*5*, *6*). The CNTs share a common dual-chain architecture: a 100 kDa heavy chain is responsible for neurospecific cell surface binding and facilitates translocation of a 50 kDa light chain into nerve cytosol (*7*−*9*). Impaired exocytosis is attributed to the proteolytic activity of the light chains (*10*−*13*). These endopeptidases exhibit exquisite specificity to hydrolyze the essential SNARE components of the neuronal fusion machinery at distinct target sites (*14*−*16*).

The SNARE proteins synaptobrevin, syntaxin, and synaptosomal-associated protein of 25 kDa (SNAP-25) are requisite for neurotransmitter release; their assembly into a low-energy ternary core complex leads to fusion of neurotransmitter-laden vesicles with the active zone in the presynaptic membrane (*17*, *18*). Proteolysis of SNAREs attenuates Ca^{2+}-dependent exocytosis either by destabilization of the ternary complex or by interference with recruitment of a Ca^{2+} sensor (*19*, *20*). TeNT and BoNT light chains differ in terms of the particular SNARE target and cleavage site, but their activity has a similar inhibitory effect on exocytosis. Notably, CNTs will block secretion from non-neuronal cells if artificially introduced into their cytosol, suggesting that the CNT light chains are general SNARE proteases, while neurospecificity is a function of the toxin heavy chains (*21*). Ultimately, the clinical manifestation of CNT intoxication depends on the subset of neurons to which the light chains are delivered. This is best illustrated by a comparison of TeNT and BoNT/B; their light chains share 52% sequence identity and hydrolyze the same peptide bond of the same SNARE target, synaptobrevin. While BoNT/B binds and acts at neuromuscular junctions, TeNT is transported away from the peripheral nervous system and instead acts at inhibitory synapses due to properties of its heavy chain (*22*). Thus, TeNT intoxication results in spastic paralysis while BoNT/B and other BoNTs cause flaccid paralysis.

The structural basis of CNT light chain SNARE specificity is only now becoming clear. To date, apo crystal structures have been determined for the light chains of BoNTs A, B, and E (*23*−*25*), and an enzyme−substrate complex has been determined for BoNT/A (*26*). Additionally, full-length structures containing both heavy and light chains have been determined for BoNT/A and BoNT/B (*27*, *28*). The substrate reported in the crystal structure of BoNT/B (*24*) has been shown to be absent (*26*, *29*). We found that BoNT/A-LC depends on a extensive array of substrate-binding exosites located around the endopeptidase's surface for proper substrate recognition (*26*). Specifically, SNAP-25 binding to BoNT/A-LC follows a path similar to the toxin's "belt" region which links the heavy and light chains prior to translocation of the light chain (*26*, *27*, *30*). The means by which the different light chains discriminate between their different SNARE targets remain unclear, however. Also, the

[†] This work was funded by National Institute of General Medical Sciences Grant 1-RO1-MH63105-01 to A.T.B.

[‡] Protein Data Bank coordinate accession code 1Z7H.

* Corresponding author. Phone: 650-736-1031. Fax: 650-736-1961. E-mail: brunger@stanford.edu.

[§] Department of Molecular and Cellular Physiology.

[||] Howard Hughes Medical Institute, Departments of Molecular and Cellular Physiology and of Neurology and Neurological Sciences, and Stanford Synchrotron Radiation Laboratory.

[1] Abbreviations: TeNT, tetanus neurotoxin; TeNT-LC, tetanus neurotoxin light chain; BoNT, botulinum neurotoxin; BoNT/A, BoNT/B, etc., botulinum neurotoxin serotypes A−G; SNAREs, soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors; MAD, multiwavelength anomalous dispersion; CNT, clostridial neurotoxin; MR, molecular replacement.

10.1021/bi050262j CCC: $30.25
Published on Web 04/27/2005

extensive substrate interface observed in BoNT/A raises the question of how *Clostridia* acquired functional SNARE-binding exosite arrays by conventional Darwinian vertical inheritance. In a continuing effort to address these questions, we describe here the crystal structure of TeNT-LC. We obtained insights for the basis for discrimination between different SNARE targets among CNTs by analysis of regions implicated in substrate recognition.

MATERIALS AND METHODS

Protein Purification. Plasmid DNA pOG-7 encoding TeNT-LC was kindly provided by Richard Scheller (Stanford University). Residues 1−442 were amplified from the plasmid template by polymerase chain reaction (PCR) with the following oligonucleotides (underlining indicates restriction sites): (a) 5′-GATATCCATATGATGCCGATCAC-CATCAACAACTTCCGTTACTCCGACCC-3′ corresponding to the 5′-end and (b) 5′-GAGCTCGGATCCTTAGATGA-TTTTTTTGCACAGACCGATCAGTTTGGAAACCAGA-CCGGAA-3′ corresponding to the 3′-end of TeNT-LC. The resulting PCR product was digested with *Nde*I and *Bam*HI (New England Biolabs) and inserted into the pET-28a expression vector (Novagen).

Recombinant TeNT-LC was expressed as an N-terminal hexa-His-tagged protein in Rosetta cells (Novagen) at 30 °C in a BIOFLO 3000 fermentor (New Brunswick, NJ) using ECPM1 medium in the presence of 50 mg/mL kanamycin sulfate and 30 μg/mL chloramphenicol. Following a 3 h induction period, cells were harvested by centrifugation. Cells were immediately flash-frozen in liquid nitrogen and stored at −80 °C.

Frozen cell pellets were resuspended in a 1:5 (w/v) ratio of lysis buffer (20 mM HEPES, pH 7.4, 100 mM NaCl, 1 mM DTT). Protease inhibitors added to the lysis buffer included 1 mM PMSF and EDTA-free complete protease inhibitor cocktail tablets (Roche). Cells were lysed via two passes through an M-110EH microfluidizer processor (Microfluidics) at 18000 psi. Insoluble debris was removed from the lysate by centrifugation in a JA-20 rotor (Beckman) at 19500 rpm for 45 min. Each 200 mL portion of clarified lysate was incubated with 25 mL of Ni-NTA agarose resin (Qiagen) in batch for 4 h at 4 °C before being washed with 10 column volumes of washing buffer (20 mM HEPES, pH 7.4, 500 mM NaCl, 1 mM DTT). Bound protein was eluted with 25 mL of elution buffer (20 mM HEPES, pH 7.4, 200 mM imidazole, 100 mM NaCl, 1 mM DTT). His tags were removed by adding bovine α-thrombin (Haematologic Technologies) directly to the elutant, which was then dialyzed overnight against 20 mM HEPES, pH 7.4, 100 NaCl, and 1 mM DTT. The thrombin-cleaved product, which retained five N-terminal residues as a cloning artifact, was further purified (>95%) by anion-exchange (Amersham-Pharmacia monoQ resin) chromatography. Purified TeNT-LC was quantified by UV/vis spectroscopy at 280 nm, based on its theoretical extinction coefficient in denaturing conditions. Proteolytic activity of the TeNT-LC was qualitatively verified by incubation with synaptobrevin residues 1−96; cleavage products were observed on an Omni-Flex MALDI mass spectrometer (Bruker).

Crystallization and Structure Determination. TeNT-LC was concentrated to 20 mg mL^{-1} in a solution of 50 mM HEPES, pH 7.4, 100 mM sodium chloride, and 1 mM DTT. Crystals of TeNT-LC were grown by hanging-drop vapor diffusion over a 0.5 mL resevoir of 20% (w/v) PEG 3350 and 200 mM magnesium nitrate at 20 °C over a period of 4−5 days. Crystals grew in clusters of needles that could be separated into individual needles approximately 0.20 × 0.02 × 0.02 mm^3 in size. Individual needles were soaked for 2 h in artificial mother liquor supplemented with 20 mM $ZnSO_4$ to maximize Zn^{2+} occupancy and were then flash-frozen after brief exposure to an 18% glycerol cryoprotectant. Extensive attempts to solve an enzyme−substrate complex for TeNT/LC were unsuccessful due to severe nonmerohedral twinning in TeNT−synaptobrevin cocrystals.

Table 1: Data Collection and Refinement Statistics for TeNT-LC

(A) cell parameters		
space group	*C*222	
a, *b*, *c* (Å)	105.4, 176.9, 57.2	
(B) data collection statistics[a]		
	λ1	λ2
wavelength (eV)	9673	10400
resolution range (Å)	40−2.2	40−2.2
unique reflections	27585	27597
multiplicity	9.2 (9.3)	9.2 (9.0)
completeness (%)	99.9 (99.8)	99.9 (99.9)
R_{sym}[b] (%)	6.60 (32.7)	6.80 (34.7)
I/σ	8.5 (1.9)	7.7 (1.7)
source/detector	SSRL 9.2/ADSC Quantum-315	
integration/scaling software	MOSFLM/SCALA	
(C) molecular replacement	1F82.pdb	
search model		
(D) Zn^{2+} MAD phases		
figure of merit (FOM)		
centrics	0.47 (0.29)	
acentrics	0.35 (0.22)	
all	0.36 (0.22)	
density-modified mean FOM	0.92	
(E) model statistics		
resolution	40−2.3	
R_{free}[c]	0.214	
$R_{working}$[c]	0.208	
rmsd, bonds (Å)	0.0066	
rmsd, angles (deg)	1.32	
rmsd, dihedrals (deg)	22.71	
rmsd, impropers (deg)	0.79	
Ramachandran plot		
most favored ϕ-ψ (%)	85.5	
additionally allowed	14.5	
generously allowed	0.0	
disallowed	0.0	
disordered residues	1, 209−220, 253−262, 428−442	
Luzatti coordinate error (cross-validated) (Å)	0.28	
Luzatti coordinate error (Å)	0.25	
σ_A coordinate error (cross-validated) (Å)	0.17	
σ_A coordinate error (Å)	0.16	
average *B*-factor (Å^2)	36	
range of *B*-factors (Å^2)	9−114	
no. of protein residues	425	
no. of Zn^{2+}	2	
no. of water molecules	328	

[a] Numbers in parentheses represent values for the highest resolution bin. [b] $R_{sym} = \sum_h\sum_i|I_{hi} - \langle I_h\rangle|/\sum_h\sum_i|I_{hi}|$. [c] $R_{working} = \sum_{hkl}|F_o - F_c|/\sum_{hkl}|F_o|$, and R_{free} is equivalent to $R_{working}$ but is calculated for a randomly chosen 5% of reflections excluded from model refinement. The low gap between $R_{working}$ and R_{free} was achieved by using the MLHL target with MAD phases and setting the occupancies of disordered residues to zero.

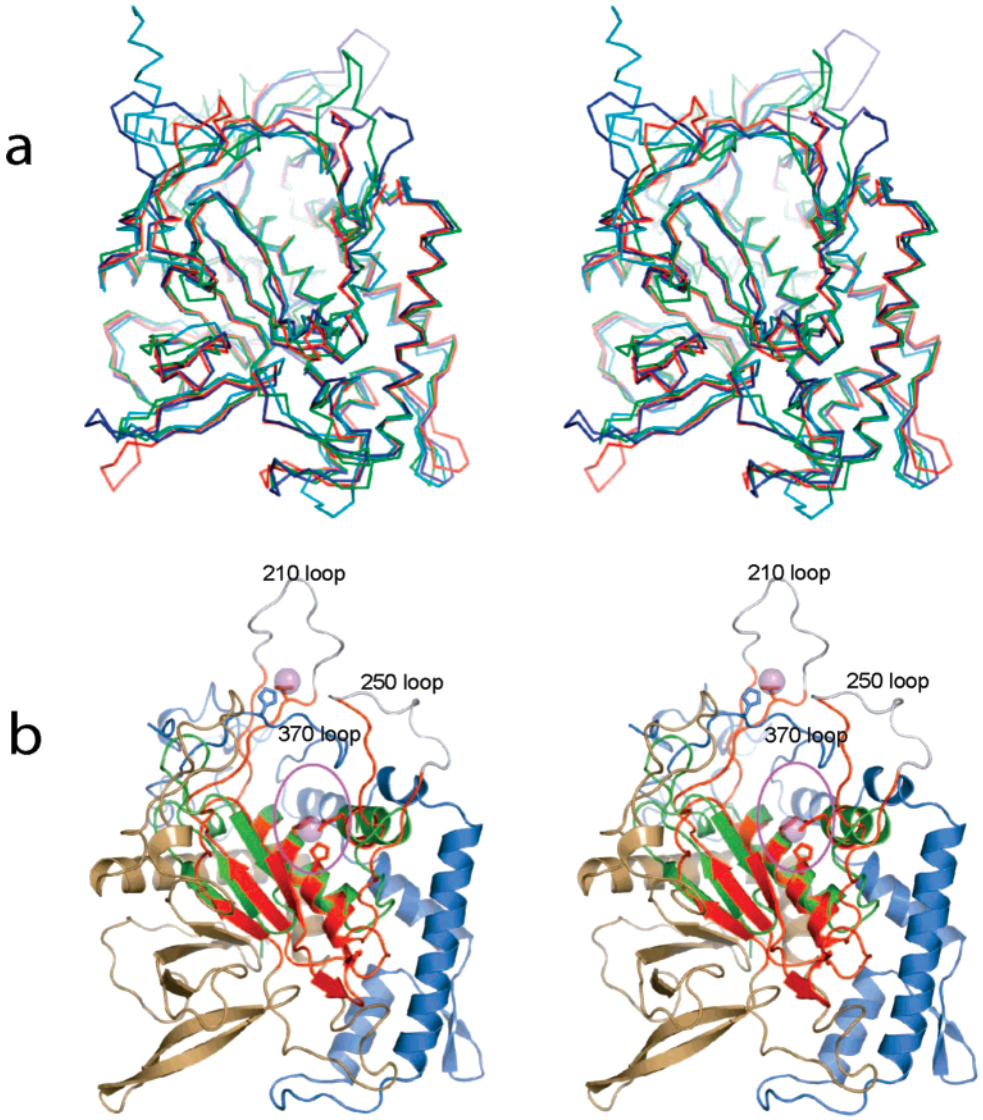

FIGURE 1: Overall TeNT-LC structure compared to thermolysin and other CNT light chains. (a) A wall-eyed stereoview of an overall C_α alignment of light chain structures from BoNT/A (green), BoNT/B (blue), and BoNT/E (cyan) with TeNT (red) demonstrates the high level of structural similarity between the CNT light chains. Disordered loop regions were excluded from the alignment. (b) A wall-eyed stereo ribbon diagram of TeNT-LC is shown aligned with thermolysin (3TMN.pdb) (*59*) residues 99−168 (green). A central, catalytic core of TeNT-LC (residues 156−274, red) shows conservation of tertiary structure with thermolysin. Primary catalytic residues (red sticks) are shared by both enzymes and coordinate a zinc ion (violet sphere) at the active site (violet oval). Loop regions (such as the 210 and 250 loops) in TeNT-LC are significantly expanded compared to the thermolysin core and are partially disordered (zero occupancy indicated gray portions of these loops). Amino-terminal (brown) and carboxy-terminal (blue) portions of TeNT-LC are distinct among known structures outside of the CNT group. A low-affinity Zn^{2+}-binding site spanning the 210 and 370 loops of TeNT-LC imparted some order to these otherwise highly flexible loops in our structure.

Anomalous diffraction data were collected on a Quantum-315 CCD detector (Area Detector Systems) at beamline 9-2 at the Stanford Synchrotron Radiation Laboratory (SSRL). Diffraction was observable to approximately $d_{min} = 2.0$ Å. Exploiting TeNT-LCs endogenous zinc, a multiwavelength anomalous dispersion (MAD) experiment around the zinc K-edge was carried out in order to obtain experimental phases (*31*). Oscillation diffraction data at two wavelengths (9673 and 10400 eV) were collected in 10° wedges over a 120° range using inverse beam geometry. Data collection statistics are shown in Table 1.

An initial MR solution was found using coordinates from BoNT/B-LC (1F82.pdb) (*24*) against TeNT-LC diffraction data at 20−4 Å resolution. The locations of two anomalously scattering Zn^{2+} were identified by anomalous difference Fourier analysis using phases derived from the molecular replacement solution. MAD phases extending to 2.3 Å were calculated and combined with available MR phases, yielding highly interpretable electron density. Phasing and refinement statistics are shown in Table 1.

Model building was performed using the program O (*32*). The initial model, derived from the MR solution, was refined with cycles of simulated annealing with torsion angle dynamics (*33*) and restrained *B*-factor refinement (*34*) using the MLHL maximum likelihood target function with native amplitudes and MAD phase probability distributions (*35*) followed by manual rebuilding. The progress of model refinement was monitored by cross-validation using R_{free} (*36*), which was computed from a randomly assigned test set comprising 5% of the data. All refinement calculations were performed using CNS (*37*). The crystal structure consists of one model in the asymmetric unit. Disordered loop regions

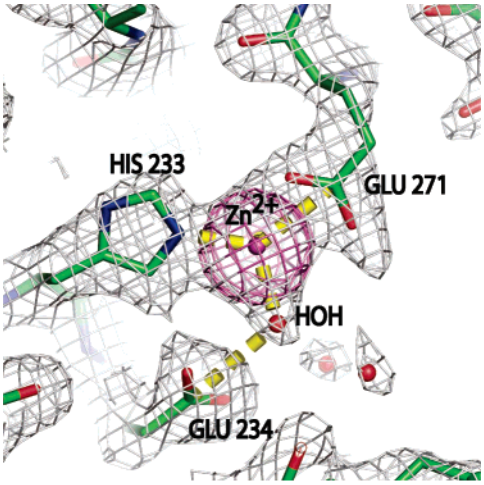

FIGURE 2: Active site Zn^{2+} coordination. A cross-validated, σ_A-weighted, phase combined electron density map contoured at 1σ (gray mesh) is shown at the active site of TeNT-LC. The map was computed by Fourier transformation of $2m|F_o|\varphi_{comb} - D|F_c|\varphi_{calc}$, where φ_{comb} and φ_{calc} are the combined and model phases, respectively, m is the combined figure of merit, and D is an estimate of model incompleteness and error, both obtained from the σ_A method (*60*). Anomalous Fourier difference density contoured at 10σ (violet mesh) highlights the position of the catalytic zinc ion, with coordination explicitly shown (yellow dashes). The fourth ligand (His 237) of the tetrahedral coordination sphere is obscured from view by the zinc ion.

included residues 204−216 and 249−258. Plausible models for these loops based on equivalent regions in other LC structures were constructed, but atomic occupancies were set to zero. Residues C-terminal to Asp 427 and N-terminal to Pro 2 were disordered and not included in the model.

Structural alignments and figures were prepared with PyMOL (*38*). Electrostatic potential values were calculated using the adaptive Poisson−Boltzmann solver (APBS) (*39*) using Chemistry at HARvard Molecular Mechanics (CHARMM) force field parameters (*40*). Surface variability analysis was performed using ConSurf (*41*) using a multiple sequence alignment of all CNT light chains performed with ClustalW (*42*). Homology searches were conducted with BLASTp (*43*), using masked CNT light chain consensus sequences. The multiple sequence alignment used in Figure 5 was created using MultAlin (*44*).

RESULTS AND DISCUSSION

The overall structure of TeNT-LC is similar to light chains of BoNTs A, B, and E (Figure 1a). Structural correlation is reflected by high DALI Z scores (*45*) obtained by C_α coordinate alignment between TeNT and the other known BoNT light chains ($Z = 55.3$ for BoNT/B, 51.8 for BoNT/E, and 45.9 for BoNT/A). The high degree of structural similarity of TeNT to the other light chains is not a result of crystallographic phase bias since experimental zinc MAD phases were used to independently confirm the model. The light chain structures of BoNT/A, BoNT/B, and BoNT/E were solved by using experimental phase information (*25*, *27*, *28*). The common architecture of these enzymes reflects their common function; each must specifically bind and hydrolyze a specific SNARE target which is largely unstructured in the absence of a binding partner (*46*, *47*). In the case of BoNT/A, we observed an array of substrate-binding

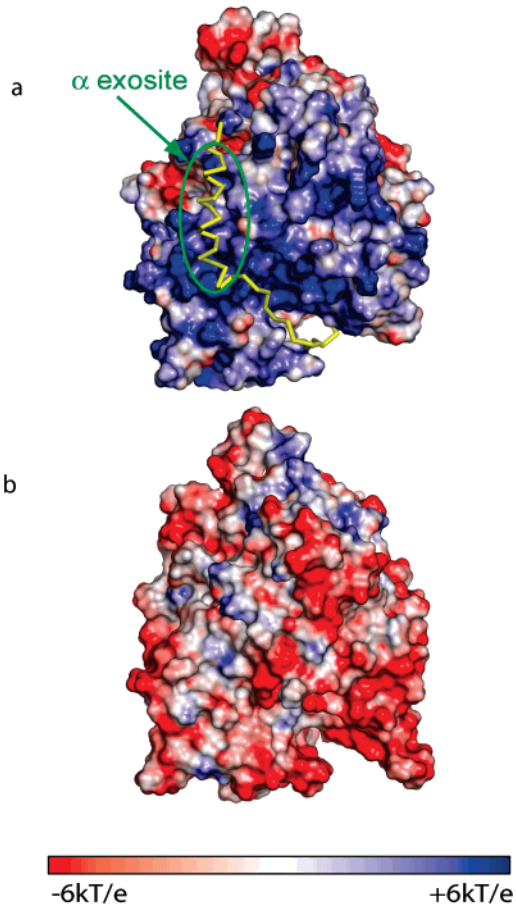

FIGURE 3: Electrostatic potential surfaces of BoNT/A and TeNT light chains. Molecular surfaces of (a) BoNT/A-LC and (b) TeNT-LC are colored by electrostatic potential values ranging from +6 kT/e (blue) to −6 kT/e (red). Equivalent views of the light chain faces opposite from their active sites are shown, highlighting the location of the BoNT/A α-exosite, which is shown bound to substrate (yellow C_α trace). The surface residues involved in BoNT/A substrate recognition near this region are considerably more basic than the equivalent residues in TeNT, despite nearly identical underlying topology (Figure 1a).

exosites that induce secondary structure elements in SNAP-25 serving to specifically position the targeted scissile bond in close proximity to the protease's attacking nucleophile. Given the extremely high degree of structural similarity between the CNT light chains, we speculate that this exosite-based strategy is common to the entire group of toxins. This hypothesis is supported by kinetic data showing that the CNT light chains as a group require long stretches of substrate peptide (as long as 50−60 residues) for optimal catalytic efficiency and can be strongly influenced by substrate amino acid substitution remote from the scissile bond (*26*, *48*−*52*).

Primary structural differences between TeNT-LC and other light chains are limited to surface features of the protease including loops and other solvent-exposed protrusions. Notably, three loop regions exhibit weak electron density and high temperature factors, suggesting conformational variability in these regions. Shown in Figure 1b, this group of loops includes residues 206−223 (the "210 loop"), residues 248−266 (the "250 loop"), and residues 368−382 (the "370 loop"). Equivalents of the 250 and 370 loops in BoNT/A-LC were shown to undergo significant conformational changes upon substrate binding (*26*). In TeNT-LC,

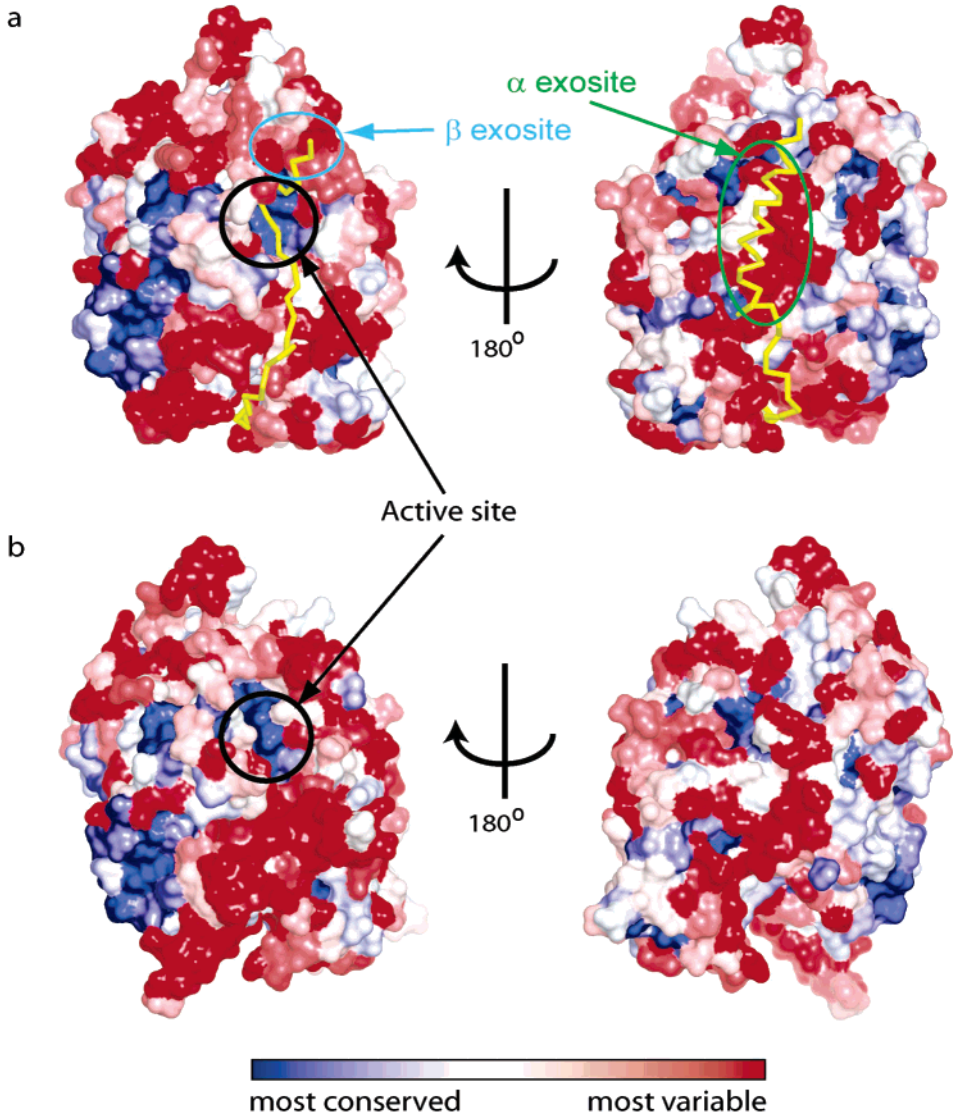

FIGURE 4: Surface variability of CNT light chains. Molecular surfaces of (a) BoNT/A-LC and (b) TeNT-LC are colored according to relative surface variability. Relative residue variability (indicated by the blue → red color ramp) was determined using ConSurf (*41*) based on a ClustalW (*42*) multiple sequence alignment of BoNTs/A−G and TeNT. The yellow C_α trace indicates the path of the SNAP-25 in complex with the BoNT/A endopeptidase. Excluding the highly conserved active site (indicated by a black ellipse), major substrate-binding sites (including the α-exosite indicated by a green ellipse and the *β*-exosite indicated by a cyan ellipse) follow a path of relatively high variability on the surface of BoNT/A. This stripe of high variability extending around CNT light chains may allow discrimination between different SNARE targets, if the substrate recognition exosites observed in BoNT/A (*26*) are similarly distributed in the other light chains.

the 210 and 370 loops are partially stabilized due to a secondary zinc-binding site identified by anomalous Fourier maps at the zinc edge (Figure 1b). This secondary zinc site is likely induced by the zinc-enriched artificial mother liquor in which the crystal was soaked prior to data collection. The secondary zinc restricted the mobility of the 210 and 370 loops relative to a structure obtained from crystals that had not been subjected to the zinc supplementation (not shown). Under physiological conditions, however, these loops are probably conformationally variable in TeNT-LC.

When the first CNT structure was determined, DALI searches revealed that its catalytic core was similar to that of the general zinc metalloprotease thermolysin (*27*). A central portion of TeNT-LC is clearly structurally related to a portion of thermolysin, as shown in Figure 1b. While loop regions of the TeNT-LC core are significantly extended compared their thermolysin counterparts, overall tertiary structure is conserved between the two enzymes. This thermolysin-like core is sandwiched between structurally distinct amino- and carboxy-terminal portions of the light chain. These N- and C-terminal elements appear to be structurally unique to the CNT group, and most significantly, they contain the SNARE-binding exosites observed in BoNT/A. The thermolysin-like core of BoNT/A (residues 150−266) does not participate in substrate-specific side-chain/side-chain contacts when in complex with its target SNARE. The 250 loop does, however, participate in a substrate-nonspecific main-chain/main-chain interaction at the *β*-exosite (*26*). The structural elements flanking the thermolysin-like core of TeNT-LC (and likely all CNT light chains) are thus responsible for imparting SNARE specificity to this unique class of proteases.

TeNT-LC active site geometry and zinc coordination are similar to those of other known light chain structures. Residues 233−237 constitute the His-Glu-X-X-His Zn^{2+}-binding motif characteristic of the light chains as well as of

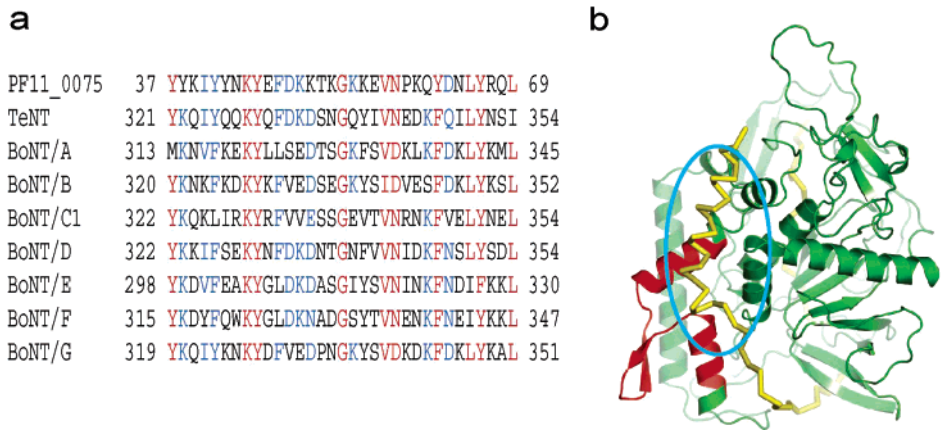

FIGURE 5: Eukaryotic protein with homology to CNT light chains. (a) A 33-residue portion of the hypothetical *P. falciparum* gene PF11_0075 aligned with TeNT and BoNT light chains reveals a high level of sequence identity and similarity. This stretch of PF11_0075 is up to 42.4% identical, above the homology threshold for a peptide of this length (*61*), to portions of the CNT light chains likely involved in SNARE binding. (b) The featured stretch of BoNT/A from this alignment is highlighted in red and constitutes two of the four helices involved in forming the α-exosite observed in the BoNT/A−SNAP-25 complex (*26*). The SNAP-25 substrate is shown in yellow bound to the BoNT/A light chain (green), where the blue ellipse indicates the approximate location of the α-exosite.

thermolysin (*53*). Glu 271 and an ordered water molecule complete the tetrahedral Zn^{2+} coordination sphere (Figure 2). Likely employing a catalytic mechanism similar to that of thermolysin, TeNT-LC is thought to deprotonate a water molecule coordinated by both Zn^{2+} and Glu 234 (Figure 2). The resulting nucleophile likely attacks the carbonyl carbon of synaptobrevin Gln 76, forming an oxyanion. However, structural equivalents of oxyanion-stabilizing thermolysin residues Tyr 157 and His 231 (*53*) do not exist in TeNT or the other CNT light chains.

While the underlying architecture of all CNT light chains is essentially identical, there are significant differences in surface composition. For example, the electrostatic profile of the TeNT-LC surface is substantially more acidic than that of BoNT/A (Figure 3). This electrostatic difference is especially pronounced for regions that are involved in SNAP-25 binding to BoNT/A-LC. Such electrostatic variation may be necessary for the different light chains to accommodate their different SNARE targets, but it is important to stress that the precise subset of TeNT-LC residues that impart specificity for synaptobrevin is not yet known. While each light chain has a unique electrostatic profile, generalizations about SNARE binding based on patterns in surface potentials cannot yet be made; BoNT/B and TeNT light chains, which share the same synaptobrevin target, have dissimilar electrostatic profiles. Until more structures of CNT light chains in complex with their SNARE substrates are available, electrostatic surface profiles only serve to illustrate that surface features are not conserved among these enzymes. An analysis of primary sequence relative variability was performed on the basis of a multiple sequence alignment of all CNT light chains (Figure 4). Remarkably, the most variable residues include the substrate-binding path for the BoNT/A−SNAP-25 complex while the active sites and interior cores of the toxins are well conserved. These regions of high sequence variability, extending from the catalytic pocket around the toxin surface, suggest that distinct SNARE-binding arrays are likely present in other CNT light chains. These variable surface features probably include exosite positions that specialize in binding to different portions of synaptobrevin, SNAP-25, or syntaxin depending on the toxin. Notably, the most variable residues also follow the path of the belt regions observed in full-length BoNT/A and BoNT/B structures (*27*, *28*), indicating that these heavy chain residues occupy exosite positions on the light chain surfaces, thus inactivating the enzyme prior to translocation. As observed in BoNT/A-LC, the exosites orient and position the substrate's target scissile bond within the toxin active site, increasing catalytic efficiency by lowering the Michaelis constant (*26*).

CONCLUSIONS

The identification of SNAP-25-binding exosites in BoNT/A and the likely presence of such sites in the other toxins have interesting evolutionary implications. The toxins function at substantially impaired (and perhaps physiologically irrelevant) efficiency for truncated substrates that exclude the remote exosites, as shown in several kinetic analyses (*50*, *51*, *54*, *55*). Due to its complexity, the gradual, stepwise evolution of a functional exosite array that could benefit the *Clostridia* is difficult to reconcile. Interestingly, we found a 33-residue stretch of a hypothetical gene from the recently deciphered *Plasmodium falciparum* genome that is homologous to a portion of the BoNT/A α-exosite (Figure 5). Thus, one possibility is that structural elements involved in SNARE interaction may have been imported into Clostridia via a horizontal gene transfer event and recombined with a general thermolysin-like enzyme to yield these SNARE-specific proteases. The *Plasmodium* gene, PF11_0075, has not been characterized for function yet. While the secretory machinery of higher order eukaryotes is relatively well studied, little is known about SNARE-binding factors in organisms such as *Plasmodium*. It is thus possible that the structurally distinct N- and C-terminal portions of the CNT light chains actually have eukaryotic origins. Horizontal gene transfer from eukaryote to bacteria has been documented in the case of glutaminyl-tRNA synthase (*56*−*58*), and such a process could account for the presence of SNARE-binding elements in *Clostridia*. While the homologous fragment from PF11_0075 constitutes only a portion of the BoNT/A SNARE-binding array, it presents an intriguing possibility that a more complete match will be discovered as more eukaryotic genomes are decoded. Clearly, the observed homology is not necessarily a result of horizontal gene transfer;

without additional evidence, convergent evolution could also explain the high level of sequence similarity. Nevertheless, a better understanding of CNT−SNARE interactions and their evolutionary origins might lead to the discovery of new components of the eukaryotic secretory fusion machinery.

ACKNOWLEDGMENT

We thank T. Fenn, S. Kaiser, Z. Panepucci, and R. B. Sutton for technical assistance and critical reading. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy (Office of Basic Energy Sciences). The SSRL Structural Molecular Biology Program is supported by the Department of Energy (Office of Biological and Environmental Research) and by the National Institutes of Health (National Center for Research Resources, Biomedical Technology Program).

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP April 27, 2005; the formula for map computation in the legend of Figure 2 was corrected and the paper reposted May 5, 2005.

REFERENCES

1. Simpson, L. L. (1986) Molecular pharmacology of botulinum toxin and tetanus toxin, *Annu. Rev. Pharmacol. Toxicol. 26*, 427−453.
2. Arnon, S. S., Schechter, R., Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Fine, A. D., Hauer, J., Layton, M., Lillibridge, S., Osterholm, M. T., O'Toole, T., Parker, G., Perl, T. M., Russell, P. K., Swerdlow, D. L., and Tonat, K. (2001) Botulinum toxin as a biological weapon: medical and public health management, *J. Am. Med. Assoc. 285*, 1059−1070.
3. Humeau, Y., Doussau, F., Grant, N. J., and Poulain, B. (2000) How botulinum and tetanus neurotoxins block neurotransmitter release, *Biochimie 82*, 427−446.
4. Pellizzari, R., Rossetto, O., Schiavo, G., and Montecucco, C. (1999) Tetanus and botulinum neurotoxins: mechanism of action and therapeutic uses, *Philos. Trans. R. Soc. London, Ser. B 354*, 259−268.
5. Kao, I., Drachman, D. B., and Price, D. L. (1976) Botulinum toxin: mechanism of presynaptic blockade, *Science 193*, 1256−1258.
6. Ashton, A. C., and Dolly, J. O. (1988) Characterization of the inhibitory action of botulinum neurotoxin type A on the release of several transmitters from rat cerebrocortical synaptosomes, *J. Neurochem. 50*, 1808−1816.
7. Turton, K., Chaddock, J. A., and Acharya, K. R. (2002) Botulinum and tetanus neurotoxins: structure, function and therapeutic utility, *Trends Biochem. Sci. 27*, 552−558.
8. Dolly, J. O., Black, J., Williams, R. S., and Melling, J. (1984) Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization, *Nature 307*, 457−460.
9. Lalli, G., Bohnert, S., Deinhardt, K., Verastegui, C., and Schiavo, G. (2003) The journey of tetanus and botulinum neurotoxins in neurons, *Trends Microbiol. 11*, 431−437.
10. Schiavo, G., Poulain, B., Rossetto, O., Benfenati, F., Tauc, L., and Montecucco, C. (1992) Tetanus toxin is a zinc protein and its inhibition of neurotransmitter release and protease activity depend on zinc, *EMBO J. 11*, 3577−3583.
11. Blasi, J., Chapman, E. R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., Sudhof, T. C., Niemann, H., and Jahn, R. (1993) Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25, *Nature 365*, 160−163.
12. Montecucco, C., and Schiavo, G. (1993) Tetanus and botulism neurotoxins: a new group of zinc proteases, *Trends Biochem. Sci. 18*, 324−327.
13. Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R., and Montecucco, C. (1992) Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin, *Nature 359*, 832−835.
14. Montecucco, C., and Schiavo, G. (1995) Structure and function of tetanus and botulinum neurotoxins, *Q. Rev. Biophys. 28*, 423−472.
15. Schiavo, G., Rossetto, O., Benfenati, F., Poulain, B., and Montecucco, C. (1994) Tetanus and botulinum neurotoxins are zinc proteases specific for components of the neuroexocytosis apparatus, *Ann. N.Y. Acad. Sci. 710*, 65−75.
16. Gerst, J. E. (1999) SNAREs and SNARE regulators in membrane fusion and exocytosis, *Cell. Mol. Life Sci. 55*, 707−734.
17. Chen, Y. A., Scales, S. J., Patel, S. M., Doung, Y. C., and Scheller, R. H. (1999) SNARE complex formation is triggered by Ca^{2+} and drives membrane fusion, *Cell 97*, 165−174.
18. Sollner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J. E. (1993) SNAP receptors implicated in vesicle targeting and fusion, *Nature 362*, 318−324.
19. Pellegrini, L. L., O'Connor, V., Lottspeich, F., and Betz, H. (1995) Clostridial neurotoxins compromise the stability of a low energy SNARE complex mediating NSF activation of synaptic vesicle fusion, *EMBO J. 14*, 4705−4713.
20. Xu, T., Binz, T., Niemann, H., and Neher, E. (1998) Multiple kinetic components of exocytosis distinguished by neurotoxin sensitivity, *Nat. Neurosci. 1*, 192−200.
21. Penner, R., Neher, E., and Dreyer, F. (1986) Intracellularly injected tetanus toxin inhibits exocytosis in bovine adrenal chromaffin cells, *Nature 324*, 76−78.
22. Li, Y., Foran, P., Lawrence, G., Mohammed, N., Chan-Kwo-Chion, C. K., Lisk, G., Aoki, R., and Dolly, O. (2001) Recombinant forms of tetanus toxin engineered for examining and exploiting neuronal trafficking pathways, *J. Biol. Chem. 276*, 31394−31401.
23. 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.
24. Hanson, M. A., and Stevens, R. C. (2000) Cocrystal structure of synaptobrevin-II bound to botulinum neurotoxin type B at 2.0 Å resolution, *Nat. Struct. Biol. 7*, 687−692.
25. 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.
26. Breidenbach, M. A., and Brunger, A. T. (2004) Substrate recognition strategy for botulinum neurotoxin serotype A, *Nature 432*, 925−929.
27. Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R., and Stevens, R. C. (1998) Crystal structure of botulinum neurotoxin type A and implications for toxicity, *Nat. Struct. Biol. 5*, 898−902.
28. Swaminathan, S., and Eswaramoorthy, S. (2000) Structural analysis of the catalytic and binding sites of Clostridium botulinum neurotoxin B, *Nat. Struct. Biol. 7*, 693−699.
29. Rupp, B., and Segelke, B. (2001) Questions about the structure of the botulinum neurotoxin B light chain in complex with a target peptide, *Nat. Struct. Biol. 8*, 663−664.
30. Lacy, D. B., and Stevens, R. C. (1999) Sequence homology and structural analysis of the clostridial neurotoxins, *J. Mol. Biol. 291*, 1091−1104.
31. Hendrickson, W. A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation, *Science 254*, 51−58.
32. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A 47* (Part 2), 110−119.
33. Rice, L. M., and Brunger, A. T. (1994) Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement, *Proteins 19*, 277−290.
34. Hendrickson, W. A. (1985) Stereochemically restrained refinement of macromolecular structures, *Methods Enzymol. 115*, 252−270.
35. Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998) Incorporation of prior phase information strengthens maximum-likelihood structure refinement, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 1285−1294.

36. Brunger, A. T. (1992) Free R-Value—A novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472−475.
37. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54* (Part 5), 905−921.
38. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, http://www.pymol.org.
39. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037−10041.
40. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586−3616.
41. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information, *Bioinformatics 19*, 163−164.
42. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs, *Nucleic Acids Res. 31*, 3497−3500.
43. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
44. Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering, *Nucleic Acids Res. 16*, 10881−10890.
45. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol. 233*, 123−138.
46. Fiebig, K. M., Rice, L. M., Pollock, E., and Brunger, A. T. (1999) Folding intermediates of SNARE complex assembly, *Nat. Struct. Biol. 6*, 117−123.
47. Fasshauer, D., Bruns, D., Shen, B., Jahn, R., and Brunger, A. T. (1997) A structural change occurs upon binding of syntaxin to SNAP-25, *J. Biol. Chem. 272*, 4582−4590.
48. Yamasaki, S., Baumeister, A., Binz, T., Blasi, J., Link, E., Cornille, F., Roques, B., Fykse, E. M., Sudhof, T. C., Jahn, R., et al. (1994) Cleavage of members of the synaptobrevin/VAMP family by types D and F botulinal neurotoxins and tetanus toxin, *J. Biol. Chem. 269*, 12764−12772.
49. Shone, C. 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.
50. Schmidt, J. J., and Bostian, K. A. (1995) Proteolysis of synthetic peptides by type A botulinum neurotoxin, *J. Protein Chem. 14*, 703−708.
51. Pellizzari, R., Rossetto, O., Lozzi, L., Giovedi, 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.
52. Cornille, F., Martin, L., Lenoir, C., Cussac, D., Roques, B. P., and Fournie-Zaluski, M. C. (1997) Cooperative exosite-dependent cleavage of synaptobrevin by tetanus toxin light chain, *J. Biol. Chem. 272*, 3459−3464.
53. Matthews, B. W. (1988) Structural basis of the action of thermolysin and related zinc peptidases, *Acc. Chem. Res. 21*, 333−340.
54. Schmidt, J. J., and Bostian, K. A. (1997) Endoproteinase activity of type A botulinum neurotoxin: substrate requirements and activation by serum albumin, *J. Protein Chem. 16*, 19−26.
55. Li, L., Binz, T., Niemann, H., and Singh, B. R. (2000) Probing the mechanistic role of glutamate residue in the zinc-binding motif of type A botulinum neurotoxin light chain, *Biochemistry 39*, 2399−2405.
56. Brown, J. R., and Doolittle, W. F. (1999) Gene descent, duplication, and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases, *J. Mol. Evol. 49*, 485−495.
57. Brown, J. R. (2003) Ancient horizontal gene transfer, *Nat. Rev. Genet. 4*, 121−132.
58. Lamour, V., Quevillon, S., Diriong, S., N'Guyen, V. C., Lipinski, M., and Mirande, M. (1994) Evolution of the Glx-tRNA synthetase family: the glutaminyl enzyme as a case of horizontal gene transfer, *Proc. Natl. Acad. Sci. U.S.A. 91*, 8670−8674.
59. Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H., and Matthews, B. W. (1987) Slow- and fast-binding inhibitors of thermolysin display different modes of binding: crystallographic analysis of extended phosphonamidate transition-state analogues, *Biochemistry 26*, 8542−8553.
60. Read, R. J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors, *Acta Crystallogr. A 42*, 140−149.
61. Sander, C., and Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment, *Proteins 9*, 56−68.

BI050262J