SNAP-25 SUBSTRATE PEPTIDE (180-183) BINDS TO BUT BYPASSES CLEAVAGE BY CATALYTICALLY ACTIVE CLOSTRIDIUM BOTULINUM NEUROTOXIN E

Rakhi Agarwal and Subramanyam Swaminathan

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

Running title: Crystal structure of BoNT E -SNAP-25 peptide complex

*Corresponding author: S. Swaminathan, Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA. Phone-1(631)344-3187, Fax: 1(631)344-3407 Email: swami@bnl.gov

Clostridium botulinum neurotoxins are the most potent toxins to humans. The recognition and cleavage of SNAREs (soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptors) is a prime event in exhibiting their toxicity. We report here the crystal structure of catalytically active full-length botulinum serotype E catalytic domain (BoNT E) in complex with SNAP-25 (a SNARE protein) substrate peptide Arg180-Ile181-Met182-Glu183 (P1-P3'). It is remarkable that the peptide spanning the scissile bond binds to but bypasses cleavage by the enzyme and inhibits the catalysis fairly with $K_i \approx 69\mu M$. The inhibitory peptide occupies the active site of BoNT E and shows well-defined electron density. The catalytic zinc and the conserved key residue Tyr350 of the enzyme facilitate the docking of Arg180 (P1) by interacting with its carbonyl oxygen that displaces the nucleophilic water. The general base Glu212 side chain interacts with the main chain amino group of P1 and P1'. Conserved Arg347 of BoNT E stabilizes the proper docking of Ile181 (P1') main chain, while the hydrophobic pockets stabilize the side chains of I181 (P1') and Met182 (P2'), and 250 loop stabilizes Glu183 (P3'). Structural and functional analysis revealed an important role for P1' residue and S1' pocket in driving substrate recognition and docking at the active site. This study is the first of its kind and rationalizes the substrate cleavage strategy of BoNT E. Also, our complex structure opens up an excellent opportunity of structure-based drug design for this fast acting and extremely toxic high priority BoNT E.

Clostridium botulinum neurotoxins (BoNTs) are produced as an inactive single chain ($\approx 150kDa$) and released as active dichains, heavy ($\approx 100kDa$, HC) and light chain ($\approx 50kDa$, LC)(1-4). BoNT serotype E is an exception since it is released as a single chain (4). The C-terminal domain of the HC mediates binding of the toxin to specific neuronal receptors and the N-terminal domain enables the catalytically active LC translocation to the cytosol (5). The light chain is a zinc dependent endopeptidase, which recognizes and cleaves one of the three SNARE proteins (SNAP-25, VAMP-2 or syntaxin) (5). SNARE proteins associate to form low energy ternary complex required for fusion and docking of neurotransmitter carrying vesicles to target membranes of peripheral neurons enabling neurotransmitter release (6). Cleavage of any one of them prevents the complex formation resulting in flaccid paralysis. BoNTs A and E cleave SNAP 25, while BoNTs B, D, F and G cleave VAMP-2. BoNT C is unique and cleaves both SNAP-25 and syntaxin (7).

BoNTs are listed as category A bioterrorism agents by the Centers for Disease Control and Prevention (CDC). Accordingly, developing inhibitors to counter their activity is a priority. BoNTs A, B and E are more potent than the other serotypes. Although BoNTs E and A cleave the same substrate SNAP-25, E blocks neurotransmission faster (8). Crystal structures of catalytic domains of all serotypes are available (9-15). However, insight to the substrate recognition and binding in each serotype, crucial for developing common and/or serotype specific inhibitors, is yet to be established.

As of now the structure of enzyme-substrate peptide (146-206) complex has been reported only for BoNT A (PDB id 1XTG) (16). However, to prevent cleavage during co-crystallization, an inactive double mutant enzyme was used. Though this structure maps the exosite interactions, the region flanking the scissile bond (about 6 or 7 residues) are disordered and does not make contact with the active site residues of the enzyme (16). Recently, crystal structures of BoNT A in complex
with peptide inhibitors: Arg-Arg-Gly-Cys (RRGC) and its variants at Cys position, and N-Ac-Cys-Arg-Ala-Thr-Lys-Met-Leu (N-Ac-CRATKML) have been reported (17-19). However, variation of substrate residues flanking the scissile bond and their subsites in BoNT E and other serotypes warrants determination of structures of substrate-enzyme complex. Structural information of substrate recognition and binding at the active site is crucial in better understanding the substrate specificity among various serotypes and for developing potent serotype specific inhibitors.

To our knowledge, structural information for BoNT-substrate/analog peptide complexes is not available for any serotype except BoNT A. Here, we present for the first time the crystal structure of the full length BoNT E catalytic domain with its substrate SNAP-25 (180-183) Arg-Ile-Met-Glu-amide (RIME, P1:Arg180, P1’:Ile181, P2’:Met182, P3’:Glu183) peptide spanning the scissile bond. The S1, S1’, S2’ and S3’ pockets of BoNT E and their crucial interactions with P1-P3’ residues are identified. We also modeled P2:Asp179 residue of the substrate and identified its possible interactions in S2 pocket. Interestingly, we found that the substrate peptide RIME itself acts as a moderate inhibitor of catalysis. This study rationalizes the substrate cleavage strategy of BoNT E, which could be extended to other serotypes too. The present complex structure is the first report of substrate binding to the active site of BoNT E and helps to better understand the mechanism of catalysis and substrate recognition and binding at and near the active site. This will serve as an important starting point for designing specific structure-based potent inhibitors/drugs for highly toxic BoNT E.

**Experimental procedures:**

**Cloning, expression and purification of BoNT E catalytic domain:** The clone for BoNT E catalytic domain in pET9c vector was used as described (20). Protein was expressed using auto-induction protocol (21) and was purified (20). This yielded >20mg protein/500ml culture and the protein was concentrated to 10mg/ml in Heps buffer (20mM, pH 7.8, 200mM NaCl).

**Kinetic Parameters and Ki determination**

Kinetic values $k_{m}$ and $k_{cat}$ were determined for SNAP25-(146–186) as substrate with BoNT E. Enzyme concentrations were adjusted (0.5nM) to cleave <20% substrate at various concentrations of substrate ranging from 35 to 500µM. The reactions were carried out in 20 µl of reaction buffer containing 20mM Heps, pH 7.4, 2mM DTT, and 10 µM Zn acetate. The reaction mixtures were incubated at 37°C for 15 min, and stopped by adding 2 µl of 1%TFA. The samples were run in HPLC (Beckman) with Discovery Bio-Wide pore C18 column with acetonitrile and aqueous phase gradient (Buffer A: 0.1%TFA in water, Buffer B: 0.1%TFA, 90% Acetonitrile). The amount of cleaved product was calculated using software provided by the Beckman. Kinetic constants were calculated from Lineweaver-Burk plot using Enzyme Kinetic Module and Sigmaplot (Systat Software, Inc., CA). IC$_{50}$ and $K_i$ determination for RIME (P1-P3’) (50-950 µM concentrations) was performed in similar assay conditions and data analyzed using GraphPad software and Cheng-Prusoff equation respectively. Peptide 2.0 Inc. synthesized this tetrapeptide.

**Crystallization and structure determination:** The catalytic domain was cocrystallized with RIME in 0.5M Ammonium sulfate, 0.1M Na citrate buffer pH 5.6, 2M Li$_2$SO$_4$ in sitting drop in 24 well plates. Protein at 7mg/ml and peptide at ~10mM concentration was used for cocrystallization. Lower concentrations of the peptide during cocrystallization led to poor occupancy of the substrate at the active site. Diamond shaped crystals appeared in 2-7 days at 22 ºC. Crystals were immersed in mother liquor containing 15% (v/v) glycerol and immediately flash frozen in liquid nitrogen. X-ray data was collected at beamline X29 of NSLS, BNL at 1.08Å wavelength and crystals diffracted to better than 2.25Å resolution. Crystals belong to space group P2$_1$2$_1$2 with 2 molecules per asymmetric unit. A rigid body refinement was performed using BoNT E catalytic domain model (PDB id 1T3A) followed by simulated annealing in CNS (22). The refinement statistics are given in Table 1. Molecules A and B are complete for 1-411 and 1-409aa respectively but molecule A has weak electron density for 235-238aa. Residues 59-60 are not modeled in both molecules due to poorly
defined electron density. The SNAP-25 substrate RIME (P1-P3') bound to molecule B could be modeled unambiguously; however the C-terminal residue P3' and P1 side chain could not be modeled for peptide bound to molecule A because of weak electron density. P1:Arg180-P2':Met182 (RIME) aligns well in both molecules of enzymes (RMSD = 0.051Å) with minor variations in interacting distances of P1:Arg180 (peptide RIME) main chain amino-group with the Zinc and Glu212 of enzyme. All our discussion in this paper is based on the substrate peptide RIME bound to B molecule.

**Results and Discussion**

**Arg180-Ile181-Met182-Glu183-amide (RIME, P1-P3') catalytically competes with SNAP-25 (146-186):** We assayed the catalytic activity of BoNT E on SNAP-25 (146-186) using HPLC and determined the kinetic values. It showed $k_{\text{in}}$ of $8.57 \times 10^{-5}$ M and $k_{\text{cat}}$ of 2830 min$^{-1}$. The tetrapeptide RIME of SNAP-25 (180-183) represents P1-P3' residue for BoNT E and fairly competes with SNAP-25 (146-186) with $K_c$ ~69μM suggesting its binding to the active site and its role as a competitive substrate inhibitor (Figure 1).

**Overall structure and crucial interactions of BoNT E with inhibitory peptide (RIME):** BoNT E recognizes and cleaves SNAP-25 substrate specifically between P1:Arg180-P1':Ile181. In our complex structure RIME binds to the active site cavity of BoNT E without being cleaved (Figure 2A and B), interacts with several residues (Figure 2C) and shows well-defined electron density in omit map in molecule B (Figure 3). The crucial interactions of various peptide residues (P1-P3') and P2:Asp179 modeling results are discussed below:

**Docking of P1: Arg180 at the active site:** P1:Arg180 makes interactions required for proper docking of P1 residue to the active site (Figure 2C and 4). Its carbonyl O interacts with both the catalytic zinc (2.69Å) and Tyr350 (3.08Å) and displaces the nucleophilic water (Figure 2C and 4). In wild type protein structure (PDB id 1T3A) the nucleophilic water is at 2.18Å from zinc. The O of P1 makes similar tetrahedral arrangement as nucleophilic water. Interactions between zinc, Tyr350 (conserved in all BoNTs) and nucleophilic water are common in all catalytic domain structures. Tyr350Ala mutant of BoNT E led to complete loss of catalytic inactivity (at ≤500nM concentration) (23). P1:Arg180 main chain amino group interacts with, both Oe1 and Oe2 of Glu212 displacing a water molecule present in the native structure. However, it interacts with another water present in the current structure. Glu212, a general base, plays a crucial role in catalysis and the Glu212Gln mutation led BoNT E to be catalytically inefficient without alteration in protein conformation (23).

In the crystal structure, P1:Arg180 side chain NH1 and NH2 form a strong salt bridge with Glu158 and stabilizes P1 at the active site (Figure 2C). Glu158 is conserved in BoNTs E, F and B, and is Asp in C1, D, G, and Gln in A. The sequence analysis of SNAP-25 and VAMP2 showed that they have either of Q/R/K at the P1 position (except BoNT G) which can complement well with E/D/Q at the equivalent position of Glu158 (BoNT E) in all serotypes (24). A triple mutation Glu158Ala/Thr159Ala/Asn160Ala in BoNT E decreased the catalytic efficiency by ~8 fold and decreased substrate affinity as compared to the wild type enzyme (23). Interestingly, Arg180Lys (P1) variant in MurineSNAP-25 still shows cleavage by BoNT E (25).

**P1':Ile181 is important for specificity and docking at the active site:** P1':Ile181 main chain N interacts with the side chain Oe1 and Oe2 of catalytic base Glu212 and also with Thr159 O (Figure 2C and 4). P1’ O interacts with the side chain NH1 and NH2 groups of the conserved residue Arg347. Arg347Ala mutation reduced the catalytic efficiency ~1000 fold relative to the wild type (23). These interactions are presumably conserved among the other serotypes with their specific substrates and help to proper position and stabilize the substrate. However, the ones which are not conserved are side chain interactions of P1’ and the integrity of S1’ pocket. Here, the hydrophobic P1’:Ile181 is stabilized in S1’ pocket comprising Thr159, Phe191, and Thr208 confirming our earlier prediction (Figure 2C and 4) (23). This seems to be typically different for...
various serotypes defining the specificity. Mutational studies on Thr159, Thr208 and Phe191 showed markedly lower \( k_{\text{cat}} \) values by \( \approx 20 \), \( \approx 30 \) and \( \approx 80 \)-fold, respectively for SNAP25 cleavage, relative to wild type (26). Phe191Asn mutation also showed a lower (\( \approx 100 \)-fold) \( k_{\text{cat}} \). Mutagenesis of P1′:Ile181 to Cys in 141-206aa long SNAP-25 led it to be completely resistant to cleavage by BoNT E (Agarwal and Swaminathan unpublished results). However, change to Ala reduced the activity by \( \approx 75 \) fold relative to wild type (27).

P1′ is the crucial residue where mutations are hardly tolerated and any mutation leads to either non or poor cleavage (25,27). The S1′ pocket residues also play a prime role in delivering substrate specificity by allowing optimal docking of, P1′ residue of specific size and charge. It is tempting to speculate that probably an improper docking of P1′ residue also affects the proper docking of the P1 residue which abrogates the phenotype. We therefore conclude that P1′ residue and its respective S1′ pocket together contribute primarily in the optimum docking of the substrate at the active site and thus in scissile bond specificity. Not surprisingly, most of the variations observed are at P1′ position in various substrates as compared to the P1 position (24).

P2′:Met182 interactions and the dual role of Phe191 of BoNT E: In the crystal structure Tyr350 OH interacts with P2′:Met182 O and helps in stabilizing its main chain (Figure 2C and 4). However, the hydrophobic side chain resides in the hydrophobic pocket formed by Phe191, Tyr354 and Tyr356. Phe191 plays a dual role of stabilizing both P1′:Ile181 and P2′:Met182 side chains (Figure 2C and 4). The above observation is supported by the fact that Phe191Ala mutation showed \( \approx 75 \) fold lower \( k_{\text{cat}} \) than the wild type enzyme (26). All the above three residues are not conserved in various serotypes except for Phe191, corresponding to Phe194 in BoNT A. But Phe194 side chain takes a different rotamer position in both native substrate-free (PDB id 1E1H) and complex structure of BoNT A (PDB id 1XTG) possibly due to presence of different surrounding residues unlike BoNT E. Mutation of P2′:Met182 to Val reduced the catalytic activity of BoNT E on SNAP-25 by \( \approx 25\% \) relative to wild type (27).

The current structure confirms that Arg347 and Tyr350 play a crucial role in transition state stabilization by allowing proper docking of the main chain of P1, P1′ and P2′ residues at the active site.

P3′:Glu183 interactions with 250-loop: In the native substrate-free protein crystal structure (PDB id 1T3A) the electron density for the region 234-244 residues (250 loop) was disordered and thus were not modeled. But in the present structure the electron density is well defined allowing this region to be modeled unambiguously in molecule B (Figure 2A). Ordering of the 250 loop could probably be due to the interactions of P3′:Glu183 with Ile240, Thr241, Asn242, Thr246 of BoNT E (Figure 2C and 4). In the present structure the carboxylate side chain of P3′:Glu183 takes two rotamer positions both involved in multiple interactions with 250 loop. Two additional water molecules and Gly352 are also contributing in stabilizing the P3′ (Figure 2C). Functional importance assay on P3′: Glu183Ala mutant showed a reduced activity by \( \approx 25\% \), suggesting the requirement of Glu for optimal activity (27).

Modeling of P2′:Asp179 in BoNT E structure: P2′:Asp179 is one of the crucial residues for better positioning of the scissile residue of the substrate. Mutation of P2′:Asp179 to Val and Ala reduced the catalytic efficiency of BoNT E by \( \approx 290 \) and 60 fold, respectively (27). Therefore, to trace the S2 pocket in BoNT E, the current and BoNT A-SNAP-25 peptide structures (PDB id 1XTG) (16) are compared. P2′:Asp179 of SNAP-25 is modeled similar to P2′:Asn196 (in BoNT A), close to P1′:Arg180 in BoNT E structure with some minor adjustments in rotation and translation. P2′:Asn196 (in BoNT A) side chain interacts with His227 (BoNT A) ND1 in 1XTG. But here P2′: Asp179 side chain will be stabilized by interactions with Lys224 NZ and His215 ND1 of BoNT E (Figure 5). These interactions are probably stronger than for Asn196 in BoNT A, as Lys224Ala BoNT E mutant showed reduced (\( <80\% \)) catalytic activity (26).

Interestingly, variation at P2 position observed in Human SNAP-23 and Murine SNAP-23 which have 186\{KRI\}187 and 183\{QK\}186 sequence in place of 179\{DRI\}181 of SNAP-25 and show resistance and
poor cleavage respectively by BoNT E (25). P2:Asp179 (SNAP-25) complements well with His215 and Lys224 whereas P2:Lys185 (Human SNAP-23) in its place will not and may affect drastically the scissile bond alignment and so the phenotype. However, P2:Gln184 (Murine SNAP-23) still may complement with Lys224 and may not affect the phenotype that severely.

**S1’ pockets in BoNTs are unique:** The BoNTs E/C/D/G crystals structures (PDB id 1T3A, 2QN0, 2FPQ, 2ZB7) are compared for the integrity of the S1 and S1’ pockets since they have positively charged side chain for P1 (except BoNT G) and hydrophobic for P1’ residue (Table 2). The S1’ pocket of E, C and D are hydrophobic but fine variations in S1’ in each serotype make them very specific for a particular residue. Pro168 in place of Thr159 and also relative narrowing of the S1’ pocket in BoNT C may not allow any bigger side chain residue at P1’. While S1’ pocket is unique in all, BoNT C and D share some similarity. The presence of Ile instead of Ala at position 226 and the flexibility of S1’ pocket (PDB id 2FPQ), might facilitate BoNT D to dock bigger side chain of P1’. However, P1’ mutations to Ala did not completely abrogate and but drastically reduced the activity of BoNTs. Since presence of Ala at P1’ may not allow a stable docking and thus the optimal alignment at the active site required for catalysis. Lower catalytic efficiency of BoNT C (scissile bond: Arg-Ala) could also be due to similar reasons.

**Substrate binding subsites vary between BoNT E and A:** In order to evaluate the variation/similarity among serotypes about substrate recognition/specificity at the active site we structurally aligned BoNT E-SNAP-25 (180-183) with the existing structures of BoNT A with inhibitors: Arg-Arg-Gly-Cys (Pdb id 3C88) and N-Ac-CRATKML (PDB id 3BOO) (Figure 6A and B)(17-19). DALILITE structural alignment showed RMSD of 2.3Å between RIME and RRGC structures for 386 Cα atoms. Although both BoNT E and A recognize SNAP-25 they cleave at different scissile bonds. BoNT E cleaves between P1:Arg180-P1’:Ile181 whereas BoNT A between P1:Gln197-P1’:Arg198. N-Ac-CRATKML has SNAP-25 sequence R198ATKML203 (P1’-P6’).

A structural alignment with the above two structures showed that the position of P1 main chain O of RIME structure aligns reasonably with O of Arg of RRGC and sulfur of cysteine of N-Ac-CRATKML. But the spatial position of amino group of the RIME structure is at variance with RRGC (Figure 6C). Also, the side chain rotamer position of P1:Arg180 is different from side chain of Arg of RRGC (the residue is in bold and underlined for clarity). In our structure P1:Arg180 side chain makes a strong salt bridge with Glu158 side chain whereas side chain of Arg of RRGC interacts with a sulfate ion and Glu164 (Asn160 in BoNT E) in BoNT A. However, in the real substrate P1 is Gln instead of Arg for BoNT A.

Comparison of P1’:I181 with Arg in RRGC reveals variation in spatial positioning as well as docking (Figure 6D). Ile has a smaller side chain and goes deep inside the cavity and is stabilized by the Phe191 (Phe194 in BoNT A), Thr159 (Phe163 in BoNT A) and Thr208 (Thr220 in BoNT A). Residue corresponding to Phe191 exists in similar position in BoNT A but has a different rotamer position in both native substrate-free and complex structures (PDB id 1E1H and 1XTG). Also Thr159 is replaced with a relatively bigger residue, Phe163 (Figure 6D). The long side chain of Arg of RRGC takes a different rotamer position making a salt bridge interaction with the Asp370 (BoNT A). Asp370 (BoNT A) corresponds to Tyr354 in BoNT E which has a different rotamer position and helps in stabilizing P2’ Met182 side chain (Figure 6D).

Interestingly, in N-Ac-CRATKML structure the P1’ position is occupied by the N-terminal acetate moiety. The methyl group of the acetate occupies the same position as CG1 of P1’:Ile181 in RIME structure and is stabilized by Phe163 (Thr159) and Phe194 (Phe191). However, the carboxylate part of acetate interacts with the main chain amino and carbonyl group of enzyme indicating a facilitated docking.

In BoNT E, P2’:Met182 is stabilized in a hydrophobic pocket formed by Phe191, Tyr354, and Tyr356. BoNT A lacks this pocket due to the presence of Asp370 (Figure 6D). Gly in RRGC and Ala in N-Ac-CRATKML do not interact with protein unlike P2’ in BoNT E. P3’:Glu183
interacts with 250 loop residues while Cys (RRGC) takes a different turn (Figure 6B) and is stabilized by the hydrophobic pocket of 250 loop. But the actual P3' residue is Thr which might not bind in a similar manner as Cys. In BoNT A the 250 loop is bigger and has different spatial position than BoNT E 250 loop. Thr in N-Ac-CRATKML stabilized by a water molecule interacting with Asp370 in BoNT A. Interestingly RATKML (P1'-P6') of N-Ac-CRATKML represents P2'-P7' and occupies the S2'-S7' subsites instead of S1'-S6' (Figure 6A). This is because the oxidized cysteine side chain occupies the place of P1 main chain and its N-terminal-acetate part goes in S1' pocket and from then on the substrate residues slide by one. Surprisingly this indicated a very interesting possibility that the docking at S1 and S1’ pocket can decide the overall docking of other residues in BoNT A by challenging the substrate specificity. The P1'-P6' residues dock in S2'-S7' pockets even though they do not match with the actual P2'-P6' residues.

Overall analysis of structures suggested a variation in the S1'-S3' pocket integrity between BoNT E and A. However, the presence of hydrophobic residues at the S1' pocket in BoNT A, Phe194 (Phe 191 in BoNT E) and Phe163 (Thr159 in BoNT E) might allow docking of hydrophobic side chains. But the hydrophobic interactions of both phenylalanines and a different rotamer position of Phe194 unlike Phe191 (BoNT E), cause a shallower S1' pocket in BoNT A than in E. Arg side chain at P1' takes a suitable rotamer position and is stabilized differently than Ile in BoNT E. Also, in BoNT E Phe191 side chain adopts different rotamer positions, opening the S1' pocket for residues like Ile of suitable size to fit in. The absence of a similar residue at Asp370 (BoNT A) position will restrict the binding of residues like Arg or charged residue at P1' and P2’ position in BoNT E.

**Substrate cleavage strategy for BoNT E:** Initially, it was a surprise in our crystal structure to have well-defined electron density for the full tetrapeptide. However, combining and comparing the results of the enzyme inhibition and structural studies led us to better understand the molecular events of binding and cleavage. We assume that the reason the peptide did not get cleaved is due to the requirement of a minimal length as mentioned earlier. However, sequence of I^{178}DRIME^{183} as a N-terminal GST fusion protein preceded by SNAP-25 (93-146) is also shown as cleavable length (27). SNAP-25 166-181 sequence that has C-terminal Arg-Ile showed inhibition (50%) of the catalytic activity of BoNT E at 400μM concentration (27).

Also, our crystal structure clearly reveals that mere presence of P1-P1’ and additional P2’ and P3’ are not sufficient for the cleavage of the substrate but is sufficient for recognition and binding at the active site. Thus both P1 and P1’ with either N- or C-terminal extensions are not sufficient enough to bring the proper cleavage. Rather, they can serve as potential candidates for substrate inhibitors.

To understand BoNT E specificity for P1:Arg181-P1’:Ile181 we analyzed the SNAP-25 sequence (1-206aa). It has at least three regions containing similar scissile bond (RI), L^{57}ERIE^{65}, I^{178}DRIME^{183} and K^{189}TRIDE^{194}. But BoNT E selects only I^{178}DRIME^{183} for cleavage. This may probably be due to the prime requirement of exosite interactions and/or cleavage region in larger peptides. Also, the size and type of P2 and P2’ may play significant roles.

Therefore, we assume that at least three residues on either side of the scissile bond are needed for precise alignment of the scissile bond required for the cleavage. However, all of them may not contribute to substrate affinity or specificity, although longer substrates serve better for high affinity binding and faster scissile bond cleavage. Also, simultaneous substrate recognition both at the exosites and the active site facilitates the high affinity binding and improved rate of cleavage in longer substrates. Together, the unique exosites and subsites at and near active site determine the substrate selectivity and scissile bond specificity, which is unique in all BoNTs.

**Implications in inhibitor design for BoNT E:** The extreme toxicity and fast action of BoNT E in blocking neurotransmission requires advancement in developing countermeasures against it. Though several inhibitors are being developed for BoNT A and B and a few for BoNT F (18,19,28-36), a knowledge gap exists for BoNT E. The
unavailability of structural information regarding substrate-enzyme interactions is the major reason for the existing gap. We for the first time report the substrate binding at the active site of full-length catalytically active enzyme on any BoNTs. Also, the identification of S1-S3’ subsites, hydrophobic nature of S1’ and S2’ pockets and the positive charge nature of S2, can be exploited in designing of better specific inhibitors for BoNT E. Interestingly, the substrate RIME itself acts as an inhibitor and specifically binds at the active site. Overall, the BoNT E-substrate peptide complex crystal structure delivers the specific required information essential for structure-based design of potent inhibitors.

**Catalytic mechanism of BoNT serotype E:** Based on our complex structure analysis and previous mutational analysis of active site residues of BoNT E we propose a mechanism of catalysis for botulinum neurotoxins (Figure 7). Here in our structure Arg181 carbonyl O displaces the nucleophilic water and the main chain N displaces the water coordinated to Glu212. Glu212 being a general base activates a water molecule, which donates OH to the carboxylate group of Arg180 and H to the leaving amino of Ile181. In our structure we do not see any water close to Glu212. Perhaps in the presence of larger substrates with P2/P3 residues and/or S4 SNARE motif (37) there is a possibility of subtle rearrangements at the active site to achieve the transition state. The transition state is often the result of strain or distortion of the reactants to form the particular electronic structure needed for the proper collision and product formation. However, due to smaller substrate length a transition state is not achieved which leads RIME to act as a substrate inhibitor.

**Acknowledgements:** Research was supported by awards from NIH (1R56AI058175), DAMD17-02-2 and DTRA BO742081 under DOE prime contract N0. DEAC02-98CH10886 with Brookhaven National Laboratory. We gratefully acknowledge data collection support from beamline X29 (NSLS).
References:
Footnotes:


Data deposition to RCSB: Coordinates and structure factors deposited to RCSB PBD id 3D3X.
**Figure legends:**

**Figure 1:** BoNT E enzyme inhibition assay with SNAP-25-RIME (P1-P3’) using 146-186aa long peptide substrate. The tetrapeptide showed IC\(_{50}\) of 340 \(\mu\)M (GraphPad software, logEC50=IC\(_{50}\)) and \(K_i\) of 69 \(\mu\)M using the Cheng-Prusoff equation: \(K_i=IC_{50}/1+(S/k_m)\). \([S]\) and \(k_m\) are substrate concentration and Michaelis-Menten constant respectively.

**Figure 2:** The SNAP-25 P1:Arg180-P1’:Ile181-P2’:Met182-P3’:Glu183-amide (RIME) binding and interactions at and near the active site of BoNT E. (A) The cartoon view of BoNT-E-RIME complex structure. The native substrate-free and complex structures vary with the addition of the 250 loop (cyan color cartoon). BoNT E is shown in blue and yellow cartoon. RIME in magenta ball and stick model, and zinc in green sphere model. Only molecule B is shown for the clarity. (B) BoNT E in blue surface model and RIME in white ball and stick model to show that the substrate peptide binds in a cavity at and near the active site. (C) The dashed line in orange shows the interactions of P1-P3’ residues of RIME with various residues of S1-S3’ pocket of BoNT E. The RIME and enzyme residues are shown in green and white color ball and stick model respectively. BoNT E secondary structure is shown in blue loop cartoon. The RIME (P1-P3’) and BoNT E residues are labeled in green and black color respectively. Figures 2,3,5 and 6 are made using PYMOL program (38).

**Figure 3:** The walleye stereo view of substrate electron density omit map contoured at 1\(\sigma\) level at RIME peptide (green ball and stick model, not labeled for clarity). BoNT E residues interacting with RIME are shown in white ball and stick model and labeled black.

**Figure 4:** The schematic representation of interactions and distances of the RIME peptide representing the P1-P1’-P2’-P3’ residues of SNAP-25 with BoNT E residues identifying the S1-S1’-S2’-S3’ pocket. The RIME peptide and BoNT E substrate subsite residues are in blue and grey color stick model. The distances (Å) are given in green colored numbers. The red colored double arrowhead points to the scissile bond of cleavage. Figure is made using Chemdraw program (CambridgeSoft).

**Figure 5:** Modeling of P2:Asp179 and identification of the S2 pocket residues. This modeling is based on P2: Asn196 of SNAP-25 as in complex structure of SNAP-25-BoNT A (PDB id 1XTG), with few adjustments. RIME and BoNT E are in green and white ball and stick model.

**Figure 6:** The superposition of BoNT E-substrate peptide RIME (P1-P3’) with the BoNT A inhibitor peptides; N-Ac-CRATKML (PDB id 3BOO) and RRGC (PDB id 3C88). Here P1, P1’, P2’, P3’ positions in RIME structure is: Arg-Ile-Met-Glu, in RRGC: Arg-Arg-Gly-Cys and in N-Ac-CRATKML: oxidized Cys-Acetate-Arg-Ala. Color code: BoNT E, RIME, BoNT A-RRGC, BoNT A-N-Ac-CRATKML structures in white, green, magenta and cyan color ball and stick model respectively. Zinc is in deep-cyan color for RIME (zinc of other structures is not shown for clarity). (A) The position of CSO, ACE, Arg and Ala in N-Ac-CRATKML structure, aligns well with the P1:Arg-, P1’:Ile-, P2:Met-’ and P3’:Glu- of BoNT E-RIME structure but for BoNT A in actual RATKML represents P1’-P6’. (B) The P1-P1’-P2’ of RRGC aligns with RIME with few differences. Cys of RRGC goes in a very different spatial position which is different from, equivalent P3’:Glu for BoNT E and Ala in N-Ac-CRATKML. (C) Superposition of P1 residues and S1 pockets in all three structures. Only BoNT E residues are labeled. (D) Superposition of P1’ residues and S1’ pockets in all three structures. Black and magenta labels represent BoNT E and A respectively.

**Figure 7:** A mechanism of catalysis and transition state interactions adopted by BoNT E. The precise substrate docking at the active site (stage A) leads to a transition state (stage B) which facilitates cleavage of the scissile bond. Figure is made using Chemdraw program (CambridgeSoft).
Table 1: Data collection and refinement statistics.

**Data collection statistics:**
- Wavelength (Å): 1.08
- Resolution (Å): 50-2.25
- Space group: P2₁2₁2₁
- Cell dimensions: a=89.41 Å, b=144.738 Å, c=83.303 Å, α=β=γ=90°
- Molecules/asymmetric unit: 2
- Redundancy (overall/outermost shell): 13.1(8.5)
- I/σ(I): 11.1 (3)
- Rmerge (overall/outermost shell): 0.06 (0.22)
- Completeness (%)(overall/outermost shell): 94.5(67.3)
- Number of Reflections: 49224

**Refinement statistics:**
- Resolution range (Å): 50-2.25
- Number of reflections: 47715
- Completeness (working+test) (%): 91.6
- Rfactor: 0.23
- Rfree: 0.28
- Number of protein atoms: 6595
- Number of sulfate ions: 8
- Number of water molecules: 251
- B Values:
  - from wilson plot (Å²): 21.2
  - mean B value (overall, Å²): 37.6
  - r.m.s.d. bonds (Å): 0.008
  - r.m.s.d. angles (°): 1.4
- Ramachandran plot analysis:
  - Most favored region (additionally allowed) (%): 85.6(13.2)
  - Generously allowed region (disallowed) (%): 1.0(0.3)
**Table 2:** The comparison of the S1 and S1’ side chain pocket residues of BoNT E with other serotypes (C/D/G) displaying similarity in P1 and P1’ residue by charge and hydrophobicity, respectively. S1 pocket also includes Zinc and Tyr (Tyr350 in BoNT E), common in all serotypes.

<table>
<thead>
<tr>
<th></th>
<th>BoNT E</th>
<th>BoNT C</th>
<th>BoNT D</th>
<th>BoNT G</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1- P1’</td>
<td>Arg-Ile</td>
<td>Arg/Lys-Ala</td>
<td>Lys-Leu</td>
<td>Ala-Ala</td>
</tr>
<tr>
<td>S1 pocket</td>
<td>Glu158</td>
<td>Asp167</td>
<td>Asp167</td>
<td>Asp169</td>
</tr>
<tr>
<td>S1’ pocket</td>
<td>Thr159</td>
<td>Pro168</td>
<td>Tyr168</td>
<td>Asn170</td>
</tr>
<tr>
<td></td>
<td>Phe191</td>
<td>Leu200</td>
<td>Leu200</td>
<td>Asn201</td>
</tr>
<tr>
<td></td>
<td>Thr208</td>
<td>Ile226</td>
<td>Ala226</td>
<td>Thr227</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Met221</td>
<td>Met221</td>
<td>Tyr376</td>
</tr>
</tbody>
</table>
Equation: \( Y = \min + \frac{(\max - \min)}{(1 + 10^{((\LogEC50 - X)\times\text{HillSlope})})} \)
Figure 2:
Figure 3:
Figure 5:
Figure 6:
Figure 7:
Snap-25 substrate peptide (180-183) binds to but bypasses cleavage by catalytically active clostridium botulinum neurotoxin E
Rakhi Agarwal and Subramanyam Swaminathan

J. Biol. Chem. published online July 25, 2008

Access the most updated version of this article at doi: 10.1074/jbc.M803756200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts