Mechanism of substrate recognition by the novel Botulinum Neurotoxin subtype F5

Jiubiao Guo1,2, Edward Wai Chi Chan1,2 & Sheng Chen1,2

Botulinum Neurotoxins (BoNTs) are the causative agents of botulism, which act by potently inhibiting the neurotransmitter release in motor neurons. Seven serotypes of BoNTs designated as BoNT/A-G have been identified. Recently, two novel types of Botulinum neurotoxins, which cleave a novel scissile bond, L54-E55, of VAMP-2 have been reported including BoNT/F subtype F5 and serotype H. However, little has been known on how these BoNTs recognize their substrates. The present study addressed for the first time the unique substrate recognition mechanism of LC/F5. Our data indicated that the optimal peptide required for efficient LC/F5 substrate cleavage is VAMP-2 (20–65). Interestingly, the overall mode of substrate recognition adopted by LC/F5 was similar to LC/F1, except that its recognition sites were shifted one helix toward the N-terminus of VAMP-2 when compared to that of LC/F1. The composition of LC/F5 pockets were found to have changed accordingly to facilitate specific recognition of these new sites of VAMP-2, including the P2′, P1′, P2, P3, B3, B2 and B1 sites. The study provides direct evidence of the evolutionary adaption of BoNT to recognize its substrate which is useful for effective antitoxin and inhibitor development.

Botulism, named after the Latin word “botulus” for sausage, was first described by Justinus Kerner after a food poisoning outbreak associated with ingestion of blood sausages1. Symptoms of botulism, which often originated from food or wound infection, are attributed to the most potent protein neurotoxin known to mankind, namely Botulinum Neurotoxins (BoNTs), which are produced by the bacterium Clostridia botulinum2. Human botulism is serious and fatal. If prompt diagnosis and immediate treatment are not given, the mortality rate of botulism can reach 5–10%; hence early detection and diagnosis is the key to preventing botulism. Moreover, due to its stability and easy-to-deliver/synthesize characteristics, BoNTs are regarded as potential biological weapons in bioterrorism attacks and listed as Category A agent by the Center for Disease Control and Prevention in the United States3,4.

BoNTs are typical neurotransmitter release blocker which inhibits the release of acetylcholine (Ach) by specifically targeting and degrading the SNARE (Soluble N-ethylmaleimide sensitive fusion protein (NSF) Attachment protein Receptor) proteins at the neuromuscular junction. BoNTs belong to the AB toxin family, in which the holotoxin is a ~150 kDa single polypeptide chain which can be functionally divided into two domains: an N-terminal ~50 kDa light chain (LC, catalytic domain) and a ~100 kDa C-terminal heavy chain (HC), both of which are covalently linked through a disulfide bond until they encounter reducing conditions in the neuronal cytosol3. The HC domain is composed of two sub-domains: translocation domain H3 which mediates translocation of LC across the endosomal membrane, and cell surface receptor-binding domain Hc,4,5. In the past decade, seven serotypes of BoNTs (designated as BoNT/A-G) and 31 BoNT subtypes/variants have been identified2,8–16. In addition, an eighth serotype, BoNT/H, has been reported13 and subsequently confirmed to be a hybrid of serotypes F and A characterized with BoNT/A antigenicity and BoNT/F5 light chain function18.

Recently, two novel BoNTs, BoNT/F subtype F5 and serotype H, which cleave a novel scissile bond of VAMP-2 (Vesicle-associated membrane protein 2), namely L54-E55, have been identified. The VAMP-2 is highly conserved and brain-specific, play key role in neuron exocytosis by forming the core SNARE complex with some other proteins19,20. BoNT/F5 exhibits highest homology to serotype F with 46–49% identity in amino acid sequence to the other six subtypes within the BoNT/F serotype21,22. It cleaves the substrate VAMP-2 at a different scissile bond

1Shenzhen Key lab for Food Biological Safety Control, Food Safety and Technology Research Center, Hong Kong PolyU Shen Zhen Research Institute, Shenzhen, P. R. China. 2State Key Lab of Chirosciences, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. Correspondence and requests for materials should be addressed to S.C. (email: sheng.chen@polyu.edu.hk)
from all the other six subtypes, suggesting that it may belong to a novel class of BoNT in terms of biochemical activity. BoNT/H has recently been reported by California Department of Public Health. The novel toxin gene contains regions of similarity to both \( \text{bont/A1} \) and \( \text{bont/F5} \). The LC domain of BoNT/H is 83% homologous to LC/F5, while the HCc domain is nearly identical (>90%) to \( \text{bont/A1} \). However, the HCc domain of this hybrid is less similar (<80%) to the HCc of either \( \text{bont/A1} \) or \( \text{bont/F5} \), which is different from other hybrid BoNT such as C/D hybrid. Consistently, the toxic effects of this hybrid-like BoNT are completely eliminated by existing serotype A antitoxin. However, little has been known on the substrate recognition mechanisms of this novel toxin. In the present study, we investigated the substrate recognition and cleavage mechanism of one of this novel toxin, LC/F5. Findings in the present work could broaden our understanding of the substrate recognition mechanism utilized by BoNTs and provide insights into the development of inhibitors of the novel BoNTs.

**Results**

Since LC/H showed over 80% homology with LC/F5, we selected LC/F5 to study the mechanisms of substrate recognition for these novel BoNTs. GST tagged LC/F5, namely GST-LC/F5 (1–450), was shown to be the most functionally active recombinant protein, it is used for this study. First, the minimal substrate for LC/F5 was determined through screening for the truncation mutations of V AMP-2. V AMP-2 (20–97) and V AMP-2 (1–65) were as efficient as WT V AMP-2 to be cleaved by LC/F5, therefore the minimal substrate was determined to be V AMP-2 (20–65). Next, saturation mutagenesis analysis of V AMP-2(22–66) was performed to dissect the role of individual residue on LC/F5 substrate recognition (Fig. 1). Several residues within the minimal substrate was shown to be critical for LC/F5 substrate cleavage including R31, E41, N49, V50, D51, K52, E55, R56 and Q58, substitution of which to Ala caused ~200-, 80-, 10-, 50-, 10-, 20-, 750-, 4000- and 25-fold reduction of LC/F5 substrate cleavage respectively (Fig. 1), with most of these residues being located within or adjacent to the V1 motif of V AMP-2 (Fig. 1, inlet).

To understand the mechanism of LC/F5 substrate recognition and specificity, the complex structure of LC/F5 and V AMP-2 needs to be determined. However, the three-dimensional complex structure of LC/F5-V AMP-2 is not available currently. Instead, the LC/F5-V AMP-2 complex structure was modeled in the present work. LC/F5 and LC/F1 display a high degree similarity (~48%) in sequence, and the detailed interactions between LC/F1 and V AMP-2 have been dissected previously which was featured by a hydrophobic interaction between a helical structure of V AMP-2 and a hydrophobic interface in LC/F1 (Fig. 2A). By utilizing these information and findings, the LC/F5 structure (Fig. 2B) and the LC/F5-V AMP-2 complex structure (Fig. 2D) were modeled. As expected, a hydrophobic surface and an N-terminus shifted helix structure were identified in LC/F5 (Fig. 2B) and V AMP-2 (Fig. 2C) respectively. Based on the modeled LC/F5-V AMP-2 complex structure, the residues that may specifically interact with the corresponding residues in V AMP-2 were identified and characterized (Fig. 3) to confirm their functional roles through mutational analysis. All LC/F5 mutants were found to exhibit a similar secondary structure profile as that of the wild type LC/F5 protein through far-UV CD analysis (Fig. 4) and partial trypsin digestion (data not shown), suggesting that the conformation of the mutant proteins remains stable.

**LC/F5 active site substrate recognition.** Based on the modeled complex structure of LC/F5-V AMP-2, four potential substrate recognition pockets, S2, S1, S2 and S3 that specifically interact with P2′, P1′, P2 and P3 sites of V AMP-2 were identified and further characterized (Fig. 3).

**LC/F5 S2, -P2, substrate recognition.** LC/F5 was found to cleave the V AMP-2 P2′ site mutant, V AMP-2 (R56A), with ~4000-fold lower efficiency (Fig. 1), indicating that the V AMP-2 P2′ site played an important role on the hydrolysis of LC/F5. Based on the LC/F5-V AMP-2 complex structure, the S2′ pocket in LC/F5 that specifically recognized the positively charged P2′ site at residue R56 of V AMP-2 comprised one negatively charged aspartic acid, D90 and D162 (Fig. 5A). LC/F5 mutations, D90A and D162A, caused a ~10-fold and ~1000-fold reduction in...
the efficiency of VAMP-2 hydrolysis respectively, mainly affecting the substrate catalysis, $k_{\text{cat}}$ (Table 1). In addition, the double mutations, LC/F5 (D70A/D161A), were found to completely abolish LC/F5 activity.

**LC/F5 P1', S1', substrate recognition.** Previous studies proved that the P1' site of VAMP-2 or SNAP-25 played an important role on recognition and hydrolysis by BoNTs25–29. In the present work, the substrate P1' site mutation, VAMP-2 (E55A), was found to reduce the LC/F5 hydrolysis efficiency by ~750-fold, indicating a very important role of the VAMP-2 P1' site in the recognition by LC/F5 (Fig. 1). The S1' pocket of LC/F5 was predicted to comprise a positively charged residue, K218 (Fig. 5A). The LC/F5 (K218A) or LC/F5 (K218D) mutations could completely abolished the catalytic activity of LC/F5, whereas the LC/F5 (K218R) mutation displayed ~8-fold reduction effect on substrate hydrolysis (Table 1). Taken together, the data indicated that the salt bridge between K218 of LC/F5 and E55 of VAMP-2 played important role on substrate recognition (Fig. 5E).

**Dual recognition of VAMP-2 P2 (V53) by the S2 pocket of LC/F5.** The P2 site of VAMP-2 plays a moderate role in LC/F5 substrate recognition, with the V53A change in VAMP-2 reducing LC/F5 substrate hydrolysis by ~10-fold (Fig. 1). Two hydrophobic residues in LC/F5, Y183 and Y239, were predicted to interact with V53 of VAMP-2 (Fig. 5A). The single alanine substitution, Y183A or Y239A, exerted ~3-fold and ~8-fold reduction on the LC/F5 substrate hydrolysis respectively, but the double mutation of LC/F5, which resulted in the Y183A/Y239A changes, displayed ~1000-fold reduction on efficiency of substrate hydrolysis, with the effect mainly on $k_{\text{cat}}$ but not $K_m$ (Table 1). These data indicated that the hydrophobic pocket formed by Y183 and Y239 exhibited the best interactive effects with V53 of VAMP-2, and that the effect of loss of one Tyrosine site could be compensated by the other (Fig. 5E). This dual recognition strategy has previously been observed in the case of LC/D substrate recognition26.
LC/F5 P3-S3 substrate recognition. The alanine substitution of the P3 site of VAMP-2, K52A, exerted ~20-fold reductive effect on the substrate hydrolysis by LC/F5 (Fig. 1). An aspartic acid, D161, in LC/F5 was predicted to partially interact with the K52 of VAMP-2 based on the modeled LC/F5-VAMP-2 complex structure (Fig. 5A). The LC/F5 (D161A) mutant displayed ~1000-fold attenuated VAMP-2 cleavage efficiency by mainly affecting $k_{cat}$ (Table 1). As shown above, the LC/F5 (D161A) mutant displayed a significant effect on the substrate hydrolysis efficiency compared to LC/F5(D70A) might be due to the fact that residue D161 of LC/F5 may exert simultaneous contribution to both of the P2' and P3 sites of VAMP-2 via salt bridge interactions (Fig. 5A,E).

LC/F5 binding pockets interactions with VAMP-2. Three binding pockets, designated as B1–B3, in LC/F5 were identified based on the modeled LC/F5-VAMP-2 complex structure (Fig. 4).
The pockets in LC/F5 that specifically recognize the B region/site of VAMP-2 were designated as B pockets. The VAMP-2 B3 region, which is located adjacent to the scissile bond, is a VAMP-2 helix with a hydrophobic surface that comprised the residues I45, V50 and D51 (Fig. 5B). The amino acid substitutions of I45A, V50A and D51A in VAMP-2 exerted ~4-, ~50- and ~10-fold reduction on LC/F5 substrate hydrolysis respectively (Fig. 1). In the B3 pocket of LC/F5, two tyrosines (Y26 and Y50) and one polar residue (T192) were predicted to interact with the B3 site of VAMP-2 in which residue Y26 of LC/F5 was likely to interact with I45 of VAMP-2 through hydrophobic interaction, and the residues Y50 and T192 were expected to recognize residue V50 of VAMP-2 via salt bridge. (B) In the B3 pocket of LC/F5, the salt bridge between D51 of VAMP-2 and K58 of LC/F5, and the hydrophobic or hydrogen bond interactions between I45 and V50 of VAMP-2 and Y26, Y50 and T192 of LC/F5, are likely to play a role in stabilizing the binding of VAMP-2 with LC/F5 and orienting the scissile bond toward the active site to prompt substrate hydrolysis. (C) Like the B3 site/region interactions, another salt bridge interaction between E41 of VAMP-2 and R133 in the B2 site/region of LC/F5 are likely to exert an effect in stabilizing the interactions of VAMP-2 and LC/F5 and orienting the scissile bond to prompt hydrolysis by LC/F5 as well. (D) In the B1 site/region interactions, the residues E147 and E308 of LC/F5 possibly interact with residue R31 of VAMP-2 through salt bridge formation, with the residue E147 playing a dominant role. The interactions may initiate the recognition and binding between VAMP-2 and LC/F5. The structure of LC/F5 in the right panel is shown as surface with 40% transparency, the residues of VAMP-2 that were found to interact with the corresponding residues of LC/F5 were shown in stick. The active site recognition and binding site interactions of LC/F5 were shown in stick and highlighted based on side chain: negatively charged (red), positively charged (blue) and hydrophobic (gray); (E) The nature and distances of interacting residues.

LC/F5 B3 pocket. The pockets in LC/F5 that specifically recognize the B region/site of VAMP-2 were designated as B pockets. The VAMP-2 B3 region, which is located adjacent to the scissile bond, is a VAMP-2 helix with a hydrophobic surface that comprised the residues I45, V50 and D51 (Fig. 5B). The amino acid substitutions of I45A, V50A and D51A in VAMP-2 exerted ~4-, ~50- and ~10-fold reduction on LC/F5 substrate hydrolysis respectively (Fig. 1). In the B3 pocket of LC/F5, two tyrosines (Y26 and Y50) and one polar residue (T192) were predicted to interact with the B3 site of VAMP-2 in which residue Y26 of LC/F5 was likely to interact with I45 of VAMP-2 through hydrophobic interaction, and the residues Y50 and T192 were expected to recognize residue V50 of VAMP-2 via hydrophobic or hydrogen bond interactions (Fig. 5B,E). The amino acid changes of LC/F5 (Y26A), LC/F5 (Y50A) and LC/F5 (T192A) reduced substrate hydrolysis efficiency by ~60-, ~10- and ~3-fold respectively.

### Figure 5. Specific interactions between VAMP-2 and LC/F5.

(A) At the active site of LC/F5, the S2’ pocket in LC/F5 that specifically recognizes the positively charged residue R56 of VAMP-2 comprises two aspartic acids, D70 and D161. A salt bridge formed between residue D161 of LC/F5 and the P2’ residue of VAMP-2 contributes partially to the interaction. The corresponding S1’ pocket that specifically recognizes the P1’ site of VAMP-2 (E55) in LC/F5 contains a positively charged residue, K218. In the corresponding interaction site of LC/F5, two hydrophobic residues, Y183 and Y239 interact complementarily and synergistically with the P2 site of VAMP-2 (Y53). In the S3-P3 interaction, the D161 of LC/F5 partially interacts with the K52 of VAMP-2 via salt bridge. (B) In the B3 pocket of LC/F5, the salt bridge between D51 of VAMP-2 and K58 of LC/F5, and the hydrophobic or hydrogen bond interactions between I45 and V50 of VAMP-2 and Y26, Y50 and T192 of LC/F5, are likely to play a role in stabilizing the binding of VAMP-2 with LC/F5 and orienting the scissile bond toward the active site to prompt substrate hydrolysis. (C) Like the B3 site/region interactions, another salt bridge interaction between E41 of VAMP-2 and R133 in the B2 site/region of LC/F5 are likely to exert an effect in stabilizing the interactions of VAMP-2 and LC/F5 and orienting the scissile bond to prompt hydrolysis by LC/F5 as well. (D) In the B1 site/region interactions, the residues E147 and E308 of LC/F5 possibly interact with residue R31 of VAMP-2 through salt bridge formation, with the residue E147 playing a dominant role. The interactions may initiate the recognition and binding between VAMP-2 and LC/F5. The structure of LC/F5 in the right panel is shown as surface with 40% transparency, the residues of VAMP-2 that were found to interact with the corresponding residues of LC/F5 were shown in stick. The active site recognition and binding site interactions of LC/F5 were shown in stick and highlighted based on side chain: negatively charged (red), positively charged (blue) and hydrophobic (gray); (E) The nature and distances of interacting residues.
The triple mutations which resulted in the formation of the mutant protein of LC/F5 (Y26A/Y50A/T192A) were found to reduce the hydrolytic activity of LC/F5 on V AMP-2 by ~600-fold (Table 1). In addition, one positively charged residue in the B3 pocket of LC/F5 (K58) (Fig. 5B) was predicted to interact with residue D51 of V AMP-2 (Table 1). The K58A substitution exerted ~250-fold reduction on LC/F5 substrate hydrolysis (Table 1). Surprisingly, the residue R31 was important for LC/F5 substrate recognition. Based on the modeled LC/F5-V AMP-2 complex structure, we conclude that the possible salt bridge formed between R31 of V AMP-2 and the residues E147 and E308 of LC/F5 (E147A/E308A) (Table 1). The charge reversal substitution of E147R, displayed ~80-fold reduction on LC/F5 activity (Table 1), suggesting that recognition at the B3 site of V AMP-2 by LC/F5 pocket facilitates the orientation and tuning of substrate for further active site recognition and hydrolysis, in addition to its primary role on substrate binding.

**Discussion**

In the present report, the detailed substrate recognition and cleavage mechanism employed by LC/F5 was uncovered for the first time. Saturation mutagenesis analysis on substrate V AMP-2 revealed that the minimal peptide region required for efficient substrate cleavage by LC/F5 was V AMP-2 (20–65), in which both the P1′ (E55) and P2′ (R56) site residues played a very significant role in substrate recognition by LC/F5 (Table 1), which is different from other serotypes of BoNT, for which the P1′ residue played the most important role on substrate hydrolysis. The general mechanism of substrate recognition by LC/F5 could be summarized as follows based on our

| LC/F5 pockets | VAMP-2 site/region (residue) | LC/F5 derivatives | Activity reduction (fold) | Km (μM) | Kcat (s⁻¹) | kcat/Km (s⁻¹ μM⁻¹) |
|---------------|-----------------------------|-------------------|--------------------------|--------|-----------|---------------------|
| **Active Site** | | | | | | |
| S2’ | P2’ (R36) | LC/F5 (1–450) | 1 | 5.13 ± 1.75 | 938.90 | 183.02 |
| | | D161A | 1000 | 1.81 ± 0.69 | 1.58 | 0.87 |
| | | D70A | 10 | 4.11 ± 1.41 | 131.1 | 31.90 |
| | | D70A/D161A | ND | ND | ND | ND |
| S1’ | P1’ (E55) | K218A | ND | ND | ND | ND |
| | | K218D | ND | ND | ND | ND |
| | | K218R | 8 | 8.27 ± 2.65 | 192.8 | 23.31 |
| S2 | P2 (V51) | Y183A | 3.2 | – | – | – |
| | | Y239A | 8 | – | – | – |
| | | Y183A/Y239A | 1000 | 4.91 ± 2.41 | 0.85 | 0.17 |
| S3 | P3 (K52) | D161A | 1000 | 1.81 ± 0.69 | 1.58 | 0.87 |
| | | D58A | K58A | 250 | 16.17 ± 9.00 | 4.5 | 0.28 |
| | | V50A | 10 | 14.52 ± 7.36 | 402.8 | 27.74 |
| | | T192A | 3 | – | – | – |
| | | Y26A | 60 | 1.29 ± 0.25 | 18.43 | 14.29 |
| | | Y26A/Y50A/T192A | 600 | 6.55 ± 2.33 | 3.2 | 0.49 |
| **B3** | | | | | | |
| | D15 | I183A | 300 | 5.80 ± 3.76 | 2.01 | 0.35 |
| | E147A/E308A | 5 | – | – | – | – |
| | E147R | 80 | 6.07 ± 2.58 | 23.49 | 3.87 |
| | E308R | 0.8 | – | – | – | – |

Table 1. Efficiency of VAMP-2 hydrolysis and kinetic constants of LC/F5 and derivatives. aThe ratio of the amount of LC/F5 derivatives/LC/F5 wt needed to cleave 50% of V AMP-2 wt. bND, not detectable. The mutant was too inactive for determination of its kinetic constants in the present study. cKinetic constants were not determined.
biochemical data. LC/F5 recognition of substrate VAMP-2 was initiated through interaction between the B1 site of VAMP-2 (R31) and the B1 pocket of LC/F5, which is composed of the residue E137. The R31-E137 interaction was thought to be important in initiating the first step of substrate recognition. LC/F5 then further interacted with the B2 site of VAMP-2 which comprised E41, a key residue recognized by the B2 pocket of LC/F5, which comprised residue R133, via the formation of a salt bridge. Interaction between these two binding sites facilitates further binding of LC/F5 to the B3 site of VAMP-2 through multiple interactions including hydrophobic interaction between I45 and V98 of VAMP-2 and Y40 and T192 of LC/F5, as well as formation of a salt bridge between D51 of VAMP-2 and K39 of LC/F5. The B3 pocket recognition event facilitates stabilization and fine tuning of the VAMP-2 structure for more efficient LC/F5 active site recognition rather than strong substrate binding. The initial substrate binding and stabilization between LC/F5 and VAMP-2 facilitate further recognition of the VAMP-2 scissile bond by the active site of LC/F5, where P3-S3, P2-S2, P1′-S′1 and P2′-S′2′ interactions and recognitions were mediated by various interaction events including salt bridge interaction between K52-D161 (P3-S3), the hydrophobic interactions between the V53, Y185/Y239 (P2-S2), and the E55-K135 (P1′-S′1) and the R56-D161/D70 (P2′-S′2′) salt bridges. Interestingly, our results indicate that amino acid D161 of LC/F5 may simultaneously interact with residues in the P3 (K52) and P2′ (R56) sites of VAMP-2 (Fig. 5A). These interactions stabilize the alignment of P1 and P1′ sites of VAMP-2 onto the zinc ion located at the active site of LC/F5 and initiate the substrate hydrolysis process.

Results of biochemical characterization of different LC/F5 residues based on the prediction of LC/F5-VAMP-2 complex structure are consistent with our modeled complex structure. Comparison of the mechanisms of substrate recognition employed by LC/F1 and LC/F5 could provide insight into the evolution of substrate recognition by LC/F subtypes. From the modeled and biochemically proven complex structure of LC/F5-VAMP-2 (Fig. 6A), the helix in VAMP-2 was formed by residues I45MRV50 whereas in the complex structure of LC/F1-VAMP-2, the N-terminal α-helix in the VAMP-2 structure in the modeled LC/F1-VAMP-2 complex was extracted and refined from reference28 with permission.

Figure 6. An overall comparison of the four-residue N-terminal shifted α-helix in the VAMP-2 structure. When comparing the modeled LC/F5-VAMP-2 (A) and the LC/F1-VAMP-2 (B) complex structures, the VAMP-2 α-helix in the later was found to be shifted four residues toward the N-terminal, the shift prompts VAMP-2 to better fit the cleft in the LC/F5 binding pockets and active site for efficient substrate cleavage. The modeled VAMP-2 structure is represented in cartoon with the shifted α-helix colored in red, and so were the corresponding residues. The VAMP-2 structure in the modeled LC/F1-VAMP-2 complex was extracted and refined from reference25 with permission.

Experimental Materials and Methods

LC/F expression and purification. To facilitate the expression of target protein, the corresponding sequence to full length LC/F residues 1–450 (GenBank: ADA79579.1) was optimized for the codons preferred in E. coli, then synthesized by Tech Dragon (Hong Kong, China) and sub-cloned into a pGEX-2T vector through the SacI/BamHI restriction sites. In addition, the human VAMP-2 (1–97) construct was synthesized as described previously28. All the mutated derivatives of LC/F5 and VAMP-2 were performed by using the QuickChange (Stratagene) commercial kit following the manufacturer’s instruction and confirmed by sequencing in BGI (Shenzhen, China). Based on our previous study of the expression and characterization of LC/F524, GST tagged full length LC/F5 (1–450) were found to display the highest solubility and activity when compared with other versions of LC/F5. Thus in the present study, all the assays were carried out by using purified GST-LC/F5 (1–450).
protein. Purification of GST-LC/F5 (1–450), VAMP-2 (1–97) and all other derivatives was performed as previously described24,30.

**Standard Linear Velocity Reaction.** Linear velocity assays were performed as previously described29. Briefly, in 10 μl reaction mixture, 5–10 μM VAMP-2 or derivatives was mixed with an indicated amount of LC/F5 in 10/20 buffer (10 mM Tris-HCl, pH 7.6, 20 mM NaCl). After 20 min incubation at 37 °C, the reactions were stopped by adding SDS-PAGE sample buffer, heated at 100 °C for 5 min, analyzed by SDS-PAGE. The amount of VAMP-2 cleaved was determined by densitometry.

**Determination of Kinetic Parameters.** As described previously26, the procedures for $k_{cat}$ and $K_m$ determination were almost the same as mentioned above, but the amount of LC/F5 or its derivatives used was adjusted to achieve <10% cleavage of VAMP-2, the concentrations of which ranged from 1 to 72 μM. Reaction velocity against substrate concentration was fitted into the Michaelis–Menten equation and kinetic constants were derived using the GraphPad program. For each protein, at least three independent assays were performed to determine the kinetic constants.

**Far-UV Circular Dichroism Analysis.** As detailed previously26, LC/F5 and derivatives were analyzed by far-UV CD for assessing the secondary structure change in a 10-mm path length quartz cuvette of 400 μl volume (containing 0.1–0.4 mg/ml protein in 10:20 buffer). A JASCO J-810 spectropolarimeter was used with the following parameters: scanning speed 50 nm/min, 1 s response time, 1 nm data pitch, 1 nm band width and accumulation times was set as 3. The wavelength range of 190–250 nm was scanned, and raw CD data were converted to molar ellipticity using Yang as reference31, and the spectrum was generated using GraphPad Prism.

**Molecular Modeling.** The complex structure of LC/F5-VAMP-2 was modeled by using SWISS-MODEL and refined with PyMoL software as detailed previously26. Briefly, the structure of LC/F5 (1–450) was modeled by SWISS-MODEL using the crystal structure of LC/F (PDB 2ABA) as searching template and the structure of VAMP-2 was extracted from the SNARE complex crystal structure (chain A, PDB 1SFC), and both structures were modified by PyMoL. The LC/F5-VAMP-2 complex structure was then modeled by aligning to the LC/F-VAMP-2 inhibitor complex structure (PDB 3FIE) and refined in the PyMoL software.

**References**

1. Erbguth, F. J. Historical notes on botulism, Clostridium botulinum, botulinum toxin, and the idea of the therapeutic use of the toxin. Mov Disord. 19, 52–66, doi: 10.1022/msd.2003 (2004).
2. Hill, K. K. & Smith, T. J. Genetic diversity within Clostridium botulinum serotypes, botulinum neurotoxin gene clusters and toxin subtypes. Curr Top Microbiol Immunol. 364, 1–20, doi: 10.1007/978-3-642-33570-9_1 (2013).
3. Arndt, J. W. Independent evolution of neurotoxin and flagellar genetic loci in proteolytic Clostridium botulinum. Infect Genomics 10, 115, doi: 10.1186/1471-2164-10-115 (2009).
4. Smith, T. J. et al. Sequence variation within botulinum neurotoxin serotypes impacts antibody binding and neutralization. Infect Immun 73, 5450–5457 (2005).
5. Arrati, J. W. et al. A structural perspective of the sequence variability within botulinum neurotoxin subtypes A1–A4. J Mol Biol 362, 733–742 (2006).
6. van Ermengem, E. Classics in infectious diseases. A new anaerobic bacillus and its relation to botulism. E. van Ermengem. Originally published as “Ueber einen neuen anaeroben Bacillus und seine Beziehungen zum Botulismus” in Zeitschrift für Hygiene und Infektionskrankheiten 26: 1–56, 1897. Reviews of infectious diseases 1, 701–719 (1979).
7. Burke, G. S. Notes on Bacillus botulinus. J Biol Chem. 4, 555–570 (1919).
8. Bengston, I. A. Standardization of Botulinum Antitoxins. Am J Public Health (NY) 11, 352–357 (1921).
9. Moller, V. & Scheibl, I. Preliminary report on the isolation of an apparently new type of Cl. botulinum. Acta Pathol Microbiol Scand. 48, 80 (1960).
10. Gimenez, D. F. & Ciccarelli, A. S. Clostridium botulinum type F in the soil of Argentina. Appl Microbiol 16, 732–734 (1968).
11. Gimenez, D. F. & Ciccarelli, A. S. Another type of Clostridium botulinum. Zentralbl Bakteriol Orig. 215, 221–224 (1970).
12. Dover, N., Barash, J. R., Hill, K. K., Xie, G. & Arnon, S. S. Molecular Characterization of a Novel Botulinum Neurotoxin Type H Gene. J Infect Dis. 145, 192–202, doi: 10.1093/infdis/jit450 (2014).
13. Kalb, S. R. et al. Functional Characterization of Botulinum Neurotoxin Serotype H as a Hybrid of Known Serotypes F and A (BoNT/A). J Biol Chem. 281, 631–643 (2006).
14. Hussian, S. & Davanger, S. Post-synaptic VAMP/Synaptobrevin Facilitates Differential Vesicle Trafficking of GluA1 and GluA2 AMPA Receptor Subunits. Plos One 10, e0140862, doi: 13.17137/journal.pone.0140862 (2015).
15. Simons, S. E. et al. A Novel Botulinum Toxin, Previously Reported as Serotype H, has a Hybrid Structure of Known Serotypes A and F that is Neutralized by Serotype A Antitoxin. J Infect Dis, doi: 10.1093/infect/dis327 (2015).
16. Wu, K. & Chen, S. Expression and biochemical characterization of light chains of botulinum neurotoxin subtypes F5 and F7. Protein Expr Purif. 111, 87–90 (2015).
17. Chen, S. & Wan, H. Y. Molecular mechanisms of substrate recognition and specificity of botulinum neurotoxin serotype F. Biochim Biophys Acta 1833, 277–284 (2011).
18. Guo, J. & Chen, S. Unique Substrate Recognition Mechanism of the Botulinum Neurotoxin D Light Chain. J Biol Chem. 288, 27881–27887, doi: 10.1074/jbc.M113.491134 (2013).
27. Chen, S., Karalewitz, A. P. & Barbieri, J. T. Insights into the different catalytic activities of Clostridium neurotoxins. *Biochemistry* **51**, 3941–3947, doi: 10.1021/bi3000998 (2012).
28. Chen, S., Hall, C. & Barbieri, J. T. Substrate recognition of VAMP-2 by botulinum neurotoxin B and tetanus neurotoxin. *J Biol Chem.* **283**, 21153–21159, doi: 10.1074/jbc.M800611200 (2008).
29. Chen, S., Kim, J. J. & Barbieri, J. T. Mechanism of substrate recognition by botulinum neurotoxin serotype A. *J Biol Chem* **282**, 9621–9627 (2007).
30. Chen, S. & Wan, H. Y. Molecular mechanisms of substrate recognition and specificity of botulinum neurotoxin serotype F. *Biochemical Journal* **433**, 277–284, doi: 10.1042/bj20101310 (2011).
31. Yang, J. T., Wu, C. S. & Martinez, H. M. Calculation of protein conformation from circular dichroism. *Methods Enzymol.* **130**, 208–269 (1986).
32. Chen, S. & Barbieri, J. T. Multiple pocket recognition of SNAP25 by botulinum neurotoxin serotype E. *J Biol Chem.* **282**, 25540–25547, doi: 10.1074/jbc.M701922200 (2007).

**Acknowledgements**

We are grateful for Mr. Sam Ka-chi Hui to create some mutations and express some proteins for this project. This work was supported by the General Research Fund from the Hong Kong University Grant Council (for SC, 560211).

**Author Contributions**

S.C. designed the research, analyzed the data and edited the manuscript, J.G. performed the research, analyzed the data and drafted the manuscript and Edward W.C.C. analyzed the data and edited the manuscript.

**Additional Information**

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Guo, J. *et al.* Mechanism of substrate recognition by the novel Botulinum Neurotoxin subtype F5. *Sci. Rep.* **6**, 19875; doi: 10.1038/srep19875 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/