

SCIENTIFIC REPORTS

OPEN

Structural characterisation of the catalytic domain of botulinum neurotoxin X - high activity and unique substrate specificity

Geoffrey Masuyer¹, Sica Zhang², Sulyman Barkho², Yi Shen², Linda Henriksson¹, Sara Košenina¹, Min Dong² & Pål Stenmark¹

Botulinum neurotoxins (BoNTs) are among the most potent toxins known and are also used to treat an increasing number of medical disorders. There are seven well-established serotypes (BoNT/A-G), which all act as zinc-dependent endopeptidases targeting specific members of the SNARE proteins required for synaptic vesicle exocytosis in neurons. A new toxin serotype, BoNT/X, was recently identified. It cleaves not only the canonical targets, vesicle associated membrane proteins (VAMP) 1/2/3 at a unique site, but also has the unique ability to cleave VAMP4/5 and Ykt6. Here we report the 1.35 Å X-ray crystal structure of the light chain of BoNT/X (LC/X). LC/X shares the core fold common to all other BoNTs, demonstrating that LC/X is a bona fide member of BoNT-LCs. We found that access to the catalytic pocket of LC/X is more restricted, and the regions lining the catalytic pocket are not conserved compared to other BoNTs. Kinetic studies revealed that LC/X cleaves VAMP1 with a ten times higher efficiency than BoNT/B and the tetanus neurotoxin. The structural information provides a molecular basis to understand the convergence/divergence between BoNT/X and other BoNTs, to develop effective LC inhibitors, and to engineer new scientific tools and therapeutic toxins targeting distinct SNARE proteins in cells.

The clostridial neurotoxins are a family of bacterial toxins including seven BoNTs and the related tetanus neurotoxin (TeNT). They are the causative agents of the severe paralytic diseases, botulism and tetanus, respectively¹. BoNTs are the most poisonous protein toxins known to man, and yet have also been successfully used clinically to treat an ever-increasing number of disorders, such as strabismus, blepharospasm, cervical dystonia, overactive bladder, and pain².

BoNTs consist of a light chain (LC, ~50 kDa) and a heavy chain (HC, ~100 kDa) linked by an inter-chain disulphide bridge³. LC is a zinc-dependent endopeptidase, while HC is composed of two functional domains that are responsible for translocation (H_N) and receptor binding (H_C)⁴. The toxin acts by first recognizing specific receptors at motor nerve terminals and enters neurons via receptor-mediated endocytosis. The acidic pH in endosomes then causes a conformational change of the toxin, resulting in translocation of LC across the endosomal membrane³. LC targets one of three members of the SNARE family. BoNT/A, /C and /E cleave the peripheral membrane protein SNAP-25. BoNT/B, /D, /F, and /G cleave the vesicle-associated membrane protein VAMP1/2/3. BoNT/C can also cleave syntaxin 1⁴. Cleavage of these SNARE proteins blocks fusion of the synaptic vesicles to the plasma membrane and thus inhibits neurotransmission release and causes paralysis.

BoNT-LCs are remarkably specific proteases and the three sets of neuronal SNARE proteins are the only known targets. Furthermore, each BoNT has its own unique cleavage site on their substrates. This specificity is due to extensive toxin-substrate interactions between regions (designated exosites) outside the conserved catalytic site in LCs and regions in the substrates beyond the cleavage site. In particular, a conserved motif in SNARE proteins known as the SNARE motif has been proposed to play crucial roles for recognition by toxins⁵.

¹Department of Biochemistry and Biophysics, Stockholm University, SE-106 91, Stockholm, Sweden. ²Department of Urology, Boston Children's Hospital, Department of Microbiology and Immunobiology and Department of Surgery, Harvard Medical School, Boston, MA, 02115, USA. Geoffrey Masuyer and Sica Zhang contributed equally to this work. Correspondence and requests for materials should be addressed to M.D. (email: min.dong@childrens.harvard.edu) or P.S. (email: stenmark@dbb.su.se)

For instance, VAMP2 contains two copies of the SNARE motif (V1 and V2) and V1 is critical for efficient cleavage of VAMP2 by BoNT/D⁷. Indeed, the co-crystal structure of LC/A in complex with SNAP-25 has defined two exosites (α - and β -) that interact with specific regions in SNAP-25⁷, and the crystal structure of LC/F in complex with peptides derived from VAMP2 also revealed three exosites that interact with VAMP2⁸.

The crystal structures of all BoNT-LCs have been resolved^{9–16}. Despite the low degree of identity at the amino acid levels, all LCs display a highly conserved fold, presenting a compact globular aspect with mixed secondary structures of α -helices and β -strands. The catalytic pocket, which contains the HEXxE zinc-dependent protease motif, showed similar composition and geometries across all BoNTs. It is likely that variations in the composition and placement of exosites determine which SNARE proteins can be cleaved, as well as the spatial location of the cleavage site on a SNARE protein. In addition, it has been shown that the residue located at the C-terminal side of the cleavage site (P1' position) in SNARE proteins is critical for efficient cleavage, and mutations at this site usually abolishes the cleavage by BoNTs. For instance, although BoNT/C cleaves both SNAP-25 and syntaxin 1, the P1' position in these two different substrates are conserved (Ala).

A new serotype of BoNT with unique substrate specificity was recently identified from the genome sequence of *C. botulinum* strain 111, and designated BoNT/X¹⁷. BoNT/X cleaves VAMP1/2/3 between Arg66-Ala67 in VAMP2, which is distinct from the sites targeted by other BoNTs¹⁷. Remarkably, BoNT/X is the first and only BoNT member that can cleave additional VAMP family proteins, VAMP4, VAMP5 and Ykt6. The cleavage site on VAMP4, VAMP5, and Ykt6 are at the homologous location as Arg66-Ala67 in VAMP2, but the residues at the cleavage site are different (Lys-Ser in VAMP4 and Ykt6, Arg-Ser in VAMP5). Thus, LC/X showed the unique capability to recognize a broad range of VAMPs and a remarkable degree of tolerance for residue changes at the cleavage site.

To understand the structural basis of LC/X action, here we resolved the high resolution X-ray crystal structure of LC/X and of its zinc-free, apo-form. LC/X presents common features with all other clostridial neurotoxins. However the detailed structural characterisation also highlights unique features that may confer LC/X with its unusual substrate specificity and enhanced cleavage efficiency.

Results and Discussion

Crystal structure of LC/X. LC/X (residues 1–439) was recombinantly produced in *E. coli*, with an N-terminal poly-His(x6) tag. Two crystal forms were obtained under distinct crystallization conditions. In the first form, data from a single crystal were collected up to 1.35 Å resolution (Table 1) and the structure of LC/X was solved using molecular replacement with LC/B as a template. Residues 1–413 were clearly defined, however no electron density was observed for the N-terminal tag and the C-terminal end (414–439), all in solvent-accessible areas. In the second crystal form, the metal chelator EDTA was included in the crystallisation condition, which resulted in the loss of the Zn²⁺ ion from LC/X. Data from a single plate-like crystal were collected up to 2.4 Å resolution and comprises residues 1–423 (Table 1).

Despite only ~30% identity at the amino acid levels with other BoNTs, the overall structures of LC/X present a fold similar to other BoNTs with largely conserved secondary structural elements (Fig. 1). This high degree of similarity between LC/X and other BoNTs are reflected by the low root mean square deviations (rmsd) values between 2.3 and 2.9 Å when compared to other BoNT-LCs (Fig. 1B, Table 2, Supplementary Fig. S1). Interestingly, two of the BoNT-LCs that showed the highest sequence identity to LC/X (LC/B, 36%; LC/Te, 34%, for the structurally aligned residues) showed the least structural resemblance (rmsd of 2.8 Å and 2.9 Å, respectively). These differences between the structures rest mainly on the higher variations in the length and position of the flexible loop regions (Fig. 1C, Supplementary Fig. S1). LC/D, which also cleaves VAMP1/2/3, is the closest structural homologue (rmsd = 2.3 Å; Z-score = 47.1). However there was no general correlation between structural similarity and substrate preference, contrary to what was observed in the primary sequence pairwise alignments where the SNAP-cleaving BoNT have the lowest similarity to LC/X (Table 2). These data demonstrate that LC/X is a bona fide member of BoNT-LCs.

Notable structural differences between LC/X and other BoNT-LCs include the loop region 195–217 (Fig. 1A). It starts and ends with anti-parallel β -strands, which are longer than the corresponding regions in LC/B and orientate the rest of the unstructured loop towards the solvent accessible surface. While the majority of the LC fold usually remains identical no matter whether the LC is an isolated protein or within the context of the full-length toxin (holotoxin), similar loop regions in other BoNT-LCs have showed flexibility to accommodate interactions with BoNT-HC^{3,18}, suggesting that the loop 195–217 might contribute to inter-domain interactions within BoNT/X. Indeed, superposition of LC/X with BoNT/B hints that loop 195–217 may interact with the central helices of the translocation domain (Fig. 2).

In addition to loop 195–217, a major difference was observed for one of the loops involved in SNAP-25 binding by LC/A known as loop-250⁷, which in LC/X makes a three-stranded β -sheet arrangement from residues 243 to 263 and extends away from the catalytic channel, in a structural feature unique to LC/X. Loop-250 was shown to take on a different conformation in LC/A where it went from an open configuration in its free form to a closed configuration upon substrate binding⁷. Interestingly this loop corresponds to a random coil in other VAMP-cleaving LCs, with the exception of LC/F where it makes a β -hairpin, all have a length of 16 to 20 amino acids (Supplementary Fig. S2). Superposition of LC/X with BoNT/B suggests that region 243–263 could interact with the translocation domain in full length BoNT/X (Fig. 2), in agreement with this loop's apparent flexibility.

Active site of LC/X and the structure of zinc-free LC/X. BoNTs are metallo-endopeptidases of the M27 family¹⁹ related to thermolysin, with a classical HEXxH+E tetrahedral zinc-binding motif. The active site of LC/X has the overall conserved geometry shared with other BoNTs (Supplementary Fig. S3). The Zn²⁺ ion is coordinated by the triad His227, Glu228 and His 231 and a fourth coordination provided by a water molecule linked to Glu266 (Fig. 3A). The water molecule provides the nucleophile base required for the proteolytic activity.

	LC/X	Apo-LC/X
PDB code	6F47	6F4E
Synchrotron source, λ	Diamond I02, 0.979 Å	SLS PXI, 1.00 Å
Resolution (Å)	1.35 (1.37–1.35)	2.4 (2.49–2.40)
Space group	$P2_12_12_1$	$P22_12_1$
Cell dimensions and angles	a = 58.3, b = 86.7, c = 93.1 Å; $\alpha = \beta = \gamma = 90^\circ$	a = 52.3, b = 94.5, c = 113.0 Å; $\alpha = \beta = \gamma = 90^\circ$
Total/Unique reflections	592,965 98,954	176,299 21,016
Completeness (%) ^a	94.9 (64.8)	93.4 (88.9)
$R_{\text{merge}}^{\text{a,b}}$	0.083 (0.98)	0.123 (0.97)
$R_{\text{pim}}^{\text{a,c}}$	0.053 (0.74)	0.062 (0.52)
$I/\sigma(I)$ ^a	9.2 (1.0)	8.9 (2.1)
$CC_{1/2}$ ^d	0.99 (0.42)	0.99 (0.69)
Multiplicity	6.0 (3.9)	8.4 (7.8)
$R_{\text{cryst}}^{\text{e}}$ (%)	14.3	21.5
$R_{\text{free}}^{\text{f}}$ (%)	17.6	26.0
Rmsd in bond lengths (Å)	0.009	0.009
Rmsd in bond angles (°)	1.34	1.31
B-factor statistics (Å²)		
Protein all atoms	20.3	53.6
Protein main chain atoms	18.4	51.4
Protein side chain atoms	22.2	55.7
Solvent atoms	36.0	46.8
Zn ²⁺ ion	18.2 (0.33 occupancy)	n/a
Ramachandran statistics (Molprobit)		
Favored	97.7%	97.1%
Outliers	0.0%	0.0%

Table 1. Crystallographic statistics of the LC/X X-ray crystal structures. ^aValues in parentheses refer to the highest resolution shell. ^b $R_{\text{merge}} = \sum_i |I_h - I_{hi}| / \sum_i I_h$, where I_h is the mean intensity for reflection h . ^c $R_{\text{pim}} = \sum_h (1/n_h - 1) \sum_i |I_{hi} - (I_h)| / \sum_h \sum_i (I_h)$. ^dCorrelation coefficient between random half datasets. ^e $R_{\text{cryst}} = \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are measured and calculated structure factors, respectively. ^f $R_{\text{free}} = \sum ||F_o| - |F_c|| / \sum |F_o|$, calculated from 5% of the reflections selected randomly and omitted during refinement.

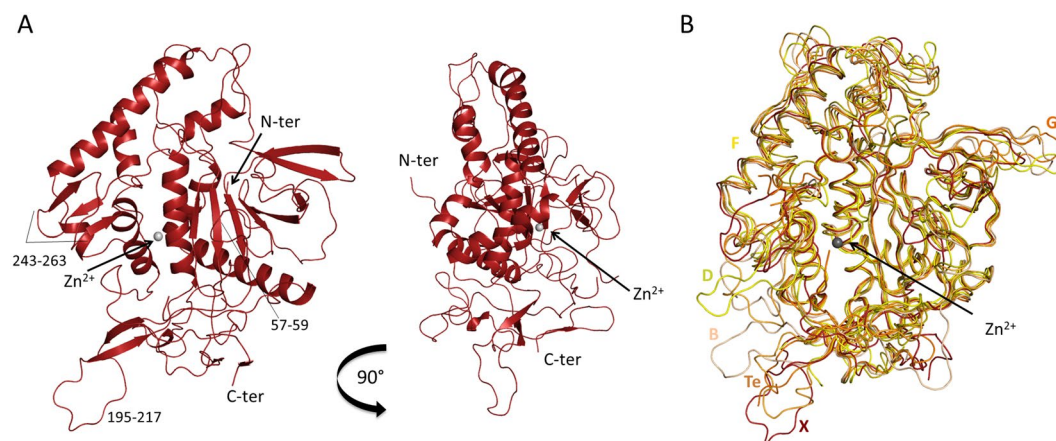


Figure 1. Crystal structure of LC/X. (A) Ribbon representation of LC/X, the zinc ion (grey sphere) and other regions of interest are highlighted. (B) Comparison of LC/X with other VAMP-cleaving toxins. LC of serotypes B (sand, PDB 1EPW), D (yellow-green, PDB 2FPQ), F (gold, PDB 2A97), G (dark orange, PDB 1ZB7) and Te (light orange, PDB 1Z7H) were superimposed with LC/X (red).

Additionally, the conserved Tyr363 is believed to stabilize catalytic intermediates¹⁵ and is within water-mediated hydrogen bond distance to the active site.

In the second crystal form, which lacks Zn²⁺, the architecture of the active site did not present any rearrangement, demonstrating that the role of zinc is functional rather than structural (Fig. 3B). Only weak electron density

	Dali					Needle			
	PDB	rmsd	No. of aligned positions	No. of residues in matched structure	% Sequence identity of aligned positions	Z score	% Sequence similarity	% Sequence identity	Score
LC/D	2fpq	2.3	377	414	31	47.1	48	28	500
LC/F	2a97	2.3	365	392	33	46.4	47	30	520
LC/C	2qn0	2.3	384	429	31	45.8	47	31	507
LC/A	1xtf	2.4	389	427	32	45.4	44	29	465
LC/G	1zb7	2.4	367	414	36	44.0	49	33	546
LC/E	1t3a	2.5	370	400	30	47.0	45	29	434
LC/B	1f82	2.8	381	424	36	45.7	48	33	590
LC/Te	1z7h	2.9	384	426	34	46.5	48	31	549

Table 2. Pairwise comparison of LC/X with other clostridial neurotoxins using Dali³⁷ and Needle (EMBOSS server).

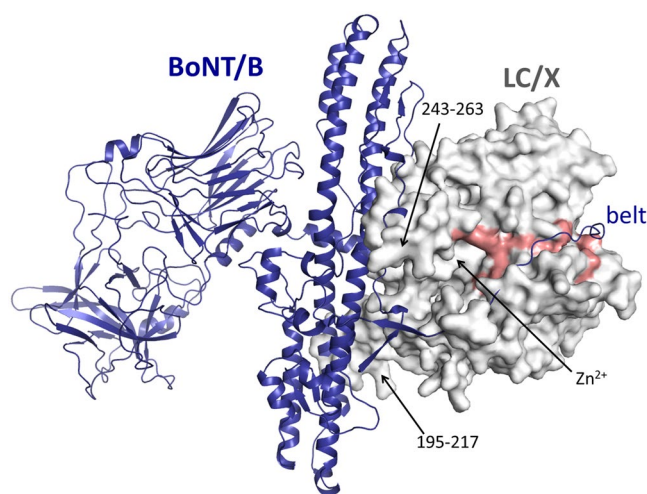


Figure 2. Superposition of LC/X with BoNT/B. LC/X was superposed to the LC of BoNT/B (PDB code 1EPW¹⁶) to illustrate the potential interaction with the translocation domain. LC/X is shown as a white surface while BoNT/B is showed as a blue ribbon (with its LC omitted from the structure for clarity). Residues flanking the groove of LC/X are highlighted in red.

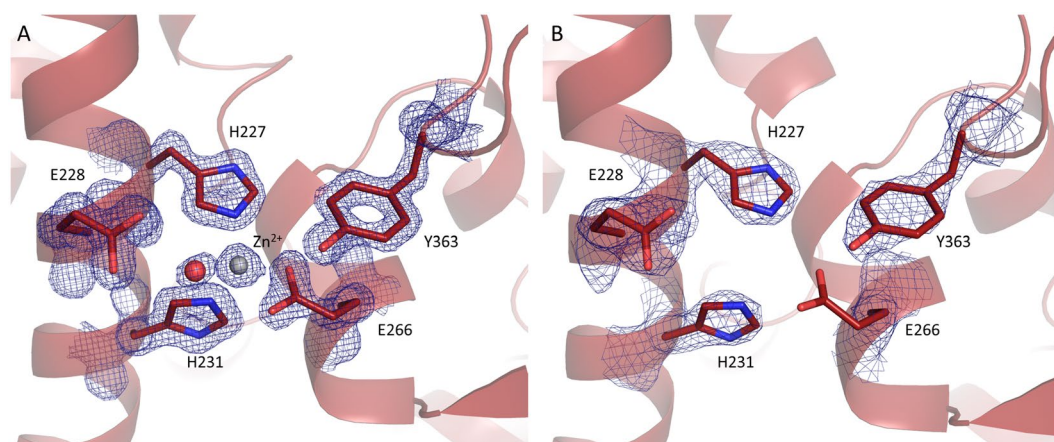


Figure 3. Catalytic site of LC/X. Close-up view of the active site of LC/X in (A) the high-resolution structure of LC/X; and (B) the Apo-LC/X structure. Residues involved in proteolysis represented in sticks, zinc ion and water molecules as grey and red spheres, respectively. The electron density ($2F_o - F_c$, contoured at 2σ in both (A) and (B)) is shown as a blue mesh.

could be observed for Glu266, which may be an indication of the side chain motility in absence of Zn^{2+} coordination. Additionally, residues 411–421 present a unique beta hairpin configuration in the apo-LC/X structure. This region could not be modelled in the high-resolution structure due to weak electron density, which suggests that the LC/X's carboxy-terminal end is flexible.

Access to the active site of LC/X is more restricted. Analysis of the surface accessible electrostatic potential highlighted the conserved negative electrostatic potential of the active site, deep inside an open cavity (Fig. 4). Interestingly, access to the catalytic pocket of LC/X appeared more restricted than in LC/B, and /F (Fig. 4). LC/D, the closest structural homologue of LC/X also presents a restrictive access to the catalytic site but contrast by its significantly more charged surface potential. The catalytic pocket of LC/X shows a narrow entrance flanked on one side by loop 57–60, composed of polar residues (Thr-Asn-Asn-Thr), and on the other side by the short 261–263 β -strand (Fig. 4A,B). Deeper within the cleft, residues 160–165 form a β -strand that borders the catalytic channel, with the S sub-pocket delineated by residues 235–242 (Fig. 4B). The S sub-pocket of LC/X is not particularly conserved compared to the other VAMP-cleaving toxins (Fig. 4C). These structural elements may be involved in substrate recognition and allow LC/X to cleave VAMPs at a specific and unique position²⁰. On the other end, the S' sub-pocket delimited by loops 192–195 and 360–365 is well conserved (Fig. 4C), and may help stabilise the interaction with the C-terminal end of the substrate upon binding. In particular, Arg360, which is strictly conserved across all clostridial neurotoxins, likely plays a key role in substrate interaction, similarly to its role described in BoNT/A²¹. Consistently, mutations of Arg360 and Tyr363 in BoNT/X abrogated the toxin's activity¹⁷.

Comparison to LC/F-VAMP structure. Superposition of LC/X with the structure of LC/F bound to a VAMP-based peptide inhibitor⁸ revealed a clear groove that extends from the catalytic site and all around LC/X (Fig. 5A). This structural feature is common to other BoNTs and is expected to be involved in extended substrate interactions. In addition, superposition of LC/X with the full length BoNT/B (Fig. 2) suggests that this groove also likely interacts with the so-called 'belt' region of BoNT/X. The belt region surrounds the light chain in the full-length toxins and was shown to play a protective chaperone role in BoNT/A²².

The comparison with the VAMP-bound LC/F also highlighted a significant difference at the loop defining exosites 1 and 2 in LC/F (residues 171–181)⁸, which corresponds to a much shorter three-residue long linker (171–173) in LC/X (Fig. 5B). This difference in exosites may explain their distinct cleavage sites on VAMP2. Assuming that VAMP binds to LC/X in the canonical direction, the LC/F peptide inhibitor is expected to fill the S subsite following a pattern where the C-terminal cysteine interacts with Zn^{2+} . Based on this assumption, we further analysed the catalytic pocket to understand the variation in substrate specificity (Fig. 5C).

An unusual feature of LC/X is its tolerance for residue changes at the cleavage site: it cleaves VAMP1/2/3 between R-A, VAMP4 and Ykt6 between K-S, and VAMP5 between R-A (Supplementary Fig. S4). We note that the residues at the P1 site (R and K) are positively charged. It is possible R/K may interact with the negatively charged or polar side chains of residues E265, S240 and N235 located within the S1 site. On the other side of the scissile bond, the P1' site (A or S) may be accommodated between Tyr363 and Asn165. At the P2' site, the negatively charged side chain of E or D may be involved in salt bridge formation with Arg350 further within the S' sub-pocket. Deeper within the catalytic cleft, positions P3'-P4' (A/V and L) of VAMP1/2/3 and Ykt6 are likely to make hydrophobic interactions towards Pro194, Phe195 and Pro219. We note that the P3' position in VAMP4/5 consists of polar residues (S and Q) that may come within hydrogen bonding distance of Lys193 and Lys365.

LC/X has a higher cleavage efficiency than LC/B and LC/Te. As the substantial differences in the accessibility to the active site in LC/X may influence the efficiency of substrate cleavage, we next compared the cleavage kinetics between LC/X versus LC/B and LC/Te using a fluorescence resonance energy transfer (FRET) based toxin sensor. This sensor was based on the toxin reporter previously developed²³, but replacing the original cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP) with a new generation of fluorescent proteins, mRuby2 and Clover, as the FRET pair. We also replaced the linker between the two fluorescent proteins from the original VAMP2 to VAMP1, as VAMP1 is the dominant form expressed at peripheral motor nerve terminals²⁴. Cleavage of VAMP1 by toxins separates mRuby2 and Clover, thus reducing the FRET signal.

As shown in Fig. 6A, incubation of the FRET sensor (500 nM) with LC/X or /B (20 nM each) resulted in loss of FRET signal over time. The traces were fitted to a one-phase single exponential decay and the rate constants were determined as a measurement for activity. LC/X showed ~3–4 fold more activity than LC/B towards the VAMP-1 substrate, with a reaction rate of 0.07 and 0.02 min^{-1} for LC/X and LC/B, respectively (Fig. 6B, Table 2).

We next estimated cleavage kinetics by using lower levels of LCs. To generate a time-dependent enzyme progress curve, the sensor concentration was increased to 7.5 μM and enzyme concentrations were optimized based on estimated relative activities (2.4 nM LC/X, 4.8 nM LC/B, 92 nM LC/T). The cleavage product concentration was estimated based on normalizing the FRET signals to the signal at the initial sensor concentration (7.5 μM , Fig. 6C). The data were then fitted to estimate kinetic parameters, which are listed in Table 3. The estimated kinetic parameters for LC/B and LC/Te are similar to previously reported values²⁵. The kinetic data confirmed that LC/X is substantially more efficient than LC/B (~10 fold) and LC/Te (~30 fold). This enhanced cleavage efficiency is largely due to a higher apparent turnover rate (270 min^{-1}) of LC/X towards VAMP1 than LC/B and LC/Te.

In summary, the structural characteristics of LC/X highlighted several unique features such as the restricted access to the catalytic pocket and its surrounding loop regions. The location and composition of the potential VAMP-binding pockets and the groove surrounding LC/X have implications with regard to substrate binding and interaction of LC/X with the rest of the holotoxin. Key regions of the light chain likely exhibit a high level of flexibility to adapt from the tight interactions with the belt and translocation domain of BoNT/X, to a different conformation upon substrate recognition when the free LC reaches its intracellular neuronal target. Our studies

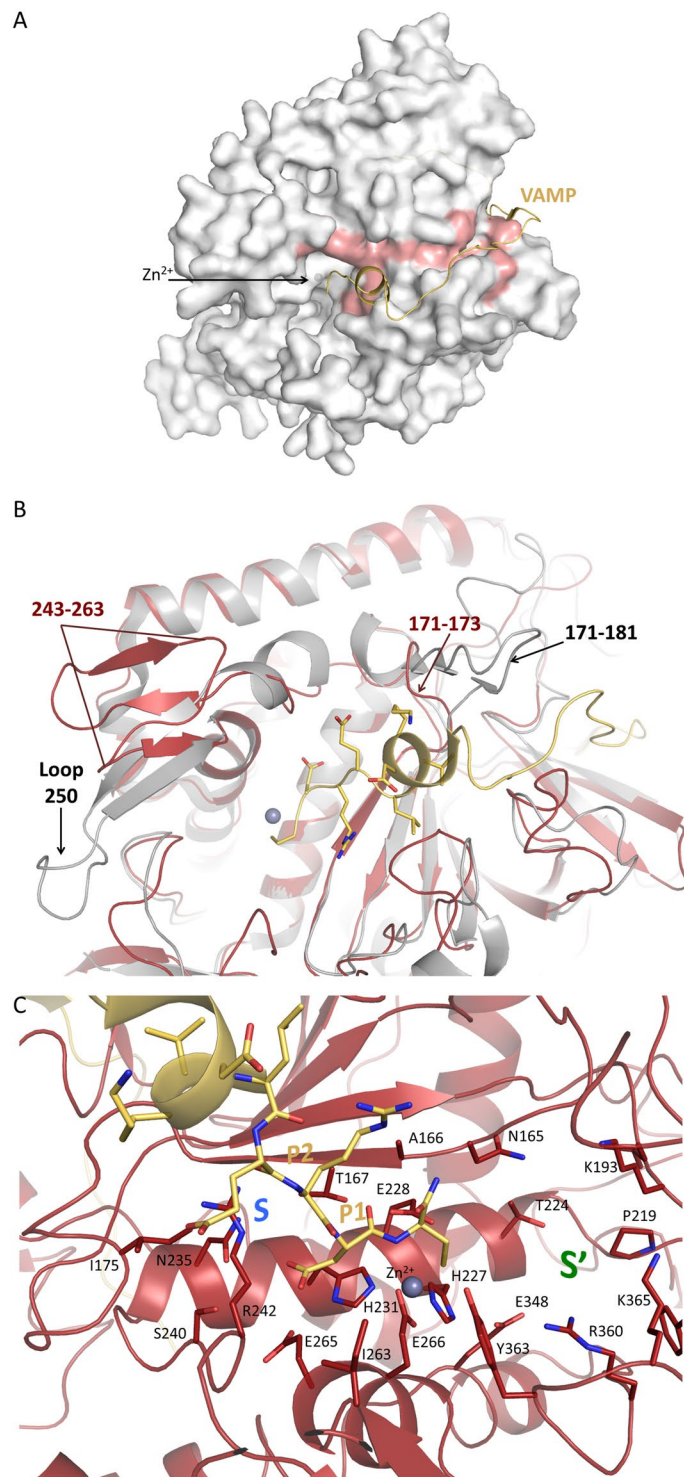


Figure 5. Comparison of LC/X to LC/F in complex with a VAMP-derived peptide inhibitor. LC/F in complex with a VAMP-derived inhibitor (PDB 3FIE) was superposed to LC/X. **(A)** The surface of LC/X is shown in white and the VAMP peptide is represented as a yellow ribbon (LC/F is not shown). Residues flanking the groove and potential exosites are highlighted in red. **(B)** Close-up view of the catalytic pocket highlighting the variations in the surrounding loops and exosites. LC/F is shown as a grey ribbon. **(C)** Close-up view of LC/X catalytic pocket and potential substrate binding sites. The bound VAMP inhibitor is shown as sticks (LC/F is not shown) and residues of LC/X potentially involved in substrate binding are presented as sticks.

activity than many BoNTs, it may serve as a promising candidate to engineer new toxin variants that can cleave SNARE proteins that currently cannot be targeted, such as VAMP7 and VAMP8, to provide novel scientific tools and therapeutic toxins.

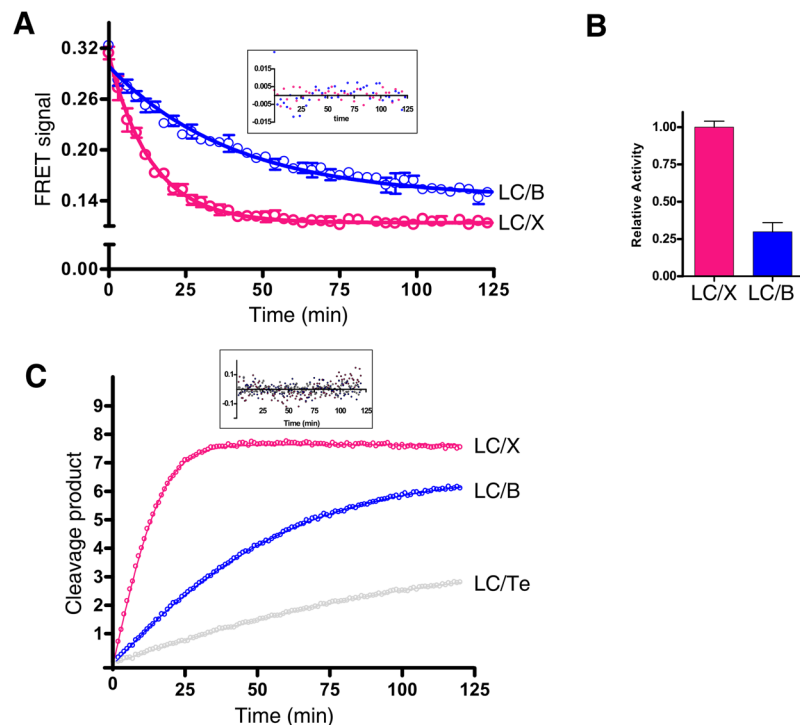


Figure 6. LC/X showed a higher level of cleavage efficiency than LC/B and /Te. **(A)** FRET-based toxin sensor (500 nM) was incubated with LC/X or LX/B (20 nM each). FRET signals were recorded with a microplate reader. Traces represent triplicates of experiments. The data were fitted to a single exponential function (Error bars are SEM; inset are the residuals of the exponential fit). The fitting data were presented in Supplementary Table S1. **(B)** Relative activity of LC/B was normalized to the activity of LC/X based on their rate constants estimated from panel A. **(C)** FRET-based toxin sensor (7.5 μ M) was incubated with LC/X (2.4 nM), LC/B (4.8 nM) or LC/Te (92 nM). The cleavage products (Y-axis, μ M) were estimated based on normalizing the FRET signal to the initial signal with 7.5 μ M toxin sensor. The data were fitted to a 3-parameter equation to estimate the kinetic constants for each enzyme (inset: residuals of fit). The kinetics data were presented in Table 3.

	V (μ M/min)	Km (μ M)	Apparent rate (min^{-1})	Apparent efficiency (min^{-1}/μ M)
LC/X	0.65 (0.03)	4.3 (0.3)	271	63
LC/B	0.22 (0.01)	8.2 (0.8)	46	6
LC/Te	0.06 (0.01)	2.8 (0.8)	7	3

Table 3. Lambert-Omega Function parameters (see Fig. 6C).

Methods

LC/X expression and purification. The gene for LC/X was synthesized and cloned into a pET/28a vector for expression in *E. coli* BL21 (DE3). Cultures were grown at 37 °C and induction carried out at 18 °C using 0.5 mM IPTG. Bacterial pellets were lysed with an Emulsiflex-C3 (Avestin, Germany) at 20 kPsi and the protein purified by affinity chromatography using a 5 ml HisTrap HP column (GE Healthcare) followed by gel filtration with a 16/60 Superdex 200 column (GE Healthcare). The final sample was stored at 13.6 mg/ml in 20 mM HEPES pH7.5, 300 mM NaCl, 10% Glycerol and 2 mM TCEP.

LC/X crystallisation and structure determination. Crystals of LC/X were obtained in two different forms. One presented prism-shaped crystals that grew within 24 h at 21 °C using the hanging drop vapour-diffusion method where 1 μ l of protein at 13.6 mg/ml was mixed with 1 μ l of reservoir solution consisting of 19% v/v PEG1500, 0.1 M MMT pH 6.0. The second form presented plate-like crystals that grew within 2–3 days at 21 °C with a similar set up and a reservoir solution consisting of 11% v/v PEG20000, 0.1 M MES pH 6.5, 10 mM EDTA. The crystals were transferred into a cryo-protectant solution consisting of their respective growing conditions supplemented with 25% v/v glycerol, and then frozen in liquid nitrogen for data collection. Diffraction data were collected at stations I02 of the Diamond Light Source (Didcot, UK), equipped with a PILATUS-6M detector (Dectris, Switzerland), and PXI of the Swiss Light Source (Villigen, Switzerland) equipped with a EIGER-16MX detector (Dectris, Switzerland). Complete datasets to 1.35 Å and 2.4 Å were collected from single crystals at 100 K. Raw data images were processed and scaled with Mosflm²⁶, or XDS²⁷ and AIMLESS²⁸ using the CCP4

suite 7.0²⁹. The structure was solved using molecular replacement with the coordinates of a truncated LC/B (PDB code 1EPW¹⁶) as a search model in PHASER³⁰. The working models were refined using REFMAC5³¹ and manually adjusted with COOT³². Water molecules were added with the help of ARP/wARP³³ at positions where $F_o - F_c$ electron density peaks exceeded 3σ , and potential hydrogen bonds could be made. Validation was performed with MOLPROBITY³⁴. Crystallographic data statistics are summarized in Table 1. All figures were drawn with PyMOL (Schrödinger, LLC, New York).

FRET-based toxin sensor. The FRET toxin sensor was constructed by fusing the Clover and mRuby2 fluorescent proteins to each ends of a human VAMP1 fragment (residues 34–87, GenBank No. CAA88760.1), and was cloned into pBAD/His-B vector using Gibson assembly. Electrocompetent *E. coli* strain DH10B was transformed with the plasmid. Transformed bacteria were cultured overnight on agar plates with 0.4 mg/ml ampicillin. A single colony was picked and grown overnight in 5 mL LB medium supplemented with 0.1 mg/ml ampicillin at 37 °C. The 5 mL culture was then used to inoculate 500 mL of LB medium with ampicillin and grown to an optical density of 0.6. Protein expression was induced with the addition of 0.02% w/v arabinose and the culture was grown at 37 °C for 4 h and then 20 °C overnight. Bacteria were then harvested and the FRET sensor protein was purified from the supernatant by Ni-NTA affinity chromatography. The buffer of the purified protein was then exchanged to 10 mM Tris-HCl, 150 mM NaCl, pH 7.2 using a desalting column.

Cleavage kinetics analysis. The cleavage assay was performed in a 50 μ L volume at 37 °C, with 500 nM FRET sensor in the assay buffer (50 mM HEPES, pH 7.1, 200 mM NaCl, 10 μ M ZnCl₂, 500 μ g/mL BSA, 2 mM DTT). The FRET signals were detected using microplate reader in 96-well plates at 620/520 nm. In Fig. 6A, the data were fit to a one-phase single exponential decay using GraphPad Prism 4 software. In Fig. 6C, FRET sensor concentration was increased to 7.5 μ M. The LC/X concentration was optimized experimentally and we selected 2.4 nM, which is sufficient to turn over all substrates within the assay time (~30 min). Higher concentrations of LC/B (4.8 nM) and LC/T (92 nM) were used as these LCs showed lower activity compared to LC/X. Data were converted to a product accumulation plot by first transforming Y-axis (dividing 1 over the FRET signal), subtracting the baseline (the value from the control sample containing LC/X, FRET sensor, and EDTA), and then the concentration was deducted by normalizing to the value of the initial substrate concentration (7.5 μ M). The data were then fitted to equation (1) to estimate kinetic parameters^{35,36}.

$$Y = W(x(t)) = W\left\{\frac{[S]_0}{KM} \cdot \exp\left(\frac{[S]_0 - V \cdot t}{KM}\right)\right\} \quad (1)$$

where W is the Lambert W(x) function approximation (2):

$$W^w(x) = \ln(1+x) \cdot \left\{1 - \frac{\ln(1+\ln(1+x))}{2 + \ln(1+x)}\right\} \quad (2)$$

All analysis and fitting were performed in Synergy KaleidaGraph 4 or GraphPad Prism 4.

Data availability. The atomic coordinates and structure factors (codes 6F47 and 6F4E) have been deposited in the Protein Data Bank (<http://wwpdb.org>).

References

- Schiavo, G., Matteoli, M. & Montecucco, C. Neurotoxins affecting neuroexocytosis. *Physiol. Rev.* **80**, 717–766 (2000).
- Chen, S. Clinical Uses of Botulinum Neurotoxins: Current Indications, Limitations and Future Developments. *Toxins* **4**, 913–939 (2012).
- Montal, M. Botulinum neurotoxin: a marvel of protein design. *Annu. Rev. Biochem.* **79**, 591–617 (2010).
- Swaminathan, S. Molecular structures and functional relationships in clostridial neurotoxins. *FEBS J.* **278**, 4467–4485 (2011).
- Rossetto, O. *et al.* SNARE motif and neurotoxins. *Nature* **372**, 415–416 (1994).
- Pellizzari, R., Mason, S., Shone, C. C. & Montecucco, C. The interaction of synaptic vesicle-associated membrane protein/syntaxin with botulinum neurotoxins D and F. *FEBS Lett.* **409**, 339–342 (1997).
- Breidenbach, M. A. & Brünger, A. T. Substrate recognition strategy for botulinum neurotoxin serotype A. *Nature* **432**, 925–929 (2004).
- Agarwal, R., Schmidt, J. J., Stafford, R. G. & Swaminathan, S. Mode of VAMP substrate recognition and inhibition of Clostridium botulinum neurotoxin F. *Nat. Struct. Mol. Biol.* **16**, 789–794 (2009).
- Agarwal, R., Eswaramoorthy, S., Kumaran, D., Binz, T. & Swaminathan, S. 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 (2004).
- Agarwal, R., Binz, T. & Swaminathan, S. Structural analysis of botulinum neurotoxin serotype F light chain: implications on substrate binding and inhibitor design. *Biochemistry* **44**, 11758–11765 (2005).
- Arndt, J. W., Yu, W., Bi, F. & Stevens, R. C. Crystal structure of botulinum neurotoxin type G light chain: serotype divergence in substrate recognition. *Biochemistry* **44**, 9574–9580 (2005).
- Arndt, J. W., Chai, Q., Christian, T. & Stevens, R. C. Structure of botulinum neurotoxin type D light chain at 1.65 Å resolution: repercussions for VAMP-2 substrate specificity. *Biochemistry* **45**, 3255–3262 (2006).
- Breidenbach, M. A. & Brünger, A. T. 2.3 Å crystal structure of tetanus neurotoxin light chain. *Biochemistry* **44**, 7450–7457 (2005).
- Jin, R. *et al.* Structural and biochemical studies of botulinum neurotoxin serotype C1 light chain protease: implications for dual substrate specificity. *Biochemistry* **46**, 10685–10693 (2007).
- Segelke, B., Knapp, M., Kadkhodayan, S., Balhorn, R. & Rupp, B. Crystal structure of Clostridium botulinum neurotoxin protease in a product-bound state: evidence for noncanonical zinc protease activity. *Proc. Natl. Acad. Sci. USA* **101**, 6888–6893 (2004).
- Swaminathan, S. & Eswaramoorthy, S. Structural analysis of the catalytic and binding sites of Clostridium botulinum neurotoxin B. *Nat. Struct. Biol.* **7**, 693–699 (2000).
- Zhang, S. *et al.* Identification and characterization of a novel botulinum neurotoxin. *Nat. Commun.* **8**, 14130 (2017).

18. Masuyer, G., Davies, J. R., Moore, K., Chaddock, J. A. & Acharya, K. R. Structural analysis of Clostridium botulinum neurotoxin type D as a platform for the development of targeted secretion inhibitors. *Sci. Rep.* **5**, 13397 (2015).
19. Rawlings, N. D., Waller, M., Barrett, A. J. & Bateman, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **42**, D503–509 (2014).
20. Sikorra, S., Henke, T., Galli, T. & Blinz, T. Substrate recognition mechanism of VAMP/synaptobrevin-cleaving clostridial neurotoxins. *J. Biol. Chem.* **283**, 21145–21152 (2008).
21. Binz, T., Bade, S., Rummel, A., Kollwe, A. & Alves, J. Arg(362) and Tyr(365) of the botulinum neurotoxin type a light chain are involved in transition state stabilization. *Biochemistry* **41**, 1717–1723 (2002).
22. Brünger, A. T. *et al.* Botulinum neurotoxin heavy chain belt as an intramolecular chaperone for the light chain. *PLoS Pathog.* **3**, 1191–1194 (2007).
23. Dong, M., Tepp, W. H., Johnson, E. A. & Chapman, E. R. Using fluorescent sensors to detect botulinum neurotoxin activity *in vitro* and in living cells. *Proc. Natl. Acad. Sci. USA* **101**, 14701–14706 (2004).
24. Peng, L. *et al.* Widespread sequence variations in VAMP1 across vertebrates suggest a potential selective pressure from botulinum neurotoxins. *PLoS Pathog.* **10**, e1004177 (2014).
25. 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 (2008).
26. Battice, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. IMosflm: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Cryst.* **D67**, 271–281 (2011).
27. Kabsch, W. 2010. XDS. *Acta Cryst.* **D66**, 125–132.
28. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr.* **D62**, 72–82 (2006).
29. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta. Crystallogr.* **D50**, 760–763 (1994).
30. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
31. Murshudov, G. N. *et al.* Refmac5 for the refinement of macromolecular crystal structures. *Acta Crystallogr.* **D67**, 355–367 (2011).
32. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr.* **D66**, 486–501 (2010).
33. Lamzin, V. S., Perrakis, A. & Wilson, K. S. The ARP/WARP suite for automated construction and refinement of protein models. In *Int. Tables for Crystallography. Vol. F: Crystallography of biological macromolecules.* (eds Rossmann, M. G., Arnold, E.) 720–722 (Dordrecht, Kluwer Academic Publishers, 2001).
34. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr.* **D66**, 12–21 (2010).
35. Goličnik, M. Exact and approximate solutions for the decades-old Michaelis-Menten equation: Progress-curve analysis through integrated rate equations. *Biochem. Mol. Biol. Educ.* **39**, 117–125 (2011).
36. Goličnik, M. On the Lambert W function and its utility in biochemical kinetics. *Biochem. Eng. J.* **63**, 116–123 (2012).
37. Hasegawa, H. & Holm, L. Advances and pitfalls of protein structural alignment. *Curr. Opin. Struct. Biol.* **19**, 341–348 (2009).

Acknowledgements

We thank the scientists at stations I02 of the Diamond Light Source (UK) and PXI of the Swiss Light Source (Switzerland) for their support during X-ray data collection. This study was partially supported by NIH grants (NS080833 and AI132387) to M.D.; the Swedish Research Council (2014–5667), the Wenner-Gren Foundation, and the Swedish Cancer Society (to P.S.). M.D. acknowledges support by the NIH-funded Harvard Digestive Disease Center (P30DK034854) and Boston Children's Hospital Intellectual and Developmental Disabilities Research Center (P30HD18655). M.D. holds the Investigator in the Pathogenesis of Infectious Disease award from the Burroughs Wellcome Fund.

Author Contributions

L.H. and S.K. performed protein purification and crystallisation. G.M. carried out crystallography and analysed the data. S.Z. and S.B. carried out kinetic analysis of the toxin LC activity. Y.S. created the FRET sensor used in this study. G.M., P.S., and M.D. wrote the manuscript with input from all co-authors.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-22842-4>.

Competing Interests: Boston Children's Hospital has filed a provisional patent application (U.S. 62/360,239) for medical use of BoNT/X, with P.S., S.Z. and M.D. as inventors.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018