Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structures of engineered *Clostridium botulinum* neurotoxin derivatives

Geoffrey Masuyer,[a] Patrick Stancombe,[b] John A. Chaddock[b] and K. Ravi Acharya[a]*

[a]Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, England, and [b]Syntaxin Limited, Units 4–10, The Quadrant, Barton Lane, Abingdon, Oxon OX14 3YS, England

Correspondence e-mail: bsskra@bath.ac.uk

Received 20 July 2011
Accepted 23 August 2011

PDB References: LC/A-SNAP23-HA, 3zus; LC/A(0)-SNAP25-HA, 3zur; LC/B-GS-Hn/B, 3zuq.

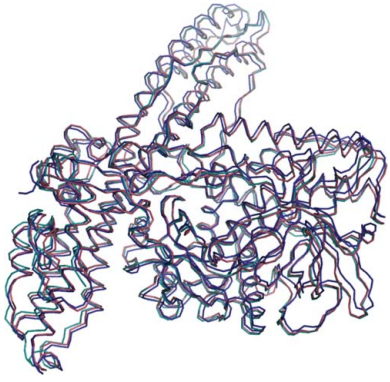
Targeted secretion inhibitors (TSIs) are a new class of engineered biopharmaceutical molecules derived from the botulinum neurotoxins (BoNTs). They consist of the metalloprotease light chain (LC) and translocation domain (Hn) of BoNT; they thus lack the native toxicity towards motor neurons but are able to target soluble *N*-ethylmaleimide-sensitive fusion protein attachment receptor (SNARE) proteins. These functional fragment (LHn) derivatives are expressed as single-chain proteins and require post-translational activation into di-chain molecules for function. A range of BoNT derivatives have been produced to demonstrate the successful use of engineered SNARE substrate peptides at the LC–Hn interface that gives these molecules self-activating capabilities. Alternatively, recognition sites for specific exoproteases can be engineered to allow controlled activation. Here, the crystal structures of three LHn derivatives are reported between 2.7 and 3.0 Å resolution. Two of these molecules are derivatives of serotype A that contain a SNARE peptide. Additionally, a third structure corresponds to LHn serotype B that includes peptide linkers at the exoprotease activation site. In all three cases the added engineered segments could not be modelled owing to disorder. However, these structures highlight the strong interactions holding the LHn fold together despite the inclusion of significant polypeptide sequences at the LC–Hn interface.

1. Introduction

The botulinum neurotoxins are the most potent protein toxins. They target pre-synaptic cholinergic neurons to inhibit neurosecretion, thereby causing the potentially fatal illness botulism. Seven serotypes (A–G) are expressed by several *Clostridium* strains and correspond to a 150 kDa multi-domain di-chain protein. The heavy chain is composed of a C-terminal 50 kDa domain (Hc) responsible for specific binding and an N-terminal 50 kDa translocation domain (Hn) that allows transport of the light chain (LC), a zinc metalloprotease, within the cells (for a review, see Turton *et al.*, 2002).

In a chain of events, the toxin binds to its receptors on the nerve terminals, leading to endocytosis of this complex. Low vesicular pH is believed to cause a conformational change inducing translocation and release of LC within the cytosol. BoNTs target one of the soluble *N*-ethylmaleimide-sensitive fusion protein attachment receptor (SNARE) proteins. These proteins form a complex necessary for the docking and fusion of synaptic vesicles in cell secretion processes. BoNT serotypes A, C and E target the synaptosome-associated protein of 25 kDa (SNAP25), while serotypes B, D, F and G cleave synaptobrevin. Serotype C is also known to cleave syntaxin. Blocking the assembly of the SNARE complex causes inhibition of neurotransmitter release and thus flaccid paralysis. For a detailed recent review, see Montal (2010).

BoNTs have been a point of focus for protein engineering since it was demonstrated that their functionality as secretion inhibitors could be extended to non-neuronal targets by engineering a new cell targeting domain onto the LHn fragment of BoNT. Proof-of-concept for the retargeting approach was first provided by the use of lectin



and growth-factor conjugates (Chaddock *et al.*, 2000, 2004; Duggan *et al.*, 2002). Subsequently, novel recombinant BoNT derivatives consisting of the catalytic LC and the Hn translocation domain, together forming the LHn fragment (Fig. 1*a*), coupled with retargeting polypeptides have been developed as a novel biological platform: targeted secretion inhibitors (TSI).

An essential step in the production of fully active BoNT is the post-translational cleavage from a single-chain polypeptide into its di-chain activated form (Wey *et al.*, 2006). Recombinant derivatives of BoNT need to be processed either with nonspecific proteases or, for a more controlled approach, with activation by specific exoproteases, such as factor Xa or enterokinase, cleaving at an engineered site (Sutton *et al.*, 2005). However, this is a delicate and limiting process in terms of protein production. Thus, use of the BoNT LC proteolytic activity towards SNARE peptides has been investigated as a tool for engineering both self-activating and LC-activating molecules (Stancombe *et al.*, 2011). This resulted in several novel derivatives in which an optimum SNARE protein region was inserted in the LHn/A backbone. This allowed the assembly of functional activated backbones when tested for proteolytic activity on SNAP25 *in vitro* and in spinal cord neurons.

The crystal structures of BoNT/A (Lacy *et al.*, 1998), BoNT/B (Swaminathan & Eswaramoorthy, 2000) and BoNT/E (Kumaran *et al.*, 2009) and more recently those of LHn/A and LHn/B (Masuyer *et al.*, 2009, 2011) have been determined. In all the available structures, the botulinum neurotoxin domains conserve their singular fold (Swaminathan, 2011). In particular, LHn/A and LHn/B retain the structural stability of the LC–Hn interactions seen in the holotoxin despite lacking the Hc binding domain.

Here, we report the crystal structures of three engineered derivatives of the LHn backbones from serotypes A and B. Two of these molecules correspond to constructs used as single-chain controls in a study of self-cleaving LC-SNARE-Hn molecules (Stancombe *et al.*, 2011), namely LC/A-SNAP23-Hn/A and LC/A(0)-SNAP25-Hn/A. SNAP23 is a noncleavable SNAP25 homologue. LC/A(0) was mutated at key residues to impair LC/A proteolytic activity. The third molecule represents an LHn/B backbone which included peptide-linker segments around the exoprotease activation site, with the aim of analysing the differences made to primary sequence changes in this essential region of the botulinum neurotoxins. Crystal structures of the LHn fragment derivatives were solved to 2.7–3.0 Å resolution. Although the engineered fragment could not be modelled owing to

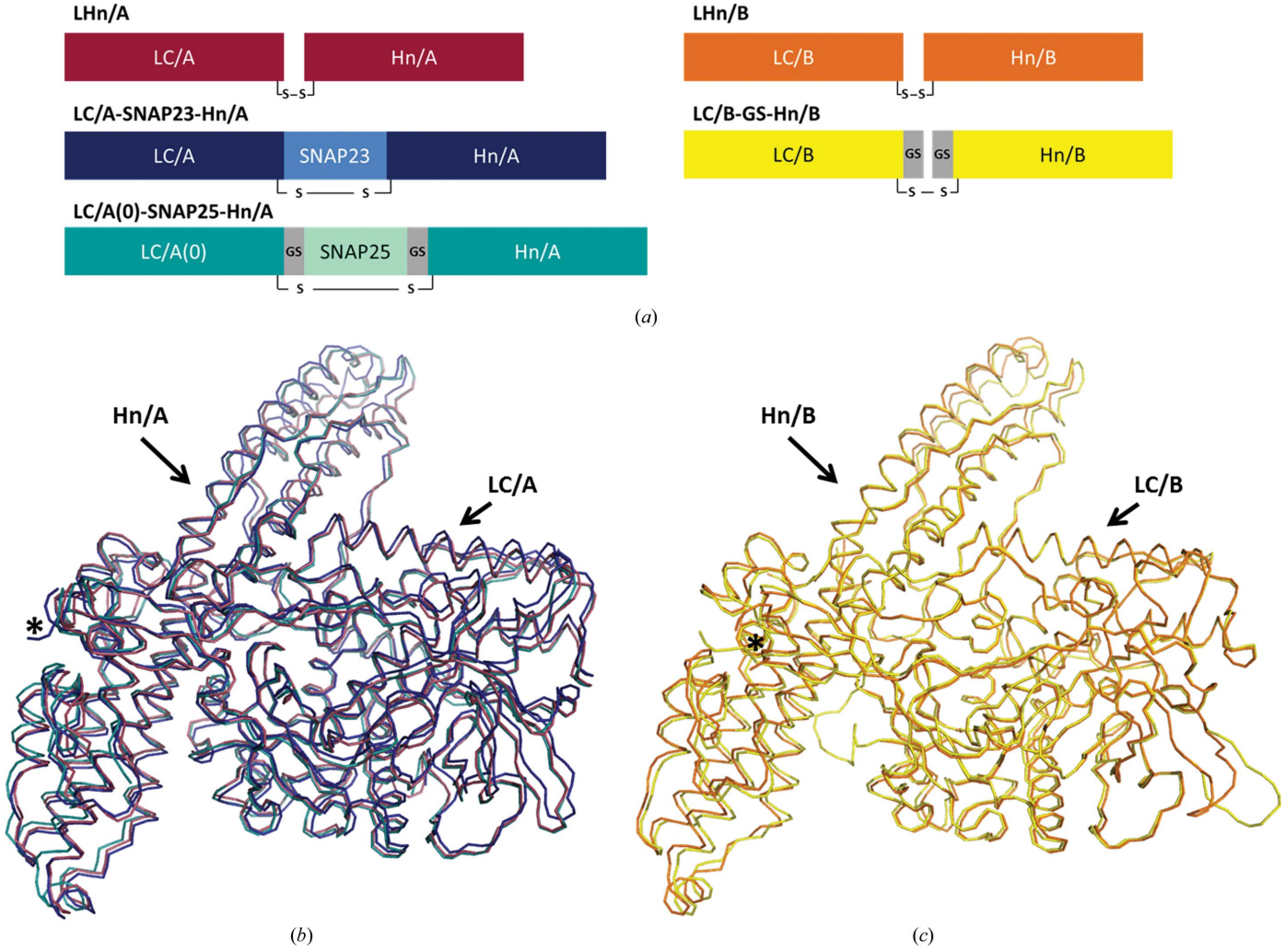


Figure 1
Structures of engineered SNARE–botulinum neurotoxin derivatives. (*a*) Schematic representation of LHn and the engineered constructs analysed by X-ray crystallography. BoNT domains and engineered fragment representation. (*b*) Ribbon diagrams of the LC/A-SNAP23-Hn/A (blue), LC/A(0)-SNAP25-Hn/A (cyan) and LHn/A crystal structures (magenta; PDB entry 2w2d). The asterisk denotes the location of the LC–Hn interface with a disulfide bridge. (*c*) Ribbon diagrams of the LC/B-GS-Hn/B (yellow) and LHn/B crystal structures (orange; PDB entry 2xhl). The asterisk denotes the location of the LC–Hn interface with a disulfide bridge.

Table 1
X-ray data-collection and refinement statistics.

Values in parentheses are for the outer shell.

	LC/A-SNAP23-Hn/A	LC/A(0)-SNAP25-Hn/A	LC/B-GS-Hn/B
Data-collection statistics			
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$
Molecules per asymmetric unit	4	2	1
Unit-cell parameters (Å, °)	a = 89.2, b = 205.0, c = 130.9, $\alpha = \gamma = 90$, $\beta = 91.9$	a = 79.0, b = 157.5, c = 209.4, $\alpha = \beta = \gamma = 90$	a = 89.4, b = 103.8, c = 115.0, $\alpha = \beta = \gamma = 90$
Resolution range (Å)	50–2.95	30–2.7	58–2.7
R_{merge}† (%)	13.0 (55.0)	10.0 (49.6)	13.4 (49.2)
Mean $I/\sigma(I)$	10.4 (3.0)	14.5 (2.9)	8.5 (3.3)
Completeness (%)	99.6 (99.2)	96.6 (79.7)	93.3 (100.0)
Total No. of reflections	515451	508979	146733
No. of unique reflections	97807	69459	28039
Multiplicity	5.3 (5.1)	7.3 (4.7)	5.2 (5.4)
Wilson B factor (Å^2)	64.6	59.5	47.0
Refinement statistics			
Resolution range (Å)	130–3.95	126–2.7	77–2.7
R_{cryst}‡ (%)	25.0	21.5	25.0
R_{free}§ (%)	29.3	26.6	28.2
No. of non-H atoms per asymmetric unit	27661	13823	6953
Protein atoms	27586	13649	6908
Ligand atoms	4 Zn atoms	5 sulfate ions	1 Zn atom
Water molecules	71	149	44
Average temperature factor (B factor) (Å^2)	54.5	41.0	36.5
R.m.s.d. in bond lengths (Å)	0.006	0.008	0.006
R.m.s.d. in bond angles (°)	0.821	1.094	0.844

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the ith and the mean measurements of the intensity of reflection hkl, respectively. ‡ $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes of reflection hkl, respectively. § R_{free} is equal to R_{cryst} for a randomly selected 5.0% subset of reflections that were not used in the refinement.

disorder, the structures demonstrated the remarkable stability of the LHn fold despite the addition of significant peptide segments at the LC–Hn interface, a key molecular property for botulinum neurotoxin activity. This supports the relevance of LHn as a safe and useful molecule for the study of botulinum neurotoxin engineering.

2. Materials and methods

2.1. LC/A-SNAP23-Hn/A and LC/A(0)-SNAP25-Hn/A cloning, expression and purification

Synthetic DNA encoding human SNAP23 (151–211) or human SNAP25 (145–206) was cloned into modified pET vector (Novagen, UK) in which the pelB signal leader was deleted. The plasmid contained an open reading frame (ORF) for LHn/A encoding 842 amino acids. In-frame fusion was achieved by the use of unique restriction sites introduced into the DNA encoding the C-terminus of the light chain, the N-terminus of the translocation domain and the SNARE peptide linker. LC/A(0)-SNAP25-Hn/A contains mutations in the active site at positions Glu224Gln and His227Tyr as well as GS linkers flanking the SNAP25 fragment. Poly-GS linkers correspond to GGGGS repeats. Expression and purification was performed as described previously (Stancombe *et al.*, 2011) with further purification necessary for crystallization. Briefly, LC/A-SNAP23-Hn/A was purified by affinity chromatography (HisTrap HP, GE Healthcare) followed by hydrophobic interaction chromatography (Phenyl Sepharose HP, GE Healthcare). LC/A(0)-SNAP25-Hn/A was purified by affinity chromatography (HisTrap HP, GE Healthcare) followed by gel filtration (Superdex 75, GE Healthcare). Final samples were concentrated to 6.1 and 5.1 mg ml^{-1}, respectively, and kept in 0.05 *M* HEPES pH 7.2, 0.2 *M* sodium chloride.

2.2. LC/B-GS-Hn/B cloning, expression and purification

The synthetic gene encoding 863 amino acids of LHn/B was cloned into modified pET vector (Novagen, UK) with a C-terminal 6×His tag and transformed into *Escherichia coli* BL21 expression cells. The LHn/B gene was engineered to encode a factor Xa cleavage site (IEGR) flanked by three GGGGS repeats between the LC and Hn domains (LC/B Asp439–Hn/B Leu477). Expression and purification of LC/B-GS-Hn/B was performed as described previously for LHn/B (Masuyer *et al.*, 2011). Briefly, expression cultures were grown in Terrific Broth and induced with IPTG (1 m*M*). After cell lysis, LHn/B in the soluble fraction was purified by affinity chromatography (HisTrap HP, GE Healthcare), activated by factor Xa cleavage (New England BioLabs) and finally purified by hydrophobic interaction chromatography (Phenyl Sepharose HP, GE Healthcare) prior to concentration to 6.5 mg ml^{-1}. The sample was stored in 0.05 *M* HEPES pH 7.2, 0.2 *M* sodium chloride.

2.3. Crystallization and structure determination

LC/A-SNAP23-Hn/A crystals were grown by the hanging-drop method in 2–3 months using 0.1 *M* imidazole malate pH 6.0, 15% PEG 3350 in a 3 µl drop (2:1 protein:mother liquor ratio) at 289 K. X-ray diffraction data were collected to 2.95 Å resolution from a single crystal cryoprotected with 25% PEG 3350 at 100 K on beamline I03 at the Diamond Light Source, UK. Data were processed and scaled in the monoclinic space group $P2_1$ using *MOSFLM* and *SCALA* (Leslie, 2006; Winn *et al.*, 2011; Table 1). Initial phases were obtained by molecular replacement using *Phaser* (McCoy *et al.*, 2007) with the coordinates of LHn/A (PDB entry 2w2d; Masuyer *et al.*, 2009).

LC/A(0)-SNAP25-Hn/A crystals were grown in 2–3 months by the hanging-drop method in 0.1 *M* Tris acetate pH 7.5, 15% PEG 3350, 0.2 *M* lithium sulfate in a 3 µl drop (2:1 protein:mother liquor ratio) at 289 K. X-ray diffraction data were collected to 2.7 Å resolution from a single crystal cryoprotected with 2 *M* lithium sulfate at 100 K on beamline I02 at the Diamond Light Source, UK. Data were processed and scaled in the orthorhombic space group $P2_12_12_1$ using *MOSFLM* and *SCALA* (Table 1). Initial phases were obtained by

molecular replacement using *Phaser* with the coordinates of the LHn/A molecule.

LC/B-GS-Hn/B crystals were grown in 2–3 weeks by the hanging-drop method in 0.1 *M* Tris acetate pH 8.5, 0.2 *M* magnesium chloride, 12% PEG 6000 in a 3 µl drop (2:1 protein:mother liquor ratio) at 289 K. Diffraction data were collected at 100 K on beamline I03 at the Diamond Light Source, UK. A complete data set to 2.7 Å resolution was obtained from a single crystal. The data were processed and scaled in the orthorhombic space group $P2_12_12_1$ using *MOSFLM* and *SCALA* (Table 1). Initial phases were obtained by molecular replacement using *Phaser* with the coordinates of LHn/B (PDB entry 2xhl; Masuyer *et al.*, 2011).

Crystallographic refinement was carried out using *REFMAC*5 (v.5.5; Murshudov *et al.*, 2011) and manual model fittings were performed using *Coot* v.0.6.1 (Emsley & Cowtan, 2004). The structures were validated using *MolProbity* (Chen *et al.*, 2011). Figures were drawn with *PyMOL* (DeLano Scientific LLC).

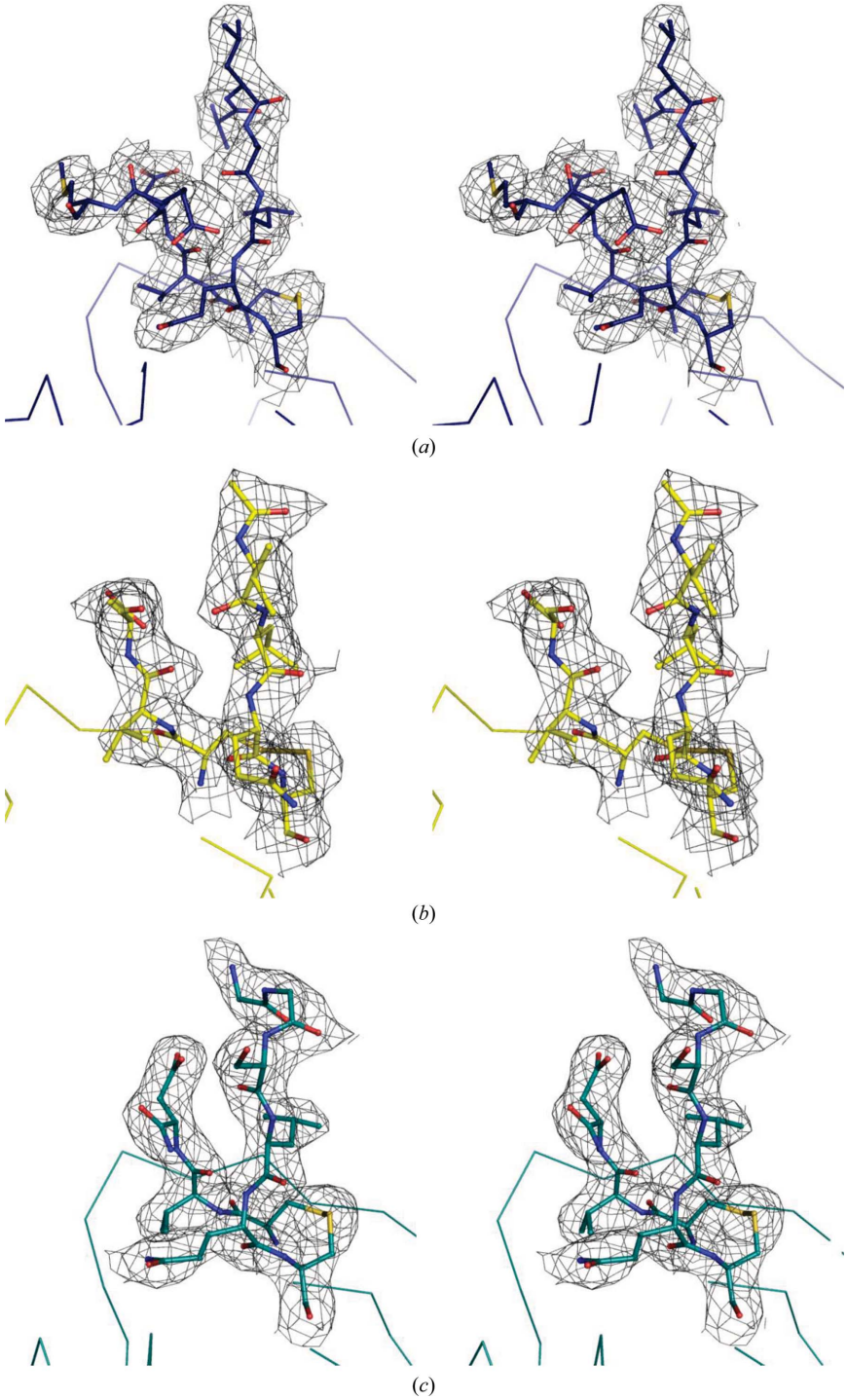


Figure 2
Missing SNARE/GS linker segments: LC–Hn interface highlighting the disulfide bridge and missing electron density above the LHn backbone. $F_o - F_c$ map at 1σ, ball-and-stick structures are shown for (*a*) LC/A-SNAP23-HnA (blue), (*b*) LC/A(0)-SNAP25-Hn/A (cyan) and (*c*) LC/B-GS-Hn/B (yellow).

3. Results and discussion

3.1. Structures of LC/A-SNAP23-Hn/A and LC/A(0)-SNAP25-Hn/A

The crystal structure of LC/A-SNAP23-Hn/A was determined at 2.95 Å resolution (Fig. 1*b*). A straightforward molecular-replacement solution was found with *Phaser* using LHn/A as a search model, with four molecules per asymmetric unit. While the whole LHn/A backbone could be fitted, no electron density was visible for SNAP23 at the LC–Hn interface, corresponding to a 63-residue peptide, owing to disorder (Fig. 2*a*). The structure was refined to a final R_{free} of 29.3% and R_{cryst} of 25.0% (Table 1), with 96.2% of amino acids in the favoured region of the Ramachandran plot and 71 water molecules in the asymmetric unit.

The LC/A(0)-SNAP25-Hn/A crystal structure was solved at 2.7 Å resolution (Fig. 1*b*). Again, LHn/A was used as a search model in *Phaser* and led to a molecular-replacement solution with two molecules per asymmetric unit. However, a 98-residue region corresponding to the engineered SNAP25 peptide flanked by GS linkers between LC and Hn could not be located owing to disorder (Fig. 2*b*). The structure was refined to a final R_{free} of 26.6% and R_{cryst} of 21.5% (Table 1), with 94.8% of amino acids in the favoured region of the Ramachandran plot. Solvent molecules (154 water molecules) and five sulfate ions were visible per asymmetric unit.

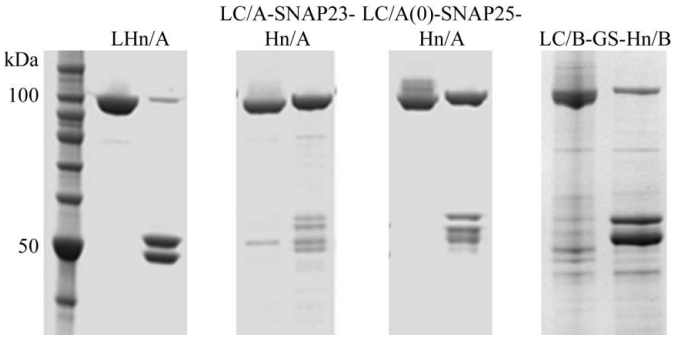


Figure 3
SDS–PAGE of LHn derivatives. For all samples, nonreduced and reduced (left and right lanes, respectively) samples are shown. The LHn/A activated control shows nonreduced di-chain at 100 kDa and two bands in the reduced lane corresponding to LC and Hn (upper and lower band, respectively). LC/A-SNAP23-Hn/A and LC/A(0)-SNAP25-Hn/A show single-chain material (Stancombe *et al.*, 2011). LC/B-GS-Hn/B is a di-chain molecule after factor Xa treatment.

Table 2
Differences in catalytic site interactions owing to mutations in LA(0)-SNAP25-HA.

Potential bonds in LC/A† (distance in Å)			Potential bonds in LC/A(0) (distance in Å)		
His223 NE2	Zinc	2.7	His223 NE2	Tyr227 OH	3.0
His227 NE2	Zinc	2.3	Gln224 OE1	Water 1	2.8
Glu224 OE2	Water	2.5	Gln224 OE1	Water 2	3.1
Water	Zinc	2.0	Thr220 CG2	Water 1	3.2
Glu262 OE1	Zinc	2.1	Glu262 OE2	Ser259 OG	3.0
Glu262 OE2	Tyr366 OH	3.3	Glu262 OE2	Ser259 N	2.6
			Glu262 OE2	Glu257 O	2.7
			Glu262 OE1	Tyr366 OH	2.7

† LC/A of LHn/A from PDB entry 2w2d (Masuyer *et al.*, 2009)

The structures present the two domains of LHn interacting strongly. Despite the inclusion of the SNARE peptides between the LC and Hn domains, each retains the fold seen in the structures of the full-length toxin and LHn/A (Fig. 1*b*), including the 'belt' region of Hn surrounding LC [Leu500–Leu595 and Leu534–Leu627 for LC/A-SNAP23-Hn/A and LA(0)-SNAP25-HA, respectively], which is stabilized by a single disulfide bridge. This is highlighted by root-mean-square deviations with LHn/A of 0.8 and 0.7 Å, respectively (over 847 C^{α} atoms). In contrast to the structure obtained for LHn/A, the SNAP engineered constructs are single-chain molecules. Analysis by SDS–PAGE (Fig. 3) confirmed that the two molecules are not subject to self-proteolysis since SNAP23 is not a natural substrate of BoNT/A and the SNAP25 construct proteolytic activity is prevented by a double mutation at the active site (Stancombe *et al.*, 2011).

In native BoNT/A the zinc ion is tetrahedrally coordinated by His223, His227, Glu262 and water-bound Glu224 (Fig. 4*a*, Table 2). This was observed in the LC/A-SNAP23-Hn/A structure, in which anomalous difference density was clearly observed at the zinc position. The mutations engineered into LC/A(0)-SNAP25-Hn/A,

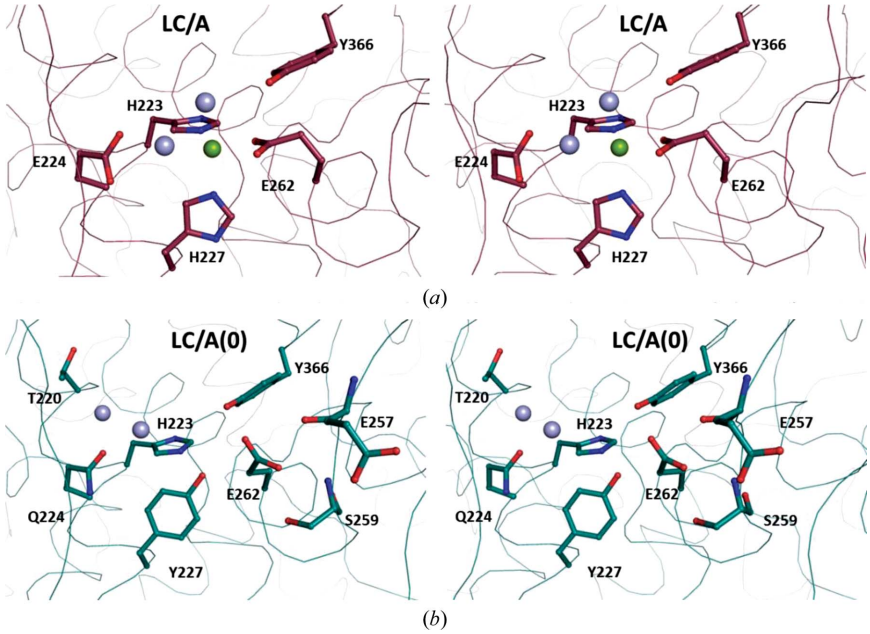


Figure 4
Structural consequences of mutations at the LC/A catalytic site. Crystal structures of (*a*) LHn/A (PDB entry 2w2d; magenta) and (*b*) LC/A(0)-SNAP25-Hn/A (cyan). Residues involved in catalytic site interactions are highlighted in ball-and-stick representation together with water molecules (light blue spheres) and zinc ion (green sphere).

although different from those in a previously reported inactive LC/A mutant structure (Glu224Gln and Tyr366Phe; Breidenbach & Brünger, 2004), prevent binding of the catalytic zinc ion necessary for proteolytic activity. More particularly, the hydroxyl group of Tyr227 of LC/A(0) is in the position held by the Zn atom and is stabilized by a hydrogen bond to His223. Interestingly, this causes the side chain of Glu262 to change orientation and form hydrogen bonds to residues Glu257 and Ser259. Mutation from Glu to Gln at position 224 prevents the interaction of this amino acid with residues previously involved in the active site. Gln224 is stabilized by direct interactions of its carboxyl group with Phe163 and two water molecules, one of which mediates a link to Thr220 (Fig. 4*b*, Table 2).

Superposition of the three LHn/A backbones (Fig. 1*b*) indicates a slight shift at the extremities of the long α-helices of the translocation domain. These areas of Hn seem to be located in solvent pockets and

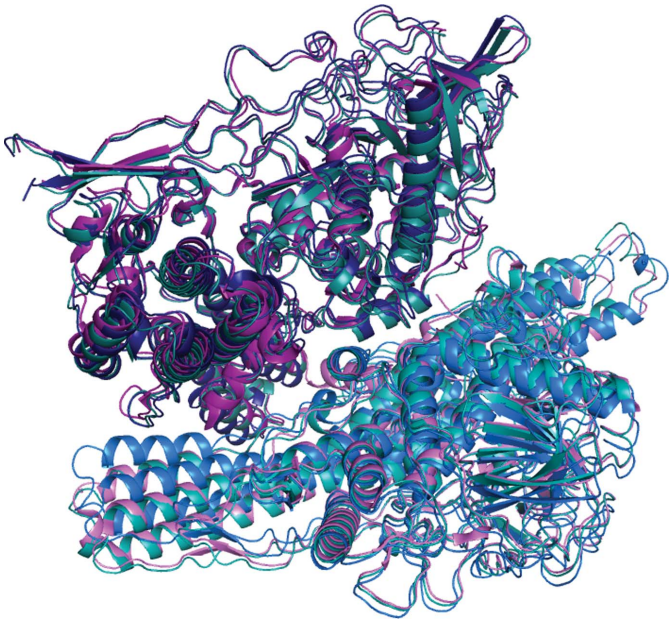


Figure 5
Crystallographic dimer in LHn/A backbone structures. Superposition of the crystallographic dimer of LHn/A in the crystal structures of LHn/A (magenta; PDB entry 2w2d), LC/A-SNAP23-HA (blue) and LC/A(0)-SNAP25-HA (cyan). The three structures present similar interactions of the dimers.

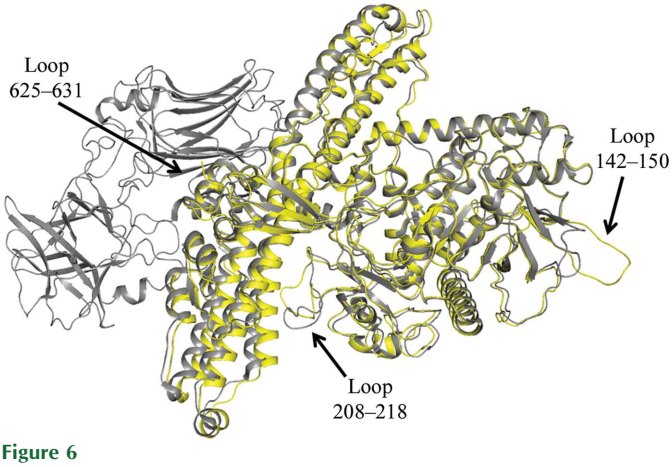


Figure 6
Comparison of LC/B-GS-Hn/B with BoNT/B. Cartoon representations of the LC/B-GS-Hn/B (yellow) and BoNT/B (PDB entry 1epw, grey; Swaminathan & Eswaramoorthy, 2000) crystal structures. Loops with different conformations are highlighted.

Table 3
Potential interactions stabilizing loop 208–218 in LC/B-GS-Hn/B.

Bonds involving helix α17 are highlighted in bold.

BoNT/B			LC/B-GS-Hn/B		
LC/B	Hn/B	Distance (Å)	LC/B	Hn/B	Distance (Å)
Ala212 O	Asn760 CG	3.0	Lys210 O	Asp840 OD1	3.3
Ala212 O	Asn760 OD1	2.6	Gly211 O	Glu833 O	2.9
Ala212 CB	Asn764 ND2	3.4	Gly211 O	Gln837 CG	3.3
Ser213 CA	Asn760 ND2	3.2	Arg217 CD	Asn836 OD1	3.3
Ser213 C	Asn760 ND2	3.4	Arg217 NH1	Asn837 O	3.2
Ile214 N	Asn760 ND2	2.7	Arg217 NH1	Asn836 OD1	3.4
Ile214 CB	**Asn723 OD1**	**2.9**	Arg217 NH1	Asp840 OD1	3.0
Phe215 CE2	**Glu730 OE2**	**3.3**	Arg217 NH2	Asn843 ND2	3.4
Phe215 CE2	Asn756 OD1	3.2	Arg217 NH2	Asp840 OD1	3.1
Phe215 CZ	Asn756 OD1	3.1	**Arg217 O**	**Tyr791 CZ**	**3.3**
Arg218 NH1	**Tyr719 CE2**	**3.3**	**Arg217 O**	**Tyr791 OH**	**2.6**
			Arg218 NH1	Asn832 OD1	3.4

are not involved with symmetry-related molecules. The most pronounced change is for LC/A-SNAP23-Hn/A, which presents the lowest solvent fraction. It is likely that some of these regions are involved in membrane insertion upon pH-mediated changes. Recent reports have pointed to conformational changes of the 659–681 region, while residues 826–835 were identified in membrane-association experiments with LHn/A (Mushrush *et al.*, 2011). The solvent-accessible loop corresponding to residues 561–564 is not visible in LC/A(0)-SNAP25-Hn/A.

Despite the different crystallization conditions and crystallographic unit cells (Table 1), the dimeric interactions of the crystallographic LHn/A backbones appear to be similar in the three LHn/A backbone structures (Fig. 5). In this packing, the orientation of the short β-strands at the LC–Hn interface is in a solvent-accessible pocket (Fig. 5). This may explain why the SNAP peptides linking the two domains are not visible. SNAP has been shown to have a flexible nature in solution (Margittai *et al.*, 2001) and to adopt a coiled-coil fold when in complex with its SNARE partners (Sutton *et al.*, 1998; Fiebig *et al.*, 1999). The localization of this peptide in the two constructs studied, together with the favoured LHn crystal packing, prevented SNAP–SNAP interaction adopting a coiled-coil structure.

SNAP25 forms partial secondary-structural arrangements when bound to LC/A (Breidenbach & Brünger, 2004), which explain the long substrate requirements for optimal activity. Brunger *et al.* (2007) demonstrated that the belt region of Hn and SNAP25 superpose well. The structures presented here show that SNAP peptides do not disturb the strong domain interactions within LHn. LC/A-SNAP23-Hn/A and LC/A(0)-SNAP25-Hn/A are single-chain proteins and crystallized under nonreducing conditions, thus favouring the stability of the two domains.

3.2. Structure of LC/B-GS-Hn/B

The crystal structure of LC/B-GS-Hn/B was determined at 2.7 Å resolution. A straightforward molecular-replacement solution was found with *Phaser* using LHn/B as a search model, with one molecule per asymmetric unit and 50.2% solvent content. The LHn/B backbone, composed of LC-Hn in its di-chain form linked by a disulfide bridge, was observable but no electron density could be seen for the GS linkers (GGGGS repeats) inserted on both sides of the factor Xa cleavage site between LC and Hn. This corresponds to 32 missing residues (excluding the 'IEGR' protease site). The structure was refined to a final R_{free} of 28.1% and R_{cryst} of 25.0% (Table 1, Fig. 1*c*), with 96.1% of amino acids in the allowed region of the Ramachandran plot and 44 water molecules per asymmetric unit.

Overall, LC/B-GS-Hn/B superposes well with the solved structures of LHn/B and BoNT/B (Figs. 1*c* and 6), with a root-mean-square deviation of 1.2 Å over 839 C^{α} atoms for LHn/B. SDS–PAGE analysis of the purified LC/B-GS-Hn/B indicated its activation by factor Xa into a di-chain molecule (Fig. 3). The addition of GS linker sequences between LC and Hn did not alter the structure of the protein or the interaction of the two domains. Indeed, the Hn belt region surrounds LC at a similar position to BoNT/B and ends in a short β-sheet arrangement that stabilizes the disulfide bridge linking the two domains. Electron density is only visible around the residues delimited by the LHn/B backbone, with nothing visible between LC Gly441 and Hn Leu477 (Fig. 2*c*). GS linkers are flexible protein segments by nature and in the context of the crystal packing are within a solvent-accessible area; it is therefore not surprising that this region of the structure is missing.

Comparison of LC/B-GS-Hn/B with previous structures highlights several noticeable differences, particularly around three loop regions (Fig. 6). Firstly, the loop corresponding to residues 142–150 protrudes from the globular fold of LC and has a different orientation to that seen in previous LC/B structures. The electron density in this area is weak, indicative of the loop's movement. Secondly, residues 208–218 downstream of the active site, which could not be modelled in the first LHn/B structure, present a different arrangement to that seen in BoNT/B, in which they were stabilized by interactions with helices α17 and α19 of Hn. In the LC/B-GS-Hn/B structure this loop interacts mainly with α19 through different potential hydrogen bonds and van der Waals interactions, as summarized in Table 3. The alternative orientation of this loop confirms its flexibility and may have some implications for substrate binding or access to the nearby catalytic site. Additionally, loop 610–615 of Hn could not be modelled owing to a lack of electron density, similarly to LHn/B. This loop is stabilized by the binding domain in BoNT/B, which is missing in this structure. On the other hand, the Hn domain showed little difference from what had been observed in previous structures, with only a slight shift in the extremities of its long helices that are not involved in crystal-packing interactions. In a parallel with BoNT/A, these regions might be involved in pH-dependent conformational changes and membrane insertion.

4. Conclusions

The crystal structures of three novel constructs have been reported. Two constructs were designed around the LHn/A scaffold and were engineered to contain an extended SNARE peptide at the LC–Hn interface. In order to maintain an intact protein assembly, these constructs were engineered to prevent self-activation by use of either a noncleavable substrate homologue (SNAP23) or an enzymatically inactive light-chain mutant. They therefore offer interesting models for the study of the stability of engineered LHn derivatives. Despite the entire molecules not being observed, the structure of the main framework could be determined. The structures demonstrated the stability of LHn/A and also its flexibility in supporting additional engineered peptide segments. Similarly, the third structure presented, of LHn/B in its activated di-chain form, confirmed the stability of the LHn fold when compared with its parent neurotoxin. Furthermore, these structures add some insights into localized flexible regions of BoNT structures by offering noticeable conformational changes.

The engineered peptides at the LC–Hn interface were not visible. The localization of such peptides in solvent pockets of the crystal, as well as the flexible structure of these peptides, may explain this phenomenon. It would be interesting to design larger peptide inserts with known secondary structures in order to study their impact on the LHn fold.

LHn fragments are safe and reliable tools for the study of the mechanism of action of botulinum neurotoxins (Fischer *et al.*, 2008; Mushrush *et al.*, 2011). The crystal structures presented in this study show that LHn provides a fantastic framework for protein engineering to which functional peptides can be added without disturbing the overall structure of the proteins.

The X-ray diffraction data for LHn/B crystals were collected on beamlines I02 and I03 at the Diamond Light Source (Oxon, UK). This work was supported by a postgraduate studentship to GM through a BBSRC–Syntaxin Limited (UK) CASE award. KRA wishes to acknowledge the Royal Society (UK) for an Industry Fellowship.

References

Breidenbach, M. A. & Brunger, A. T. (2004). *Nature (London)*, **432**, 925–929.
Brunger, A. T., Breidenbach, M. A., Jin, R., Fischer, A., Santos, J. S. & Montal, M. (2007). *PLoS Pathog.* **3**, 1191–1194.
Chaddock, J. A. *et al.* (2004). *Mov. Disord.* **19**, Suppl. 8, S42–S47.
Chaddock, J. A., Purkiss, J. R., Duggan, M. J., Quinn, C. P., Shone, C. C. & Foster, K. A. (2000). *Growth Factors*, **18**, 147–155.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Duggan, M. J. *et al.* (2002). *J. Biol. Chem.* **277**, 34846–34852.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fiebig, K. M., Rice, L. M., Pollock, E. & Brunger, A. T. (1999). *Nature Struct. Biol.* **6**, 117–123.
Fischer, A., Mushrush, D. J., Lacy, D. B. & Montal, M. (2008). *PLoS Pathog.* **4**, e1000245.
Kumaran, D., Eswaramoorthy, S., Furey, W., Navaza, J., Sax, M. & Swaminathan, S. (2009). *J. Mol. Biol.* **386**, 233–245.
Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R. & Stevens, R. C. (1998). *Nature Struct. Biol.* **5**, 898–902.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Margittai, M., Fasshauer, D., Pabst, S., Jahn, R. & Langen, R. (2001). *J. Biol. Chem.* **276**, 13169–13177.
Masuyer, G., Beard, M., Cadd, V. A., Chaddock, J. A. & Acharya, K. R. (2011). *J. Struct. Biol.* **174**, 52–57.
Masuyer, G., Thiyagarajan, N., James, P. L., Marks, P. M., Chaddock, J. A. & Acharya, K. R. (2009). *Biochem. Biophys. Res. Commun.* **381**, 50–53.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Montal, M. (2010). *Annu. Rev. Biochem.* **79**, 591–617.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Mushrush, D. J., Koteiche, H. A., Sammons, M. A., Link, A. J., McHaourab, H. S. & Lacy, D. B. (2011). *J. Biol. Chem.* **286**, 27011–27018.
Stancombe, P. R., Masuyer, G., Birch-Machin, I., Beard, M., Foster, K. A., Chaddock, J. A. & Acharya, K. R. (2011). Submitted.
Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. (1998). *Nature (London)*, **395**, 347–353.
Sutton, J. M., Wayne, J., Scott-Tucker, A., O'Brien, S. M., Marks, P. M., Alexander, F. C., Shone, C. C. & Chaddock, J. A. (2005). *Protein Expr. Purif.* **40**, 31–41.
Swaminathan, S. (2011). *FEBS J.*, doi:10.1111/j.1742-4658.2011.08183.x.
Swaminathan, S. & Eswaramoorthy, S. (2000). *Nature Struct. Biol.* **7**, 693–699.
Turton, K., Chaddock, J. A. & Acharya, K. R. (2002). *Trends Biochem. Sci.* **27**, 552–558.
Wey, J.-J., Tang, S.-S. & Wu, T.-Y. (2006). *Acta Pharmacol. Sin.* **27**, 1238–1246.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.