Supplementary Methods

Protein Purification

BoNT/A-LC and BoNT/A E224Q/Y366F

Plasmid DNA pBN3 encoding wt-BoNT/A-LC was provided by Thomas Binz (Medizinische Hoschule Hannover, Germany). Truncation of the LC C-terminus was achieved by amplifying residues 1-420 from the original construct by polymerase chain reaction (PCR) using the primers a) 5'-ACAGAATTCGCAATTAAGGAGATAATAGGTATG-3' corresponding to the 5' end and b) 5'-GCTCCCGGGAGTAAAATTTTTTAGTTTAGTAAAATTCATATTATTAATTTCTGTATTTTGAC-3' corresponding to the 3' end. The resulting PCR product was digested with EcoRI and SmaI and re-inserted into the original pBN3 vector that had been digested with the same restriction enzymes. A variant of this construct with a thrombin-cleavable affinity tag was prepared by digesting the plasmid with SmaI and inserting the following sequence: 5'-CTGGTTCCGCGTGGATCT-3' paired with 5'-AGATCCACGCGGAACCAG-3'. Constructs encoding point mutant E224Q and double mutant E224Q/Y366F were prepared using the Quickchange protocol (Stratagene).

All BoNT/A-LC constructs were expressed at 20° in M15[pREP4] cells (Qiagen) and affinity purified according to protocols previously described, with minor modifications[1]. Cells were lysed via two passes through a M-110EH Microfluidizer Processor (Microfluidics) at 18,000 psi. Insoluble debris was removed from the lysate by centrifugation in a JA-20 rotor (Beckman) at 19,500 rpm for 45 min. Protease inhibitors

added to the lysate include 1mM phenylmethylsulfonyl fluoride (PMSF) and EDTA-free Complete Protease Inhibitor Cocktail Tablets (Roche). The affinity tag was removed with bovine α-thrombin (Haematologic Technologies), and additional purification (>95%) was achieved with cation-exchange (Amersham-Pharmacia mono-S resin) chromatography. BoNT/A-LC was quantified by UV/Vis spectroscopy at 280 nm, based on its theoretical extinction coefficient in denaturing conditions. Proteolytic activity of the BoNT/A-LC was qualitatively verified by incubation with SNAP-25 residues 141-204; cleavage products were observed with an Omni-Flex Matrix-Assisted Laser Desorption/Ionization (MALDI) mass spectrometer (Bruker).

SNAP-25

Plasmid DNA encoding human SNAP-25a residues 141-204 (SN2) in a pET-28a expression vector was obtained from J. Ernst (Stanford University). The peptide was expressed and purified as previously described[2]. Purified SNAP-25 peptide for crystallization was quantified by the BCA colorimetric assay (Pierce). Plasmid DNA encoding sn2 point mutants I156E, M167E, S187E, and M202Y were prepared using the Quickchange protocol (Stratagene). Mutants were expressed and purified by the same method as wild-type.

Crystallization and data collection

BoNT/A-LC E224Q/Y366F SNAP-25 complex

Concentrated stocks of BoNT/A E224Q/Y366F and SNAP-25 in 20 mM HEPES pH 7.4 were diluted into a crystallization mix to final concentrations of 242 μM and 484 μM, respectively. Crystals were grown from this mix by hanging-drop vapour diffusion

over 1 mL reservoirs of 10% (w/v) PEG 8000, 200 mM magnesium acetate, and 100 mM sodium cacodylate pH 6.5 at 4° over the course of 2-3 weeks. Crystals typically grew in clusters of blade-shapes with approximate dimensions of 0.3 x 0.05 x 0.05 mm^3. Thoroughly washed crystals were confirmed to contain both enzyme and uncleaved substrate by both MALDI and analysis on a Phastgel SDS-PAGE system (Amersham-Pharmacia). Nucleation events were unusually rare and it was necessary to propagate crystal growth by macroseeding to obtain a large a large population of high-quality single crystals. Individual blades were harvested and flash-frozen after brief exposure to 25% ethylene glycol cryoprotectant.

Crystals were characterized at beamline 8.2.2 of the Advanced Light Source (ALS). Data indexed in space group $P4_32_12$ with unit cell dimensions a=b=86.0 Å, c=165.4 Å, and diffraction extended to d_{min}=2.0 Å. A native high-resolution dataset was collected at 11,500 eV. A second, low-resolution dataset was collected on a different crystal at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-2. This inverse-beam dataset was collected at a lower energy (7,500 eV) in 1º wedges with high multiplicity to record the weak anomalous signal from the sulphur atoms. Both datasets were collected on CCD Quantum-315 detectors (Area Detector Systems). All diffraction datasets were integrated and scaled using MOSFLM and SCALA in the CCP4 computational suite[3]. Low-resolution reflections (30-2.6 Å) from the anomalous dataset were scaled with high-resolution reflections (3.2-2.1 Å) from the native dataset. Data collection statistics are shown in supplementary table 1.

apo BoNT/A-LC E224Q/Y666F

BoNT/A-LC 1-420 E224Q/Y336F with C-terminal hexa-histidine tag attached was concentrated to 10 mg mL^{-1} in 20 mM HEPES pH 7.4. Crystals were grown by hanging-drop vapour diffusion over 1 mL reservoirs of 15-18% PEG 3350 and 200 mM di-sodium hydrogen phosphate dihydrate. Crystals appeared over the course of 4-5 days, typically assuming a cubic morphology with approximately 0.25 mm edges. Individual crystals were flash-frozen in a 20% glycerol cryoprotectant.

Diffraction data was collected at 11,500 eV on a Quantum-210 CCD detector (Area Detector Systems) at beamline 8.2.1 at ALS. Reflections extended to approximately d_{min}=2.2 Å, and indexed in space group P2 with near-orthorhombic unit cell dimensions a=57.90 Å, b=40.49 Å, c=195.89 Å, and β=90.25°. A total of 150° of data were collected in 0.5° oscillations before high-resolution reflections were no longer detectable due to radiation damage. Data collection statistics are shown in supplementary table 2.

Crystallographic refinement and structure determination

All phasing and refinement calculations were performed using the Crystallography and NMR System version 1.1 (CNS)[4]. Progress of refinement for both apo and holo structures was monitored by the cross-validated R_{free} value which was computed from a randomly assigned test set comprising 5% of the data[5]. Model building was performed using the program O[6]. Hydration sites were assigned by automated inspection of peaks larger than 2.5 standard deviations above the mean in σ_A-weighted F_o-F_c electron density maps. If a site exhibited chemically feasible hydrogen-binding

partners and distances, and its B-factor refined to <60 $Å^2$, it was included in the model as an ordered water. Model stereochemical quality was evaluated with PROCHECK[7]. Figures were prepared using PyMOL[8].

BoNT/A-LC E224Q/Y366F SNAP-25 complex

A molecular replacement (MR) solution was found with coordinates of the BoNT/A-LC extracted from 3BTA.pdb[9]. Anomalously scattering sulphur atoms (as well as the Zn^{2+} and one chloride) were located with an anomalous difference Fourier map calculated using reflections from 20 to 3.2 Å resolution. Experimental single-wavelength anomalous dispersion (SAD) phases were computed from these low-resolution reflections and, after density modification, used as an independent confirmation of the sn2 trace. Electron density for both enzyme and sn2 substrate was easily interpretable in a σ_A-weighted MR ($2m|F_o|\varphi_{calc}-D|F_c|\varphi_{calc}$) map. The register of the sn2 peptide was double-checked by matching anomalous difference peaks to methionine residues (figure 3a). The model was refined with alternate cycles of simulated annealing with torsion angle dynamics and restrained B-factor refinement using the maximum likelihood target function with amplitudes and phase probability distributions followed by manual rebuilding[10-12]. In addition to the MR map described above, a MR/SAD phase-combined ($2m|F_o|\varphi_{comb}-D|F_c|\varphi_{calc}$) map was also used to further reduce model bias; little change was observed between the two maps. Weak electron density was observed in three regions of the complex including BoNT/A residues 200-205, and sn2 residues 183-190 and 197-199. The fragmented electron density allowed for approximate main-chain trace for these

regions, but side chain occupancies were set to zero. Model and phasing statistics are summarized in Table 1.

apo BoNT/A-LC E224Q/Y366F

MR solutions were found for two molecules of BoNT/A LC per asymmetric unit using light chain coordinates extracted from 3BTA.pdb[9]. A near-perfect two-fold screw axis related the two molecules. A plausible orientation for one molecule was located and fixed before an additional translation search could locate the second molecule.

The initial model, derived from the MR solution, was refined with alternate cycles of simulated annealing with torsion angle dynamics and restrained B-factor refinement using the maximum likelihood target function followed by manual rebuilding[10,12,13]. Poorly ordered regions of the endopeptidase included residues 201-205 and 247-254 in both molecules of the asymmetric unit. Fragmented electron density allowed for approximate main-chain trace for these regions, but side chain occupancies were set to zero. Model and phasing statistics are summarized in Table 2. The substrate-bound structure has a significantly lower R_{free} than the apo-structure presumably due to stabilization of some of the loops (especially the 250 and 370 loops) by bound substrate.

Enzyme Kinetics

Kinetic parameters K_m and k_{cat} were determined at 37 ºC for wt-BoNT/A 1-420 and wt-SN2 as well as the substrate point mutants I156E, M167E, S187E, and M202Y. Quantitative BoNT/A-LC activity assays were conducted using a method based on reverse-phase high-performance liquid chromatography (RP-HPLC) similar to previously described protocols[14,15]. Digestions were performed in activity assay buffer (30 mM

HEPES pH 7.2, 50 μM $ZnCl_2$, 200 μM DTT) and included a fixed concentration wt-BoNT/A (5 nM) and various concentrations of substrates (between 7 μM and 1.2 mM). Incubation times were adjusted such that <10% of total substrate was hydrolyzed to maintain steady-state velocity and reactions were quenched with 1% trifluoracetate (TFA). C-terminal proteolysis products from these digestions were quantified following fractionation on a 218TP54 column (Vydac) equilibrated in 0.1% TFA. A 0-90% acetonitrile gradient was used to separate the small C-terminal cleavage product from the N-terminal and uncut moieties. Initial substrate concentrations, as well as the relation between A_{214} peak area and C-terminal protein yield were determined by quantitative amino-acid analysis performed on a Beckman 7300 Analyzer by the W.M. Keck Facility (Yale University). Assays were conducted at a minimum of six different substrate concentrations for each kinetic characterization. Plots of reaction velocity as a function of substrate concentration were fit to the hyperbolic Michaelis-Menten equation using the Origin Scientific Graphing and Analysis software package (OriginLab) to yield K_m, k_{cat}, and error estimates.

Supplementary References

1. Li, L. & Singh, B. R. High-level expression, purification, and characterization of recombinant type A botulinum neurotoxin light chain. *Protein Expr Purif* **17**, 339-44 (1999).

2. Ernst, J. A. & Brunger, A. T. High resolution structure, stability, and synaptotagmin binding of a truncated neuronal SNARE complex. *J Biol Chem* **278**, 8630-6 (2003).

3. Bailey, S. The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallographica Section D-Biological Crystallography* **50**, 760-763 (1994).

4. Brunger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 (Pt 5)**, 905-21 (1998).

5. Brunger, A. T. Free R-Value - a Novel Statistical Quantity for Assessing the Accuracy of Crystal-Structures. *Nature* **355**, 472-475 (1992).

6. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**, 110-9 (1991).

7. Laskowski, R. A., Moss, D. S. & Thornton, J. M. Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049-67 (1993).

8. DeLano, W. L. The PyMOL Molecular Graphics System. (2002).

9. Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R. & Stevens, R. C. Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nat Struct Biol* **5**, 898-902 (1998).

10. Hendrickson, W. A. Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol* **115**, 252-70 (1985).

11. Pannu, N. S., Murshudov, G. N., Dodson, E. J. & Read, R. J. Incorporation of prior phase information strengthens maximum-likelihood structure refinement. *Acta Crystallogr D Biol Crystallogr* **54**, 1285-94 (1998).

12. Rice, L. M. & Brunger, A. T. Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277-90 (1994).

13. Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proc Natl Acad Sci U S A* **94**, 5018-23 (1997).

14. Sukonpan, C. et al. Synthesis of substrates and inhibitors of botulinum neurotoxin type A metalloprotease. *J Pept Res* **63**, 181-93 (2004).

15. Schmidt, J. J. & Bostian, K. A. Proteolysis of synthetic peptides by type A botulinum neurotoxin. *J Protein Chem* **14**, 703-8 (1995).

16. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-402 (1997).

17. Hanson, M. A. & Stevens, R. C. Cocrystal structure of synaptobrevin-II bound to botulinum neurotoxin type B at 2.0 A resolution. *Nat Struct Biol* **7**, 687-92 (2000).

18. Rupp, B. & Segelke, B. Questions about the structure of the botulinum neurotoxin B light chain in complex with a target peptide. *Nat Struct Biol* **8**, 663-4 (2001).

19. Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H. & Matthews, B. W. Slow- and fast-binding inhibitors of thermolysin display different modes of binding: crystallographic analysis of extended phosphonamidate transition-state analogues. *Biochemistry* **26**, 8542-53 (1987).

20. 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 U S A* **101**, 6888-93 (2004).

Supplementary Table 1

Data collection and refinement statistics for BoNT/A-LC/SNAP-25 complex

A. Cell Parameters		
Space group	$P4_32_12$	
a,b,c (Å)	86.07, 86.07, 164.52	
B. Data Collection Statistics[a]		
	Low Resolution	**High Resolution**
Wavelength (eV)	7,500	11,500
Resolution range (Å)	30.0-2.6	3.2-2.1
Source/detector	SSRL 9.2/ADSC Quantum-315	ALS 8.2.2/ADSC Quantum-315
Unique reflections	37,046	
Multiplicity	19.3(4.7)	
Completeness (%)	99.9(100.0)	
I/σ	5.9(1.9)	
R_{sym}[b](%)	10.6(36.3)	
Integration/Scaling software	MOSFLM/SCALA	
C. Molecular replacement search model:		3BTA.pdb (LC only)

D. Sulphur SAD Phases			
	Centrics	Acentrics	All
Figure of merit		0.12	
Density-modified figure of merit			0.89

E. Model statistics			
R_{free}(%)	24.9		
$R_{working}$(%)	21.9		
Rms deviation, bonds (Å)	0.0064	Luzatti coordinate error (cross-validated) (Å)	0.31
Rms deviation, angles (deg)	1.23	Luzatti coordinate error (Å)	0.26
Rms deviation, dihedrals (deg)	22.4	σ_A coordinate error (cross-validated) (Å)	0.20
Rms deviation, impropers (deg)	0.84	σ_A coordinate error (Å)	0.18
		Average B-factor (Å^2)	32.5
Ramachandran Plot		Range of B-factors (Å^2)	11.3-83.6
Most-favored phi-psi (%)	87.0		
Additionally allowed (%)	12.5	# protein residues	483
Generously allowed (%)	0.5[c]	# ions	2
Disallowed (%)	0.0	# water molecules	371

a) Numbers in parenthesis indicate values for highest resolution bin.

b) $R_{sym} = \Sigma_h\Sigma_i|I_{hi} - \langle I_h \rangle|/\Sigma_h\Sigma_i|I_{hi}|$

c) Stereochemically strained sn2 residues R198 and T200 are both within the catalytic pocket of BoNT/A.

Supplementary Table 2

Data collection and refinement statistics for BoNT/A-LC E224Q/Y366F

A. Cell Parameters	
Space group	P2
a,b,c (Å)	57.90,40.49,195.88
β (degrees)	90.25
B. Data Collection Statistics[a]	
Wavelength (eV)	11,500
Resolution range (Å)	50-2.2
Unique reflections	44,297
Multiplicity	2.9(2.5)
Completeness(%)	95.0(94.5)
R_{sym}[b](%)	5.2(19.2)
I/σ	9.1(2.4)
Source/detector	ALS 8.2.1/ADSC Quantum-210
Integration/Scaling software	MOSFLM/SCALA
C. Molecular replacement search model:	3BTA.pdb (LC only)

D. Model Statistics			
R_{free}(%)	27.3		
$R_{working}$(%)	22.0		
Rms deviation, bonds (Å)	0.011	Luzatti coordinate error (cross-validated) (Å)	0.37
Rms deviation, angles (deg)	1.66	Luzatti coordinate error (Å)	0.27
Rms deviation, dihedrals (deg)	22.9	σ_A coordinate error cross-validated (Å)	0.40
Rms deviation, impropers (deg)	1.03	σ_A coordinate error (Å)	0.31
Ramachandran Plot		Average B-factor (Å^2)	47.2
Most-favored phi-psi (%)	86.1	Range of B-factors	13.4-108.2
Additionally allowed (%)	12.3	# protein residues	854
Generously allowed (%)	1.6[c]	# ions	2
Disallowed (%)	0.0	# water molecules	312

a) Numbers in parenthesis indicate values for highest resolution bin.

b) $R_{sym} = \Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma_h\Sigma_i|I_{hi}|$

c) Stereochemically strained residues fall within partially disordered loop regions of apo-BoNT/A.

Supplementary Figure 1.

Summary of SNAP-25/BoNT/A contacts

A schematic of the sn2 (green) interface with BoNT/A (light blue) shows the extensive network of contacts used in substrate recognition. Both the α and β exosites are indicated. Residues involved in side-chain/side-chain interactions are explicitly shown. Blue dashes indicate polar side-chain contacts, with water mediation indicated by red dots. Yellow wavy lines indicate regions of hydrophobic side-chain interaction. Orange wavy lines indicate regions of backbone polar contacts. Masked protein-protein BLAST searches of sn2 residues mediating the BoNT/A interface indicate that BoNT/A is specific for SNAP-25 in mammals, birds, amphibians, and both bony and cartilaginous fishes[16].

Supplementary Figure 2.

Questionable electron density from a reported BoNT/B-synaptobrevin complex.

A CNS[4] cross-validated, σ_A-weighted ($2m|F_o|\varphi_{calc}-D|F_c|\varphi_{calc}$) OMIT map (blue mesh) with simulated annealing calculated using deposited structure factors from 1F83.pdb[17] is shown contoured at 1σ. The view of the BoNT/B LC is approximately the same as that for BoNT/A in figure 2 (left panel) of the main text. Synaptobrevin residues (red) were omitted while all residues from BoNT/B-LC (brown) were included in the electron density map calculation. The bias-reduced electron density in the region near the substrate is indistinguishable from noise, so placement of the substrate is not supported by the experimental data as previously shown[18]. Near the active site, the orientation of the published model is also reversed with respect to that in our BoNT/A-sn2 complex

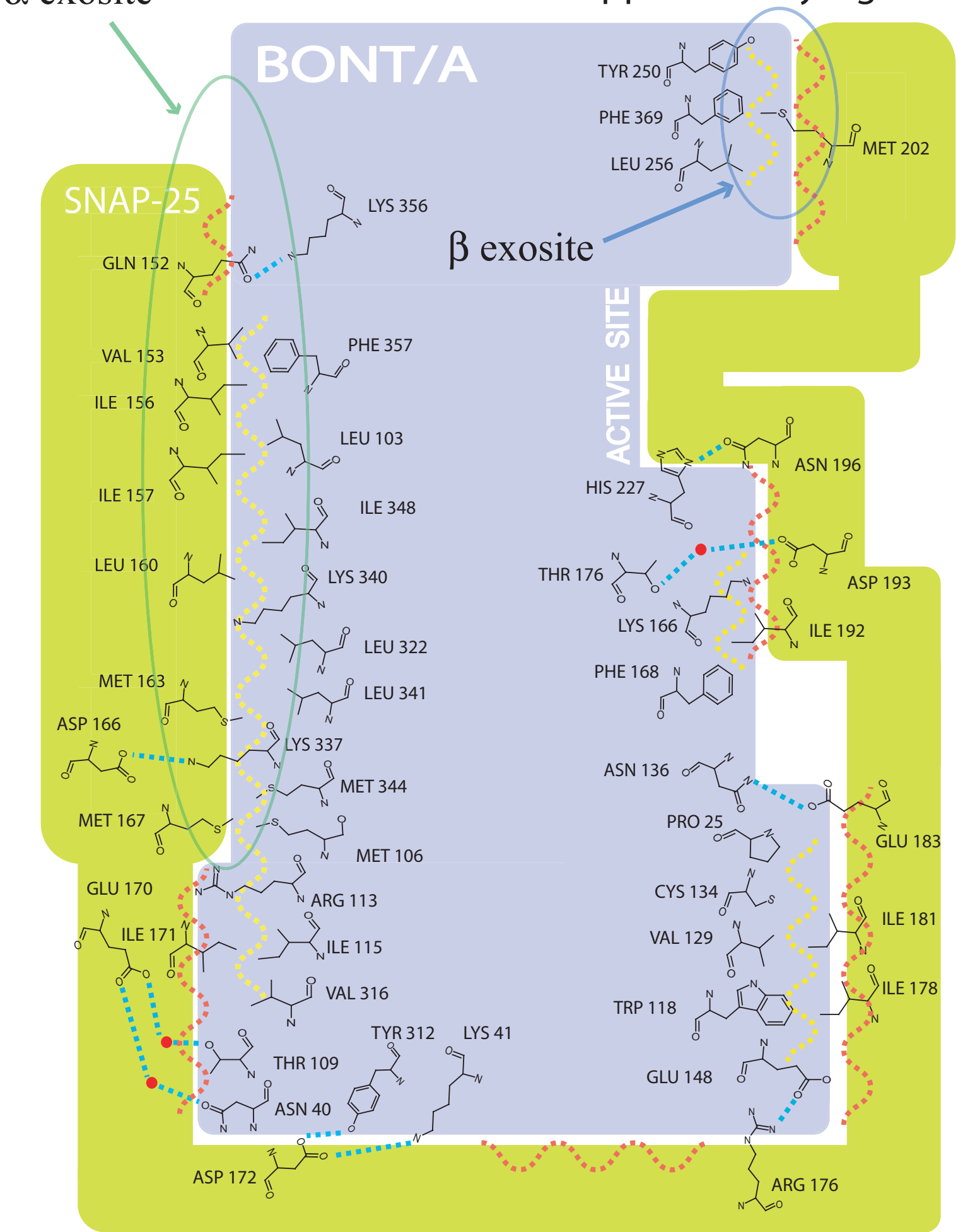
structure (*cf.* figure 2 in main text) and no exosite interactions were found, further questioning the validity of the 1F83.pdb[17] structure.

Supplementary Figure 3.

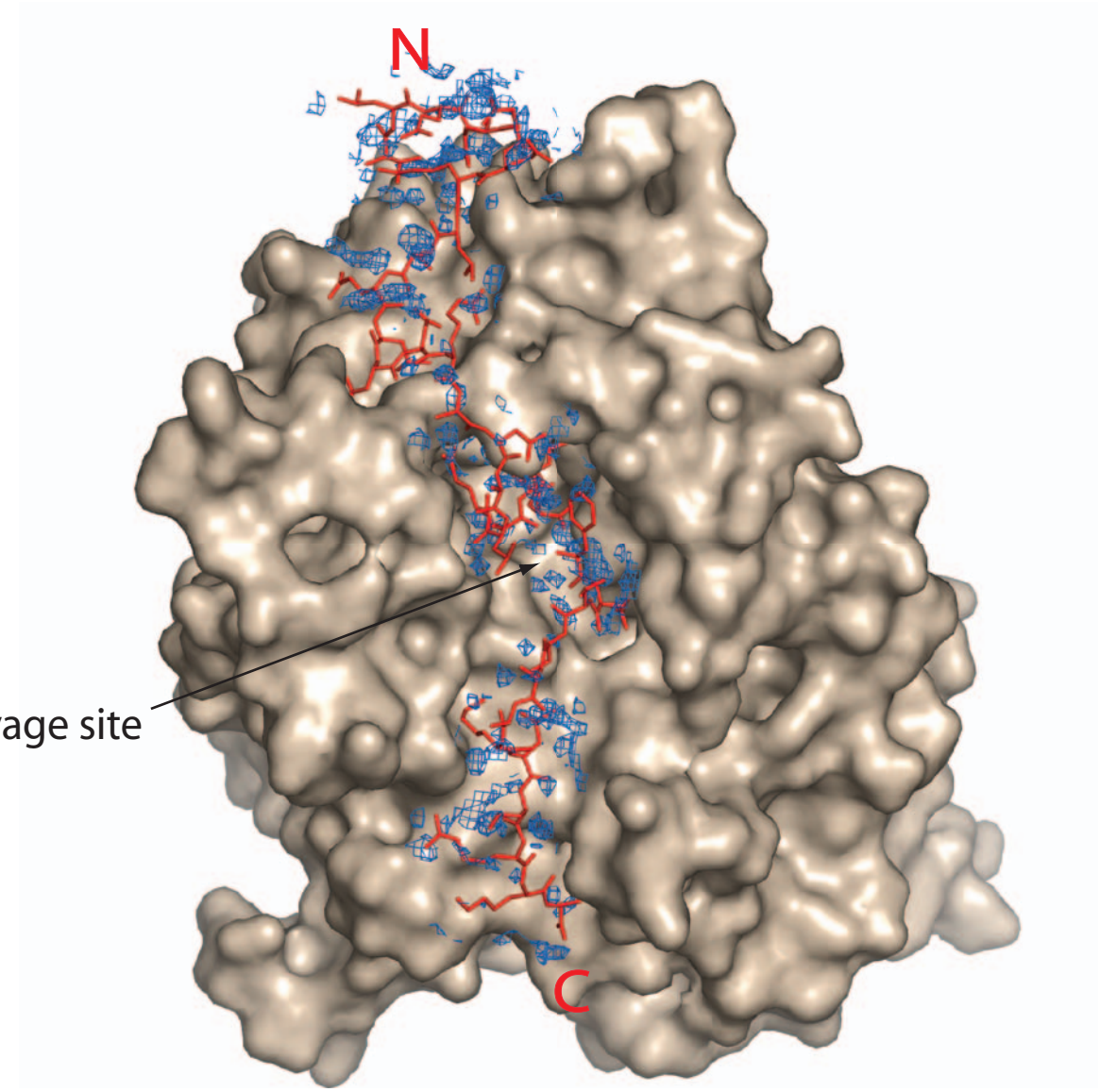
A comparison of BoNT/A and thermolysin substrate orientation.

Inhibitor-bound thermolysin[19] (4tmn.pdb, green) is shown aligned with substrate-bound BoNT/A (red) at the conserved HEXXH zinc-binding motif. The orientation of the thermolysin phosphonamidate peptide inhibitor PheP-L-Leu-L-Ala (green tube) closely matches that of sn2 (brown tube) with respect to the Zn^{2+} ion (yellow) and catalytic water (blue). The shared substrate orientation between these two Zn^{2+}-metalloproteases provides the first structural evidence that they may employ similar mechanisms for initiating proteolytic attack on their substrates.

Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3