

High-resolution crystal structures of the botulinum neurotoxin binding domains from subtypes A5 and A6

Jonathan R. Davies^{1,2} , Amy Britton¹, Sai Man Liu³  and K. Ravi Acharya¹ 

¹ Department of Biology and Biochemistry, Claverton Down, University of Bath, Bath, UK

² Department of Biochemistry and Biophysics, Stockholm University, Sweden

³ Ipsen Bioinnovation Limited, Milton Park, Abingdon, UK

Keywords

binding domain structure; botulinum neurotoxin; *Clostridium botulinum*; subtypes; X-ray crystallography

Correspondence

K. Ravi Acharya, Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK
Tel: +44-1225-386238
E-mail: bsskra@bath.ac.uk

Clostridium botulinum neurotoxins (BoNTs) cause flaccid paralysis through inhibition of acetylcholine release from motor neurons; however, at tiny doses, this property is exploited for use as a therapeutic. Each member of the BoNT family of proteins consists of three distinct domains: a binding domain that targets neuronal cell membranes (H_C), a translocation domain (H_N) and a catalytic domain (LC). Here, we present high-resolution crystal structures of the binding domains of BoNT subtypes/A5 ($H_C/A5$) and/A6 ($H_C/A6$). These structures show that the core fold identified in other subtypes is maintained, but with subtle differences at the expected receptor-binding sites.

(Received 18 June 2020, revised 7 July 2020, accepted 8 July 2020)

doi:10.1002/2211-5463.12931

Clostridium botulinum neurotoxins (BoNTs) are responsible for causing the deadly condition, botulism, in vertebrates [1–4]. There are seven distinct serotypes termed BoNT/A through BoNT/G, of which serotypes /A, /B, /E and /F5]. Each BoNT serotype can be further categorised into subtypes based on amino acid sequence identity. For example, there are currently eight known subtypes of BoNT/A (/A1–/A8), which share between 84% and 97% sequence identity [6]. While BoNTs are the most toxic biological molecules known to science, they are used in human therapy, especially BoNT/A1 [7].

The BoNTs contain three major functional domains, a binding domain located in the C-terminal half of the heavy chain (H_C), a translocation domain located in the N-terminal half of the heavy chain (H_N) and a Zn^{2+} -dependent protease domain located in the light chain (LC). The H_C is responsible for targeting the BoNT to the neuronal cell membrane by binding to specific gangliosides and protein receptors on the

neuronal cell surface. The H_N facilitates entry of the LC into the cytosol where it cleaves a target SNARE protein(s), which inhibits exocytosis. Although there are currently more than 46 different BoNT subtypes, there is limited structural information available for the majority of these natural variants. Many of these subtypes have been found to contain beneficial properties when compared to the commercially available toxins.

The BoNT subtypes from within the same serotype display a high degree of amino acid sequence identity and similarity; however, several studies have found distinct differences in their properties [8–12] (Fig. 1). Although the molecular basis of intoxication is not yet fully understood, the LC appears to define the length of intoxication (duration of action), while both H_N and H_C appear to be responsible for the spread and speed of cellular entry (onset of action). Considering the toxic nature of BoNTs, they are classed as tier 1 select agents due to their potential misuse in bioterrorism or as a bioweapon. From this perspective,

Abbreviations

BoNT, botulinum neurotoxin; GBS, ganglioside-binding site; H_C , cell-binding domain; H_N , translocation domain; LC, light chain.

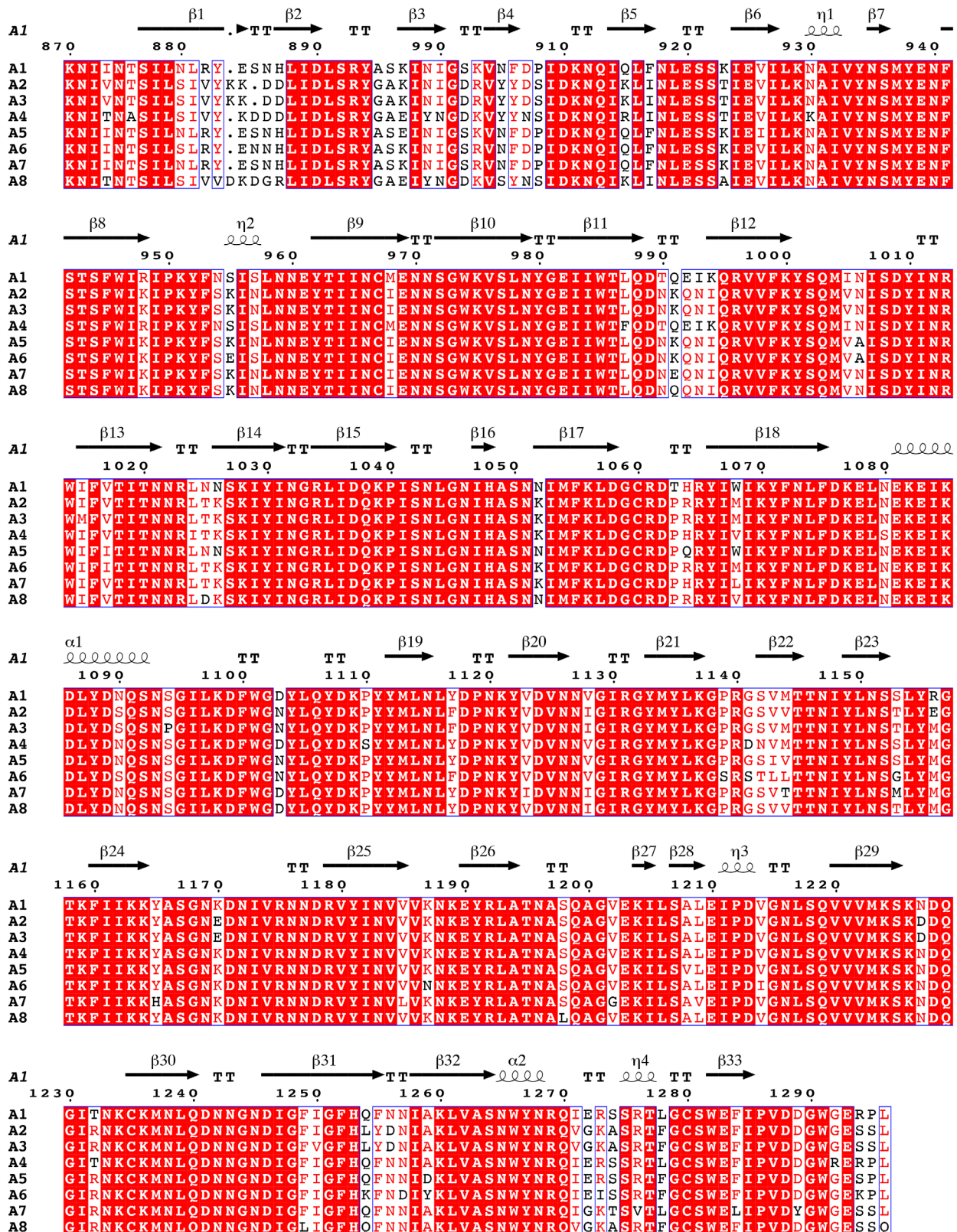


Fig. 1. Alignment of the binding domain sequences from BoNT/A1 to A8. BoNT/A1 numbering and secondary structure used for annotation. Figure generated using ESPript [34]

structural details of each subtype may aid the design of broadly BoNT-neutralising antibodies.

Previously, we and others have determined the crystal structures of the binding domains from BoNT subtypes /A1, /A2, /A3 and /A4, and the related /HA alone [13–17], and in complex with various receptors: H_C/A1-GT1b [18], H_C/A1-SV2C [19,20], H_C/A2-SV2C [14,21] and H_C/A3-GD1a [22]. Here, we report the crystal structures of the BoNT/A5 and BoNT/A6 receptor-binding domain and compare the binding sites with previous crystal structures of the BoNT/A subtype.

Materials and methods

All reagents used were purchased from Sigma-Aldrich (Dorset, UK) or Fisher Scientific (Leicestershire, UK) unless otherwise specified.

Protein expression and purification

The binding domain (residues 871–1296) of BoNT/A5 and BoNT/A6 was cloned into the pJ401 vector (Atum Bio, California, USA) from their respective full-length sequences (UniProtKB: C7BEA8 and C9WWY7) with an N-terminal 6xHis tag. Constructs were expressed and purified as described previously [17]. The N-terminal 6xHis tag was not removed from the proteins prior to crystallisation.

Protein crystallisation

Crystallisation conditions were screened using commercially available 96-well screens from Molecular Dimensions (Sheffield, UK) at 16 °C. H_C/A5 (4 mg·mL⁻¹) and H_C/A6 (6 mg·mL⁻¹) were dispensed using an Art Robbins Phoenix crystal screening nano-dispenser into 96-well 3-drop Intelliwell plates (Molecular Dimensions, UK). Multiple screening kits from Molecular Dimensions were used. Crystals of H_C/A5 were obtained using the sitting-drop vapour diffusion method with 0.1 M sodium formate, 0.1 M ammonium acetate, 0.1 M sodium citrate tribasic dihydrate, 0.1 M sodium potassium tartrate tetrahydrate, 0.1 M sodium oxamate, 0.1 M imidazole, 0.1 M 2-[N-morpholino]ethanesulfonic acid, pH 6.5, 10% (v/v) ethylene glycol, 10% (w/v) PEG 8000 from (MOR-PHEUS screen, condition G2) and flash-cooled in liquid nitrogen. Crystals of H_C/A6, however, were obtained using the hanging-drop vapour diffusion method and 0.2 M sodium acetate trihydrate, 0.1 M Bis-Tris propane-HCl pH 7.5, 22% (w/v) PEG 3350 (based on condition G7 of the PACT Premier screen) and flash-cooled in liquid nitrogen after cryoprotection with 1 : 1 50% (v/v) glycerol in reservoir solution.

X-ray data collection and structure determination

Complete X-ray diffraction data sets were collected from single crystals of H_C/A5 and H_C/A6 (3600 images each) using 0.1° oscillations and a wavelength of 0.98 Å at beamlines IO3 and IO4 (Diamond Light Source, Didcot, UK). Raw images were processed using DIALS [23], and integrated data were scaled and merged using Aimless [24] from the CCP4 suite [25]. The 3D structures of both proteins were solved by molecular replacement with PHASER [26] using the coordinates from Phyre2 web server homology models [27] as search models. Both models were manually built COOT [28] and refined with REFMAC [29] in the CCP4 suite of programs [25]. The structures were validated with PDB_REDO [30], MOLPROBITY [31] and WWPDB VALIDATION [32]. Crystallographic data processing and refinement statistics are given in Table 1. Structure-based figures were generated with either PyMOL (Schrödinger, LLC, New York, NY, USA) or MOE (Chemical Computing Group, Quebec).

Results and Discussion

Structure of the BoNT/A5-binding domain (H_C/A5)

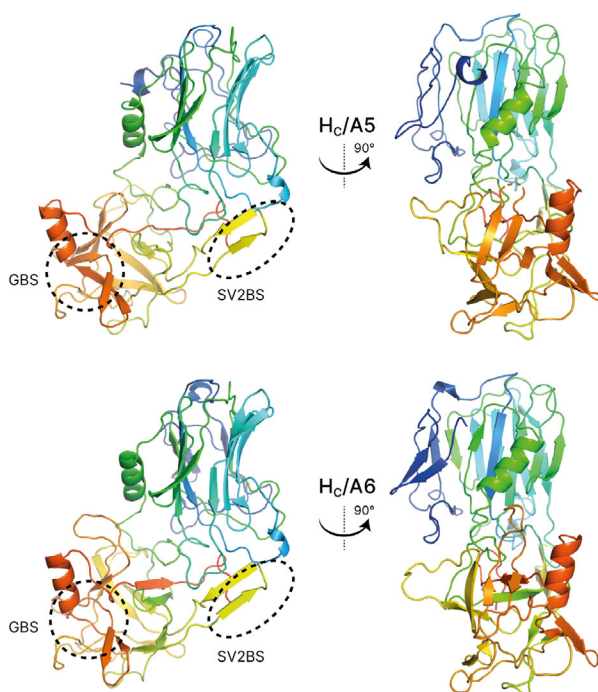
The crystal of H_C/A5 belonged to the orthorhombic space group *P*2₁2₁2₁, and it diffracted to a resolution of 1.15 Å (Table 1). Electron density was excellent throughout, with all H_C/A5 residues (except the N-terminal 6xHis tag and Lys871) being easily observed. The structure closely resembles the structures of other BoNT-binding domains [6] with an N-terminal jelly roll-like fold and C-terminal modified β-trefoil fold containing a conserved ganglioside-binding site (SxWY) (Fig. 2). However, compared to the structure of BoNT/A1 in complex with GT1b (PDB: 2VU9), the loop of residues 1260–1280, which contains ganglioside-interacting residues, adopts a different arrangement (Fig. 3a,b). It is possible that upon ganglioside binding, the loop changes conformation to allow S1275 and R1276 to take part in the interaction. In comparison with the unbound GD1a-binding site of BoNT/A1 (PDB: 3BTA) and BoNT/A3 (PDB: 6F0O), the corresponding site in H_C/A5 perhaps more resembles that of the latter rather than the former, which is consistent with a higher sequence identity between the sites at residues corresponding to positions 1117, 1254 and 1278 (Figs 1 and 4). Either way, considering that both BoNT/A1 and BoNT/A3 are able to bind to GD1a (PDB: 5TPB and 6THY, respectively), this suggests that BoNT/A5 is able to do so too.

In close proximity to the ganglioside binding loop was observed an unusual feature – a methylene bridge between the S_γ of Cys1280 and N_ε of Lys1236

Table 1. X-ray data collection and refinement parameters. Outer shell statistics are shown in parentheses

	H _C /A5	H _C /A6
Beamline	I03, DLS	I04-1, DLS
Wavelength (Å)	0.9763	0.9159
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit-cell parameters		
a,b,c (Å)	43.55, 60.27, 185.15	39.54, 105.59, 112.41
α = β = γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	92.57–1.15 (1.17– 1.15)	112.41–1.35 (1.37– 1.35)
R _{merge} (%)	0.087 (1.098)	0.094 (1.725)
R _{meas} (%)	0.090 (1.158)	0.099 (1.839)
R _{pim} (%)	0.026 (0.363)	0.030 (0.631)
CC _{1/2} (%)	0.998 (0.447)	0.999 (0.463)
Mean <I/σ(I)>	11.1 (1.9)	10.7 (0.9)
Completeness (%)	100 (99.5)	100 (99.7)
No. of observed reflections	2,107,443 (84,519)	1,099,418 (43,053)
No. of unique reflections	173,797 (8,450)	104,434 (5,129)
Multiplicity	12.1 (10.0)	10.5 (8.4)
Refinement statistics		
R _{work} /R _{free}	0.135/0.161	0.147/0.168
RMSD bond lengths (Å)	0.02	0.01
RMSD bond angles (°)	2.20	1.66
Ramachandran statistics (%)		
Favoured	96.3	97.0
Allowed	3.7	3.0
Outliers	0	0
Wilson B-factor (Å ²)	11.4	13.5
Average B-factors (Å ²)		
Protein	16.8	20.3
Water	32.6	31.5
No. of atoms		
Protein	3827	3633
Water	522	423
PDB code	6TWP	6TWO

(Fig. 5a), rather than a disulfide bond with a nearby cysteine residue (Cys1235). During refinement of the H_C/A5 structure, clear electron density was observed between the side chains of Cys1280 and Lys1236, into which a methylene group could be fitted. Weak anomalous data recorded at the start of data collection were used to generate a low-resolution anomalous difference map. Despite the noise, large peaks were observed at the location of sulfur atoms, which confirmed the location of each cysteine residue (Fig. 5b). This specific methylene bond between a lysine and cysteine side chain is unusual, and the mechanism surrounding the formation of a methylene-bridged lysine

**Fig. 2.** Crystal structures of H_C/A5 and H_C/A6. Overall structure of the binding domain (H_C) of BoNT/A5 (top) and BoNT/A6 (bottom). Putative ganglioside- and SV2-binding sites are indicated by a dashed ellipse labelled GBS and SV2BS, respectively.

and cysteine is not fully understood [33]. Whether this bond is biologically relevant remains to be established. While there are indications of this bond in the electron density maps of other BoNT crystal structures, it is possible that this may be an artefact of exposure to synchrotron radiation.

Inspection of the H_C/A5 structure corresponding to the BoNT/A1 SV2C-binding site (¹¹³⁹PRGSVMTT¹¹⁴⁶ + Arg1156) reveals the presence of perhaps a slightly shortened β-hairpin (Fig. 3d,e). The three different residues at positions 1143, 1144 and 1156 (V → I, M → V and R → M, respectively) do not appear to preclude the possibility of SV2C binding. Indeed, the related binding domain of BoNT/HA possesses the same residues at the corresponding location and is still able to bind to SV2C [20]. However, inspection of the accompanying SV2C glycan-binding site reveals one residue (Gln1064) potentially hindering the binding of glycan (Fig. 3g,h). This residue, corresponding to His1064 in BoNT/A1, has been shown to drastically decrease the binding affinity to SV2C [20]. Although this suggests that SV2C may not be the protein receptor for BoNT/A5, it should be noted that there exists a second BoNT/A5 sequence that differs by this one residue (UniProtKB: C11PK2).

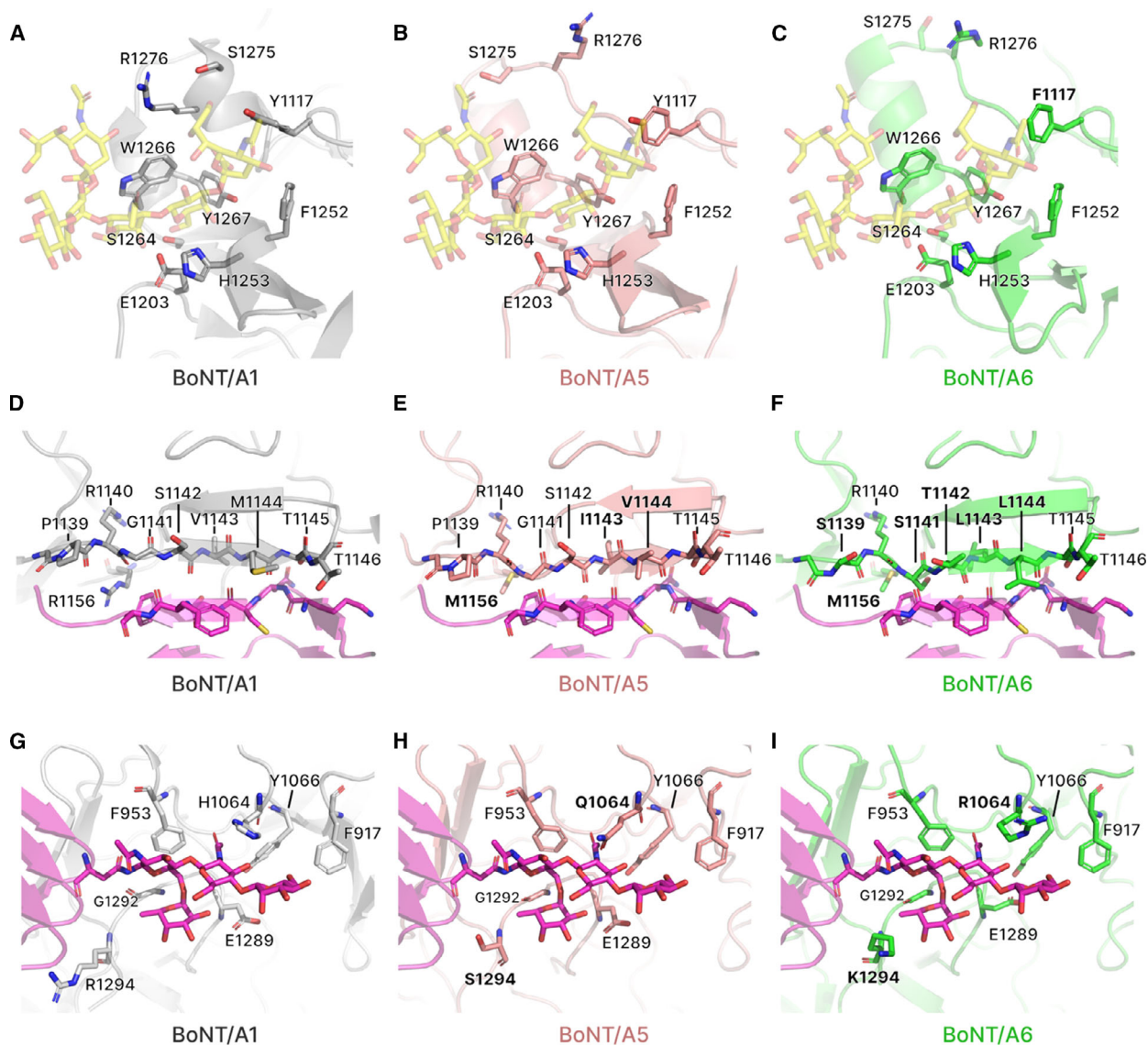


Fig. 3. Comparison of receptor-binding sites of BoNT/A5 and BoNT/A6. (A–C) Comparison of the ganglioside-binding site of $H_C/A1$ (PDB: 2VU9), $H_C/A5$ and $H_C/A6$. Note that the ganglioside, GT1b (yellow), has been superposed into the putative site of $H_C/A5$ and $H_C/A6$. (D–F) Comparison of the SV2-binding site of $H_C/A1$ (PDB: 5JLV), $H_C/A5$ and $H_C/A6$. Note that the protein receptor, SV2C (purple), has been superposed into the putative site of $H_C/A5$ and $H_C/A6$. (G–I) Comparison of the SV2 glycan-binding site of $H_C/A1$ (PDB: 5JLV), $H_C/A5$ and $H_C/A6$. Note that the SV2C glycans (purple) have been superposed into the putative site of $H_C/A5$ and $H_C/A6$. Residues involved in receptor binding are indicated and those which differ from BoNT/A1 are highlighted in bold.

Structure of the BoNT/A6-binding domain ($H_C/A6$)

The crystals of $H_C/A6$ belong to orthorhombic space group $P2_12_12_1$ and diffracted to a resolution of 1.35 Å. Electron density was excellent, with all but the first six residues of $H_C/A6$ being clearly observed, and like $H_C/A5$, the overall protein fold was highly similar to other BoNT H_C structures. The ganglioside-binding site was identical to that of $H_C/A5$ except for residue 1117, which was a Phe rather than a Tyr (Fig. 3b,c).

Although the absence of the hydroxyl group would result in the loss of hydrogen bonding with the terminal sialic acid of GT1b, the side chain can still continue to interact with the carbon ring. Compared to the unbound GD1a-binding site of BoNT/A1 (PDB: 3BTA) and BoNT/A3 (PDB: 6F0O), the corresponding site in $H_C/A6$ also more resembles that latter rather than the former, even though there is no greater sequence identity between the sites at residues

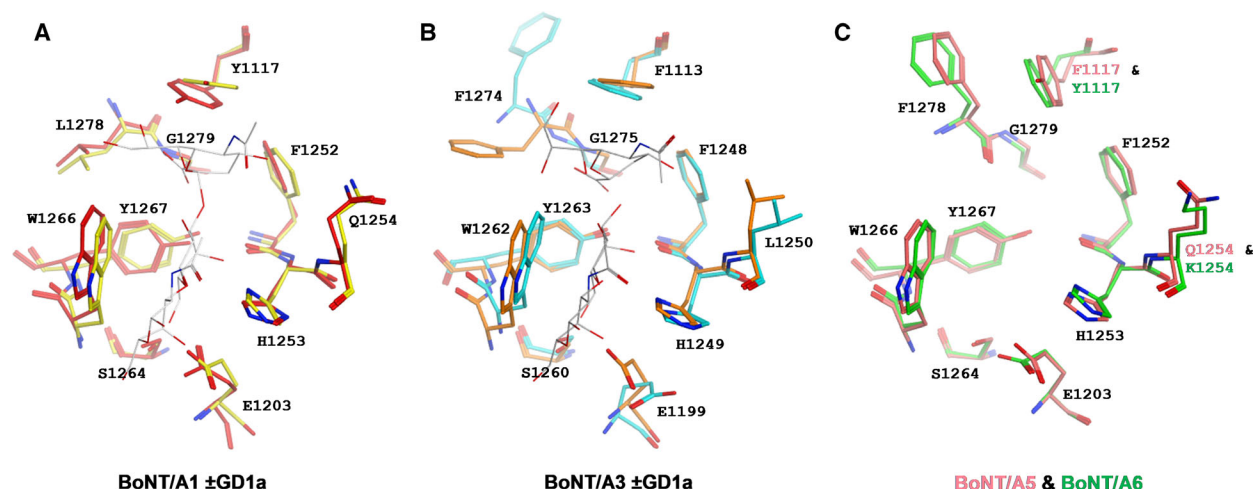


Fig. 4. Comparison of the GD1a-binding site. (A) The GD1a-binding site of BoNT/A1 is shown in the bound (red; PDB: 5TPB) and unbound (yellow; PDB: 3BTA) conformation (the interacting sugar moieties are shown in white). (B) The GD1a-binding site of BoNT/A3 is shown in the bound (orange; PDB: 6THY) and unbound (cyan; PDB: 6F0O) conformation (the interacting sugar moieties are shown in grey). (C) The putative GD1a-binding site of BoNT/A5 (pink) and BoNT/A6 (green) is shown in the unbound conformation. Note that residues at positions 1117, 1254 and 1278 either match the corresponding residue in BoNT/A1, BoNT/A3 or neither.

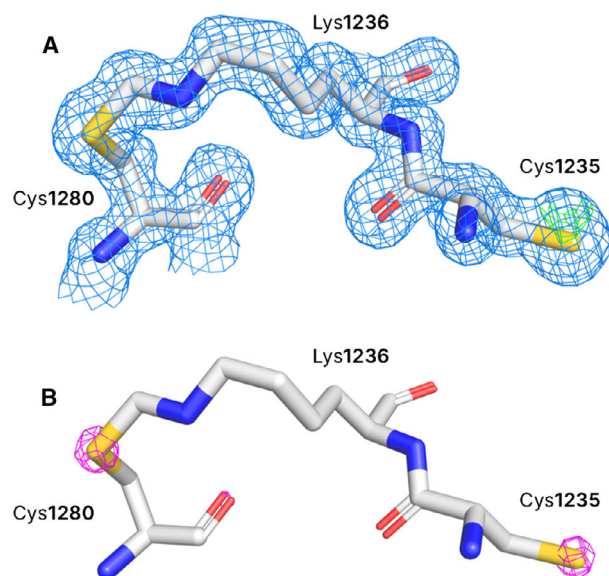


Fig. 5. Electron density around Cys1280 and Lys1236 in $H_C/A5$. (A) The electron density around Cys1280 and Lys1236 is clear and continuous, indicating the presence of a covalent bond between the two side chains. The $2mF_o-F_c$ map (blue) is contoured to $0.5 \text{ e}/\text{\AA}^3$, and the F_o-F_c (negative: red; positive: green) map is contoured to $0.46 \text{ e}/\text{\AA}^3$. (B) A weak sulfur-SAD signal within the diffraction data enabled the calculation of an anomalous map (contoured to $0.06 \text{ e}/\text{\AA}^3$; magenta), which indicates the precise position of each sulfur atom for each Cys residue.

corresponding to positions 1117, 1254 and 1278 (Figs 1 and 4). Like that for the $H_C/A5$ structure, BoNT/A6 is predicted to be able to bind to GD1a as well.

For the corresponding BoNT/A1-SV2C-binding site in $H_C/A6$, a larger sequence variation is observed: $^{1139}\text{SRSTLLTT}^{1146} + \text{Met1156}$ rather than $^{1139}\text{PRGSVMTT}^{1146} + \text{Arg1156}$ for BoNT/A1. Despite these differences, the β -hairpin remains available to bind to SV2C via mostly backbone-backbone hydrogen bonding (Fig. 3d,f). Like $H_C/A5$, $H_C/A6$ possesses a different residue in the glycan-binding site at position 1064 (Arg) compared to that of BoNT/A1 (His, Fig. 3g,i), and this has also been reported to significantly reduce binding of glycosylated SV2C [20]. This would suggest that BoNT/A6 may have a lower affinity for SV2C than BoNT/A1. Interestingly, BoNT/A2 also contains an Arg at position 1064 and it has previously been reported that both BoNT/A2 and BoNT/A6 are capable of entering hiPSC-derived neurons faster than BoNT/A1 [8].

Conclusions

The BoNT/A subtypes are believed to bind to the target cell surface via a dual-receptor complex involving a ganglioside and protein receptor. For BoNT/A1, they are GT1b (preferentially) and SV2C, respectively, but for most of the others, the exact identities of these receptors have not yet been determined. Structural analysis of the expected binding sites has revealed some differences with that of BoNT/A1, suggesting either an altered binding affinity to each receptor or a different receptor specificity altogether. Our high-resolution structures further add to the body of

knowledge around BoNT receptor binding and enhance the available molecular information for engineering novel therapeutic BoNTs and BoNT-binding moieties.

Acknowledgements

We thank Diamond Light Source (UK) for the use of beamlines IO3 and IO4-1 (proposal mx17212). J.R.D. was supported by a joint postgraduate studentship between University of Bath (UK) and Ipsen Bioinnovation Limited.

Data accessibility

Accession codes: The atomic coordinates and structure factors (codes 6TWP and 6TWO) have been deposited in the Protein Data Bank (<http://www.pdb.org>).

Author contributions

JRD performed purification and crystallisation of H_C/A5, supervised AB, processed and analysed the data, and drafted the manuscript. AB performed purification and crystallisation of H_C/A6. SML analysed the data and edited the manuscript. KRA supervised the study, analysed the data, and edited the manuscript.

Conflict of interest

The authors JRD, AB and KRA from the University of Bath declare no competing financial interests. SML is an employee of Ipsen Bioinnovation Limited.

References

- Schiavo G, Matteoli M and Montecucco C (2000) Neurotoxins affecting neuroexocytosis. *Physiol Rev* **80**, 717–766.
- Johnson EA and Montecucco C (2008) Botulism. *Handb Clin Neurol* **91**, 333–368.
- Rossetto O, Pirazzini M and Montecucco C (2014) Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat Rev Microbiol* **12**, 535–549.
- Dong M, Masuyer G and Stenmark P (2019) Botulinum and tetanus neurotoxins. *Annu Rev Biochem* **88**, 811–837.
- Coffield JA, Bakry N, Carlson J, Gomella LG and Simpson LL (1997) In vitro characterization of botulinum toxin types A, C and D action on human tissues: combined electrophysiologic, pharmacologic and molecular biologic approaches. *J Pharmacol Exp Ther* **280**, 1489–1498.
- Davies JR, Liu SM and Acharya KR (2018a) Variations in the botulinum neurotoxin binding domain and the potential for novel therapeutics. *Toxins (Basel)* **10**, E421.
- Fonfria E, Maignel J, Lezmi S, Martin V, Splevins A, Shubber S, Kalinichev M, Foster K, Picaut P and Krupp J (2018) The expanding therapeutic utility of botulinum neurotoxins. *Toxins (Basel)* **10**, E208.
- Moritz MS, Tepp WH, Bradshaw M, Johnson EA and Pellett S (2018) Isolation and characterization of the novel botulinum neurotoxin A subtype 6. *mSphere* **3**, 1–12.
- Wang D, Krilich J, Pellett S, Baudys J, Tepp WH, Barr JR, Johnson EA and Kalb SR (2013) Comparison of the catalytic properties of the botulinum neurotoxin subtypes A1 and A5. *Biochim Biophys Acta* **834**, 2722–2728.
- Whitemarsh RCM, Tepp WH, Bradshaw M, Lin G, Pier CL, Scherf JM, Johnson EA and Pellett S (2013) Characterization of botulinum neurotoxin A subtypes 1 through 5 by investigation of activities in mice, in neuronal cell cultures, and in vitro. *Infect Immun* **81**, 3894–3902.
- Pellett S, Tepp WH, Whitemarsh RC, Bradshaw M and Johnson EA (2015) In vivo onset and duration of action varies for botulinum neurotoxin A subtypes 1–5. *Toxicon* **107**, 37–42.
- Pellett S, Bradshaw M, Tepp WH, Pier CL, Whitemarsh RCM, Chen C, Barbieri JT and Johnson EA (2018) The light chain defines the duration of action of botulinum toxin serotype A subtypes. *MBio* **9**, e00089-18.
- Lacy DB, Tepp W, Cohen AC, DasGupta BR and Stevens RC (1998) Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nat Struct Biol* **5**, 898–902.
- Benoit RM, Schärer MA, Wieser MM, Li X, Frey D and Kammerer RA (2017) Crystal structure of the BoNT/A2 receptor-binding domain in complex with the luminal domain of its neuronal receptor SV2C. *Sci Rep* **7**, 43588.
- Davies JR, Hackett GS, Liu SM and Acharya KR (2018b) High resolution crystal structures of the receptor-binding domain of *Clostridium botulinum* neurotoxin serotypes A and FA. *PeerJ* **6**, e4552.
- Yao G, Lam KH, Perry K, Weisemann J, Rummel A and Jin R (2017) Crystal structure of the receptor-binding domain of botulinum neurotoxin type HA, also known as type FA or H. *Toxins (Basel)* **9**, E93.
- Davies JR, Rees J, Liu SM and Acharya KR (2018c) High resolution crystal structures of *Clostridium botulinum* neurotoxin A3 and A4 binding domains. *J Struct Biol* **202**, 113–117.
- Stenmark P, Dupuy J, Imamura A, Kiso M and Stevens RC (2008) Crystal structure of botulinum

- neurotoxin type A in complex with the cell surface co-receptor GT1b-insight into the toxin-neuron interaction. *PLoS Pathog* **4**, e1000129.
- 19 Benoit RM, Frey D, Hilbert M, Kevenaar JT, Wieser MM, Stirnimann CU, McMillan D, Ceska T, Lebon F, Jaussi R *et al.* (2014) Structural basis for recognition of synaptic vesicle protein 2C by botulinum neurotoxin A. *Nature* **505**, 108–111.
- 20 Yao G, Zhang S, Mahrhold S, Lam KH, Stern D, Bagramyan K, Perry K, Kalkum M, Rummel A, Dong M *et al.* (2016) N-linked glycosylation of SV2 is required for binding and uptake of botulinum neurotoxin A. *Nat Struct Mol Biol* **23**, 656–662.
- 21 Gustafsson R, Zhang S, Masuyer G, Dong M and Stenmark P (2018) Crystal structure of botulinum neurotoxin A2 in complex with the human protein receptor SV2C reveals plasticity in receptor binding. *Toxins (Basel)* **10**, E153.
- 22 Gregory KS, Liu SM and Acharya KR (2020) Crystal structure of botulinum neurotoxin subtype A3 cell binding domain in complex with GD1a co-receptor ganglioside. *FEBS Open Bio* **10**, 298–305.
- 23 Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, Fuentes-Montero L, Vollmer M, Michels-Clark T, Young ID *et al.* (2018) DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr* **D74**, 85–97.
- 24 Evans PR and Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr* **D69**, 1204–1214.
- 25 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW and McCoy A (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr* **D67**, 235–242.
- 26 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC and Read RJ (2007) Phaser crystallographic software. *J Appl Cryst* **40**, 658–674.
- 27 Kelley LA, Mezulis S, Yates CM, Wass MN and Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845–858.
- 28 Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr* **D66**, 486–501.
- 29 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F and Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr* **D67**, 355–367.
- 30 Joosten RP, Long F, Murshudov GN and Perrakis A (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCr J* **1**, 213–220.
- 31 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS and Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst* **D66**, 12–21.
- 32 Gore S, Sanz García E, Hendrickx PMS, Gutmanas A, Westbrook JD, Yang H, Feng Z, Baskaran K, Berrisford JM, Hudson BP *et al.* (2017) Validation of structures in the Protein Data Bank. *Structure* **25**, 1916–1927.
- 33 Ruzsalkowski M and Dauter Z (2016) On methylene-bridged cysteine and lysine residues in proteins. *Protein Sci* **25**, 1734–1736.
- 34 Robert X and Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucl Acids Res* **42**, W320–W324.