

The cryo-EM structure of the BoNT/Wo-NTNH complex reveals two immunoglobulin-like domains

Sara Košenina¹ , Jana Škerlová^{1,2} , Sicai Zhang^{3,4,5}, Min Dong^{3,4,5} and Pål Stenmark¹ 

1 Department of Biochemistry and Biophysics, Stockholm University, Sweden

2 Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague, Czech Republic

3 Department of Urology, Boston Children's Hospital, MA, USA

4 Department of Microbiology, Harvard Medical School, Boston, MA, USA

5 Department of Surgery, Harvard Medical School, Boston, MA, USA

Keywords

BoNT-NTNH complex; botulinum neurotoxins; cryo-EM structure; *Weissella oryzae*

Correspondence

P. Stenmark, Department of Biochemistry and Biophysics, Stockholm University, Svante Arrhenius väg 16C SE-106 91, Stockholm, Sweden
 Tel: +46 8 16 37 29
 E-mail: stenmark@dbb.su.se

(Received 3 July 2023, revised 7 September 2023, accepted 20 September 2023)

doi:10.1111/febs.16964

The botulinum neurotoxin-like toxin from *Weissella oryzae* (BoNT/Wo) is one of the BoNT-like toxins recently identified outside of the *Clostridium* genus. We show that, like the canonical BoNTs, BoNT/Wo forms a complex with its non-toxic non-hemagglutinin (NTNH) partner, which in traditional BoNT serotypes protects the toxin from proteases and the acidic environment of the hosts' guts. We here report the cryo-EM structure of the 300 kDa BoNT/Wo-NTNH/Wo complex together with pH stability studies of the complex. The structure reveals molecular details of the toxin's interactions with its protective partner. The overall structural arrangement is similar to other reported BoNT-NTNH complexes, but NTNH/Wo uniquely contains two extra bacterial immunoglobulin-like (Big) domains on the C-terminus. Although the function of these Big domains is unknown, they are structurally most similar to bacterial proteins involved in adhesion to host cells. In addition, the BoNT/Wo protease domain contains an internal disulfide bond not seen in other BoNTs. Mass photometry analysis revealed that the BoNT/Wo-NTNH/Wo complex is stable under acidic conditions and may dissociate at neutral to basic pH. These findings established that BoNT/Wo-NTNH/Wo shares the general fold of canonical BoNT-NTNH complexes. The presence of unique structural features suggests that it may have an alternative mode of activation, translocation and recognition of host cells, raising interesting questions about the activity and the mechanism of action of BoNT/Wo as well as about its target environment, receptors and substrates.

Introduction

Botulinum neurotoxins (BoNTs) are the most toxic proteins known. They are typically produced by spore-forming anaerobic bacteria of the genus *Clostridium*. Since the first description of botulism, the disease caused by botulinum neurotoxins, in 1897 [1] seven

canonical serotypes of the toxin have been identified – BoNT/A-G [2–7]. In addition the BoNT-like toxins BoNT/X, BoNT/En, PMP1 and BoNT/Wo have recently been identified [8–12]. All of them specifically target neuromuscular junctions where they cleave the

Abbreviations

Big, bacterial immunoglobulin-like; *bont*, gene for BoNT; BoNTs, botulinum neurotoxins; *ha*, gene for hemagglutinin protein; HA, hemagglutinin protein; H_C, binding domain; HC, heavy chain – translocation and binding domain; H_N, translocation domain; Ig, immunoglobulin; LC, light chain – catalytic domain; L-PTC, large functional progenitor toxin complex; M-PTC, minimally functional progenitor toxin complex; *ntnh*, gene for NTNH; NTNH, non-toxic non-hemagglutinin protein; *orfX*, gene for OrfX protein.

proteins from the SNARE family, thereby blocking the release of the neurotransmitter acetylcholine, which leads to paralysis and if not treated, even death [13]. Interestingly, despite being classified as a potential bio-terrorism threat [14], BoNTs are being used for the treatment of a large and increasing number of medical conditions, including movement disorders [15], overactive bladder [16], and severe pain [17,18].

BoNTs most often enter the target organism by consumption of contaminated food and then cross the intestinal epithelial barrier in order to reach the general circulation and travel to the peripheral cholinergic nerve terminals [19]. BoNTs utilize a dual receptor binding mechanism to target neuromuscular junctions selectively. Then they use synaptic vesicles to enter motor neurons [20]. From the synaptic vesicle, the catalytic domain is translocated to the cytosol using a mechanism driven by the transmembrane pH gradient generated by the ATPase proton pump. Upon disulfide bond reduction, the catalytic metalloprotease domain is released in the cytosol where it cleaves one of the SNARE proteins, causing flaccid paralysis [21,22].

The *bont* gene is typically located next to a gene for the non-toxic non-hemagglutinin protein (*ntnh*), and depending on the serotype, genes for either the hemagglutinin proteins (*ha*) or the OrfX proteins (*orfX*). BoNT and its corresponding non-toxic non-hemagglutinin (NTNH) partner protein form a heterodimer, also known as the minimally functional progenitor toxin complex (M-PTC) [23]. Both sequence and structural conservation of BoNT and NTNH suggest that the two genes are a result of a gene duplication event after which NTNH evolved to have a protective function [24]. NTNH has been shown to protect BoNT from external damage, such as pH denaturation and protein-modifying agents [24], thereby shielding it from the acidic pH and proteolytic degradation in the gastrointestinal tract of the target species. Moreover, in BoNT/A, the interaction with its corresponding NTNH/A is important for BoNT to cross the intestinal barrier and reach the nervous system, as the n-loop of NTNH/A is the attachment point for the HA protein complex [25–27]. The HA proteins together with M-PTC form a large progenitor toxin complex (L-PTC) which facilitates the cross-intestinal transport of BoNT [25–27]. The function of OrfX proteins is still unknown and it is not clear whether they form a complex with M-PTC. However, all OrfX proteins have been shown to belong to the tubular lipid-binding family of proteins [28–31].

The assembly of the M-PTC is pH-dependent [32] and greatly enhances the potency of the neurotoxins when ingested orally [33]. To date, the structural information available for M-PTC complexes is the crystal

or cryo-EM structures of BoNT/A, BoNT/E and BoNT/X in complex with their respective NTNHS [34–36]. NTNH has a highly similar structure to BoNT, despite a low sequence similarity. It displays a three-domain organization that corresponds to the binding domain (H_C), the translocation domain (H_N), and the catalytic domain (LC). However, NTNH lacks several of the key features of BoNTs, such as the catalytic zinc-binding motif, the activation loop, and the disulfide bridge that is involved in the translocation of the LC, and they contain none of the ganglioside-binding motifs [27,34].

Recently the first BoNT homologs outside the *Clostridium* genus were reported [8–12]. A BoNT-like protein was identified in the genome of the Gram-positive bacterium *Weissella oryzae* and designated as BoNT/Wo. The sample was isolated from fermented rice grain and the toxin has been suggested to target insects or worms [11,37]. While BoNT/Wo shows clear homology to BoNTs and the conserved HExxH metalloprotease motif [38], it presents a low sequence identity (18%) compared to other BoNTs, which is below the typical range for the BoNT family members (28–64%) [39]. Interestingly, the interchain disulfide bond, essential for the release of the LC [21], is missing in BoNT/Wo, suggesting a different mode of action. The order of the *bont-ntnh* genes is conserved among BoNT gene clusters, while for BoNT/Wo it is reversed. Moreover, genes for HA or OrfX proteins are missing [11,12]. So far, the only available structural information about BoNT/Wo is the crystal structure of the LC [40]. BoNT/Wo is evolutionarily distinct from all the other toxins and even from the other BoNT-like toxins BoNT/X, BoNT/En and PMP1 that form their own evolutionary branch [8–10]. It will be interesting to see how the evolutionary relationships within this toxin family unravel and additional new toxins are likely going to be discovered with the rapidly growing sequence databases.

Here, we report the cryo-EM structure of the BoNT/Wo M-PTC at a resolution of 2.97 Å, together with pH stability studies of the complex. The structure reveals molecular details of the toxin's interactions with its protective partner, providing the basis to understand its toxicity mechanism, which could also help to shed light on the evolution of the botulinum neurotoxins. The structure also reveals two additional bacterial immunoglobulin (Ig)-like (Big) domains, observed for the first time as parts of an NTNH, suggesting that NTNH/Wo might contribute to receptor recognition at some stage of BoNT/Wo intoxication. This could potentially also be the case for other members of this toxin family and warrants revisiting the role of NTNH also for canonical BoNTs.

Results and Discussion

The overall structure of the BoNT/Wo M-PTC

We present here the cryo-EM structure of the M-PTC-complex of BoNT/Wo at a nominal resolution of

2.97 Å, obtained from 247 143 particles selected from a total of 6232 recorded movies. Figure 1A shows an assessment of the local resolution throughout the entire map.

BoNT/Wo forms a tight, interlocked complex with the NTNH/Wo. The domains involved in complex

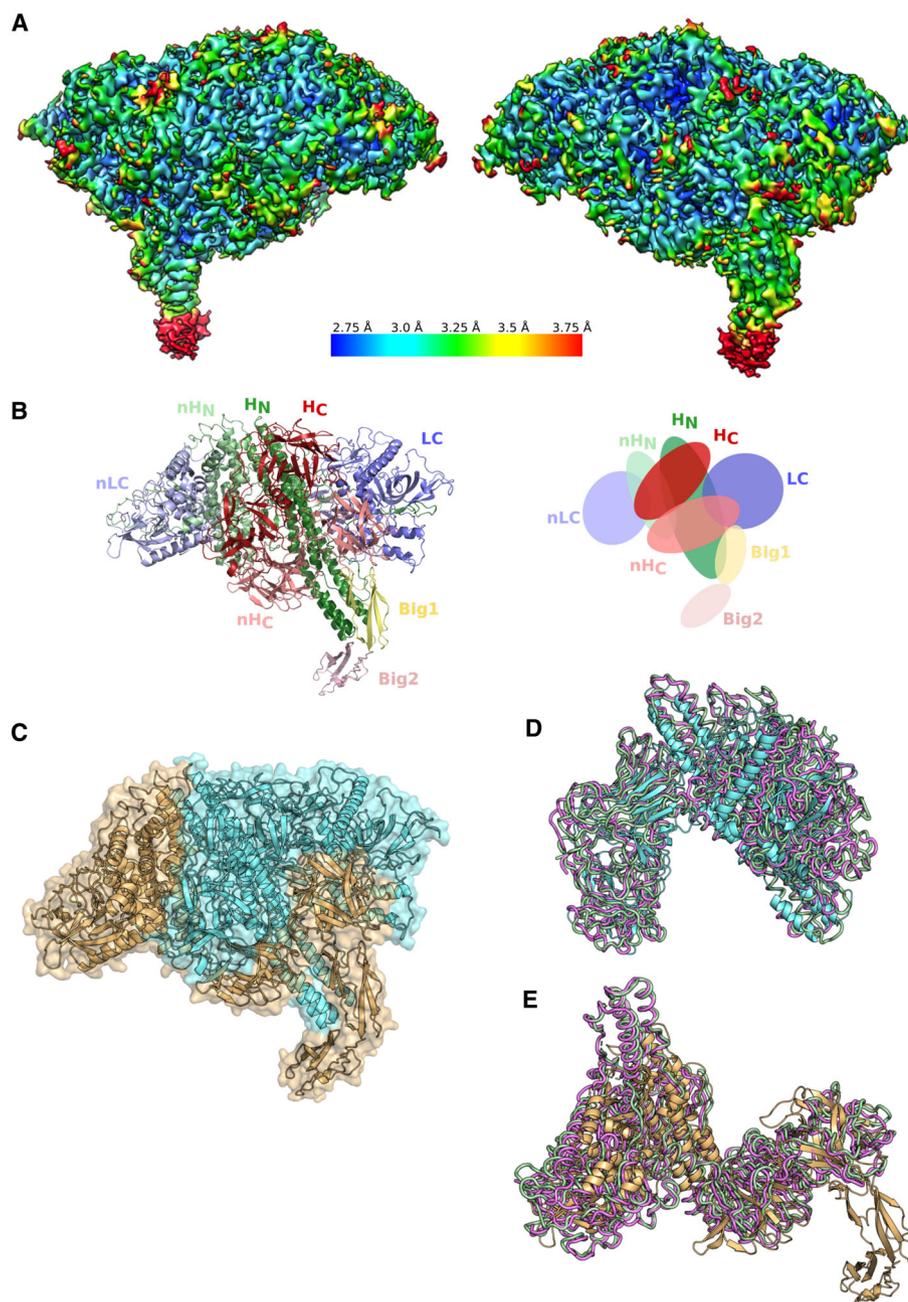


Fig. 1. (A) Local resolution estimates for the cryo-EM map of the BoNT/Wo M-PTC. (B) Domain organization of the BoNT/Wo M-PTC is shown as a cartoon representation (left) and as an illustration (right), with the BoNT/Wo and NTNH/Wo shown in dark and light shades, respectively. (C) The overall structure of the BoNT/Wo M-PTC with BoNT/Wo in cyan and NTNH/Wo in orange. (D) Comparison of BoNT/Wo (cyan) with BoNT/A (violet tube, PDB code 3V0A) and BoNT/E (green tube, PDB code 4ZKT). (E) Comparison of NTNH/Wo (orange) with NTNH/A (violet tube, PDB code 3V0A) and NTNH/E (green tube, PDB code 4ZKT).

Table 1. Pairwise comparison of BoNT/Wo and NTNH/Wo with BoNT/A M-PTC and BoNT/E M-PTC using Dali [36].

PDB	r.m.s.d (Å)	No. of aligned positions	No. of residues in matched structure	Sequence identity of aligned positions (%)	Z-score
BoNT/Wo					
3V0A	5.9	735	1280	14	24.5
4ZKT	5.6	787	1235	13	26.7
NTNH/Wo					
3V0A	4.2	959	1150	16	30.9
4ZKT	4.3	875	1114	16	26.9

formation are the H_N and H_C domains of BoNT/Wo and the corresponding domains of NTNH/Wo, while both LCs are solvent-exposed and not involved in this interaction (Fig. 1B,C). The overall structure of BoNT/Wo M-PTC is similar to the structures of M-PTC complexes of BoNT/A [34] (PDB code 3V0A) and BoNT/E [35] (PDB code 4ZKT) (Fig. 1D,E) as indicated by the high Z-scores and low r.m.s.d. values (Table 1) [41].

NTNH/Wo bacterial Ig-like domains

A major difference between BoNT/Wo M-PTC and other such complexes is the presence of additional Big domains (residues 1049–1120 and 1128–1199) on the C-terminus of the NTNH/Wo, which explains why NTNH/Wo is 10 kDa larger than NTNHs of other serotypes [11,12]. Interestingly, the Big domains extend into the solvent-accessible region and interact with the tip of the H_N/Wo helix bundle, creating a Y-shaped formation (Fig. 2A–C). Moreover, the characteristic helix bundle of the translocation domain in NTNH/Wo is unusually short in comparison with other translocation domains, measuring only ~70 Å, while the translocation domains of NTNH/A and /E are both ~105 Å long (Fig. 2E). This is a dramatic difference, likely connected to the unique presence of the Big domains in the NTNH/Wo. The functional consequences of these distinct features of the NTNH/Wo are however unknown.

In comparison to each other, the two NTNH/Big domains are almost identical (Fig. 2C) with an r.m.s.d. value of 1.4 Å. The Big domains belong to the immunoglobulin-like domain superfamily that has been identified in a variety of surface proteins. For example, *Escherichia coli* intimin and *Yersinia pseudotuberculosis* invasin, which contain a Big domain, have been shown to play important roles in invading host cells [42,43]. Additionally, Big domains in *Leptospira interrogans*

Ig-like proteins have been indicated to regulate the adhesion of pathogenic leptospire to host cells [44]. Moreover, a Dali [41] search against the PDB (Table 2, Fig. 2D) showed that the NTNH/Wo Big domain is structurally the most similar to proteins that are involved in adhesion to host cells and infection and immune evasion, such as SraP, a surface-exposed serine-rich repeat glycoprotein [45] and the bacterial surface protein Rib domain [46], respectively. All such domains extend into the solvent and have a flexible linker between separate Big domains, enabling interaction and binding to other molecules. Therefore, the NTNH/Wo Big domain could potentially be involved in the adhesion to the epithelial cells in the gut or airway and in the process of toxin delivery to the target cells. The *bont* and *ntnh* genes are typically accompanied in a gene cluster with genes that encode accessory proteins of the HA or OrfX type [47]. The HA proteins are known to interact with the NTNH n-loop and facilitate transport across the intestinal wall [25–27]. Since no such accessory proteins accompany BoNT/Wo in the *W. oryzae* genome [11], it is tempting to speculate that the Big domains take over their functions.

BoNT translocation domain

Canonical BoNTs have a conserved PYxGxAL motif, which functions in toxin translocation [11]. Moreover, the LC and the heavy chain (HC) in BoNTs are crosslinked by a disulfide bond between the two chains, which has been shown to be critical for the translocation of the LC across the endosomal membrane. The disulfide bond is reduced in the cytosol, resulting in LC liberation [21]. Interestingly, in BoNT/Wo this conserved motif and the disulfide bond have previously been shown to be absent [11,12]. Moreover, our cryo-EM structure shows that the activation loop is intact and significantly shorter compared to those of BoNT/A [34] (PDB code 3V0A) and BoNT/E [35] (PDB code 4ZKT) (Fig. 3A). Such a short loop is likely to be more resistant to proteolytic activation and suggests that BoNT/Wo may use a unique, unidentified mechanism for the translocation of the LC to the cytosol.

Structural analysis has revealed an unusual, highly negatively charged patch on the H_N, which interacts with a positively charged patch on the LC. A comparison of surface charges between BoNT/Wo and BoNT/A [34] (PDB code 3V0A) shows that BoNT/Wo is considerably more charged in the areas where the H_N and the LC interact (Fig. 3B–E). This strong electrostatic interaction could represent an alternative stabilization mechanism compensating for the absence

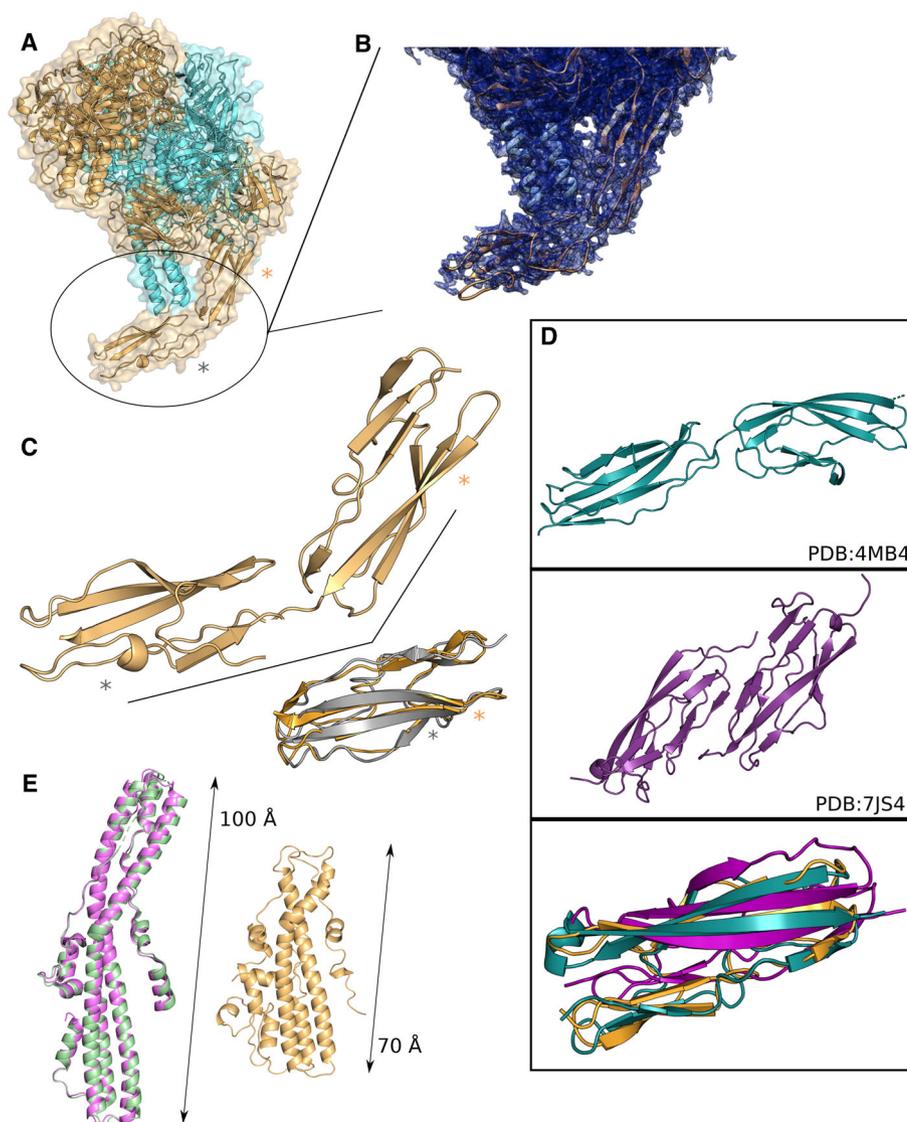


Fig. 2. (A) The Big domains of NTNH/Wo (orange) create a Y-shaped formation with the translocation domain of the toxin (cyan). (B) Cryo-EM map for the second Big domain extends into the solvent. (C) A zoom-in of the NTNH/Wo Big domain and a superposition of both NTNH/Wo Big domains (the star symbols are helping with the distinction between the two domains). (D) Some of the closest structural homologs of NTNH/Wo Big domain: a chitinase (deep teal, PDB code 4MB4), and O-glycopeptidase (purple, PDB code 7JS4), as well as a superposition of these Big domains with the N-terminal NTNH/Wo Big domain. (E) Comparison of the NTNH/Wo translocation domain (orange) with NTNH/A (violet, PDB code 3V0A) and NTNH/E (green, PDB code 4ZKT).

of the disulfide bridge and, in case the activation loop is cleaved, possibly helps to keep the LC and HC associated.

BoNT catalytic domain

The LC of BoNT/Wo is a catalytically active zinc-dependent metalloprotease previously reported to cleave the SNARE protein VAMP2 [37] and we have previously determined its crystal structure [40]. In

contrast to X-ray crystallography, it is more difficult to precisely locate protein-bound metal ions at a moderate resolution using cryo-EM, where images are generated based on bright-field phase contrast, being less sensitive to atomic number [48]. In agreement with this, the zinc ion was not well resolved in the structure of the BoNT/Wo M-PTC and we chose not to include it in the final model. We have determined the Zn^{2+} coordination in the free LC in great detail previously [40].

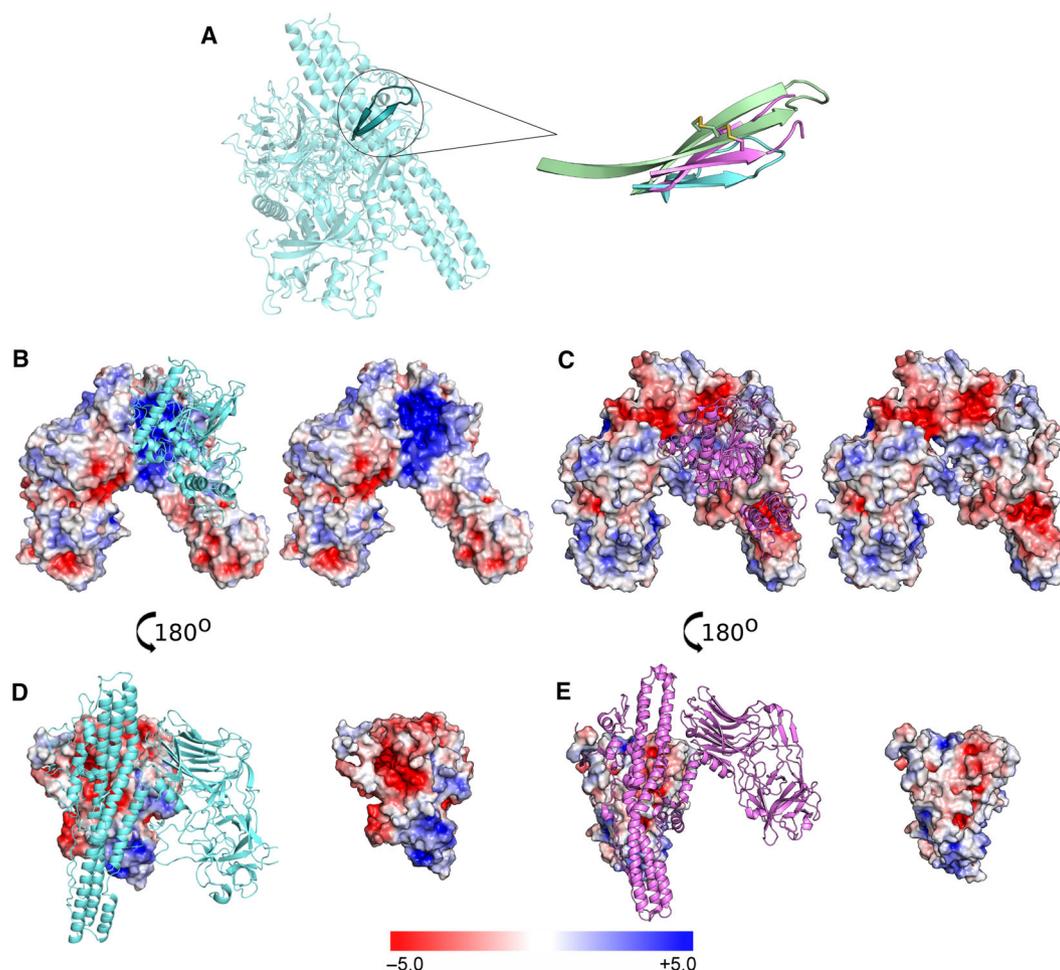


Fig. 3. (A) A comparison of the activation loop in BoNT/Wo (cyan) with activation loops in BoNT/A (violet, PDB code 3V0A) and BoNT/E (green, PDB code 4ZKT). (B–E) Comparison of HC and LC surface charge of BoNT/Wo (B, D) and BoNT/A (C, E). On the left side of each panel, a part of the toxin is shown as a cartoon for ease of orientation. (B) BoNT/Wo HC is shown as a surface representation and the LC is shown as a cyan cartoon. (C) BoNT/A HC is shown as a surface representation and the LC is shown as a violet cartoon. (D) BoNT/Wo LC is shown as a surface representation and HC is shown as a cyan cartoon. (E) BoNT/A LC is shown as a surface representation, and HC is shown as a violet cartoon. The surfaces are colored according to electrostatic potential (values in kT/e).

Table 2. Top 10 hits of Dali [36] comparison of NTNH/Wo Big domain to the whole PDB.

PDB	r.m.s.d (Å)	No. of aligned positions	No. of residues in matched structure	Sequence identity of aligned positions (%)	Z-score
4MB4	1.9	74	531	30	9.1
7JS4	2.8	73	953	21	7.5
2KPN	2.3	73	97	30	7.4
7JRL	3.2	74	457	20	7.3
4L3A	2.1	66	469	15	6.6
5FQ8	2.6	70	212	29	6.6
6S5Z	3.1	64	82	13	6.4
4M00	2.3	67	501	25	6.3
6S5Y	3.2	67	501	25	6.3
2C26	2.6	66	250	18	5.3

The catalytic pocket of BoNT/Wo has been shown to be wider and more negatively charged compared to other serotypes [40]. Remarkably, in the BoNT/Wo M-PTC the belt of BoNT/Wo dives deeper into the active site and interacts directly with the zinc ion and residues of the active site (Fig. 4A). Typically, the zinc ion is coordinated by the HExxH+E tetrahedral motif and a conserved Tyr stabilizes the catalytic intermediates [49,50]. However, in BoNT/Wo the Glu251 and the Tyr412 shift away, and instead the Asp551 from the belt takes part in the tetrahedral coordination of the zinc ion, replacing the Glu251. This interaction likely has an inhibitory function, blocking the active site of the LC until the toxin reaches its target site of action (Fig. 4B).

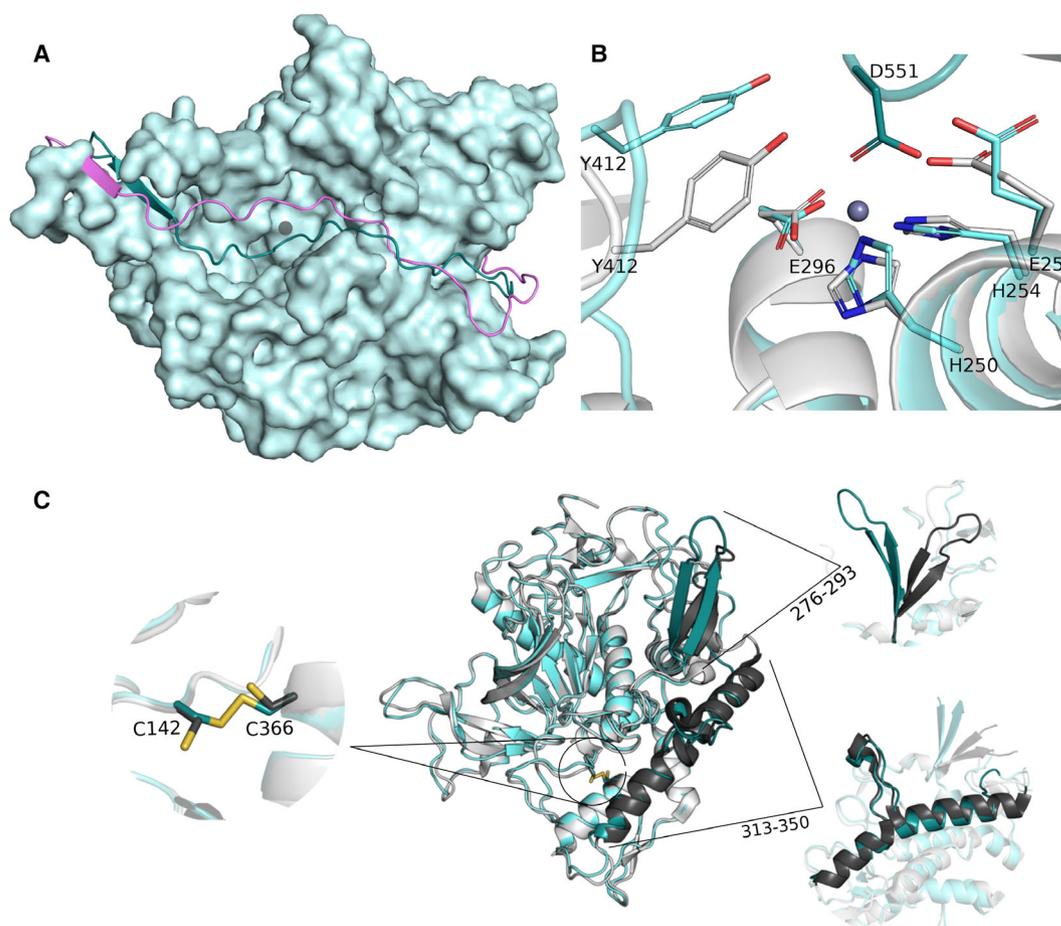


Fig. 4. (A) The surface of BoNT/Wo LC is shown in pale cyan and the belt region of BoNT/Wo is shown in dark gray. The location of the zinc ion is indicated by a gray circle. (B) Comparison with the active pocket from the LC crystal structure (light gray, PDB code 6RIM). (C) Highlighted sites of interest in the BoNT/Wo LC in the cryo-EM structure (cyan), as compared to the crystal structure (gray, PDB code 6RIM).

The BoNT/Wo LC exhibits a unique intrachain disulfide bridge (Fig. 4C), which is not present in any BoNT serotype, involving non-conserved amino acid residues Cys142 and Cys366. This bond was not observed in the crystal structure of BoNT/Wo LC (PDB 6RIM) [40], probably because of the reducing agent TCEP in the storage buffer, or potentially due to radiation damage. The disulfide bridge provides additional stabilization to the LC and could therefore make the potential translocation of the LC across the membrane more difficult since the process likely involves LC unfolding [51].

Comparison of the BoNT/Wo LC in the cryo-EM structure to the crystal structure of the LC (PDB 6RIM) [40] shows that the beta-hairpin including residues 276–293 in the M-PTC complex moves away from the LC to interact with the BoNT/Wo H_N , which forms a pocket that accommodates the loop (Fig. 4C). Additionally, the loop insertion (residues 325–337)

inside alpha-helix 313–350 observed in the crystal structure, which is unique to BoNT/Wo among all the LCs, has a similar conformation in the complex and also interacts with the BoNT/Wo H_N (Fig. 4C).

BoNT binding domain

BoNT/Wo does not display typical receptor binding sites. It lacks a conserved SxWY ganglioside binding motif, as well as the beta-hairpin motifs known to bind SV2. However, there are some surface regions on the binding domain of BoNT/Wo that exhibit a strong electrostatic potential and could be involved in the binding of yet unidentified receptors.

Interactions between BoNT/Wo and NTNH/Wo

The buried surface area between BoNT/Wo and NTNH/Wo is $\sim 7215 \text{ \AA}^2$, which is much larger than

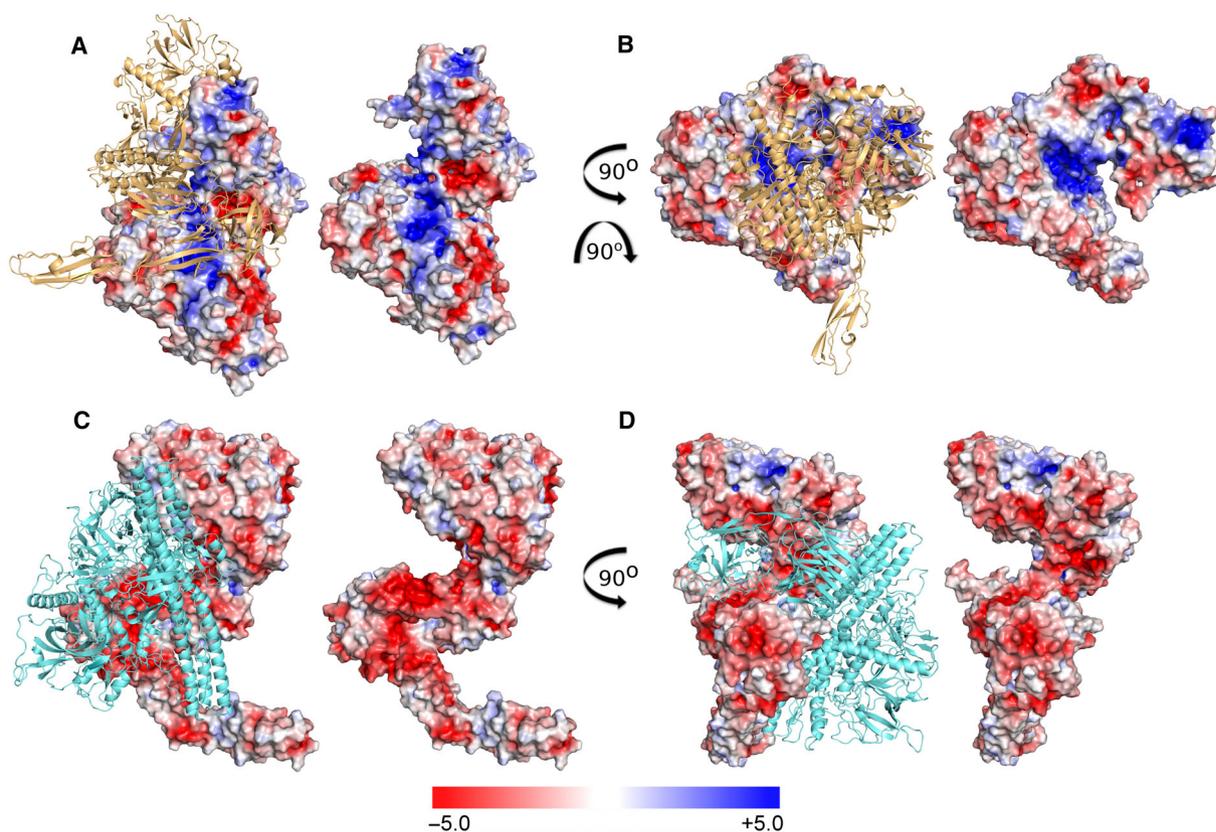


Fig. 5. Electrostatic interactions across the BoNT/Wo M-PTC interface. On the left side of each panel, a part of the complex is shown as a cartoon for ease of orientation. (A, B) BoNT/Wo is shown as a surface representation, highlighting the areas that form electrostatic interactions with NTNH/Wo (orange cartoon). (C, D) NTNH/Wo is shown as a surface representation, highlighting the areas that form electrostatic interactions with BoNT/Wo (cyan cartoon). The surfaces are colored according to electrostatic potential (values in kT/e).

the interaction interfaces of M-PTC complexes of BoNT/A, BoNT/E and BoNT/X, which measure ~ 3663 , 3404 and 4778 \AA^2 , respectively [52]. The extensive interaction surface is likely to be making the BoNT/Wo M-PTC more stable. The BoNT/Wo M-PTC interface is stabilized by a number of electrostatic interactions between charged residues of both proteins. Areas of BoNT/Wo taking part in complex formation are mostly positively charged (Fig. 5A,B), while those of NTNH/Wo are negatively charged (Fig. 5C,D). These areas include the following interactions: Residues 898–905 of the H_N and 302–303, 314–315 and 399–402 of the LC interact with residues 854–868 of the nH_C and 1046–1082 of the Big domain (Fig. 5A, C). Furthermore, residues 1123–1133 of the H_C and 893–905 of the H_N interact with residues 857–868 and 982–986 of the nH_C (Fig. 5B,D).

pH-dependent stability of the complex

It was very difficult to study the BoNT/Wo M-PTC using size-exclusion chromatography. The complex

interacted with the column material and we believe that the strength of this interaction was pH dependent, resulting in uninterpretable results. To study the disassembly of BoNT/Wo M-PTC, we instead utilized mass photometry [53]. 75 nm BoNT/Wo M-PTC was analyzed in buffers of different pH (5.5, 6.5, 7.5 and 8.5). The two peaks correspond to the BoNT/Wo M-PTC and the mixture of free BoNT/Wo and free NTNH/Wo, which are roughly the same size (Fig. 6). It is evident that the complex is stable at pH 5.5, where more than 70% of the proteins are forming a complex. The complex starts to disassemble between pH 6.5 and pH 7.5. At pH 7.5 only about 40% of the proteins are still forming a complex and at pH 8.5 the complex is almost completely disassembled. This agrees with the data for BoNT/A M-PTC and BoNT/X M-PTC, which disassemble at pH 7.5 and between 7.5 and 8.5, respectively [34,36]. However, we noticed that pH stability is concentration-dependent, so it is possible, that at higher concentrations the complex is stable even at a more basic pH.

The M-PTC of BoNT/A has been shown to have a patch of pH-sensing residues in the H_C that are

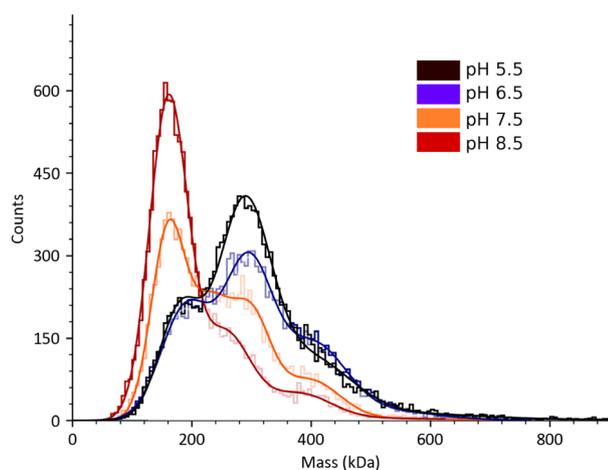


Fig. 6. Mass distribution for 75 nm BoNT/Wo M-PTC at several different pHs.

responsible for the neutral pH-driven disassembly of the M-PTC and the release of the toxin molecule [34]. Interactions between BoNT/Wo and NTNH/Wo are extensive; however, the residues involved in them are not localized as in BoNT/A M-PTC, so the pH sensing in M-PTC BoNT/Wo likely does not occur locally, but rather on a global level.

Conclusions

The cryo-EM structure of BoNT/Wo M-PTC reveals unique features of the BoNT/Wo M-PTC. BoNT/Wo lacks the essential disulfide bridge between the LC and HC, but has a non-conserved disulfide bond within the LC. In addition, it has an unusually short activation loop. These observations suggest an alternative mode of activation, translocation and liberation of the catalytic domain. Two unique Big domains are incorporated in NTNH/Wo and could be involved in the delivery of the toxin to the target cells. These findings raise interesting questions about the activity and the mechanism of action of BoNT/Wo as well as about its target environment, receptors and substrates.

Materials and methods

Cloning, protein expression and purification

DNA sequences encoding *W. oryzae bont* and *W. oryzae ntnh* were cloned into a pRSFDuet vector, where the C-terminus of *ntnh* contained a His-tag. The plasmid was transformed into *E. coli* BL21(DE3). Cultures were grown in TB media using the LEX bioreactor system (Harbinger Biotech, Toronto, ON, Canada) at 37 °C until an OD₆₀₀ value of 0.8, after which the temperature was reduced to

18 °C and protein expression was induced with 1 mM IPTG. The cells were harvested approximately 15 h post-induction by centrifugation.

Cells were resuspended in a buffer containing 20 mM MES pH 5.5, 200 mM NaCl, 5% glycerol and then lysed by sonication. The lysate was clarified by centrifugation at 45 000 *g* for 1 h before loading onto a pre-equilibrated 5 mL HisTrap HP column (GE Healthcare, Uppsala, Sweden). Purified protein was eluted with 500 mM imidazole in 20 mM MES pH 5.5, 200 mM NaCl, and 5% glycerol. Fractions containing BoNT/Wo M-PTC were pooled and further purified via size-exclusion chromatography (SEC) using a Superdex 200 16/60 column (GE Healthcare), pre-equilibrated with 20 mM MES pH 5.5, 200 mM NaCl, and 5% glycerol.

Mass photometry

Mass photometry was utilized to study the pH-dependent disassembly of the complex. A microscope coverslip was assembled with a sample well cassette. To find focus, fresh buffer was flown into the chamber, the focus position was identified and secured in place with an autofocus system based on total internal reflection for the entire measurement. For each acquisition, 5 μ L of the complex was diluted in 15 μ L of the corresponding buffer to the final concentration of 75 ng·mL⁻¹ and movies were recorded for 60 s. All measurements were performed using mass photometer Refyn Two_{MP} (Refyn Ltd, Oxford, UK), calibrated with NativeMark Unstained protein standard (Thermo Fischer Scientific, Waltham, MA, USA), and evaluated using software AcquireMP (Refyn Ltd). The mass photometry buffers used were 50 mM MES pH 5.5, 150 mM NaCl, 0.5 mM TCEP; 50 mM MES pH 6.5, 150 mM NaCl, 0.5 mM TCEP; 50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP and 50 mM TRIS pH 8.5, 150 mM NaCl, 0.5 mM TCEP.

Cryo-EM sample preparation

Quantifoil R2/2 Cu 300 mesh holey carbon grids (Quantifoil Micro Tools, Jena, Germany) were glow discharged (20 mA, 60 s, GloQube® Plus Glow Discharge System; Quorum Technologies, Laughton, UK). The sample was diluted in 25 mM MES pH 5.5, 150 mM NaCl to 0.7 mg·mL⁻¹, and 3 μ L were applied onto a grid using the Vitrobot blotting robot (FEI, Hillsboro, OR, USA) at 100% humidity and 4 °C, waiting 15 s for the sample to equilibrate before blotting for 1.5 s. Alternatively, Quantifoil R1.2/1.3 Cu 300 mesh holey carbon grids (Quantifoil Micro Tools) were glow discharged (40 mA, 90 s, GloQube® Plus Glow Discharge System; Quorum Technologies), coated with 0.2 mg·mL⁻¹ graphene oxide suspension using a drop-blot-wash technique [54] and 0.1 mg·mL⁻¹ BoNT/Wo M-PTC were applied. Grids were clipped, stored in liquid nitrogen, and screened on a Talos Arctica microscope operating at 200 kV and

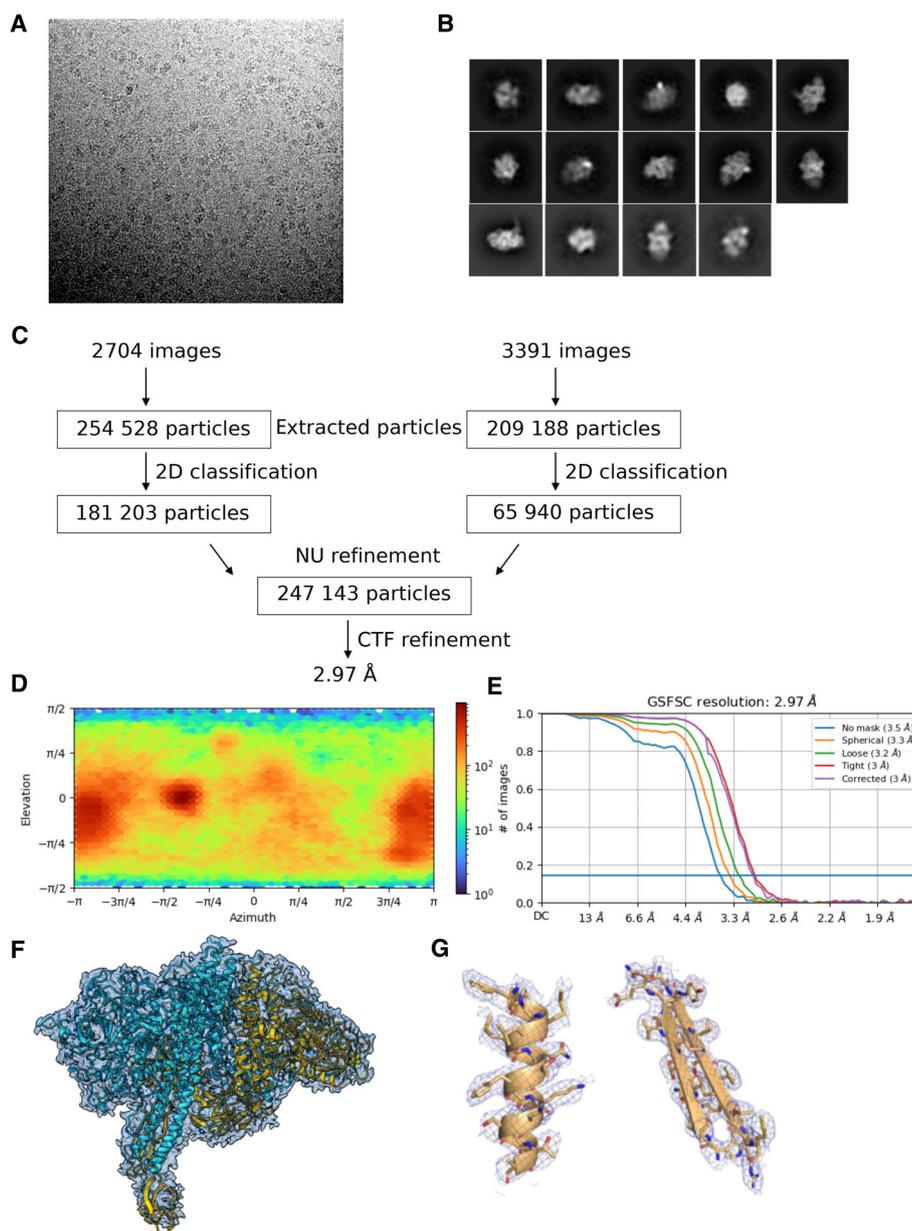


Fig. 7. Cryo-EM reconstruction of BoNT/Wo M-PTC. (A) A representative micrograph. (B) 2D classes that were selected for the 3D reconstruction and refinement. (C) Image processing pipeline, numbers indicate the number of particles selected for each reconstruction step. (D) Reconstruction and corresponding angular distribution map. (E) Fourier shell correlation (FSC) curves. (F) Fitting of the final atomic model into the final sharpened cryo-EM map. (G) Cryo-EM map around an alpha-helix and a beta-sheet.

equipped with a Falcon III direct electron detector (FEI) at the Stockholm node of the Swedish National Cryo-EM facility (SciLifeLab, Stockholm, Sweden).

Cryo-EM data acquisition and processing

Two data sets were acquired using the same settings at the Stockholm node of the Swedish National Cryo-EM facility (SciLifeLab) on a Titan Krios microscope (FEI) operating

at 300 kV and equipped with a K2 Summit direct electron detector (Gatan, Pleasanton, CA, USA). Movies were acquired at 165 000 \times nominal magnification with a pixel size of 0.82 $\text{\AA}\cdot\text{px}^{-1}$, for a total exposure time of 6 s over 40 frames; the total dose was 39.6 electrons $\cdot\text{\AA}^{-2}$. The first data set (Quantifoil R2/2 Cu 300 grids, 0.7 mg $\cdot\text{mL}^{-1}$ BoNT/Wo M-PTC) comprised 2732 recorded movies, and the second data set (graphene-oxide-coated Quantifoil R1.2/1.3 Cu 300, 0.1 mg $\cdot\text{mL}^{-1}$ BoNT/Wo M-PTC)

Table 3. Cryo-EM data collection and refinement.

	BoNT/Wo M-PTC
Data collection and processing	
Microscope	FEI Titan Krios
Nominal magnification	165 000×
Voltage (kV)	300
Defocus range (μm)	−1.6 to −3.2
Pixel size (Å)	0.82
Exposure time (s)	6
Number of frames	40
Total flux (e [−] ·Å ^{−2})	39.6
Number of images	2732 + 3500
Symmetry imposed	C1
Initial particle images (no.)	254 528 + 209 188
Final particle images (no.)	181 203 + 65 940
Map resolution (Å)	2.97
FSC threshold	0.143
B-factor applied	−100
Refinement	
Model composition	Chain A-BoNT/Wo, chain B-NTNH/Wo
Protein residues	2530
Ligands	0
B-factor (Å ²)	53.90
r.m.s.d. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	0.91
Validation	
MolProbity score	1.86
Clashscore	6.69
Poor rotamers (%)	0.88
Ramachandran statistics	
Favored (%)	91.81
Allowed (%)	7.95
Outliers (%)	0.24
PDB/EMDB ID	8C8G/16475

comprised 3500 recorded movies. Patch motion correction and patch CTF estimation were performed in cryoSPARC [55,56]. All downstream processing was also carried out in cryoSPARC (Fig. 7). In the first data set, a total of 254 528 particles were picked by blob picker followed by template picker from 2704 selected movies and after 2D classification 181 203 particles were selected. In the second data set, 65 940 particles out of 209 188 picked particles from 3391 selected movies were selected. Finally, the particle sets from both datasets were combined for 3D refinement, which included 247 143 particles. Global and local CTF refinements yielded the final map at 2.97 Å resolution, calculated based on the gold-standard FSC of 0.143 [57].

Model building

The BoNT/Wo M-PTC model with the exception of the second bacterial immunoglobulin-like domain was manually

built into the map using COOT [58]. The second Big domain model was produced in AlphaFold [59] and docked into the map as a rigid body. The whole structure excluding the second Big domain was refined with real-space refinement in PHENIX [60]. Protein structure validation was performed with MOLPROBITY [61]. Data statistics are summarized in Table 3. The atomic coordinates (PDB ID 8C8G) and the cryo-EM map (EMD-16475) were deposited in the Protein Data Bank (PDB) and the Electron Microscopy Data Bank (EMDB). Protein structure figures were rendered with PYMOL (Schrödinger, LLC, New York, NY, USA) and UCSF CHIMERA [62].

Acknowledgements

This work was supported by the Swedish Research Council (2022-03681, 2018-03406), the Swedish Cancer Society (20 1287 PjF) and Novo Nordisk Foundation (NNF20OC0064789) to PS. JŠ.'s work was supported by ERDF/ESF, OP RDE, the project "IOCB Mobility" (no. CZ.02.2.69/0.0/0.0/16_027/0008477) granted to the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences. MD was partially supported by grants from the National Institute of Health (NIH) (R01NS080833 and R01NS117626). MD holds the Investigator in the Pathogenesis of Infectious Disease award from the Burroughs Wellcome Fund. The data was collected at the Cryo-EM Swedish National Facility funded by the Knut and Alice Wallenberg, Family Erling Persson and Kempe Foundations, SciLifeLab, Stockholm University and Umeå University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank the Swedish National cryo-EM facility staff Dr. Marta Carroni, Dr. Julian Conrad and Dr. Michael Hall for their help during the acquisition of the cryo-EM datasets. Cryo-EM sample screening, optimization, and data collection were performed at the Cryo-EM Swedish National Facility, funded by the Knut and Alice Wallenberg, Family Erling Persson and Kempe Foundations, SciLifeLab, Stockholm University and Umeå University, at the SciLifeLab node in Stockholm and Umeå. We thank Dr. Markel Martinez Carranza for help with cryo-EM experiments, Dr. Anders Hofer and Dr. Saber Anooosheh for their help with mass photometry measurements and Dr. Geoffrey Masuyer for revision of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

PS and MD conceived the project and supervised the study. SK, JŠ and PS designed experiments; SK, JŠ and SZ performed experiments; SK, JŠ and PS analyzed data; SK wrote the manuscript; JŠ and PS revised the manuscript.

Peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.16964>.

Data availability statement

The atomic coordinates (PDB code [8C8G](#)) and cryo-EM reconstruction (code EMD-16475) have been deposited in the Protein Data Bank ([rcsb.org](#)) and the Electron Microscopy Database (<https://www.ebi.ac.uk/emdb/>), respectively.

References

- van Ermengem E (1897) Ueber einen neuen anaëroben *Bacillus* und seine Beziehungen zum Botulismus. *Z Hyg Infectiönsk* **26**, 1–56.
- Burke GS (1919) Notes on *Bacillus botulinus*. *J Bacteriol* **4**, 555–570.1.
- Bengtson IA (1922) Preliminary note on a toxin-producing anaerobe isolated from the larvae of *Lucilia caesar*. *Public Health Rep* **37**, 164–170.
- Meyer KF & Gunnison JB (1928) *Cl. botulinum* type D Sp. N. *Sci Proc* **26**, 88–89.
- Gunnison JB, Cummings JR & Meyer KF (1936) *Clostridium botulinum* type E. *Proc Soc Exp Biol Med* **35**, 278–280.
- Moller V & Scheibel I (1959) Preliminary report on the isolation of an apparently new type of *Cl. botulinum*. *Acta Pathol Microbiol Scand* **44**, 554–559.
- Gimenez D & Ciccarelli AS (1970) Another type of *Clostridium botulinum*. *Zentralbl Bakteriöl Orig* **215**, 221–224.
- Zhang S, Masuyer G, Zhang J, Shen Y, Henriksson L, Miyashita SI, Martínez-Carranza M, Dong M & Stenmark P (2017) Identification and characterization of a novel botulinum neurotoxin. *Nat Commun* **8**, 14130.
- Zhang S, Lebreton F, Mansfield MJ, Miyashita SI, Zhang J, Schwartzman JA, Tao L, Masuyer G, Martínez-Carranza M, Stenmark P et al. (2018) Identification of a botulinum neurotoxin-like toxin in a commensal strain of *Enterococcus faecium*. *Cell Host Microbe* **23**, 169–176.e6.
- Contreras E, Masuyer G, Qureshi N, Chawla S, Dhillon HS, Lee HL, Chen J, Stenmark P & Gill SS (2019) A neurotoxin that specifically targets *Anopheles* mosquitoes. *Nat Commun* **10**, 1–10.
- Mansfield MJ, Adams JB & Doxey AC (2015) Botulinum neurotoxin homologs in non-*Clostridium* species. *FEBS Lett* **589**, 342–348.
- Doxey AC, Mansfield MJ & Montecucco C (2018) Discovery of novel bacterial toxins by genomics and computational biology. *Toxicon* **147**, 2–12.
- Montal M (2010) Botulinum neurotoxin: a marvel of protein design. *Annu Rev Biochem* **79**, 591–617.
- Atlas RM (2002) Bioterrorism: from threat to reality. *Annu Rev Microbiol* **56**, 167–185.
- Jabbari B (2016) History of botulinum toxin treatment in movement disorders. *Tremor Other Hyperkinet Mov (N Y)* **6**, 394.
- Rapp DE, Lucioni A, Katz EE, O'Connor RC, Gerber GS & Bales GT (2004) Use of botulinum – a toxin for the treatment of refractory overactive bladder symptoms: an initial experience. *Urology* **63**, 1071–1075.
- Cheshire WP, Abashian SW & Mann JD (1994) Botulinum toxin in the treatment of myofascial pain syndrome. *Pain* **59**, 65–69.
- Silberstein S, Mathew N, Saper J & Jenkins S (2000) Botulinum toxin type a as a migraine preventive treatment. For the BOTOX migraine clinical research group. *Headache* **40**, 445–450.
- Bonventre PF (1979) Absorption of botulinal toxin from the gastrointestinal tract. *Rev Infect Dis* **1**, 663–667.
- Binz T & Rummel A (2009) Cell entry strategy of clostridial neurotoxins. *J Neurochem* **109**, 1584–1595.
- Fischer A & Montal M (2007) Crucial role of the disulfide bridge between botulinum neurotoxin light and heavy chains in protease translocation across membranes. *J Biol Chem* **282**, 29604–29611.
- Pantano S & Montecucco C (2014) The blockade of the neurotransmitter release apparatus by botulinum neurotoxins. *Cell Mol Life Sci* **71**, 793–811.
- Collins MD & East AK (1998) Phylogeny and taxonomy of the food-borne pathogen *Clostridium botulinum* and its neurotoxins. *J Appl Microbiol* **84**, 5–17.
- Rossetto O, Pirazzini M & Cesare M (2014) Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat Rev Microbiol* **12**, 535–549.
- Matsui T, Gu S, Lam KH, Carter LG, Rummel A, Mathews II & Jin R (2014) Structural basis of the pH-dependent assembly of a botulinum neurotoxin complex. *J Mol Biol* **426**, 3773–3782.
- Lee K, Lam KW, Kruegel AM, Mahrhold S, Perry K, Cheng LW, Rummel A & Jin R (2015) Inhibiting oral intoxication of botulinum neurotoxin A complex by carbohydrate receptor mimics. *Toxicon* **107**, 43–49.
- Lam KH & Jin R (2015) Architecture of the botulinum neurotoxin complex: architecture_LAB molecular

- machine for protection and delivery. *Curr Opin Struct Biol* **31**, 89–95.
- 28 Gustafsson R, Bertsson RPA, Martínez-Carranza M, El Tekle G, Odegrip R, Johnson EA & Stenmark P (2017) Crystal structures of OrfX2 and P47 from a botulinum neurotoxin OrfX-type gene cluster. *FEBS Lett* **591**, 3781–3792.
 - 29 Lam KH, Qi R, Liu S, Kroh A, Yao G, Perry K, Rummel A & Jin R (2018) The hypothetical protein P47 of *Clostridium botulinum* E1 strain Beluga has a structural topology similar to bactericidal/permeability-increasing protein. *Toxicon* **147**, 19–26.
 - 30 Košenina S & Stenmark P (2022) Crystal structure of the OrfX1 – OrfX3 complex from the PMP1 neurotoxin gene cluster. *FEBS Lett* **597**, 515–523.
 - 31 Gao L, Lam KH, Liu S, Przykopanski A, Lübke J, Qi R, Krüger M, Nowakowska MB, Selby K, Douillard FP *et al.* (2023) Crystal structures of OrfX1, OrfX2 and the OrfX1–OrfX3 complex from the orfX gene cluster of botulinum neurotoxin E1. *FEBS Lett* **597**, 524–537.
 - 32 Eisele KH, Fink K, Vey M & Taylor HV (2011) Studies on the dissociation of botulinum neurotoxin type A complexes. *Toxicon* **57**, 555–565.
 - 33 Ohishi I, Sugii S & Sakaguchi G (1977) Oral toxicities of *Clostridium botulinum* toxins in response to molecular size. *Infect Immun* **16**, 107–109.
 - 34 Gu S, Rumpel S, Zhou J, Strotmeier J, Bigalke H, Perry K, Shoemaker CB, Rummel A & Jin R (2012) Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. *Science* **335**, 977–981.
 - 35 Eswaramoorthy S, Sun J, Li H, Singh BR & Swaminathan S (2015) Molecular assembly of clostridium botulinum progenitor M complex of type E. *Sci Rep* **5**, 1–9.
 - 36 Martínez-Carranza M, Škerlová J, Lee P-G, Zhang J, Burgin D, Elliott M, Philippe J, Donald S, Hornby F, Henriksson L *et al.* (2023) Structure and activity of botulinum neurotoxin X. *bioRxiv*. doi: [10.1101/2023.01.11.523524](https://doi.org/10.1101/2023.01.11.523524)
 - 37 Zornetta I, Azarnia Tehran D, Arrigoni G, Anniballi F, Bano L, Leka O, Zanotti G, Binz T & Montecucco C (2016) The first non Clostridial botulinum-like toxin cleaves VAMP within the juxtamembrane domain. *Sci Rep* **6**, 1–7.
 - 38 Schiavo G, Rossetto O, Santucci A, DasGupta BR & Montecucco C (1992) Botulinum neurotoxins are zinc proteins. *J Biol Chem* **267**, 23479–23483.
 - 39 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.
 - 40 Košenina S, Masuyer G, Zhang S, Dong M & Stenmark P (2019) Crystal structure of the catalytic domain of the *Weissella oryzae* botulinum-like toxin. *FEBS Lett* **593**, 1403–1410.
 - 41 Hasegawa H & Holm L (2009) Advances and pitfalls of protein structural alignment. *Curr Opin Struct Biol* **19**, 341–348.
 - 42 Luo Y, Frey EA, Pfuetzner RA, Creaght AL, Knoechel DG, Haynes CA, Finlay BB & Strynadka NCJ (2000) Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* **405**, 1073–1077.
 - 43 Hamburger ZA, Brown MS, Isberg RR & Bjorkman PJ (1999) Crystal structure of invasins: a bacterial integrin-binding protein. *Science* **286**, 291–295.
 - 44 Raman R, Rajanikanth V, Palaniappan RUM, Lin YP, He H, McDonough SP, Sharma Y & Chang YF (2010) Big domains are novel Ca²⁺-binding modules: evidences from big domains of *Leptospira* immunoglobulin-like (Lig) proteins. *PLoS One* **5**, e14377.
 - 45 Yang YH, Jiang YL, Zhang J, Wang L, Bai XH, Zhang SJ, Ren YM, Li N, Zhang YH, Zhang Z *et al.* (2014) Structural insights into SraP-mediated *Staphylococcus aureus* adhesion to host cells. *PLoS Pathog* **10**, 4–15.
 - 46 Whelan F, Lafita A, Griffiths SC, Cooper REM, Whittingham JL, Turkenburg JP, Manfield IW, St John AN, Paci E, Bateman A *et al.* (2019) Defining the remarkable structural malleability of a bacterial surface protein rib domain implicated in infection. *Proc Natl Acad Sci USA* **116**, 26540–26548.
 - 47 Hill KK & Smith TJ (2013) Genetic diversity within *Clostridium botulinum* serotypes, botulinum neurotoxin gene clusters and toxin subtypes. *Curr Top Microbiol Immunol* **364**, 1–20.
 - 48 Elad N, Bellapadrona G, Houben L, Sagi I & Elbaum M (2017) Detection of isolated protein-bound metal ions by single-particle cryo-STEM. *Proc Natl Acad Sci USA* **114**, 11139–11144.
 - 49 Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A & Finn RD (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* **46**, D624–D632.
 - 50 Segelke B, Knapp M, Kadkhodayan S, Balhorn R & Rupp B (2004) 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.
 - 51 Koriazova LK & Montal M (2003) Translocation of botulinum neurotoxin light chain protease through the heavy chain channel. *Nat Struct Biol* **10**, 13–18.
 - 52 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **148**, 148–162.

- 53 Young G, Hundt N, Cole D, Fineberg A, Andrecka J, Tyler A, Olerinyova A, Ansari A, Marklund EG, Collier MP *et al.* (2018) Quantitative mass imaging of single molecules HHS public access. *Science* **360**, 423–427.
- 54 Martin TG, Boland A, Fitzpatrick AWP & Scheres SHW (2016) Graphene oxide grid preparation. *Figshare*. doi: [10.6084/m9.figshare.3178669.v1](https://doi.org/10.6084/m9.figshare.3178669.v1)
- 55 Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296.
- 56 Punjani A, Brubaker MA & Fleet DJ (2017) Building proteins in a day: efficient 3D molecular structure estimation with electron cryomicroscopy. *IEEE Trans Pattern Anal Mach Intell* **39**, 706–718.
- 57 Scheres SHW & Chen S (2016) Prevention of overfitting in cryo-EM structure determination. *Nat Methods* **9**, 853–854.
- 58 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* **66**, 486–501.
- 59 Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Žídek A, Nelson AWR, Bridgland A *et al.* (2020) Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710.
- 60 Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung L-W, Jain S, McCoy AJ *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D* **75**, 861–877.
- 61 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12–21.
- 62 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.