The Structure and Classification of Botulinum Toxins

Min Dong and Pål Stenmark

Contents

Abstract

Botulinum neurotoxins (BoNTs) are a family of bacterial protein toxins produced by various *Clostridium* species. They are traditionally classified into seven major serotypes (BoNT/A-G). Recent progress in sequencing microbial genomes has led to an ever-growing number of subtypes, chimeric toxins, BoNT-like toxins, and remotely related BoNT homologs, constituting an expanding BoNT superfamily. Recent structural studies of BoNTs, BoNT progenitor toxin complexes,

M. Dong (✉)
Department of Urology, Boston Children's Hospital, Boston, MA, USA

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

Department of Surgery, Harvard Medical School, Boston, MA, USA
e-mail: min.dong@childrens.harvard.edu

P. Stenmark (✉)
Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

Department of Experimental Medical Science, Lund University, Lund, Sweden
e-mail: pal.stenmark@med.lu.se

S. M. Whitcup, M. Hallett (eds.), *Botulinum Toxin Therapy*,
Handbook of Experimental Pharmacology 263, https://doi.org/10.1007/164_2019_342

tetanus neurotoxin (TeNT), toxin-receptor complexes, and toxin-substrate complexes have provided mechanistic understandings of toxin functions and the molecular basis for their variations. The growing BoNT superfamily of toxins present a natural repertoire that can be explored to develop novel therapeutic toxins, and the structural understanding of their variations provides a knowledge basis for engineering toxins to improve therapeutic efficacy and expand their clinical applications.

Keywords
Bacterial toxins · BoNT · BoNT-like toxins · Botox · Botulinum neurotoxin · Botulinum toxin · Tetanus neurotoxin · X-ray crystal structure

Botulinum neurotoxins (BoNTs) are a family of bacterial protein toxins that cause the human and animal disease botulism (Fig. 1) (Dong et al. 2019; Pirazzini et al. 2017; Montal 2010; Rossetto et al. 2014). Together with the related tetanus neurotoxin (TeNT), they are known as clostridial neurotoxins. These toxins are composed of two chains and three functional domains (Fig. 2a): the light chain (LC, ~50 kDa), which is a zinc-dependent metalloprotease that cleaves the target proteins in neurons, and the heavy chain (HC), which can be further divided into the N-terminal membrane translocation domain (H_N, ~50 kDa) and the C-terminal receptor-binding domain (H_C, ~50 kDa). These toxins are initially produced as a single polypeptide known as the pro-toxin. The linker region between the LC and HC needs to be

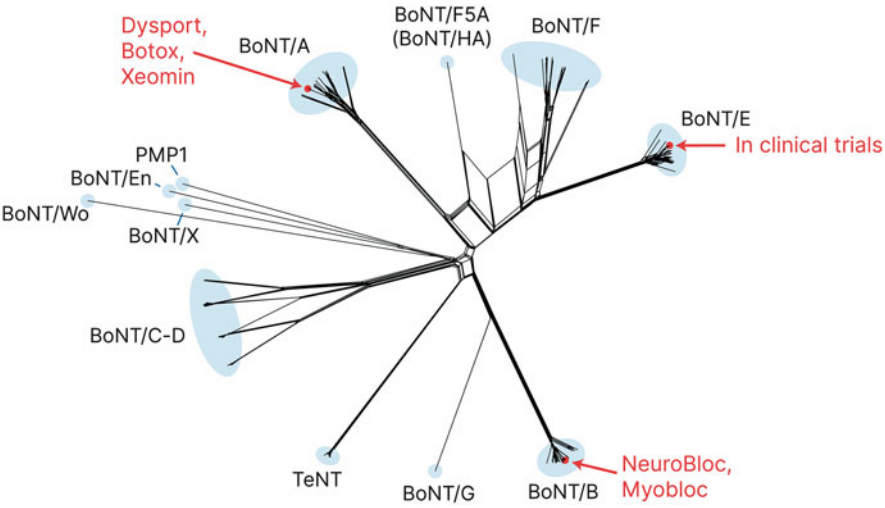

Fig. 1 A phylogenetic split network of BoNT and BoNT-like toxins. The diagram illustrates the potential evolutionary relationships based on comparing protein sequences of all known BoNT subtypes, chimeric toxins, BoNT-like toxins, and BoNT/Wo. BoNT/A1 (with such brand names as Dysport, Botox, and Xeomin from different companies) and BoNT/B1 (with the brand names NeuroBloc or Myobloc) have been approved by the FDA for medical and cosmetic uses, while BoNT/E1 is under clinical trials. These three toxins are marked in red

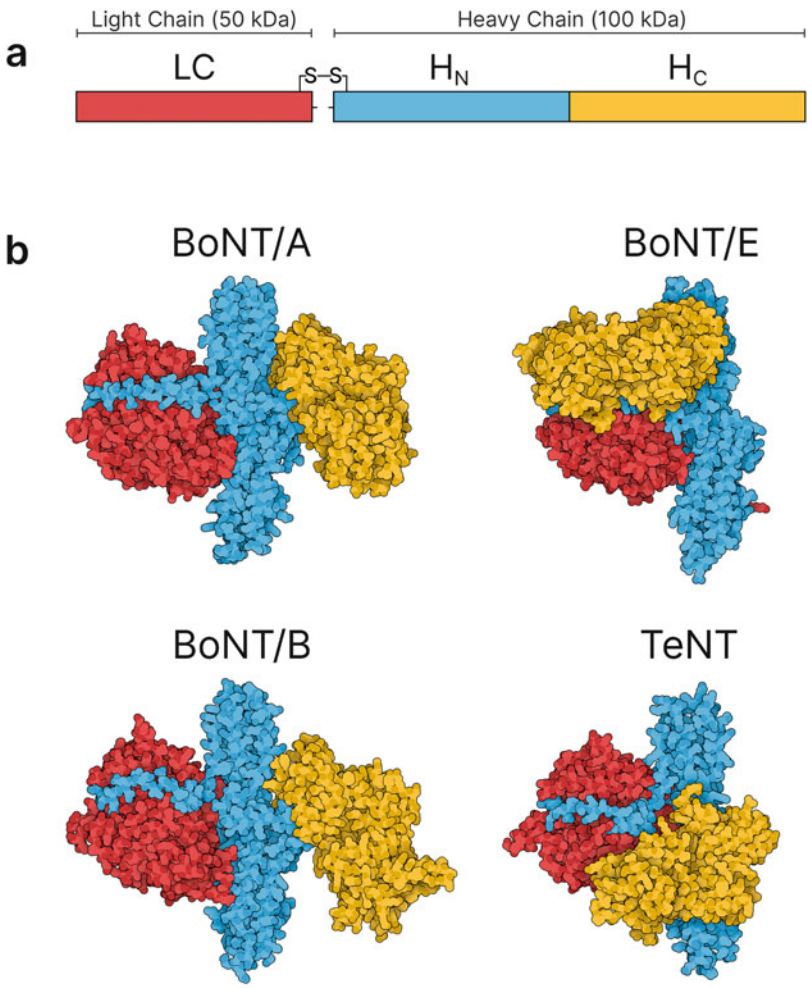

Fig. 2 The three-domain architecture of BoNTs. (**a**) A schematic drawing of the di-chain and three-domain architecture of BoNTs. The light chain (LC) is colored red, the translocation domain (H_N) blue, and the receptor-binding domain (H_C) yellow. The two chains are connected via a disulfide bond. (**b**) The crystal structures of full-length BoNT/A, BoNT/B, BoNT/E, and TeNT. LC is colored red, the H_N blue, and the H_C yellow. The protein structures are shown in a space filling representation. BoNT/A and BoNT/B display a linear arrangement for the three domains, with the LC and the H_C on each side of the H_N. BoNT/E and TeNT have their LC and H_C located on the same side of the H_N. PDB: 3BTA, 3FFZ, 1S0D, and 5N0B

cleaved by bacterial or host proteases, which converts the inactive pro-toxin to a di-chain active form. The LC and HC remain connected via a single disulfide bond. Once the H_C recognizes the receptors on nerve terminals, the toxin enters neurons via receptor-mediated endocytosis. The H_N then mediates translocation of the LC across endosomal membranes into the cytosol. The LC cleaves neuronal substrate

proteins, including Syntaxin 1, SNAP-25, and VAMP1, 2, 3, which are required for neurotransmitter release (Jahn and Scheller 2006; Sudhof and Rothman 2009), thus blocking neurotransmission.

1 BoNT Serotypes, Chimeric Toxins, and Subtypes

The classifications of BoNTs are traditionally based on their antigenicity and are known as serotypes, meaning that anti-sera generated against one toxin cannot recognize and neutralize another toxin (Fig. 1). The first BoNT was identified in 1897 (van Ermengem 1897). A serologically distinct BoNT was recognized in 1904, and hence serotypes A and B were designated to differentiate these two toxins (Burke 1919a, b; Leuchs 1910). This was followed by recognition of serotypes C in 1922, D in 1928, E in 1936, and F in 1960. The latest serotype, BoNT/G, was isolated from soils in Argentina and reported in 1970 (Gimenez and Ciccarelli 1970). This traditional serological classification has provided a way to distinguish diverse BoNT members and played a key role in developing vaccines and neutralizing antibodies against BoNTs.

The DNA and protein sequences for the prototypes of the seven BoNTs, as well as related TeNT, were resolved by the early 1990s, revealing ~37–70% variation in protein sequences among different serotypes. A phylogenic tree can be constructed based on protein sequences (Fig. 1). The two pairs BoNT/B versus BoNT/G and BoNT/E versus BoNT/F show the highest sequence identity (57% and 63%, respectively) among BoNTs. BoNT/A and BoNT/B have been approved by the FDA for use in humans (Schantz and Johnson 1992; Johnson 1999), and BoNT/C and BoNT/F have been investigated for potential medical use (Eleopra et al. 1997, 2006). BoNT/E is currently under clinical trials, which has a faster onset and shorter paralysis duration than BoNT/A (Fig. 1) (Eleopra et al. 1998).

The limitations of serological classification were recognized as early as the 1920s, when inconsistent neutralization efficacy was observed while serotyping "type C" toxins from different bacterial strains. Sequence information later revealed that this was due to the existence of naturally occurring chimeric toxins. For instance, there is a chimeric BoNT/CD with its LC-H_N derived from BoNT/C and its H_C from BoNT/D (Moriishi et al. 1996b). Anti-sera raised against BoNT/CD can neutralize BoNT/C, but anti-sera against BoNT/C are not effective in neutralizing BoNT/CD (Pfenninger 1924). There is also a chimeric toxin BoNT/DC, which is composed of a LC-H_N that is 98% identical with the corresponding region in BoNT/D and a H_C that shares 77% identity with BoNT/C-H_C (Moriishi et al. 1996a). Serologically, because this toxin can be recognized and neutralized by anti-sera against BoNT/D, it has been considered a BoNT/D (the bacteria strain is known as the strain D-5995, D-SA, or D-4947) and has been supplied as BoNT/D by a commercial vendor (Metabiologics Inc. Madison, WI, USA); however, its H_C is clearly distinct from BoNT/D-H_C.

Sequencing toxin genes has also revealed a growing number of subtype toxins with significant protein sequence variations from known toxin sequences (Hill et al. 2007; Peck et al. 2017). These variations could significantly reduce the efficacy of the standard anti-sera. For instance, there are at least eight BoNT/A subtypes (A1-A8).

The prototype is referred to as BoNT/A1, which is the only FDA-approved BoNT/A type for use in humans (Peck et al. 2017). Among BoNT/A subtypes, BoNT/A3 contains the greatest sequence variations from BoNT/A1 (15.4%). BoNT/F subtypes contain the most variation among all seven serotypes from their prototype BoNT/F1: as high as 30.2% for BoNT/F5 and 26.3% for BoNT/F7. The difference between BoNT/F5 and BoNT/F7 is 36.2%, the highest variation among all subtypes. These sequence differences help explain the significant variations in neutralization efficacy observed when BoNT/A and BoNT/F were serotyped from different bacterial strains.

The limitations of the traditional serotyping approach are further illustrated by the recent controversial naming of BoNT/H. This BoNT was identified in 2013 from a bacterial strain isolated from an infant botulism case (Dover et al. 2014; Barash and Arnon 2014). This toxin was not neutralized by anti-sera against other BoNTs following the established serotyping protocol. Thus, it was proposed as a new serotype. However, sequencing the toxin gene revealed that its LC shares ~80% identity with the LC of BoNT/F5, while its H_C shares ~84% identity with BoNT/A1-H_C (Maslanka et al. 2016). As the LC of BoNT/F5 (F5-LC) has a relatively high sequence variation from BoNT/F1-LC (only ~47% sequence identity) (Kalb et al. 2012), it is not surprising that anti-sera against BoNT/F1 failed to neutralize this toxin. Later studies showed that this toxin can be neutralized by antibodies against BoNT/A1, albeit a higher antibody titer is required than the standard serotyping protocol (Maslanka et al. 2016). Thus, this toxin is also considered a chimeric toxin BoNT/FA or more precisely BoNT/F5-A. To make the matter even more complicated, the H_N of this toxin does not appear to be close to either BoNT/F5-H_N or BoNT/A1-H_N. Thus, it has also been speculated that the LC-H_N of this toxin might be derived from a yet-to-be-identified BoNT.

Most subtypes likely cleave the same substrate protein at the same site and utilize the same receptors as their prototypes. However, exceptions with altered functional specificity have been reported. For instance, BoNT/F5 cleaves a site on VAMP1/2/3 that is distinct from all other BoNTs, indicating that its sequence variation is large enough to shift its cleavage site (Kalb et al. 2012). Similarly, sequence variations between the H_Cs of BoNT/DC and BoNT/C result in BoNT/DC utilizing a protein receptor that is not the receptor for BoNT/C (Peng et al. 2012).

Interestingly, even relatively low levels of sequence variations among subtypes, which may not alter the cleavage site on their substrates or switch receptors, could have measurable impacts on in vivo efficacy and pharmacological properties. For instance, BoNT/A2, which has 90% sequence identity with BoNT/A1, showed faster onset than BoNT/A1 on cultured neurons and in animal models (Torii et al. 2011; Pier et al. 2011; Whitemarsh et al. 2013; Pellett et al. 2015). It has been suggested that this faster onset time is because BoNT/A2 has an overall faster translocation process across the membrane than BoNT/A1 (Pier et al. 2011; Whitemarsh et al. 2013). As faster onset is clinically beneficial, BoNT/A2 has been explored for clinical uses. Additionally, it was recently reported that BoNT/A6, which has 4.3% sequence variation from BoNT/A1, also showed faster entry into neurons in culture (Moritz et al. 2018). Another example is that sequence variations in the LC of BoNT/A3 from BoNT/A1-LC result in a shorter duration of paralysis induced by BoNT/A3 compared with that induced by BoNT/A1 (Pellett et al. 2018). Thus,

sequence variations among the growing number of subtypes provide valuable resources for developing a new generation of therapeutic toxins and for optimizing toxin sequences to improve their pharmacological properties.

While traditional serotyping has value as a framework for categorizing toxins and identifying their distinct antigenic properties, describing toxins with seven serotypes is clearly insufficient to capture the growing diversity among BoNTs. Given the current ease of determining the exact toxin sequences, it will be important to note the specific subtype information when discussing a particular BoNT. To manage the naming of the growing number of subtypes, a guideline was proposed in 2017, which officialized the previous proposed threshold for a toxin sequence to be considered a new subtype as >2.6% variations at protein sequence levels from any known BoNT sequences (Peck et al. 2017). A few previous defined subtypes with <2.6% variations are grandfathered in, such as BoNT/B2, B3, and B6, which encompass only 1.5–1.9% variations, and BoNT/E2, E3, and E7, which differ from BoNT/E1 by only 1.0 to 2.1%.

To avoid duplication in numbering new subtypes, an email address has been set up at the Centers for Disease Control and Prevention (CDC, bontsubtype@cdc.gov) to receive requests for designation of new subtypes (Peck et al. 2017). There are also other efforts to develop a unified reporting system and database. One such a database, BoNTbase (https://bontbase.org) developed by Dr. Jonathan Davies in the laboratory of Prof. Stenmark, contains all reported BoNT subtypes as well as BoNT-like sequences along with associated research publications.

2 TeNT

TeNT, produced by *Clostridium tetani*, shares the same overall domain structures and mode of actions with other BoNTs; in fact, sequence alignment places TeNT in the middle of the family (Fig. 1). However, TeNT is not classified as a BoNT because it causes tetanus, a disease that is clinically distinct from botulism. TeNT and BoNTs both target and enter peripheral motor neurons. Unlike BoNTs, which block neurotransmitter release from motor neurons, thus causing muscle relaxation (flaccid paralysis), TeNT undergoes retrograde transport and transcytosis: it moves along the axons of motor neurons into the cell body in the spinal cord and is then released from motor neurons and enters the connecting inhibitor neurons where TeNT blocks neurotransmitter release (Lalli et al. 2003; Surana et al. 2018). Loss of inhibitory input leads to overactivity of motor neurons, resulting in spastic paralysis. Interestingly, it has been suggested that at least a small fraction of some BoNTs such as BoNT/A1 may also undergo long-range transport and transcytosis along peripheral neuronal axons into connecting neurons (Restani et al. 2011, 2012; Antonucci et al. 2008; Bomba-Warczak et al. 2016). The molecular basis for the different traffic pathways utilized by TeNT versus BoNTs remains unknown.

3 BoNT-Like Toxins

Rapid progress in sequencing microbial genomes in recent years has fundamentally changed how novel toxins are discovered. In 2015, a new toxin gene was recognized through bioinformatic analysis of the genome of a *Clostridium botulinum* strain. It encodes a protein containing the same three functional domains and key motifs found in BoNTs, with ~28–30% of sequence identity compared with the seven BoNTs (Fig. 1) (Zhang et al. 2017). This toxin was named BoNT/X because of varying opinions on what naming convention to utilize. Subsequent functional characterization confirmed that BoNT/X is capable of cleaving VAMP1, 2, 3 at a novel cleavage site distinct from all known cleavage sites for BoNTs. Interestingly, BoNT/X is a unique toxin that can also cleave VAMP family members VAMP4, VAMP5, and Ykt6, although the physiological consequences of these noncanonical cleavage events remain to be determined. Because BoNT/X was not recognized by antisera raised toward any of the seven BoNTs, it could be considered a novel serotype. However, unlike the seven classic BoNTs, BoNT/X showed only a low level of toxicity in mice. These findings suggest that BoNT/X may not naturally target mice and other vertebrates. The host species targeted by BoNT/X remains to be established.

In 2017, sequencing the genome of an *Enterococcus faecium* strain collected from cow feces revealed another BoNT-like toxin, designated BoNT/En (Zhang et al. 2018; Brunt et al. 2018). It too shares the same three-domain arrangement and key motifs found in BoNTs, with 24–27% protein sequence identity with the seven classic BoNTs. It is most closely related to BoNT/X, sharing 37% sequence identity. Functional validation demonstrated that BoNT/En is capable of cleaving VAMP1, 2, 3 and SNAP-25 in cultured neurons. Because BoNT/En is not recognized by any anti-sera against the seven BoNTs and BoNT/X, it too can be considered a new serotype, but BoNT/En showed no toxicity in mice. This is largely due to the lack of appropriate receptors in mice, as a chimeric toxin containing the LC-H_N of BoNT/En fused with the H_C of BoNT/A showed high neuronal toxicity and induced muscle paralysis. Thus, BoNT/En does not appear to target mouse motor neurons, and the host species naturally targeted by BoNT/En remains unknown.

In 2019, another BoNT-like toxin, PMP1 (paraclostridial mosquitocidal protein 1), was reported (Contreras et al. 2019). It was identified by screening and analyzing bacteria that can kill *anopheles* mosquito larvae. The toxin gene is located on a plasmid found in two strains with mosquitocidal activity: *Paraclostridium bifermentans malaysia* isolated from a mangrove swamp in Malaysia and *Paraclostridium bifermentans Paraiba* isolated in Brazil. PMP1 shares 36% sequence identity with BoNT/X and 34% identity with BoNT/En. These three toxins form a distinct branch in the BoNT superfamily (Fig. 1). Functional analysis showed that PMP1 is capable of cleaving mosquito Syntaxin 1 and has no toxicity in mice. PMP1 is the first known neurotoxin that naturally targets *anopheles* mosquito larvae. Its insecticidal toxicity and selectivity have the potential to be harnessed for developing novel mosquito control agents.

The crystal structure of the PMP1-H_C has been solved, revealing features distinct from the classic BoNT-H_Cs (Contreras et al. 2019). For instance, there are a dozen

aromatic residues exposed on the surface of PMP1-H_C, forming unique hydrophobic patches. Mutations at these hydrophobic patches reduced toxicity, suggesting that they may contribute to receptor binding. The receptors for PMP1 and other BoNT-like toxins remain unknown, which likely dictate the species targeted by each toxin, although other barriers may also exist. It is possible that BoNT/X and BoNT/En may also target insects or other invertebrates, and we expect that additional members of this group will continue to be discovered, which may form a group of neurotoxins targeting invertebrates. As the H_C can be switched between BoNT and BoNT-like toxins, chimeric toxins utilizing the LC-H_N part of BoNT-like toxins may provide an additional toolbox for designing new therapeutic toxins with unique properties.

4 BoNT Homologs

Bioinformatic analysis also revealed a growing number of sequences bearing various degrees of homology to BoNT, defined as BoNT homologs. The first was discovered in the genome of *Weissella oryzae*, a gram-positive anaerobe isolated from fermented rice in Japan (Mansfield et al. 2015). The protein was later named BoNT/Wo (Zornetta et al. 2016). The protein sequence of BoNT/Wo can be divided into LC, H_N, and H_C based on homology analysis with BoNTs, and it contains a few key conserved moieties found in BoNTs. However, BoNT/Wo is significantly different from BoNTs and BoNT-like toxins. First, the sequence identity of BoNT/Wo to other BoNTs and BoNT-like toxins is only 14–16% (Fig. 1). Second, there is no cysteine located at the linker region between BoNT/Wo-LC and HC. Third, while all BoNTs and BoNT-like toxins are located within a similar gene cluster (discussed in Sect. 6), BoNT/Wo is not in such a cluster. Thus, BoNT/Wo is only a distant homolog of BoNTs. It has been reported that BoNT/Wo-LC is capable of cleaving VAMP2 in vitro, but its physiological function remains to be established (Zornetta et al. 2016).

Three more BoNT homologs were recently reported in the genome of *Chryseobacterium piperi* (Mansfield et al. 2019). They showed low levels of sequence identity to BoNTs. For instance, one of these proteins, designated Cp1, shares ~17% identity with BoNT/A1. Cp1 can be divided into LC, H_N, and H_C based on homology analysis with BoNTs, and there are two cysteine residues located at the linker region between its LC and HC, suggesting an inter-chain disulfide bond. The function of these BoNT homologs remains to be fully characterized.

5 Three-Domain Architecture

The full-length crystal structures of BoNT/A, BoNT/B, BoNT/E, and TeNT have been determined, clearly demonstrating a three-domain architecture, composed of the LC, H_N, and H_C (Fig. 2b) (Lacy et al. 1998; Swaminathan and Eswaramoorthy 2000; Kumaran et al. 2009; Masuyer et al. 2017). The overall fold of each domain is largely conserved across these toxins, despite their rather low levels of amino acid sequence identity. BoNT/A and BoNT/B both showed a linear domain arrangement,

with the LC and H_C located on each side of the H_N, while the LC and H_C in BoNT/E are located on the same side of H_N and interact with each other. Thus, BoNT/E has an overall more-compact globular shape than BoNT/A and BoNT/B. The structure of TeNT has been investigated using multiple approaches: small-angle X-ray scattering analysis showed that TeNT is in a linear domain arrangement (open state) at neutral pH and changes into a compact globular form (closed state) under acidic pH (Masuyer et al. 2017). An intermediate semi-open state was also observed by low-resolution cryogenic electron microscopy (Cryo-EM). The high-resolution X-ray crystal structure of TeNT showed a closed state, with all three domains interacting with each other. Within TeNT, the LC-H_N forms a relatively stable core, while the H_C alters its position under different experimental conditions. The physiological relevance of the domain rearrangement in TeNT and whether similar flexibility exists in BoNTs remain to be determined.

5.1 Translocation Domain

The crystal structures of BoNT/A, B, E, and TeNT all reveal that the LC forms extensive contacts with the H_N. Particularly, the N-terminal region composed of ~50 residues of the H_N, designated the "belt" region, wraps around the LC (Fig. 2b). Because the belt region partially covers the active site of the LC, the LC reaches its full activity only once it is dissociated from the H_N after the disulfide bond connecting the LC and HC is broken (reduced).

The H_N is responsible for translocating the LC across the endosomal membrane. It is well established that the low pH within endosomes triggers conformational changes in BoNTs, leading to translocation of the LC, but the molecular mechanism for this translocation process remains to be elucidated. The H_N domain prominently features two long α-helices of ~105 Å. It remains unclear how these helices may alter their conformations upon encountering the low pH within endosomes.

A potential transmembrane region has been proposed based on analyzing hydrophobicity (e.g., residues 659–681 in BoNT/A), and a similar region (residues 593–686) in BoNT/A has been suggested to contribute to forming a channel in membranes (Montal et al. 1992; Lebeda and Olson 1995; Fischer et al. 2012). Bioinformatic analysis comparing the H_N of BoNTs and BoNT-like toxins revealed similarities between this region and the proposed transmembrane helix of diphtheria toxins, and also identified a conserved K/R. . .PxxG motif (Mansfield et al. 2019).

A recent study reported that the isolated H_N fragment of BoNT/A lacking the belt region can be produced as a soluble protein, and its crystal structure under acidic pH conditions has been resolved (Lam et al. 2018a). The structure highlights major conformational changes in the region from residues 620 to 667. This region is termed the BoNT-switch and contains disordered loops and short helices under neutral pH but switches to β-hairpins containing five β-strands under acidic pH. Interestingly, the sequence of this region, particularly the β2/β3 loop, is highly conserved across all BoNTs and bears an "aromatic-hydrophobic-glycine" tripeptide motif flanked by proline residues, which is similar to the lipid-binding peptide found in viral fusion proteins such as the internal fusion loop of Ebola virus glycoprotein 2. Thus, it was

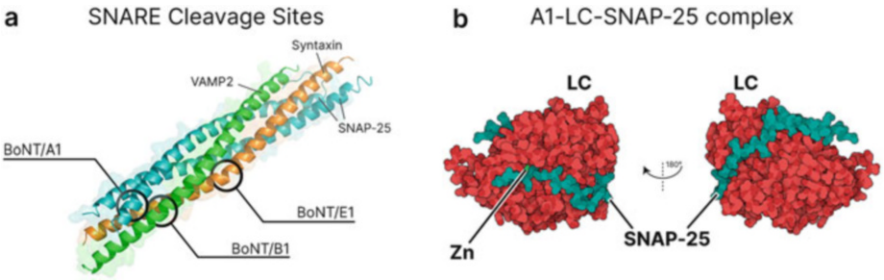

Fig. 3 BoNT-LCs cleave SNARE proteins. (**a**) The three SNARE proteins, Syntaxin 1, SNAP-25, and VAMP1/2/3 form a complex of four alpha helix bundles, which is essential for fusion of synaptic vesicle membranes to the plasma membrane of neurons. Cleavage of any one of these three SNARE proteins is sufficient to block vesicle exocytosis and neurotransmitter release. The cleavage sites for BoNT/A1, B1, and E1 are marked. PDB: 1N7S. (**b**) The crystal structure of a SNAP-25 fragment (colored dark green) in complex with BoNT/A1-LC (red), showing extensive interactions of SNAP-25 with the BoNT/A1-LC (right panel: rotated 180°). PDB: 1XTG

suggested that the BoNT-switch region is responsible for sensing the pH change and initiating membrane penetration via a mechanism similar to that used by viral fusion peptides. These results represent a major advance in our understanding of pH-induced conformational change in H_N. How the changes in the BoNT-switch region leads to further conformational changes in the rest of H_N and the eventual translocation process remains to be determined.

5.2 The Structure of the LC

By aligning the protein sequence of five BoNTs and TeNT, Giampietro Schiavo and Cesare Montecucco recognized a conserved HEXXH motif that is the key feature of metalloproteases, suggesting that BoNTs and TeNT act as proteases (Schiavo et al. 1992b). In their following seminal work published in 1992, they identified BoNT/B and TeNT as zinc-dependent proteases that cleave the synaptic vesicle protein VAMP2 (Schiavo et al. 1992a). Within a few years, it was fully established that BoNT/B, D, F, and G cleave homologous VAMP1, 2, and 3, while BoNT/A, C, and E cleave the peripheral membrane protein SNAP-25 (Fig. 3a). In addition, BoNT/C can also cleave the plasma membrane protein Syntaxin 1. BoNT/B and TeNT both share the same cleavage site on VAMP1, 2, and 3, while all other toxins have their own unique cleavage sites. These three toxin substrates are members of SNARE family proteins. They form the core complex that mediates fusion of synaptic vesicle membranes to plasma membranes, which is essential for releasing neurotransmitters (Jahn and Scheller 2006; Sudhof and Rothman 2009).

The crystal structures of all seven BoNT-LCs have been resolved, revealing an overall conserved globular fold (Jin et al. 2007; Arndt et al. 2005, 2006). The

catalytic site with the signature motif HEXXH is conserved in both composition and geometry across all BoNTs. BoNT-LCs are zinc-dependent proteases with remarkable substrate specificity. As their catalytic sites are similar, the specific recognition and cleavage of different substrates must involve regions outside the catalytic site. Indeed, co-crystal structure of an inactive form of BoNT/A-LC (A-LC, containing two-point mutations that abolish its protease activity) in complex with a fragment of its substrate SNAP-25 (residues 141–204) reveals that SNAP-25 wraps around A-LC, forming extensive interactions particularly via an α-exosite bound by the N-terminal region of the SNAP-25 fragment as well as a β-exosite bound by the C-terminal region of SNAP-25 (Fig. 3b) (Breidenbach and Brunger 2004). This requirement of "long stretch" of SNAP-25 to be properly docked into A-LC ensures specificity.

Each BoNT-LC likely possesses its own distinct exosites, whose location and composition determine the selection of the substrate SNARE proteins and the specific cleavage site. The co-crystal structure of BoNT/F-LC (F-LC) in complex with a VAMP2 fragment containing a point mutation that renders it resistant to BoNT/F is the only other toxin-substrate complex that has been resolved (Agarwal et al. 2009). This structure also demonstrated extensive interactions between F-LC and the VAMP2 fragment, with VAMP2 docked onto F-LC through at least three exosites distinct from the exosites in A-LC. The precise locations of exosites in other BoNT-LCs remain to be established. The crystal structure of BoNT/X-LC (X-LC) has been solved (Masuyer et al. 2018). Despite only ~30% sequence identity with other BoNTs, X-LC display a typical BoNT-LC fold with many conserved secondary structural features. The structure further demonstrates that X-LC is a bona fide member of the BoNT-LC family. The crystal structure of the BoNT/Wo-LC has recently been solved as well, showing that it shares a common core fold found in other BoNT-LCs but also revealing several distinct features including an unusually wide and open catalytic site (Kosenina et al. 2019).

Notably, A-LC has been shown to maintain its activity in cultured neurons for several months, which is the major reason for BoNT/A's ability to induce persistent paralysis that lasts several months in humans (Keller et al. 1999; Whitemarsh et al. 2014; Tsai et al. 2017; Pellett et al. 2015; Foran et al. 2003). This is a key pharmacological property that contributes to the success of BoNT/A as a therapeutic agent. Among the seven BoNTs, BoNT/E showed the shortest half-life, with only a few weeks in humans (Foran et al. 2003), a key feature differentiating it from BoNT/A. The molecular basis for the extremely long half-life of BoNT/A remains to be fully established. A-LC has been shown to bind the cytoskeleton protein septin complex, which may shield A-LC from degradation (Vagin et al. 2014). It has also been suggested that A-LC recruits deubiquitinase to reduce its ubiquitination (Tsai et al. 2017). The structural basis for those interactions remains to be solved.

5.3 Structural Basis for Receptor Recognition

The structure of the 50 kDa H_C showed two distinct sub-domains roughly equal in size. BoNTs have extreme specificity toward nerve terminals, which is achieved by recognizing at least two receptor components in a "double-receptor" model (Montecucco 1986). One is a family of glycolipids on cell membranes known as gangliosides, which are comprised of a lipid tail and a glycan headgroup containing various numbers of negatively charged sialic acids (Simpson and Rapport 1971; Hamark et al. 2017). Gangliosides are abundant at nerve terminals and serve as low-affinity receptors to enrich the toxin onto the cell surface. A ganglioside-binding site (GBS) has been identified and is conserved in BoNT/A, B, E, F, and G (Rummel et al. 2003, 2004; Fotinou et al. 2001). This GBS is at the C-terminal region of the H_C and contains the signature residues SXWY. The crystal structures of the H_C in complex with the headgroup of gangliosides have been solved for BoNT/A, B, and F, showing that GBS interacts with the GalNAc-Gal motif as well as sialic acids within gangliosides (Stenmark et al. 2008; Berntsson et al. 2013; Benson et al. 2011).

Besides gangliosides, many BoNTs also require specific neuronal protein receptors. Two sets of synaptic vesicle membrane proteins, synaptic vesicle glycoprotein 2 (SV2) and Synatotagmin I and II (Syt I/II), serve as receptors for multiple BoNTs. For instance, BoNT/A utilizes SV2 (including all three isoforms SV2A, SV2B, and SV2C) as its receptors, while BoNT/B, BoNT/DC, and BoNT/G all utilize homologous Syt I and Syt II as receptors (Dong et al. 2003, 2006; Nishiki et al. 1994; Mahrhold et al. 2006). These synaptic vesicle membrane proteins travel to cell surfaces only transiently, and this entry pathway is thus activity facilitated, as neuronal activity promotes synaptic vesicle exocytosis and endocytosis, leading to enhanced binding and entry of BoNTs.

Syt I and II are single-pass transmembrane proteins, with a short luminal domain (the region inside vesicles). BoNT/B, DC, and G all recognize the same short section of the luminal domain of Syt I/II, located next to the transmembrane domain. The co-crystal structure of the H_C of BoNT/B (B-H_C) in complex with the Syt II fragment containing the toxin-binding site revealed that the toxin-binding segment is induced to form an amphipathic α-helix and dock into a hydrophobic groove within B-H_C (Fig. 4a) (Jin et al. 2006; Chai et al. 2006). The complex is stabilized by highly specific side-chain-to-side-chain interactions. It has been reported that Syt II is expressed in most motor neurons in diaphragm neuromuscular junctions, while Syt I is detectable in only ~40% of motor neurons in mice (Pang et al. 2006). These findings suggest that Syt II is the dominant receptor at diaphragm motor nerve terminals. On the other hand, bladder tissues express mainly Syt I, but not Syt II (Elliott et al. 2019). Since bladder tissues are smooth muscles controlled by autonomic nerves, it is possible that Syt I is the dominant receptor in autonomic nerves.

It has long been clinically observed that higher doses of BoNT/B are required to achieve the same level of paralysis produced by BoNT/A in humans. This is because human Syt II happens to contain a residue change from phenylalanine, which is commonly found in most mammalian species, to leucine at a key position within the toxin-binding region. This single residue change drastically reduces the binding

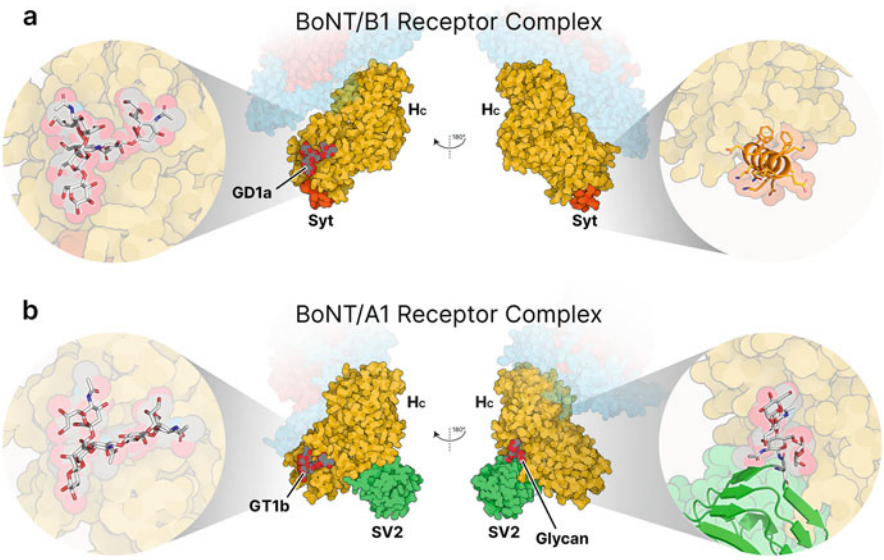

Fig. 4 BoNT/A1 and BoNT/B1 in complex with their receptors. (**a**) The crystal structure of BoNT/B1-H_C (yellow) in complex with its ganglioside coreceptor (colored according to chemical element) and the toxin-binding region of its protein receptor Syt II (orange). PDB: 4KBB. (**b**) BoNT/A1-H_C (yellow) in complex with its ganglioside coreceptor (colored according to chemical element) and the L4 of its protein receptor SV2C (green). PDB: 2VU9 and 5JLV

affinity of BoNT/B, DC, and G to human Syt II compared with Syt II from mice (Peng et al. 2012; Strotmeier et al. 2012).

This "defect" in human Syt II creates the need to engineer BoNT/B to improve its efficacy in humans. This has been achieved by structure-assisted mutagenesis approaches, which eventually identified that mutating residue 1191 (glutamic acid) to methionine, cysteine, glutamine, or valine, in combination with mutating serine 1199 to tryptophan or tyrosine, creates B-H_C mutants that can bind robustly to both mouse and human Syt II (Tao et al. 2017). Two of these toxin variants, E1191M/S1199Y and E1191Q/S1199W, have been produced recombinantly as full-length active toxins in *E. coli*. They were tested on a "humanized" transgenic mouse model, in which the mouse Syt II luminal domain has been replaced with the human version (Elliott et al. 2019). While natural BoNT/B showed a drastically reduced potency in this model, modified BoNT/B mutants showed the same level of potency in humanized mice and control mice. Therefore, these modified BoNT/B are expected to have better therapeutic efficacy in humans compared with natural BoNT/B.

While BoNT/B binds specifically to Syt I/II through side-chain-mediated interactions, BoNT/A recognition of SV2 involves not only the protein part of SV2 but also carbohydrate moieties at a glycosylation site of SV2 (Fig. 4b) (Yao et al. 2016). The three members of SV2 (SV2A, SV2B, and SV2C) are 12-trans-membrane-domain proteins, with both the N- and C-termini located in the cytosol.

The fourth luminal domain (SV2-L4) is the longest among all luminal domains, and BoNT/A recognizes its middle portion. The crystal structure of H_C of BoNT/A1 (A1-H_C) in complex with human SV2C-L4 expressed and purified in *E. coli* has been solved, revealing that SV2C-L4 folds into a right-handed, quadrilateral β-helix pattern, similar to pentapeptide-repeat proteins (Benoit et al. 2014). The overall architecture of SV2C-L4 is similar to the structure of amyloid fibrils, forming a stack of β-strands. BoNT/A1 recognizes the top of this structure by stacking two of its own β-strands in the middle of A1-H_C (Fig. 4b). This is an unusual type of toxin-receptor recognition, as most interactions are through backbone-backbone hydrogen bonds. The crystal structure of BoNT/A2 in complex with human SV2C L4 has also been solved (Benoit et al. 2017; Gustafsson et al. 2018). The complexes are highly similar to BoNT/A1 – SV2C L4.

SV2 is heavily glycosylated, and there are three conserved N-linked glycosylation sites within the L4, with one located right in the middle of the BoNT/A1-SV2 interface. These three sites are not glycosylated in SV2C-L4 purified from *E. coli*. Mutagenesis studies showed that abolishing the glycosylation site at the BoNT/A1-SV2 interface reduced the potency of BoNT/A1 on neurons, suggesting that glycosylation at this site contributes to toxin-SV2 interaction (Yao et al. 2016; Dong et al. 2008).

In 2016, the crystal structure of A1-H_C in complex with a glycosylated SV2C-L4 (expressed in eukaryotic cells) was elucidated (Yao et al. 2016). The overall structure of glycosylated SV2C-L4 and the protein-protein interaction interface between SV2C-L4 and A1-H_C are the same as shown in the structure containing un-glycosylated SV2C-L4. The new discovery is that the base of the N-linked glycan, including two GlcNAc, a mannose, and a fucose, is directly recognized by A1-H_C. These interactions with glycans significantly reduce the dissociation constant and enhance the overall binding affinity between A1-H_C and SV2C-L4 (Mahrhold et al. 2016; Yao et al. 2016). The location of this N-linked glycosylation site is highly conserved across all three SV2s as well as in different vertebrate species.

BoNT/E also utilizes SV2A and SV2B as its receptors, and the recognition is even more dependent on the presence of glycosylation at the same site of L4 than it is with BoNT/A (Dong et al. 2008; Mahrhold et al. 2013). Though structural basis for BoNT/E-SV2 interactions remains to be established, it has been shown that BoNT/E cannot recognize SV2C in neurons, which is a major difference from BoNT/A. SV2B and SV2C are both widely expressed at motor nerve terminals, while SV2A is expressed only in a subset of motor neurons controlling slow muscle fibers (Chakkalakal et al. 2010). It has also been reported that trigeminal sensory neurons are insensitive to BoNT/E, as they express only SV2C. This insensitivity can be overcome by an engineered chimeric BoNT/E-A toxin in which the H_C of BoNT/E is replaced with A1-H_C (Meng et al. 2009).

6 BoNT Gene Cluster and Progenitor Toxin Complex

BoNTs exist naturally in protein complexes when produced from *Clostridium* bacteria. BoNT/A1 was first purified in 1946 in crystalline form (Lamanna et al. 1946). It was later discovered that this purified form can be further separated into a toxic and a nontoxic component, with the latter capable of inducing aggregation of erythrocytes (hemagglutinin activity). Further studies revealed that BoNT/A1 can exist in three different complexes, known as 12S, 16S, and 19S based on ultracentrifugation analysis of their molecular weight. The 12S form contains a BoNT/A and a nontoxic molecule of similar size known as NTNHA (nontoxic non-hemagglutinin protein, also known as NTNH). The 16S contains the 12S plus a nontoxic protein complex with hemagglutinin activity (known as HA complex). The 19S appears to be a dimer of 16S, and this is the crystalline form first put into clinical use. Current generation of therapeutic BoNT/A and BoNT/B are both in complex form, containing NTNHA and HA proteins. The exception is that the product from Merz is the BoNT/A molecule alone, isolated through additional purification from the original complex form.

The genes encoding BoNTs are located within two kinds of gene clusters, both containing a gene encoding NTNHA next to the toxin gene (Fig. 5a). One cluster, designated as the HA cluster, encodes three proteins (HA17, HA33, and HA70) that form a complex with hemagglutinin activity. BoNT/A1, B, C, D, and G all contain the HA cluster and produce both 12S and 16S complexes. The other cluster is known as the OrfX cluster, encoding four proteins (OrfX1, OrfX2, OrfX3, and P47) with unknown functions. BoNT/A2, E, and F contain the OrfX cluster, and only 12S complexes were purified for these toxins.

The reason why BoNTs are produced in complexes can be understood from their mode of action: they are oral toxins. They naturally enter human and animal bodies via ingestion of contaminated food, and the toxin must be able to pass through the harsh environment of the gastrointestinal tract and be absorbed into the circulatory system as an intact molecule. Because the isolated toxin can easily be degraded by proteases, its oral toxicity is low, whereas the BoNT-NTNHA complex is extremely resistant to proteases. Thus, BoNTs are always combined with their corresponding NTNHA into a minimally functional progenitor toxin complex (M-PTC, Fig. 5b), whose formation is pH dependent. For instance, BoNT/A1 and its NTNHA form a complex under acidic pH conditions, and the toxin dissociates from NTNHA at pH 7.3. This provides a mechanism to release the toxin once the complex enters the circulatory system. In clinical applications, BoNT/A1 is expected to dissociate from the complex once it is injected into the tissue; the toxin itself is responsible for the therapeutic effect.

The structural basis for BoNT-NTNHA complexes has been established by recent work revealing the crystal structures of BoNT/A1 with its NTNHA (NTNHA-A) and BoNT/E with its NTNHA (NTNHA-E) (Gu et al. 2012; Eswaramoorthy et al. 2015). The two complex structures share a high degree of similarity and confirm that the toxins and the NTNHAs form interlocked "handshake" complexes, protecting a large portion of the solvent-accessible areas (Fig. 5b). The NTNHAs also contain three domains, termed nLC, nH_N, and nH_C, which are homologous to the LC, H_N,

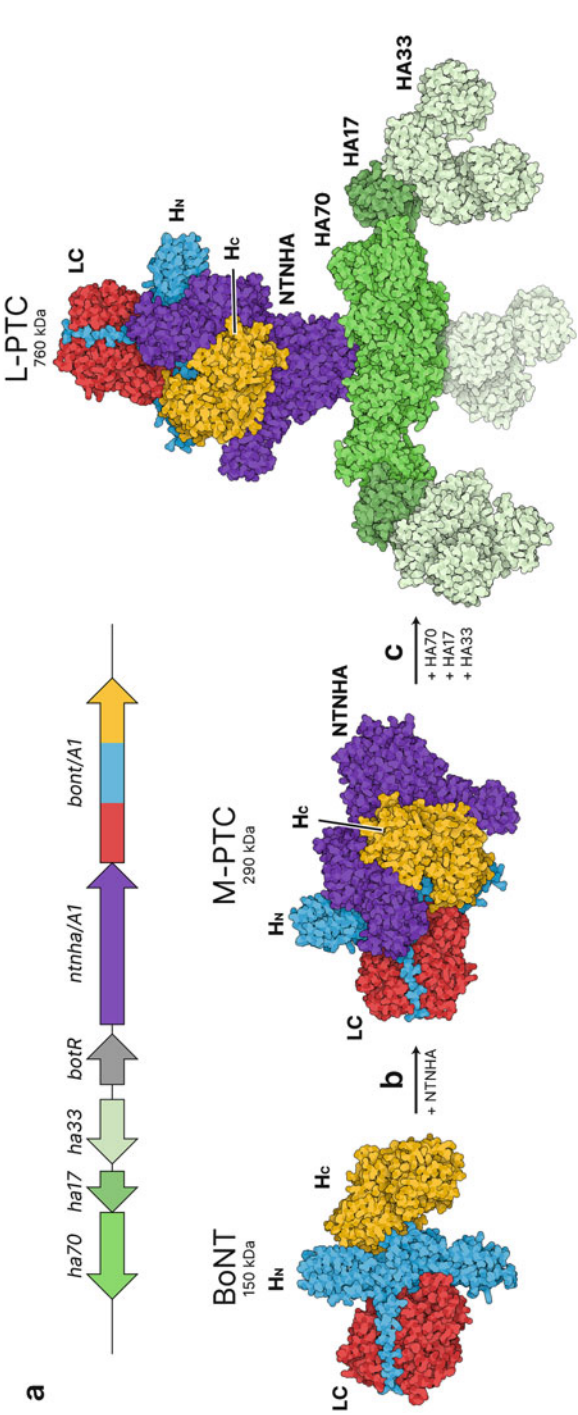

Fig. 5 The BoNT gene cluster and progenitor toxin complex. (**a**) Schematic drawing illustrating the HA gene cluster of BoNT/A1. (**b**) Structure of the M-PTC composed of a BoNT in complex with its NTNHA. The BoNT-LC is colored red, the H_N blue, H_C yellow, and NTNHA purple. PDB: 3BTA and 3V0A. (**c**) Structural model of the large progenitor toxin complex (L-PTC) composed of a BoNT, a NTNHA, three HA70, six HA33, and three HA17. HA proteins are colored in different shades of green. PDB: 4LO7 and 3V0A

and H_C in BoNTs, but NTNHAs do not contain the protease motif HEXXH in their nLCs, and there is no disulfide bond between nLC and nH_N. The H_C is at the center of the complex, forming interactions with all three domains of NTNHA, suggesting that H_C is the key region being protected. On the other hand, the LC and nLC are both pointing outward from the complex, and the LC of BoNT/A1 does not form any interactions with NTNHA-A. It has been proposed that key clusters of acidic residues at the interface of BoNT and NTNHA dictate the pH dependency of the complex formation. These acidic residues are not charged at low pH but become negatively charged (deprotonated) in neutral or basic conditions, creating a repulsive force that disassembles the complex.

The major difference between NTNHA-A and NTNHA-E is that NTNHA-A contains an extra loop within the nLC (termed nLoop), which is conserved in the NTNHAs from BoNT/A1, B, C, D, and G, but missing from all toxins associated with the OrfX cluster. This nLoop binds to the HA proteins and serves as a linker to dock the BoNT-NTNHA onto a complex of HA proteins, which comprises three HA70, three HA17, and six HA33 molecules (Fig. 5c). Together, the entire complex is defined as large progenitor toxin complex (L-PTC). The crystal structure of the HA complex of BoNT/B has been determined, showing a triskelion shape with three HA70 forming a central hub (Amatsu et al. 2013). HA17 serves as a linker, binding to one HA70 on one side and two HA33 on the other side. A similar structure has also been constructed for the HA complex of BoNT/A (Lee et al. 2013). By docking crystal structures into the shape of the entire toxin complex observed from negative-stain electron microscopy, a structural model of BoNT/A1 L-PTC has been constructed (Fig. 5c) (Lee et al. 2013).

The major function of the HA complex appears to be facilitating the absorption of the toxin complex in the intestine. First, the complex contains up to nine carbohydrate binding sites, one on each HA70 and HA33 (Lee et al. 2013). HA-carbohydrate interactions have been shown to contribute to the initial absorption of the toxin complex through microfold cells (M-cells) in the intestine (Matsumura et al. 2015). Second, the HA complexes of both BoNT/A and BoNT/B recognize cell surface adhesion molecule E-cadherin (Sugawara et al. 2010). Structural studies revealed that one HA complex binds to three E-cadherin molecules. The interaction blocks the trans-dimerization of two E-cadherin on neighboring cells, which is essential for forming cell-cell junctions, thus opening the tight junction in the intestinal epithelial barrier for the large toxin complex to pass through (Lee et al. 2014).

The function of the OrfX proteins and whether they form complexes with each other and/or with BoNT-NTNHA remain unknown. The crystal structures of OrfX2 and P47 show that both proteins contain a tubular lipid-binding (TULIP) fold, suggesting lipid-binding activity (Gustafsson et al. 2017; Lam et al. 2018b). All three BoNT-like toxins, BoNT/X, BoNT/En, and PMP1, are located in gene clusters homologous to the OrfX cluster. It has been shown that deletion of OrfX proteins reduces the toxicity of PMP1 to *anopheles* mosquito larvae, demonstrating that these OrfX proteins contribute to the oral toxicity of PMP1 (Contreras et al. 2019).

7 Concluding Remarks

Botulinum neurotoxins have captivated researchers for well over 100 years. The extreme potency of these toxins, and the impressive number of medical conditions they can be used to treat, continues to fascinate both scientists and clinicians. Our growing understanding of these toxins, with rapid advances in solving their structures and discovering new toxin variations and homologs, will lead to development of a new generation of therapeutic toxins with improved efficacy and expanded clinical applications.

Acknowledgments We thank Jonathan Davies for preparing all figures and for valuable discussions during the preparation of this chapter. This work was supported by the Swedish Research Council and the Swedish Cancer Society to P.S. and by grants from NIH (R01NS080833, R01AI132387, R01AI139087, and R21NS106159), Intelligence Advanced Research Projects Activity (IARPA, grant number W911NF-17-2-0089), and the Investigator in the Pathogenesis of Infectious Disease award from the Burroughs Wellcome Fund to M.D.

References

Agarwal R, Schmidt JJ, Stafford RG, Swaminathan S (2009) Mode of VAMP substrate recognition and inhibition of Clostridium botulinum neurotoxin F. Nat Struct Mol Biol 16(7):789–794

Amatsu S, Sugawara Y, Matsumura T, Kitadokoro K, Fujinaga Y (2013) Crystal structure of Clostridium botulinum whole hemagglutinin reveals a huge triskelion-shaped molecular complex. J Biol Chem 288(49):35617–35625

Antonucci F, Rossi C, Gianfranceschi L, Rossetto O, Caleo M (2008) Long-distance retrograde effects of botulinum neurotoxin A. J Neurosci 28(14):3689–3696

Arndt JW, Yu W, Bi F, Stevens RC (2005) Crystal structure of botulinum neurotoxin type G light chain: serotype divergence in substrate recognition. Biochemistry 44(28):9574–9580

Arndt JW, Chai Q, Christian T, Stevens RC (2006) Structure of botulinum neurotoxin type D light chain at 1.65 Å resolution: repercussions for VAMP-2 substrate specificity. Biochemistry 45 (10):3255–3262

Barash JR, Arnon SS (2014) A novel strain of Clostridium botulinum that produces type B and type H botulinum toxins. J Infect Dis 209(2):183–191

Benoit RM, Frey D, Hilbert M, Kevenaar JT, Wieser MM, Stirnimann CU, McMillan D, Ceska T, Lebon F, Jaussi R, Steinmetz MO, Schertler GF, Hoogenraad CC, Capitani G, Kammerer RA (2014) Structural basis for recognition of synaptic vesicle protein 2C by botulinum neurotoxin A. Nature 505(7481):108–111

Benoit RM, Scharer MA, Wieser MM, Li X, Frey D, 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

Benson MA, Fu Z, Kim JJ, Baldwin MR (2011) Unique ganglioside recognition strategies for clostridial neurotoxins. J Biol Chem 286(39):34015–34022

Berntsson RP, Peng L, Dong M, Stenmark P (2013) Structure of dual receptor binding to botulinum neurotoxin B. Nat Commun 4:2058

Bomba-Warczak E, Vevea JD, Brittain JM, Figueroa-Bernier A, Tepp WH, Johnson EA, Yeh FL, Chapman ER (2016) Interneuronal transfer and distal action of tetanus toxin and botulinum neurotoxins A and D in central neurons. Cell Rep 16(7):1974–1987

Breidenbach MA, Brunger AT (2004) Substrate recognition strategy for botulinum neurotoxin serotype A. Nature 432(7019):925–929

Brunt J, Carter AT, Stringer SC, Peck MW (2018) Identification of a novel botulinum neurotoxin gene cluster in Enterococcus. FEBS Lett 592(3):310–317

Burke GS (1919a) Notes on Bacillus botulinus. J Bacteriol 4(5):555–570

Burke GS (1919b) The occurrence of Bacillus botulinus in nature. J Bacteriol 4(5):541–553

Chai Q, Arndt JW, Dong M, Tepp WH, Johnson EA, Chapman ER, Stevens RC (2006) Structural basis of cell surface receptor recognition by botulinum neurotoxin B. Nature 444 (7122):1096–1100

Chakkalakal JV, Nishimune H, Ruas JL, Spiegelman BM, Sanes JR (2010) Retrograde influence of muscle fibers on their innervation revealed by a novel marker for slow motoneurons. Development 137(20):3489–3499

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):2869

Dong M, Richards DA, Goodnough MC, Tepp WH, Johnson EA, Chapman ER (2003) Synaptotagmins I and II mediate entry of botulinum neurotoxin B into cells. J Cell Biol 162 (7):1293–1303

Dong M, Yeh F, Tepp WH, Dean C, Johnson EA, Janz R, Chapman ER (2006) SV2 is the protein receptor for botulinum neurotoxin A. Science 312(5773):592–596

Dong M, Liu H, Tepp WH, Johnson EA, Janz R, Chapman ER (2008) Glycosylated SV2A and SV2B mediate the entry of botulinum neurotoxin E into neurons. Mol Biol Cell 19 (12):5226–5237

Dong M, Masuyer G, Stenmark P (2019) Botulinum and tetanus neurotoxins. Annu Rev Biochem 88:811–837

Dover N, Barash JR, Hill KK, Xie G, Arnon SS (2014) Molecular characterization of a novel botulinum neurotoxin type H gene. J Infect Dis 209(2):192–202

Eleopra R, Tugnoli V, Rossetto O, Montecucco C, De Grandis D (1997) Botulinum neurotoxin serotype C: a novel effective botulinum toxin therapy in human. Neurosci Lett 224(2):91–94

Eleopra R, Tugnoli V, Rossetto O, De Grandis D, Montecucco C (1998) Different time courses of recovery after poisoning with botulinum neurotoxin serotypes A and E in humans. Neurosci Lett 256(3):135–138

Eleopra R, Tugnoli V, Quatrale R, Rossetto O, Montecucco C, Dressler D (2006) Clinical use of non-A botulinum toxins: botulinum toxin type C and botulinum toxin type F. Neurotox Res 9 (2–3):127–131

Elliott M, Favre-Guilmard C, Liu SM, Maignel J, Masuyer G, Beard M, Boone C, Carre D, Kalinichev M, Lezmi S, Mir I, Nicoleau C, Palan S, Perier C, Raban E, Zhang S, Dong M, Stenmark P, Krupp J (2019) Engineered botulinum neurotoxin B with improved binding to human receptors has enhanced efficacy in preclinical models. Sci Adv 5(1):eaau7196

Eswaramoorthy S, Sun J, Li H, Singh BR, Swaminathan S (2015) Molecular Assembly of Clostridium botulinum progenitor M complex of type E. Sci Rep 5:17795

Fischer A, Sambashivan S, Brunger AT, Montal M (2012) Beltless translocation domain of botulinum neurotoxin A embodies a minimum ion-conductive channel. J Biol Chem 287 (3):1657–1661

Foran PG, Mohammed N, Lisk GO, Nagwaney S, Lawrence GW, Johnson E, Smith L, Aoki KR, Dolly JO (2003) Evaluation of the therapeutic usefulness of botulinum neurotoxin B, C1, E, and F compared with the long lasting type A. Basis for distinct durations of inhibition of exocytosis in central neurons. J Biol Chem 278(2):1363–1371

Fotinou C, Emsley P, Black I, Ando H, Ishida H, Kiso M, Sinha KA, Fairweather NF, Isaacs NW (2001) The crystal structure of tetanus toxin Hc fragment complexed with a synthetic GT1b analogue suggests cross-linking between ganglioside receptors and the toxin. J Biol Chem 276 (34):32274–32281

Gimenez DF, Ciccarelli AS (1970) Another type of Clostridium botulinum. Zentralbl Bakteriol Orig 215(2):221–224

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 (6071):977–981

Gustafsson R, Berntsson RP, Martinez-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(22):3781–3792

Gustafsson R, Zhang S, Masuyer G, Dong M, Stenmark P (2018) Crystal structure of botulinum neurotoxin A2 in complex with the human protein receptor SV2C reveals plasticity in receptor binding. Toxins 10(4):E153

Hamark C, Berntsson RP, Masuyer G, Henriksson LM, Gustafsson R, Stenmark P, Widmalm G (2017) Glycans confer specificity to the recognition of ganglioside receptors by botulinum neurotoxin A. J Am Chem Soc 139(1):218–230

Hill KK, Smith TJ, Helma CH, Ticknor LO, Foley BT, Svensson RT, Brown JL, Johnson EA, Smith LA, Okinaka RT, Jackson PJ, Marks JD (2007) Genetic diversity among botulinum neurotoxin-producing clostridial strains. J Bacteriol 189(3):818–832

Jahn R, Scheller RH (2006) SNAREs--engines for membrane fusion. Nat Rev Mol Cell Biol 7 (9):631–643

Jin R, Rummel A, Binz T, Brunger AT (2006) Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity. Nature 444(7122):1092–1095

Jin R, Sikorra S, Stegmann CM, Pich A, Binz T, Brunger AT (2007) Structural and biochemical studies of botulinum neurotoxin serotype C1 light chain protease: implications for dual substrate specificity. Biochemistry 46(37):10685–10693

Johnson EA (1999) Clostridial toxins as therapeutic agents: benefits of nature's most toxic proteins. Annu Rev Microbiol 53:551–575

Kalb SR, Baudys J, Webb RP, Wright P, Smith TJ, Smith LA, Fernandez R, Raphael BH, Maslanka SE, Pirkle JL, Barr JR (2012) Discovery of a novel enzymatic cleavage site for botulinum neurotoxin F5. FEBS Lett 586(2):109–115

Keller JE, Neale EA, Oyler G, Adler M (1999) Persistence of botulinum neurotoxin action in cultured spinal cord cells. FEBS Lett 456(1):137–142

Kosenina 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(12):1403–1410

Kumaran D, Eswaramoorthy S, Furey W, Navaza J, Sax M, Swaminathan S (2009) Domain organization in Clostridium botulinum neurotoxin type E is unique: its implication in faster translocation. J Mol Biol 386(1):233–245

Lacy DB, Tepp W, Cohen AC, DasGupta BR, Stevens RC (1998) Crystal structure of botulinum neurotoxin type A and implications for toxicity. Nat Struct Biol 5(10):898–902

Lalli G, Bohnert S, Deinhardt K, Verastegui C, Schiavo G (2003) The journey of tetanus and botulinum neurotoxins in neurons. Trends Microbiol 11(9):431–437

Lam KH, Guo Z, Krez N, Matsui T, Perry K, Weisemann J, Rummel A, Bowen ME, Jin R (2018a) A viral-fusion-peptide-like molecular switch drives membrane insertion of botulinum neurotoxin A1. Nat Commun 9(1):5367

Lam KH, Qi R, Liu S, Kroh A, Yao G, Perry K, Rummel A, Jin R (2018b) The hypothetical protein P47 of Clostridium botulinum E1 strain Beluga has a structural topology similar to bactericidal/permeability-increasing protein. Toxicon 147:19–26

Lamanna C, Mc EO, Eklund HW (1946) The purification and crystallization of Clostridium botulinum type A toxin. Science 103(2681):613

Lebeda FJ, Olson MA (1995) Structural predictions of the channel-forming region of botulinum neurotoxin heavy chain. Toxicon 33(4):559–567

Lee K, Gu S, Jin L, Le TT, Cheng LW, Strotmeier J, Kruel AM, Yao G, Perry K, Rummel A, Jin R (2013) Structure of a bimodular botulinum neurotoxin complex provides insights into its oral toxicity. PLoS Pathog 9(10):e1003690

Lee K, Zhong X, Gu S, Kruel AM, Dorner MB, Perry K, Rummel A, Dong M, Jin R (2014) Molecular basis for disruption of E-cadherin adhesion by botulinum neurotoxin A complex. Science 344(6190):1405–1410

Leuchs J (1910) Beitraege zur kenntnis des toxins und antitoxins des *Bacillus botulinus*. Z Hyg Infekt 76:55–84

Mahrhold S, Rummel A, Bigalke H, Davletov B, Binz T (2006) The synaptic vesicle protein 2C mediates the uptake of botulinum neurotoxin A into phrenic nerves. FEBS Lett 580 (8):2011–2014

Mahrhold S, Strotmeier J, Garcia-Rodriguez C, Lou J, Marks JD, Rummel A, Binz T (2013) Identification of the SV2 protein receptor-binding site of botulinum neurotoxin type E. Biochem J 453(1):37–47

Mahrhold S, Bergstrom T, Stern D, Dorner BG, Astot C, Rummel A (2016) Only the complex N559-glycan in the synaptic vesicle glycoprotein 2C mediates high affinity binding to botulinum neurotoxin serotype A1. Biochem J 473(17):2645–2654

Mansfield MJ, Adams JB, Doxey AC (2015) Botulinum neurotoxin homologs in non-Clostridium species. FEBS Lett 589(3):342–348

Mansfield MJ, Wentz TG, Zhang S, Lee EJ, Dong M, Sharma SK, Doxey AC (2019) Bioinformatic discovery of a toxin family in Chryseobacterium piperi with sequence similarity to botulinum neurotoxins. Sci Rep 9(1):1634

Maslanka SE, Luquez C, Dykes JK, Tepp WH, Pier CL, Pellett S, Raphael BH, Kalb SR, Barr JR, Rao A, Johnson EA (2016) A novel botulinum neurotoxin, previously reported as serotype H, has a hybrid-like structure with regions of similarity to the structures of serotypes A and F and is neutralized with serotype A antitoxin. J Infect Dis 213(3):379–385

Masuyer G, Conrad J, Stenmark P (2017) The structure of the tetanus toxin reveals pH-mediated domain dynamics. EMBO Rep 18(8):1306–1317

Masuyer G, Zhang S, Barkho S, Shen Y, Henriksson L, Kosenina S, Dong M, Stenmark P (2018) Structural characterisation of the catalytic domain of botulinum neurotoxin X – high activity and unique substrate specificity. Sci Rep 8(1):4518

Matsumura T, Sugawara Y, Yutani M, Amatsu S, Yagita H, Kohda T, Fukuoka S, Nakamura Y, Fukuda S, Hase K, Ohno H, Fujinaga Y (2015) Botulinum toxin A complex exploits intestinal M cells to enter the host and exert neurotoxicity. Nat Commun 6:6255

Meng J, Ovsepian SV, Wang J, Pickering M, Sasse A, Aoki KR, Lawrence GW, Dolly JO (2009) Activation of TRPV1 mediates calcitonin gene-related peptide release, which excites trigeminal sensory neurons and is attenuated by a retargeted botulinum toxin with anti-nociceptive potential. J Neurosci 29(15):4981–4992

Montal M (2010) Botulinum neurotoxin: a marvel of protein design. Annu Rev Biochem 79:591–617

Montal MS, Blewitt R, Tomich JM, Montal M (1992) Identification of an ion channel-forming motif in the primary structure of tetanus and botulinum neurotoxins. FEBS Lett 313(1):12–18

Montecucco C (1986) How do tetanus and botulinum toxins bind to neuronal membranes? Trens Biochem Sci 11(8):314–317

Moriishi K, Koura M, Abe N, Fujii N, Fujinaga Y, Inoue K, Ogumad K (1996a) Mosaic structures of neurotoxins produced from Clostridium botulinum types C and D organisms. Biochim Biophys Acta 1307(2):123–126

Moriishi K, Koura M, Fujii N, Fujinaga Y, Inoue K, Syuto B, Oguma K (1996b) Molecular cloning of the gene encoding the mosaic neurotoxin, composed of parts of botulinum neurotoxin types C1 and D, and PCR detection of this gene from Clostridium botulinum type C organisms. Appl Environ Microbiol 62(2):662–667

Moritz MS, Tepp WH, Bradshaw M, Johnson EA, Pellett S (2018) Isolation and characterization of the novel Botulinum neurotoxin A subtype 6. mSphere 3(5):e00466

Nishiki T, Kamata Y, Nemoto Y, Omori A, Ito T, Takahashi M, Kozaki S (1994) Identification of protein receptor for Clostridium botulinum type B neurotoxin in rat brain synaptosomes. J Biol Chem 269(14):10498–10503

Pang ZP, Melicoff E, Padgett D, Liu Y, Teich AF, Dickey BF, Lin W, Adachi R, Sudhof TC (2006) Synaptotagmin-2 is essential for survival and contributes to Ca2+ triggering of neurotransmitter release in central and neuromuscular synapses. J Neurosci 26(52):13493–13504

Peck MW, Smith TJ, Anniballi F, Austin JW, Bano L, Bradshaw M, Cuervo P, Cheng LW, Derman Y, Dorner BG, Fisher A, Hill KK, Kalb SR, Korkeala H, Lindstrom M, Lista F, Luquez C, Mazuet C, Pirazzini M, Popoff MR, Rossetto O, Rummel A, Sesardic D, Singh BR, Stringer SC (2017) Historical perspectives and guidelines for botulinum neurotoxin subtype nomenclature. Toxins 9(1):38

Pellett S, Tepp WH, Whitemarsh RC, Bradshaw M, Johnson EA (2015) In vivo onset and duration of action varies for botulinum neurotoxin A subtypes 1-5. Toxicon 107(Pt A):37–42

Pellett S, Bradshaw M, Tepp WH, Pier CL, Whitemarsh RCM, Chen C, Barbieri JT, Johnson EA (2018) The light chain defines the duration of action of botulinum toxin serotype A subtypes. MBio 9(2). pii: e00089-18

Peng L, Berntsson RP, Tepp WH, Pitkin RM, Johnson EA, Stenmark P, Dong M (2012) Botulinum neurotoxin D-C uses synaptotagmin I and II as receptors, and human synaptotagmin II is not an effective receptor for type B, D-C and G toxins. J Cell Sci 125(Pt 13):3233–3242

Pfenninger W (1924) Toxico, immunologic and serologic relationship of *B. botulinus*, type C and *B. parabotulinus*. J Infect Dis 35:347–352

Pier CL, Chen C, Tepp WH, Lin G, Janda KD, Barbieri JT, Pellett S, Johnson EA (2011) Botulinum neurotoxin subtype A2 enters neuronal cells faster than subtype A1. FEBS Lett 585(1):199–206

Pirazzini M, Rossetto O, Eleopra R, Montecucco C (2017) Botulinum neurotoxins: biology, pharmacology, and toxicology. Pharmacol Rev 69(2):200–235

Restani L, Antonucci F, Gianfranceschi L, Rossi C, Rossetto O, Caleo M (2011) Evidence for anterograde transport and transcytosis of botulinum neurotoxin A (BoNT/A). J Neurosci 31 (44):15650–15659

Restani L, Giribaldi F, Manich M, Bercsenyi K, Menendez G, Rossetto O, Caleo M, Schiavo G (2012) Botulinum neurotoxins A and E undergo retrograde axonal transport in primary motor neurons. PLoS Pathog 8(12):e1003087

Rossetto O, Pirazzini M, Montecucco C (2014) Botulinum neurotoxins: genetic, structural and mechanistic insights. Nat Rev Microbiol 12(8):535–549

Rummel A, Bade S, Alves J, Bigalke H, Binz T (2003) Two carbohydrate binding sites in the H (CC)-domain of tetanus neurotoxin are required for toxicity. J Mol Biol 326(3):835–847

Rummel A, Mahrhold S, Bigalke H, Binz T (2004) The HCC-domain of botulinum neurotoxins A and B exhibits a singular ganglioside binding site displaying serotype specific carbohydrate interaction. Mol Microbiol 51(3):631–643

Schantz EJ, Johnson EA (1992) Properties and use of botulinum toxin and other microbial neurotoxins in medicine. Microbiol Rev 56(1):80–99

Schiavo G, Benfenati F, Poulain B, Rossetto O, Polverino de Laureto P, DasGupta BR, Montecucco C (1992a) Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. Nature 359(6398):832–835

Schiavo G, Poulain B, Rossetto O, Benfenati F, Tauc L, Montecucco C (1992b) Tetanus toxin is a zinc protein and its inhibition of neurotransmitter release and protease activity depend on zinc. EMBO J 11(10):3577–3583

Simpson LL, Rapport MM (1971) Ganglioside inactivation of botulinum toxin. J Neurochem 18 (7):1341–1343

Stenmark P, Dupuy J, Imamura A, Kiso M, 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(8):e1000129

Strotmeier J, Willjes G, Binz T, Rummel A (2012) Human synaptotagmin-II is not a high affinity receptor for botulinum neurotoxin B and G: increased therapeutic dosage and immunogenicity. FEBS Lett 586(4):310–313

Sudhof TC, Rothman JE (2009) Membrane fusion: grappling with SNARE and SM proteins. Science 323(5913):474–477

Sugawara Y, Matsumura T, Takegahara Y, Jin Y, Tsukasaki Y, Takeichi M, Fujinaga Y (2010) Botulinum hemagglutinin disrupts the intercellular epithelial barrier by directly binding E-cadherin. J Cell Biol 189(4):691–700

Surana S, Tosolini AP, Meyer IFG, Fellows AD, Novoselov SS, Schiavo G (2018) The travel diaries of tetanus and botulinum neurotoxins. Toxicon 147:58–67

Swaminathan S, Eswaramoorthy S (2000) Structural analysis of the catalytic and binding sites of Clostridium botulinum neurotoxin B. Nat Struct Biol 7(8):693–699

Tao L, Peng L, Berntsson RP, Liu SM, Park S, Yu F, Boone C, Palan S, Beard M, Chabrier PE, Stenmark P, Krupp J, Dong M (2017) Engineered botulinum neurotoxin B with improved efficacy for targeting human receptors. Nat Commun 8:53

Torii Y, Kiyota N, Sugimoto N, Mori Y, Goto Y, Harakawa T, Nakahira S, Kaji R, Kozaki S, Ginnaga A (2011) Comparison of effects of botulinum toxin subtype A1 and A2 using twitch tension assay and rat grip strength test. Toxicon 57(1):93–99

Tsai YC, Kotiya A, Kiris E, Yang M, Bavari S, Tessarollo L, Oyler GA, Weissman AM (2017) Deubiquitinating enzyme VCIP135 dictates the duration of botulinum neurotoxin type A intoxication. Proc Natl Acad Sci U S A 114(26):E5158–E5166

Vagin O, Tokhtaeva E, Garay PE, Souda P, Bassilian S, Whitelegge JP, Lewis R, Sachs G, Wheeler L, Aoki R, Fernandez-Salas E (2014) Recruitment of septin cytoskeletal proteins by botulinum toxin A protease determines its remarkable stability. J Cell Sci 127(Pt 15):3294–3308

van Ermengem E (1897) Ueber Einem Neuen anaeroben Bacillus und seine Beziehungen zum Botulismus. Z Hyg Infekt 26:1–56

Whitemarsh RC, Tepp WH, Bradshaw M, Lin G, Pier CL, Scherf JM, Johnson EA, 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(10):3894–3902

Whitemarsh RC, Tepp WH, Johnson EA, Pellett S (2014) Persistence of botulinum neurotoxin a subtypes 1-5 in primary rat spinal cord cells. PLoS One 9(2):e90252

Yao G, Zhang S, Mahrhold S, Lam KH, Stern D, Bagramyan K, Perry K, Kalkum M, Rummel A, Dong M, Jin R (2016) N-linked glycosylation of SV2 is required for binding and uptake of botulinum neurotoxin A. Nat Struct Mol Biol 23(7):656–662

Zhang S, Masuyer G, Zhang J, Shen Y, Lundin D, Henriksson L, Miyashita SI, Martinez-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, Martinez-Carranza M, Stenmark P, Gilmore MS, Doxey AC, Dong M (2018) Identification of a botulinum neurotoxin-like toxin in a commensal strain of *Enterococcus faecium*. Cell Host Microbe 23(2):169–176 e166

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 6:30257