HHS Public Access

Author manuscript

Cell Host Microbe. Author manuscript; available in PMC 2019 February 14.

Published in final edited form as:

Cell Host Microbe. 2018 February 14; 23(2): 169–176.e6. doi:10.1016/j.chom.2017.12.018.

Identification of a botulinum neurotoxin-like toxin in a commensal strain of *Enterococcus faecium*

Sicai Zhang[1,*], **Francois Lebreton**[3,4,*], **Michael J. Mansfield**[2,*], **Shin-Ichiro Miyashita**[1], **Jie Zhang**[1], **Julia A. Schwartzman**[4,5], **Liang Tao**[1], **Geoffrey Masuyer**[6], **Markel Martínez Carranza**[6], **Pål Stenmark**[6], **Michael S. Gilmore**[3,4,5], **Andrew C. Doxey**[2,#], and **Min Dong**[1,#]

[1]Department of Urology, Boston Children's Hospital, Department of Microbiology and Immunobiology and Department of Surgery, Harvard Medical School, Boston, MA 02115, USA

[2]Department of Biology, University of Waterloo, 200 University Ave. West, Waterloo, Ontario, N2L 3G1, Canada

[3]Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

[4]Infectious Disease & Microbiome Program, Broad Institute, Cambridge, MA 02142, USA

[5]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02114, USA

[6]Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden

Summary

Botulinum neurotoxins (BoNTs), produced by various *Clostridium* strains, are a family of potent bacterial toxins and potential bioterrorism agents. Here we report that an *Enterococcus faecium* strain isolated from cow feces carries a BoNT-like toxin, designated BoNT/En. It cleaves both VAMP2 and SNAP-25, proteins that mediate synaptic vesicle exocytosis in neurons, at sites distinct from known BoNT cleavage sites on these two proteins. Comparative genomic analysis determines that the *E. faecium* strain carrying BoNT/En is a commensal-type, and the BoNT/En gene is located within a typical BoNT gene cluster on a 206-kb putatively conjugative plasmid.

[#]Correspondence to: A.C.D. (acdoxey@uwaterloo.ca) or M.D. (min.dong@childrens.harvard.edu).
[*]These authors contributed equally.
Lead Contact: M.D. (min.dong@childrens.harvard.edu)

Author Contributions

A.C.D. and P.S. identified BoNT/En independently. F.L., J.S., M.S.G. established the collection, sequenced IDI0629, and carried out comparative genome analysis. A.C.D., P.S., M.M., and F.L. carried out bioinformatic analysis. S.Z. carried out all other experiments. J.Z. assisted with DAS assays. S.M. assisted with protein purification. M.M.C. and P.S. generated the BoNT/X antibody. M.D., F.L., and A.C.D. wrote the manuscript with input from all co-authors.

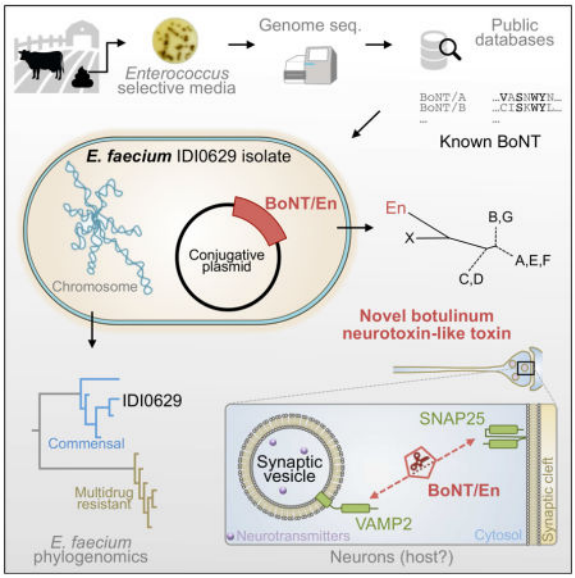
Conflicts of Interest

A provisional patent application on application of BoNT/En has been filed by Boston Children's Hospital, with A.C.D, P.S., S.Z., M.D., F.L. and M.S.G. as co-inventors. Other authors declare no conflicts of interest.

Although the host species targeted by BoNT/En remains to be determined, these findings establish an extended member of BoNTs and demonstrate the capability of *E. faecium*, a commensal organism ubiquitous in humans and animals and a leading cause of hospital-acquired multidrug resistant (MDR) infections, to horizontally acquire, and possibly disseminate, a unique BoNT gene cluster.

eTOC Blurb

Botulinum neurotoxins (BoNTs) are potent toxins produced by diverse bacteria in the *Clostridium* genus. Zhang et al. report that a commensal strain of *Enterococcus faecium* carries a conjugative plasmid encoding a BoNT-like toxin gene. Thus, a commensal organism can acquire and possibly disseminate BoNT genes.



BoNTs are one of the most dangerous potential bioterrorism agents (Category A and Tier 1 select agents) (Arnon et al., 2001). They have also been utilized to treat many medical conditions as well as for cosmetic applications (Montecucco and Molgo, 2005). There are seven well-established serotypes of BoNTs (BoNT/A-G). They are composed of a light chain (LC) and a heavy chain (HC) (Montal, 2010; Rossetto et al., 2014; Schiavo et al., 2000), connected via an inter-chain disulfide bond. The LC is a zinc-dependent metalloprotease. The HC contains the translocation domain (H_N) and the receptor-binding domain (H_C). BoNTs target neurons and block neurotransmission by cleaving host proteins VAMP1/2/3 (BoNT/B, D, F, and G), SNAP-25 (BoNT/A, C, E), or syntaxin 1 (Syx 1, BoNT/C). These three proteins mediate fusion of synaptic vesicles to plasma membranes and are the prototype of the SNARE family proteins (soluble NSF attachment protein receptor) (Jahn and Scheller, 2006; Sudhof and Rothman, 2009).

BoNT genes reside within two types of gene clusters (Hill et al., 2015). Both include a gene encoding NTNHA (non-toxic non-hemagglutinin protein), which forms a complex with BoNTs and protects them in the gastrointestinal (GI) tract (Gu et al., 2012). One type of

gene cluster expresses additional proteins HA17, HA33, and HA70, which facilitate the absorption of toxins across the epithelial barrier (Lee et al., 2014; Sugawara et al., 2010). The other type encodes proteins with unknown functions designated OrfX1, OrfX2, OrfX3, and P47 (Hill et al., 2015). Multiple mechanisms contribute to horizontal gene transfer and the recombination of BoNT clusters, including being located on plasmids or phages and the presence of transposases. Recent genomic studies revealed a growing number of subtypes and mosaic toxins (Barash and Arnon, 2014; Dover et al., 2014; Hill et al., 2007; Kalb et al., 2015; Maslanka et al., 2015; Montecucco and Rasotto, 2015). A new serotype, BoNT/X, was also recently identified in a *Clostridium botulinum* strain (Zhang et al., 2017).

The evolutionary origin of BoNTs remains a mystery. Recent studies reported a homolog of BoNT in a gram-positive bacterium *Weissella oryzae*, designated BoNT/Wo (Mansfield et al., 2015; Zornetta et al., 2016). However, BoNT/Wo is quite distinct from BoNTs. First, the sequence identity between BoNT/Wo versus other BoNTs is ~14–16%, below the normal range for the members of the BoNT family (~28–65%). Second, the two cysteines that form the essential inter-chain disulfide bond in BoNTs are not conserved in BoNT/Wo, suggesting a distinct mode of action. Third, the BoNT/Wo gene is not in a typical BoNT gene cluster.

Enterococcus faecium is a core commensal member in the human gut and widespread in most terrestrial animals (Lebreton et al., 2017; Lebreton et al., 2013; Schloissnig et al., 2013; Van Tyne and Gilmore, 2014). Since the 1970s, *E. faecium* has become a leading cause of hospital acquired multi-drug resistant (MDR) infection of the bloodstream, urinary tract, and surgical wounds (Arias and Murray, 2012; Gilmore et al., 2013). Compounding the problem, the enterococci serve as collection and distribution points for mobile elements, exemplified by acquiring and transmitting a variety of antibiotic resistance to gram-positive and gram-negative species (Courvalin, 1994). As a part of an ongoing diversity study, we have collected and sequenced a growing number of enterococcal strains. One strain, IDI0629, was recently isolated from cow feces in South Carolina in the U.S. Genomic sequencing revealed that it contains a BoNT-like toxin gene (GenBank: OTO22244.1), tentatively designated BoNT/En, which shows 29%–38.7% identity with the other BoNTs and is most closely related to BoNT/X (Fig. 1A). All key BoNT motifs are conserved in BoNT/En (Fig. 1B), including the zinc-dependent protease motif HExxH (residues H225 – H229) in the LC (Schiavo et al., 1992), two cysteines that may form an inter-chain disulfide bond (C424 and C438), and a ganglioside-binding motif SxWY in the H_C (residues S1250 to Y1253) (Rummel et al., 2004).

We first examined whether the LC of BoNT/En (En-LC) cleaves SNARE proteins. The most obvious change after incubation of rat brain detergent extracts (BDE) with recombinantly produced En-LC was disappearance of VAMP2 immunoblot signals (Fig. 1C), which was inhibited by the metalloprotease inhibitor EDTA. The LCs of BoNT/X (X-LC), which cleaves VAMP2, and BoNT/A (A-LC), which cleaves SNAP-25, were assayed in parallel as controls. Consistent with this finding, En-LC cleaved a purified VAMP2 fragment (residues 1–93) (Fig. 1D).

Using the liquid chromatography-tandem mass spectrometry approach (LC-MS/MS), we identified a single dominant peptide peak after incubation of VAMP2 (1–93) with En-LC,

which matched the sequence of D68-L93 (Fig. 1E). The corresponding N-terminal fragment was also identified. To confirm this result, we further analyzed a glutathione *S*-transferase (GST)-tagged VAMP2 (residues 33–86) (Fig. S1A). Incubation with En-LC generated a single dominant peptide that matched D68-R86 (Fig. S1B). Thus, En-LC cleaves VAMP2 between A67 and D68 (Fig. 1F).

We further examined whether other members of VAMP family (VAMP1, 3, 4, 5, 7, 8, Sec22b, and Ykt6), SNAP-25 family (SNAP-25, SNAP-23, and SNAP-29), and Syx family (Syx 1A, 1B, 2, 3, 4) can be cleaved by En-LC. These proteins were expressed in HEK293 cells via transient transfection. Cell lysates were then incubated with En-LC (Fig. S1C). VAMP1 and VAMP3 are highly homologous to VAMP2, and both were cleaved. No other VAMPs were cleaved. Among Syx members, Syx 1B and Syx 4 were cleaved. Within SNAP-25 family, SNAP-25 and SNAP-23 were cleaved. The finding that Syx 1B can be cleaved is consistent with the observation that En-LC reduced Syx 1 immunoblot signals in BDE, which include both Syx 1A and 1B (Fig. 1C). The finding that SNAP-25 was cleaved was a surprise, as SNAP-25 in BDE was hardly cleaved by En-LC (Fig. 1C). The reason for the difference between BDE and HEK cell lysates is unknown.

We next compared the cleavage efficacy of En-LC (0.1 μM) toward purified cytosolic fragments of Syx 1B and Syx 4, full-length SNAP-25 and SNAP-23, and GST-VAMP2 (33–86). While VAMP2 was fully cleaved within a few minutes, only minor cleavage of Syx 1B and SNAP-25, and no cleavage of Syx 4 or SNAP-23 were observed within 20 minutes (Fig. S1D, S2A), suggesting that En-LC cleaves VAMP2 far more efficiently than Syx and SNAP-25 family *in vitro*.

Increasing the concentration of En-LC (6 μM) and incubation time enhanced the cleavage of Syx 1B and also resulted in detectable cleavage of Syx 4 (Fig. S1E). By mass spectrometry analysis, the cleavage sites were mapped to M182-D183 in Syx 1B and K191-D192 in Syx 4, which are homologous locations between Syx 1B and Syx 4 (Fig. S1F–H). Identification of the cleavage site in SNAP-25 and SNAP-23 turned out to be problematic, as increasing En-LC concentrations and incubation time resulted in multiple smear bands. As we are concerned about the relevance of these low-efficiency cleavage events observed *in vitro*, we decided to focus on determining which SNARE proteins can be cleaved in neurons.

Incubation of nanomolar concentrations of LC-H_N of BoNTs often results in uptake of low levels of LC-H_N into neurons (Chaddock et al., 2002), which provides a convenient way to examine the action of En-LC in a more physiologically relevant manner in neurons. We first examined whether the En-LC is connected to its HC via an inter-chain disulfide bond. Sequence alignment showed two cysteine residues at expected locations (Fig. 1G). As the endogenous protease that can activate BoNT/En remains unknown, we inserted a thrombin cleavage site in the BoNT/En linker region. This modified En-LC-H_N behaves as expected after incubation with thrombin: it remains a single band without reducing agent (− DTT), and separated into two bands in the presence of DTT (Fig. 1H), confirming the existence of the inter-chain disulfide bond.

Exposing cultured rat cortical neurons to En-LC-H_N in medium for 12 h resulted in a loss of both VAMP2 and SNAP-25 immunoblot signals (Fig. 1I). Proteolytic activation of En-LC-H_N increased its potency (Fig. 1I). Interestingly, SNAP-25 was cleaved efficiently by BoNT/En in neurons (Fig. 1I). This is not without precedent: BoNT/C also cleaves SNAP-25 efficiently in cells, but inefficiently *in vitro* (Foran et al., 1996). It has been proposed that the optimal cleavage of SNAP-25 by BoNT/C requires a proper membrane environment (Foran et al., 1996), which could also be a requirement for BoNT/En. In contrast to VAMP2 and SNAP-25, Syx 1 (including Syx 1A and 1B) was not cleaved in neurons (Fig. 1I). This is likely because cleavage of Syx 1 is inefficient and requires high levels of En-LC, which can be achieved only *in vitro*. Thus, VAMP2 and SNAP-25, but not Syx 1, are relevant targets for BoNT/En in neurons.

While the SNAP-25 antibody (CI 71.1), which recognizes the N-terminal residues 20 – 40 (designated SNAP-N), did not detect any cleavage product, a second SNAP-25 antibody raised against the C-terminal residues 195–206 (designated SNAP-C) detected a cleavage product (Fig. 1I, lower panel). To determine the cleavage site, we isolated this cleavage product from neuronal lysates by immunoprecipitation using the SNAP-C antibody (Fig. S2B). By comparing the full-length SNAP-25 versus the cleavage fragment using mass spectrometry analysis, the starting residue of the fragment was determined to be D70 (Fig. S2C). We also used the SNAP-C antibody to pull-down the cleavage product from lysates of HEK293 cells transfected with HA-tagged SNAP-25, after incubation with En-LC (Fig. S2B). Mass spectrometry analysis confirmed that the C-terminal fragment generated under this condition also starts with D70 (Fig. S2C, D). Thus, the cleavage site is between K69-D70, which is distinct from all known BoNT cleavage sites on SNAP-25 (Fig. 1J). It is puzzling that there was no corresponding N-terminal fragment detectable by the SNAP-N antibody. The HA antibody also did not detect the N-terminal fragment of HA-tagged SNAP-25 cleaved by En-LC and analyzed on 4–20% gradient gels (Fig. S2E). It remains unknown whether the N-terminal fragment is further cleaved by En-LC or it is degraded.

We next assessed the toxicity of full-length BoNT/En. Due to biosafety considerations, we decided not to clone the full-length toxin gene. Instead, we produced a limited amount of full-length BoNT/En utilizing the sortase-mediated ligation of two non-toxic fragments: En-LC-H_N and En-H_C (Fig. 1K) (McCluskey and Collier, 2013; Popp et al., 2007; Zhang et al., 2017). En-H_C showed poor solubility once cleaved from the GST tag, which limited the efficiency of sortase-mediated ligation. Nevertheless, a low level of full-length BoNT/En can be generated using this approach (Fig. S2F). We found that ligation by sortase only slightly increased the cleavage of VAMP2 and SNAP-25 in cultured rat cortical neurons compared to the mixture without sortase. This enhanced activity diminished when the ligation mixture was diluted by only 1.25-fold, suggesting that ligated full-length BoNT/En has a rather low level of activity over En-LC-H_N alone (Fig. 1L). Consistently, injecting as much as 1 μg of ligated BoNT/En did not induce any paralysis in mice in a well-established non-lethal assay known as the Digit Abduction Score (DAS) assay (Aoki, 2001), which measures the degree of local paralysis (the ability to spread toes during startle response) following injection of toxins into the mouse hind limb.

To determine whether lack of activity of BoNT/En is due to its H_C, we generated a chimeric toxin by ligating En-LC-H_N with the H_C of BoNT/A (A-H_C) (Fig. S2G). The presence of A-H_C greatly facilitated entry into neurons, as a 1:1000 dilution of the chimeric toxin mixture still cleaved more VAMP2 and SNAP-25 than the undiluted mixture of En-LC-H_N and A-H_C without sortase (Fig. 1M). Consistently, injecting 1 ng of the ligated chimeric toxin paralyzed the mouse limb muscle in DAS assays (Fig. 1N). These results suggest that rat/mouse neurons do not contain high-affinity receptors for BoNT/En, although we cannot exclude the possibility that En-H_C purified in isolation has folding problems. It will be important to produce native full-length BoNT/En to properly evaluate its toxicity.

We also carried out dot blot assays using antisera raised against other BoNTs, including four horse antisera (trivalent anti-BoNT/A, B, and E, anti-BoNT/C, anti-BoNT/F) and two goat antisera (anti-BoNT/D and anti-BoNT/G) previously characterized (Zhang et al., 2017), as well as a rabbit polyclonal antibody developed specifically against BoNT/X. These antisera recognized their corresponding BoNTs, but none recognized BoNT/En (a mixture of En-LC-H_N and En-H_C), confirming that BoNT/En is a unique BoNT serotype (Fig. 1O).

We then turned our attention to understanding the genetic background of the IDI0629 strain harboring BoNT/En. The existence of distinct subpopulations of *E. faecium* has been described (Lebreton et al., 2013), with MDR epidemic hospital-adapted isolates occurring in Clade A, and most antibiotic susceptible community-derived isolates constituting a distant branch, Clade B. Comparison of nucleotide polymorphisms in 1,824 genes shared by IDI0629 and strains of Clades A and B, places the IDI0629 within the community-associated commensal strains (Clade B, Fig. 2A). It was most closely related to strains EnGen0263 and IDI0518. EnGen0263 is a rare example of a MDR Clade B strain isolated from a hospitalized patient, while IDI0518 was isolated from the feces of a wild bird in Europe and, similar to IDI0629, remains susceptible to all antibiotics except those to which *E. faecium* is intrinsically resistant (Fig. 2A and Table S1).

There are 333 genes uniquely found in IDI0629 (Fig. 2B). Except for its prophage content and an integrated conjugative element, the chromosome of IDI0629 is closely related to those of strains EnGen0263 and IDI0518 (Fig. 2C). Most of IDI0629 unique genes are on one scaffold (Genbank: NGLI01000004), which appears to be a plasmid belonging to *rep*US15 family. The *rep*US15 conjugative plasmids are prevalent, circulating in >80% of *E. faecium*, and have been associated with the transfer and dissemination of vancomycin resistance (Freitas et al., 2016). Initial sequencing left 11 gaps in this scaffold, we thus further performed long-reads MinION sequencing and hybrid assemblies, which circularized a 206-kb plasmid containing BoNT/En (termed pBoNT/En, Fig. 2C, Table S2). A 60-kb backbone of pBoNT/En, which includes the origin of replication (*repA*), a bacteriocin operon, and genes encoding a type IV secretion system and pilus assembly, is conserved among *rep*US15 plasmids in all three *E. faecium* strains (IDI0629, EnGen0263 and IDI0518) and likely confer conjugative ability (Fig. 2C). In contrast, a 90-kb region encompassing the BoNT/En operon and two putative immunoglobulin A (IgA) proteases has been acquired by this plasmid in IDI0629 (Fig. 2C). This region has lower GC content (31%) than the plasmid backbone (34.7%) and the chromosome (38% GC, Fig. 2C), suggesting that these regions were likely acquired from a source outside of the *Enterococcus*

genus. Of note, *Clostridium botulinum* has a low GC-content (28%) similar to other *Clostridium* species (Sebaihia et al., 2007).

In addition, pBoNT/En and the plasmid in EnGen0263 also share a low GC region (30%) that encodes an operon annotated as conferring urea metabolism. The presence of a urease gene cluster is not frequent in *E. faecium* but is enriched in strains isolated from domestic and wild ruminants (Laukova and Koniarova, 1995). IDI0629 was indeed isolated from cow feces. Although there is no report of cattle botulism cases in the farm where IDI0629 was isolated, cattle are susceptible to botulism (usually caused by type C and D toxins), and healthy cattle could also be intermittent carrier of *C. botulinum* (Abdel-Moein and Hamza, 2016). In such context, the GI tract of ruminants would provide an environment for ureolytic *E. faecium* and toxin-producing *Clostridium* to coexist and exchange genes.

The BoNT/En gene is located within a typical OrfX gene cluster, preceded by a NTNHA gene and containing putative *orfX2*, *orfX3*, and *p47* genes (Fig. 2D). A gene located 5′ to *orfX2* showed a relatively low degree of sequence similarity to *orfX1* and was therefore designated as an *orfX1*-like gene. The BoNT/En gene cluster is flanked by a 1719-bp direct repeat sequence (90.1 % nucleotide identity), with two truncated non-functional copies of a *repB* gene on each side. This region also contains a putative phage endolysin, an insertion element (IS204), three putative site-specific recombinases, and additional hypothetical genes. There is also a putative Phd-Doc cassette within this region, an addiction module usually utilized to maintain mobile elements. The occurrence of direct repeats flanking the pBoNT/En cluster suggests that this region was acquired by a *rep*US15 plasmid precursor through homologous recombination within conserved *repB* sequences, potentially mediated by associated putative recombinases. This also suggests that the BoNT/En cluster may be mobile by additional mechanisms beyond conjugation.

BoNT-like gene clusters have not previously been identified in any bacterial species outside of *Clostridium* and no toxins of *E. faecium* have been reported before now. It is disconcerting to find a member of potent neurotoxins in this widely distributed gut microbe, which is a leading cause of hospital acquired infections (Lebreton et al., 2017; Gilmore et al., 2013). The rarity of BoNT/En producing *E. faecium* in strains sequenced so far may reflect its recent acquisition, or may be due to the relatively limited sampling of Clade B strains from wild ecologies. To know the scope of the natural diversity of genes harbored by enterococci, and to monitor the emergence of new strains, it will be critical to survey the enterococci beyond lineages that commonly cause infection now. Many important questions remain unknown including the evolutionary origin of BoNT/En and the host species/cell types targeted by BoNT/En. Nevertheless, the capability of *E. faecium* to acquire a BoNT gene cluster could create emerging strains with severe consequences. Furthermore, the possibility of introducing a BoNT cluster into MDR *E. faecium* strains could pose a significant biosecurity threat.

Star methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Min Dong (min.dong@childrens.harvard.edu).

Data Availability

Whole genome shotgun sequencing data for IDI0629 (GenBank: NZ_NGLI00000000), EnGen0263 (GenBank: NZ_AJAD00000000), and IDI0518 (GenBank: NZ_NGLS00000000) were deposited into GenBank. The accession number for BoNT/En is OTO22244.1. Accession numbers for other *E. faecium* strains and other genes on pBoNT/En are listed in Table S1 and S2.

Experimental Model and Subject Details

Enterococcal strains—The sources of bacterial strains and genome sequences used in this study are listed in Table S1. All strains were routinely grown in brain heart infusion (BHI) at 37°C unless noted otherwise, and stored frozen at −80°C in BHI supplemented with 20% glycerol. Strain IDI0629 was isolated from a swab sample of cow feces collected in a local farm in South Carolina (USA) in 2016, by selection on bile esculin azide agar plates (Enterococcosel®, BBL).

Cell lines—HEK293T cells were purchased from ATCC and propagated in the lab. Cells were cultured in DMEM media plus 10% fetal bovine serum (FBS) and 100 U penicillin / 0.1 mg/ml streptomycin at 37 °C.

Mice and Rats—All animal studies were conducted in accord with ethical regulations under protocols approved by the Institute Animal Care and Use Committee (IACUC) at Boston Children's Hospital (#3030). Adult CD-1 strain mice (both male and female were examined randomly) and pregnant Sprague Dawley rats (CD IGS) were purchased from Charles River.

Quantification and Statistical Analysis

DAS assays utilized 4 mice per group (n = 4, Figure 1N).

Method Details

Materials—Mouse monoclonal antibodies for Syx 1 (HPC-1), SNAP-25 (C171.2), VAMP2 (C169.1) were generously provided by E. Chapman (Madison, WI) and are available from Synaptic Systems (Goettingen, Germany). Rabbit polyclonal antibody against VAMP4, Sec22b, Syx 2, Syx 3 and Syx 4 were purchased from Synaptic Systems (Cat. No. 136002, No. 186003, No. 110022, No. 110032 and No. 110042, respectively). The following mouse monoclonal antibodies were purchased from indicated vendors: actin (Sigma, AC-15); anti-HA (Covance, 16B12); anti-Myc (Millipore, 9E10); anti-SNAP-25 (Abcam, ab5666). Equine polyclonal antisera against BoNT/A/B/E, BoNT/C, BoNT/DC, BoNT/F, and goat polyclonal antisera against BoNT/G were generously provided by S. Sharma (FDA). Goat

polyclonal antibody against BoNT/D was purchased from Novus Biologicals (NB10062469). BoNTs were purchased from Metabiologics (Madison, WI).

cDNA and constructs—The cDNAs encoding En-LC (residues 1–434), En-HC (residues 863–11279), En-HN (residues 436–862), and X-LC (residues 1–439, GenBank No. WP045538952.1) were synthesized by GenScript (New Brunswick, NJ). The cDNA encoding En-LC-HN was generated in-house using the Gibson assembly method with a thrombin protease cleavage site inserted between Q432 and L435. En-LC, X-LC, A-LC (residues 1–425) were cloned into pET28 vectors with His6-tag on their N-termini. En-HC and A-HC (residues 875–1297, GenBank No. AF488749) were cloned into pGEX4T to express as GST-tagged proteins. En-LC-HN was cloned into a pET22b vector, with the peptide sequence LPETGG fused to their C-termini, followed by a His6-tag, and were purified as His6-tagged proteins. VAMP2 (1–93) was cloned into pET28 vector with a His6-tag on the N-terminus and pGEX4T vector and expressed as a GST-tagged protein. Full-length mouse VAMP1, 3, and rat VAMP7, 8 were cloned into modified pcDNA3.1 vectors, with a HA tag fused to their C-termini. Constructs expressing full-length rat Ykt6 and mouse Sec22b, in pcDNA3.1 vector with an N-terminal Myc tag, were generously provided by J. Hay (Missoula, MT). Full-length Syx 1A, Syx 1B, Syx 2, Syx 3, Syx 4, SNAP-23, and SNAP-29 were cloned into Syn-lox vector between BamHI/NotI with the exception that syntaxin 1B is fused with a HA tag to their N-termini. Full-length SNAP-25 was cloned into pcDNA3.1 vectors between BamHI/NotI, with a HA tag fused to their N-termini. Syx 1A (1–265), Syx 1B (1–251), Syx 4 (1–273), SNAP-23, and SNAP-25 were cloned between NheI/NotI sites in pET28a and expressed as His6-tagged proteins. The construct encoding His6-tagged sortase (SrtA*) was generously provided by B. Pentelute (Boston, MA) and has been described previously (McCluskey and Collier, 2013).

Bioinformatic analysis—BoNT/En was discovered using blastp with BoNT/X as a query sequence against the nr database with default parameters (BLOSUM62, gap existence = 11, gap extension = 1, with conditional compositional score matrix adjustment). As of April 2017, this search space covered a total of 231,827,651,552 bases and 200,877,884 sequences. Domains were annotated using the hmmsearch command of the HMMER package against the Pfam database (v31.0). Genomic architecture visualized using genoplotR (v0.8.6) in R (v3.4.1). BoNT sequences representing all major lineages (A–G, F5A, and X) were aligned in a multiple alignment using ClustalO (v1.2.1), then pairwise identity between BoNT/En and the others was calculated in a 50 amino acid sliding window across the length of the multiple alignment with a step of 1. Regression splines were calculated using the splines base package in R.

Illumina and MinION genome sequencing—For short read Illumina sequencing of strains IDI0629 (formerly named DIV0629), EnGen0263 and IDI0518 (formerly named DIV0518), following overnight culture in BHI broth, total DNA was isolated using the Qiagen DNeasy Blood & Tissue Kit and quantified using the Qubit dsDNA HS assay. Both jumping libraries and 180-bp paired fragment libraries were prepared using commercial kits (Kapa Biosystems). For obtaining longer sequencing reads to close gaps in the IDI0629 scaffold containing BoNT/En, total DNA was isolated using Wizard Genomic DNA

Purification Kit. The obtained DNA was then prepared for sequencing according to the MinION device protocol (March 2016, Oxford Nanopore Technologies), and ~60 ng was sequenced using an Oxford Nanopore MinION MkI (Oxford Nanopore Technologies). The sequencer was run for a total of 48 h, with a Pre-Sequencing Mix top up of a second ~60 ng after the initial 24 h. Albacore (v2.1.2) was used to extract fasta-format files from the fast5-format read files obtained. Hybrid assembly of these reads with quality-trimmed 2×150bp NextSeq Illumina reads, was then performed using SPAdes 3.8.0 with default options, except for --nanopore --only-assembler --k 25,35,45,55,65,75. Scaffolds ≤ 1000 bp were removed from the assembly.

Genome assembly and annotation—NCBI accession numbers for each genome are listed in Table S1. The genomes of previously sequenced and published E. faecium included in our phylogenetic analysis were downloaded from Genbank, bringing the total number of genomes included in the analysis to 18 (Table S1). To assure consistency and to reduce artifacts among the genomes being analyzed, all genomes, including those from Genbank, were annotated or re-annotated in uniform manner using the Broad Institute's prokaryotic pipeline, with the Enterococcus-specific approach described previously (Lebreton et al., 2017). To investigate the genomic diversity of the different enterococcal species, orthologous genes were identified in all 18 genomes using Synerclust (https://github.com/SynerClust/SynerClust). Orthogroups contain orthologs, which are vertically inherited genes that likely have the same function, and also possibly paralogs, which are duplicated genes that may have different function. There were 1824 single copy core orthogroups across our set of 18 strains. The presence of bacterial immunity (i.e. CRISPR/cas and restriction-modification systems), drug resistance, plasmid and prophages was determined using available online tools (CRISPRfinder: http://crispr.u-psud.fr/Server/; PlasmidFinder http://cge.cbs.dtu.dk/services; ResFinder: http://cge.cbs.dtu.dk/services/ResFinder/; Phast: https://cge.cbs.dtu.dk/services/Phast/).

Phylogenetic tree and comparative genomics—The phylogenetic tree of the Clade B E. faecium was constructed by applying RAxML (Stamatakis, 2006) to a concatenated alignment of 1824 single-copy core orthogroups across all organisms. The 1000 bootstrap iterations were calculated using RAxML's rapid bootstrapping algorithm. To visualize the gene content differences between strains IDI0629 and its closest phylogenetic neighbors EnGen0263 and IDI0518, the draft chromosome sequence of IDI0629 or the finished (i.e. closed and circularized) plasmid pBoNT/En sequences were used as reference for a comparative circular alignment using the synerclust orthology results and the DNAplotter software (v10.2).

Production of BoNT/X polyclonal antibody—An inactive full-length BoNT/X mutant (R360A/Y363F, BoNT/XRY) was expressed and purified as described previously (Zhang et al., 2017), using pET22b vector encoding BoNT/XRY with the His6-tag on its C-terminus for expression in E. coli BL21 (DE3). Antibody production was carried out by EZBiolab Inc. (USA). The immunogen (BoNT/XRY) was diluted with physiological saline and then mixed with the corresponding adjuvant (Freund's complete adjuvant) 1: 1. Antigens and adjuvants were completely mixed to form a stable emulsion and then the emulsion was

injected into New Zealand white rabbits using a back multipoint injection method, 0.1 ml per point. After 2 weeks, subsequent immunizations were performed with Freund's incomplete adjuvant at different points. The amount of antigen per each of four immunizations was 100 μg. The titer was tested with ELISA after four immunizations. Rabbit blood was collected from the carotid artery. Sera were purified by protein A-chromatography, dialyzed against PBS buffer, and then lyophilized.

Protein purification—*E. coli* BL21 (DE3) was utilized for protein expression. In general, induction of expression was carried out with 0.1 mM IPTG at 22 °C overnight. Bacterial pellets were disrupted in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl) by sonication, and supernatants were collected after centrifugation at 20,000g for 30 min at 4°C. Protein purification was carried out using AKTA Prime FPLC system (GE), and purified proteins were further desalted with a PD-10 column (GE, 17-0851-01).

Cleavage of SNARE proteins in rat brain detergent extracts (BDE)—Rat brain was homogenized in 15 ml 320 mM sucrose buffer, followed by centrifugation at 5000 rpm for 2 min at 4°C. Supernatants were collected and centrifuged at 11,000 rpm for 12 min. The pellet was collected and solubilized for 30 min in 15 ml Tris-buffered saline (TBS: 20 mM Tris, 150 mM NaCl) plus 2% of Triton X-100 and a cocktail of protease inhibitors (Roche, CA). Samples were subsequently centrifuged at 17,000 rpm for 20 min to remove insoluble materials. The final BDE concentration was ~ 2 mg/ml proteins. BDE (60 μl) were incubated with En-LC (2 μM), X-LC (2 μM), or A-LC (2 μM), for 1 hour at 37 °C, and then analyzed by immunoblot using the enhanced chemiluminescence (ECL) method (Pierce). As controls, LCs were pre-incubated with 20 mM EDTA for 20 minutes at room temperature (RT) prior to adding into BDE.

Cleavage of recombinant VAMP, Syx, and SNAP-25 families by En-LC—VAMP2 (1–93) was expressed and purified as a His6-tagged protein and also GST-tagged proteins. Syx 1A (1–265), Syx 1B (1–251), Syx 4 (1–273), SNAP-23, and SNAP-25 were expressed and purified as His6-tagged proteins. These proteins (0.3 mg/ml) were incubated with 0.1 or 6 μM En-LC in TBS buffer for indicated times at 37 °C. Samples were either analyzed by SDS-PAGE gels and Coomassie Blue staining, or subjected to mass spectrometry analysis.

Cleavage of VAMP, Syx, and SNAP-25 families in cell lysates—Full-length HA-tagged VAMP1, 3, 7, 8, Syx 1B, SNAP-25 and Myc-tagged Sec22b, and Ykt6 and Syx 1A, Syx 2, Syx 3, Syx 4, SNAP-23 and SNAP-29 in syn-lox vector without any tag were transfected into HEK293T cells using PolyJet transfection reagents (SignaGen, MD) following the manufacturer's instructions. Cell lysates were harvested 48 h later in RIPA buffer (50 mM Tris, 1% NP40, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 400 μl per 10-cm dish) plus a protease inhibitor cocktail (Sigma-Aldrich). Cell lysates (250 μl) were incubated with X-LC (2 μM) for 1 hour at 37 °C. Samples were then analyzed by immunoblot.

Identification of cleavage sites in VAMPs by LC-MS/MS—Samples were analyzed at Taplin Biological Mass Spectrometry Core Facility at Harvard Medical School. For VAMP2, whole-protein samples were loaded onto a 100-μm internal diameter C18 reverse-

phase HPLC column packed with 3cm of beads off-line using a pressure cell. The column was re-attached to an Accela 600 Pump (Thermo Fisher Scientific). A rapid gradient of increasing acetonitrile was used to elute the protein/peptide from the HPLC column. As peptides eluted, they were subjected to electrospray ionization and then placed into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer to acquire a high-resolution FTMS scan at 60,000 resolution, a second scan at low resolution in the ion trap, and a final scan to perform data-dependent MS/MS. The charge state envelopes were de-convoluted manually to obtain mono-isotopic masses when possible or average masses for the proteins. Peptide and protein identity were determined by matching protein databases with the acquired fragmentation pattern using the software program Sequest (Thermo Fisher Scientific). All databases include a reversed version of all the sequences, and the data were filtered to 1–2% percent peptide false-discovery rate.

For Syx 1B, Syx 4, SNAP-23, and SNAP-25, samples were first separated on SDS-PAGE. Proteins bands were excised and cut into approximately 1 mm3 pieces. Gel pieces were incubated with 50 mM ammonium bicarbonate solution containing 12.5 ng/μl modified sequencing-grade chymotrypsin or trypsin (Roche Diagnostics). Samples were digested overnight at room temperature. Peptides were then extracted and separated with reverse-phase HPLC. As peptides were eluted they were subjected to electrospray ionization and transferred into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher Scientific). Eluted peptides were detected, isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide.

Immunoprecipitation of the cleaved SNAP-25 fragments—HEK293T cells were transfected rat SNAP-25 with HA-tag at the N-terminus. Cell lysates were collected after 24 h and incubated with En-LC (at the final concentration of 2 μM) at 37°C for 1 h. SNAP-25 antibody (Abcam, ab5666) was added and incubated at 4°C overnight. The mixtures were then incubated at 4°C for 7 h with Protein G agarose beads. Afterwards, Protein G beads were pelleted, washed three times, and SDS loading buffer was added. Samples were heated at 55°C for 10 min and analyzed by SDS-PAGE gels. The bands corresponding to full-length SNAP-25 and the cleaved product were cut and analyzed by Mass Spectrometry. To purify cleaved SNAP-25 from neurons, rat cortical neurons (14 days in vitro) were exposed to En-LC-HN ligated with BoNT/A-HC (0.3 μl ligation mixture) for 24 h. Neuron lysates were collected and subjected to immunoprecipitation assays and mass spectrometry analysis as described above for HEK293T cell lysates.

Neuron culture and immunoblot analysis—Primary rat cortical neurons were prepared from E18–19 embryos using a papain dissociation kit (Worthington Biochemical), as described previously (Peng et al., 2011). Pups were collected from pregnant Sprague Dawley rats (CD IGS) (Charles River, Cambridge). Neurons (Day 12) were exposed to BoNT/En fragments or sortase ligation mixture added to culture medium for 12 h. Neurons were then lysed with RIPA buffer (50 mM Tris, 1% NP40, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS) plus a protease inhibitor cocktail (Sigma-Aldrich). Lysates were centrifuged for 10 min at maximum speed using a microcentrifuge at 4°C. Supernatants were subjected to SDS-PAGE and immunoblot analysis.

Dot blot assay—BoNTs (0.2 μg in 1 μl) were spotted onto nitrocellulose membranes and dried (10 minutes at RT). The membranes were blocked with 5% milk in TBST (TBS plus 0.05% Tween20) for 30 min and then incubated with appropriate antisera (1:500 dilution) for 30 min. The membranes were then washed three times with TBST and incubated with HRP (horseradish peroxidase)-conjugated secondary antibodies for 30 min, washed three more times with TBST, and analyzed by the ECL method. We note that the BoNT/En sample was composed of En-LC-HN and GST-En-HC at a 1:1 ratio.

Sortase-mediated ligation—GST-En-HC or GST-A-HC was cleaved overnight at 4°C by thrombin before being added into the ligation reaction mixture. The ligation reaction was set up in 50 μl TBS buffer with En-LC-HN (5 μM) pre-treated with or without thrombin, En-HC (15 μM) or A-HC (15 μM), Ca2+ (10 mM), and sortase (2 μM), for 40 min at RT.

DAS assay—En-FL and En-A chimeric toxin (En-LC-HN-A-HC) were generated by sortase-mediated ligation. Mice (CD-1 strain, 21–25g, n = 4 mice) were anesthetized with isoflurane (3–4%) and injected with En-FL (1 μg) and En-A (1 ng) using a 30-gauge needle attached to a sterile Hamilton syringe into the gastrocnemius muscles of the right hind limb. Muscle paralysis and the spread of hind paw in the startle response were examined 24 h after injection.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Edwin R. Chapman (Univ. of Wisconsin – Madison), Shashi Sharma (FDA), Chuan Hu (Univ. of Louisville), Bradley L. Pentelute (MIT), and Jesse C. Hay (University of Montana) for providing antibodies, cDNA, and other critical reagents; Daria Van Tyne and Katarina Schaufler for the collection of enterococcal isolates; Anthony Gaca for MinION sequencing, Ashlee Earl for bioinformatic analyses, and Linda Henriksson for assisting the production of BoNT/X polyclonal antibody.

This study was partially supported by National Institute of Health (NIH) grants (NS080833 and AI132387) and a grant (FunGCAT) from IARPA (The Intelligence Advanced Research Projects Activity) to M.D.; NIH/NIAID grants (AI072360, AI108710, and the Harvard-wide Antibiotic Resistance Program, AI083214) to M.S.G.; the Swedish Research Council (2014-5667), the Wenner-Gren Foundation, and the Swedish Cancer Society (to P.S.). J.A.S. was supported by a fellowship from NIH (F32GM121005). A.C.D. acknowledges funding from the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant), and an Ontario Early Researcher Award. M.D. acknowledges support by the NIH-funded Harvard Digestive Disease Center (P30DK034854) and Boston Children's Hospital Intellectual and Developmental Disabilities Research Center (P30HD18655). M.D. holds the Investigator in the Pathogenesis of Infectious Disease award from the Burroughs Wellcome Fund.

References

Abdel-Moein KA, Hamza DA. Occurrence of human pathogenic Clostridium botulinum among healthy dairy animals: an emerging public health hazard. Pathog Glob Health. 2016; 110:25–29. [PubMed: 27077311]

Aoki KR. A comparison of the safety margins of botulinum neurotoxin serotypes A, B, and F in mice. Toxicon. 2001; 39:1815–1820. [PubMed: 11600142]

Arias CA, Murray BE. The rise of the Enterococcus: beyond vancomycin resistance. Nat Rev Microbiol. 2012; 10:266–278. [PubMed: 22421879]

Arnon SS, Schechter R, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Hauer J, Layton M, et al. Botulinum toxin as a biological weapon: medical and public health management. Jama. 2001; 285:1059–1070. [PubMed: 11209178]

Barash JR, Arnon SS. A novel strain of Clostridium botulinum that produces type B and type H botulinum toxins. J Infect Dis. 2014; 209:183–191. [PubMed: 24106296]

Chaddock JA, Herbert MH, Ling RJ, Alexander FC, Fooks SJ, Revell DF, Quinn CP, Shone CC, Foster KA. Expression and purification of catalytically active, non-toxic endopeptidase derivatives of Clostridium botulinum toxin type A. Protein Expr Purif. 2002; 25:219–228. [PubMed: 12135553]

Courvalin P. Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. Antimicrob Agents Chemother. 1994; 38:1447–1451. [PubMed: 7979269]

Dover N, Barash JR, Hill KK, Xie G, Arnon SS. Molecular characterization of a novel botulinum neurotoxin type H gene. J Infect Dis. 2014; 209:192–202. [PubMed: 24106295]

Foran P, Lawrence GW, Shone CC, Foster KA, Dolly JO. Botulinum neurotoxin C1 cleaves both syntaxin and SNAP-25 in intact and permeabilized chromaffin cells: correlation with its blockade of catecholamine release. Biochemistry. 1996; 35:2630–2636. [PubMed: 8611567]

Freitas AR, Tedim AP, Francia MV, Jensen LB, Novais C, Peixe L, Sanchez-Valenzuela A, Sundsfjord A, Hegstad K, Werner G, et al. Multilevel population genetic analysis of vanA and vanB Enterococcus faecium causing nosocomial outbreaks in 27 countries (1986–2012). J Antimicrob Chemother. 2016; 71:3351–3366. [PubMed: 27530756]

Gilmore MS, Lebreton F, van Schaik W. Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. Curr Opin Microbiol. 2013; 16:10–16. [PubMed: 23395351]

Gu S, Rumpel S, Zhou J, Strotmeier J, Bigalke H, Perry K, Shoemaker CB, Rummel A, Jin R. Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. Science. 2012; 335:977–981. [PubMed: 22363010]

Hill KK, Smith TJ, Helma CH, Ticknor LO, Foley BT, Svensson RT, Brown JL, Johnson EA, Smith LA, Okinaka RT, et al. Genetic diversity among Botulinum Neurotoxin-producing clostridial strains. J Bacteriol. 2007; 189:818–832. [PubMed: 17114256]

Hill KK, Xie G, Foley BT, Smith TJ. Genetic diversity within the botulinum neurotoxin-producing bacteria and their neurotoxins. Toxicon. 2015; 107:2–8. [PubMed: 26368006]

Jahn R, Scheller RH. SNAREs--engines for membrane fusion. Nat Rev Mol Cell Biol. 2006; 7:631–643. [PubMed: 16912714]

Kalb SR, Baudys J, Raphael BH, Dykes JK, Luquez C, Maslanka SE, Barr JR. Functional characterization of botulinum neurotoxin serotype H as a hybrid of known serotypes F and A (BoNT F/A). Anal Chem. 2015; 87:3911–3917. [PubMed: 25731972]

Laukova A, Koniarova I. Survey of urease activity in ruminal bacteria isolated from domestic and wild ruminants. Microbios. 1995; 84:7–11. [PubMed: 8569526]

Lebreton F, Manson AL, Saavedra JT, Straub TJ, Earl AM, Gilmore MS. Tracing the Enterococci from Paleozoic Origins to the Hospital. Cell. 2017; 169:849–861. e813. [PubMed: 28502769]

Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, Corander J, Cheng L, Saif S, Young S, et al. Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. MBio. 2013:4.

Lee K, Zhong X, Gu S, Kruel AM, Dorner MB, Perry K, Rummel A, Dong M, Jin R. Molecular basis for disruption of E-cadherin adhesion by botulinum neurotoxin A complex. Science. 2014; 344:1405–1410. [PubMed: 24948737]

Mansfield MJ, Adams JB, Doxey AC. Botulinum neurotoxin homologs in non-Clostridium species. FEBS Lett. 2015; 589:342–348. [PubMed: 25541486]

Maslanka SE, Luquez C, Dykes JK, Tepp WH, Pier CL, Pellett S, Raphael BH, Kalb SR, Barr JR, Rao A, et al. 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. 2015; 213:379–385. [PubMed: 26068781]

McCluskey AJ, Collier RJ. Receptor-directed chimeric toxins created by sortase-mediated protein fusion. Mol Cancer Ther. 2013; 12:2273–2281. [PubMed: 23945077]

Montal M. Botulinum neurotoxin: a marvel of protein design. Annu Rev Biochem. 2010; 79:591–617. [PubMed: 20233039]

Montecucco C, Molgo J. Botulinal neurotoxins: revival of an old killer. Curr Opin Pharmacol. 2005; 5:274–279. [PubMed: 15907915]

Montecucco C, Rasotto MB. On botulinum neurotoxin variability. MBio. 2015; 6:e02131–14. [PubMed: 25564463]

Peng L, Tepp WH, Johnson EA, Dong M. Botulinum neurotoxin D uses synaptic vesicle protein SV2 and gangliosides as receptors. PLoS Pathog. 2011; 7:e1002008. [PubMed: 21483489]

Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Sortagging: a versatile method for protein labeling. Nat Chem Biol. 2007; 3:707–708. [PubMed: 17891153]

Rossetto O, Pirazzini M, Montecucco C. Botulinum neurotoxins: genetic, structural and mechanistic insights. Nat Rev Microbiol. 2014; 12:535–549. [PubMed: 24975322]

Rummel A, Mahrhold S, Bigalke H, Binz T. The HCC-domain of botulinum neurotoxins A and B exhibits a singular ganglioside binding site displaying serotype specific carbohydrate interaction. Mol Microbiol. 2004; 51:631–643. [PubMed: 14731268]

Schiavo G, Benfenati F, Poulain B, Rossetto O, Polverino de Laureto P, DasGupta BR, Montecucco C. Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. Nature. 1992; 359:832–835. [PubMed: 1331807]

Schiavo G, Matteoli M, Montecucco C. Neurotoxins affecting neuroexocytosis. Physiol Rev. 2000; 80:717–766. [PubMed: 10747206]

Schloissnig S, Arumugam M, Sunagawa S, Mitreva M, Tap J, Zhu A, Waller A, Mende DR, Kultima JR, Martin J, et al. Genomic variation landscape of the human gut microbiome. Nature. 2013; 493:45–50. [PubMed: 23222524]

Sebaihia M, Peck MW, Minton NP, Thomson NR, Holden MT, Mitchell WJ, Carter AT, Bentley SD, Mason DR, Crossman L, et al. Genome sequence of a proteolytic (Group I) Clostridium botulinum strain Hall A and comparative analysis of the clostridial genomes. Genome Res. 2007; 17:1082–1092. [PubMed: 17519437]

Sudhof TC, Rothman JE. Membrane fusion: grappling with SNARE and SM proteins. Science. 2009; 323:474–477. [PubMed: 19164740]

Sugawara Y, Matsumura T, Takegahara Y, Jin Y, Tsukasaki Y, Takeichi M, Fujinaga Y. Botulinum hemagglutinin disrupts the intercellular epithelial barrier by directly binding E-cadherin. J Cell Biol. 2010; 189:691–700. [PubMed: 20457762]

Van Tyne D, Gilmore MS. Friend turned foe: evolution of enterococcal virulence and antibiotic resistance. Annu Rev Microbiol. 2014; 68:337–356. [PubMed: 25002090]

Zhang S, Masuyer G, Zhang J, Shen Y, Lundin D, Henriksson L, Miyashita SI, Martinez-Carranza M, Dong M, Stenmark P. Identification and characterization of a novel botulinum neurotoxin. Nat Commun. 2017; 8:14130. [PubMed: 28770820]

Zornetta I, Azarnia Tehran D, Arrigoni G, Anniballi F, Bano L, Leka O, Zanotti G, Binz T, Montecucco C. The first non Clostridial botulinum-like toxin cleaves VAMP within the juxtamembrane domain. Sci Rep. 2016; 6:30257. [PubMed: 27443638]

Highlights

- A Botulinum neurotoxin-like toxin (BoNT/En) identified in a commensal *E. faecium* strain
- The gene cluster encoding BoNT/En is located on a conjugative plasmid
- BoNT/En cleaves both VAMP2 and SNAP-25 required for synaptic transmission in neurons
- The cleavage sites on VAMP2 and SNAP25 are distinct from the sites for other known BoNTs

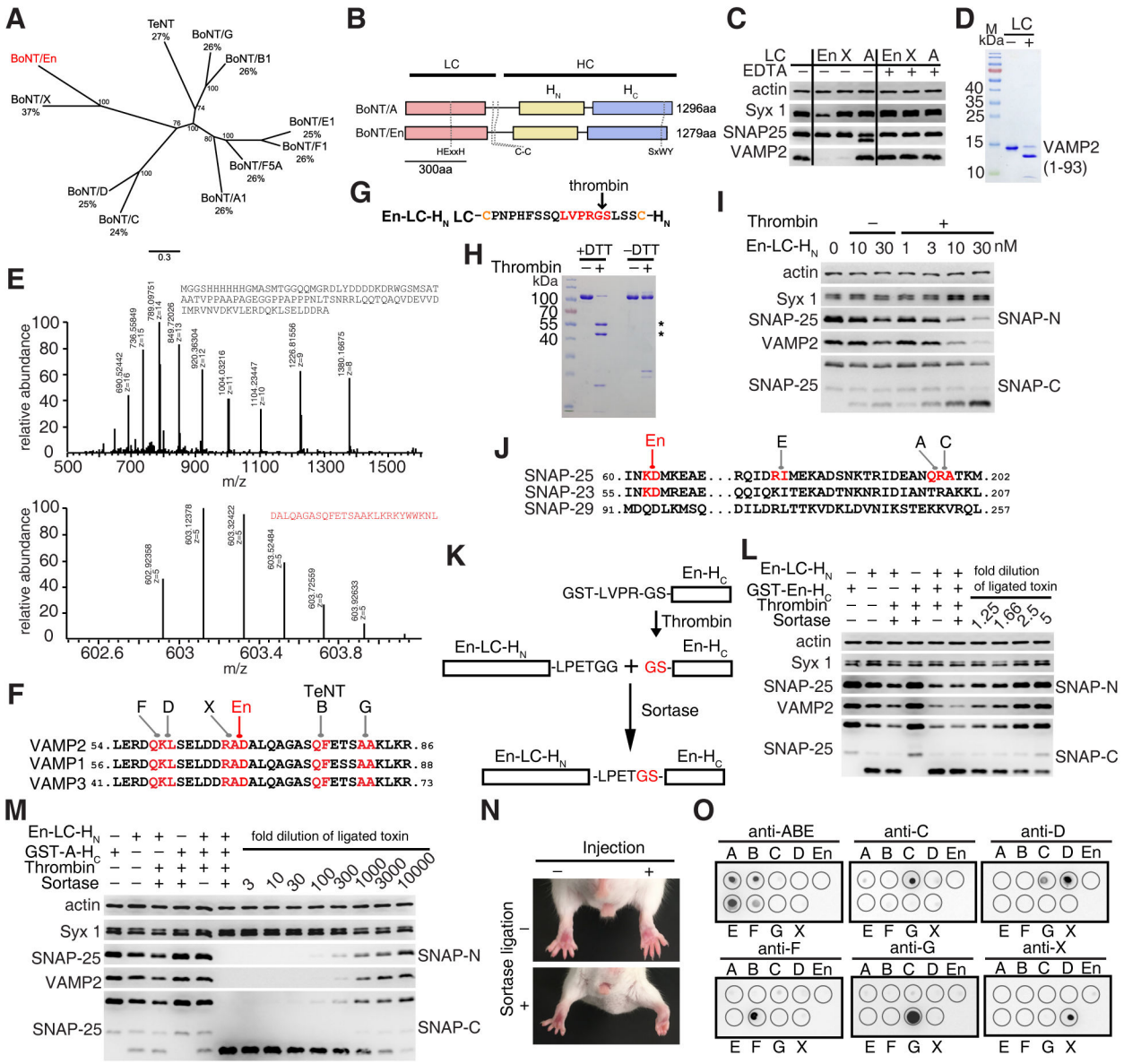


Figure 1. BoNT/En is a unique BoNT serotype and cleaves VAMP2 and SNAP-25 in neurons

(A) The maximum likelihood phylogeny of BoNT serotypes demonstrates that BoNT/En forms a distinct lineage, grouping most closely with BoNT/X. The percentages of protein sequence identity for each toxin with BoNT/En are noted. The scale bar represents mean number of amino acid substitutions per site.

(B) A schematic drawing of the three domains of BoNT/En in comparison with BoNT/A.

(C) En-LC (2 μM), with or without EDTA, was incubated with BDE (1 h). Immunoblot analysis was carried out to detect Syx 1, SNAP-25, and VAMP2. A-LC and X-LC were analyzed in parallel. Cleavage of VAMP2 by X-LC results in loss of immunoblot signals, while cleavage of SNAP-25 by A-LC generates a smaller fragment of SNAP-25 that can still be detected on immunoblot. Incubation with En-LC resulted in loss of VAMP2 immunoblot

signals. It also reduced the signal of Syx 1. EDTA blocked the activity of En-, A-, and X-LCs.

(D) His6-tagged VAMP2 (residues 1–93) was incubated with En-LC (0.1 μM, 1 h). Samples were analyzed by SDS-PAGE and Coomassie Blue staining. En-LC cleaved VAMP2 (1–93) into two smaller fragments.

(E) VAMP2 (1–93) was incubated with En-LC. Whole-protein samples were then analyzed by mass spectrometry (LC-MS/MS) to determine the precise molecular weight of cleaved fragments. The mass spectrometry data for the two cleavage products are shown, with mass-to-charge ratio (m/z) noted for each signal. The molecular weight is deducted by multiplying m with z, followed by subtracting z.

(F) The cleavage sites for BoNT/B, D, F, G, X, En, and tetanus neurotoxin (TeNT) are marked in red in VAMP1/2/3.

(G) To achieve better proteolytic activation of the LC-H_N domain of BoNT/En, a thrombin cleavage site was inserted into its linker region between two cysteines.

(H) LC-H_N of BoNT/En was treated with thrombin and then analyzed by SDS-PAGE and Coomassie Blue staining, with or without DTT. Asterisks marked the separated LC and H_N.

(I) Cultured rat cortical neurons were exposed to LC-H_N of BoNT/En for 12 h. Cell lysates were harvested and immunoblot analysis carried out to detect Syx 1, SNAP-25, and VAMP2. Actin served as a loading control. Thrombin-activated En-LC-H_N cleaved SNAP-25 and VAMP2 more efficiently than non-activated protein. SNAP-25 was detected with CI71.1 that recognizes the N-terminal region (SNAP-N), which did not detect any cleavage products. To further confirm the cleavage of SNAP-25, a second antibody that recognizes the C-terminus of SNAP-25 (residues 195–206, SNAP-C) was used to detect the C-terminal cleavage fragment (lower panel).

(J) Sequence alignment of SNAP-25, SNAP-23 and SNAP-29 around the cleavage sites (marked in red) for BoNT/En, A, C, and E.

(K) A schematic drawing of the synthesis of full-length BoNT/En using sortase ligation method.

(L) Cultured rat cortical neurons exposed to the same amount (5 μl) of sortase ligation mixture or indicated control components for 12 h in medium. Cell lysates were analyzed by immunoblot. Ligating En-LC-H_N and En-H_C together by sortase enhanced cleavage of VAMP2 and SNAP-25 only slightly compared to En-LC-H_N/En-H_C mixture without sortase and En-LC-H_N alone. Titrating the sortase-ligated mixture by only 1:1.25 abolished any enhancement effect.

(M) Rat cortical neurons were exposed to indicated control components or sortase-ligated En-A mixture (1 μl) for 12 h in media. Cell lysates were analyzed by immunoblot. Ligated En-A resulted in a greatly enhanced cleavage of VAMP2 and SNAP-25 by > 1,000 folds compared to the mixture of En-LC-H_N and A-H_C without sortase.

(N) En-A linked by sortase reaction (1 ng) was injected into the gastrocnemius muscles of the right hind limb of mice (n = 4 mice). The injected limb developed typical flaccid paralysis, and the toes failed to spread within 12 h.

(O) BoNT/A-G, BoNT/X, and BoNT/En were subjected to dot blot analysis (0.2 μg per toxin, spotted on nitrocellulose membranes), using four horse antisera (trivalent anti-BoNT/A, B, and E, anti-BoNT/C, anti-BoNT/DC, and anti-BoNT/F), two goat antisera (anti-BoNT/G and anti-BoNT/D), and a rabbit polyclonal antisera against BoNT/X. BoNT/En is

composed of purified En-LC-H_N and En-H_C at 1:1 ratio. These antisera recognized their corresponding target toxins, yet none of them recognized BoNT/En.

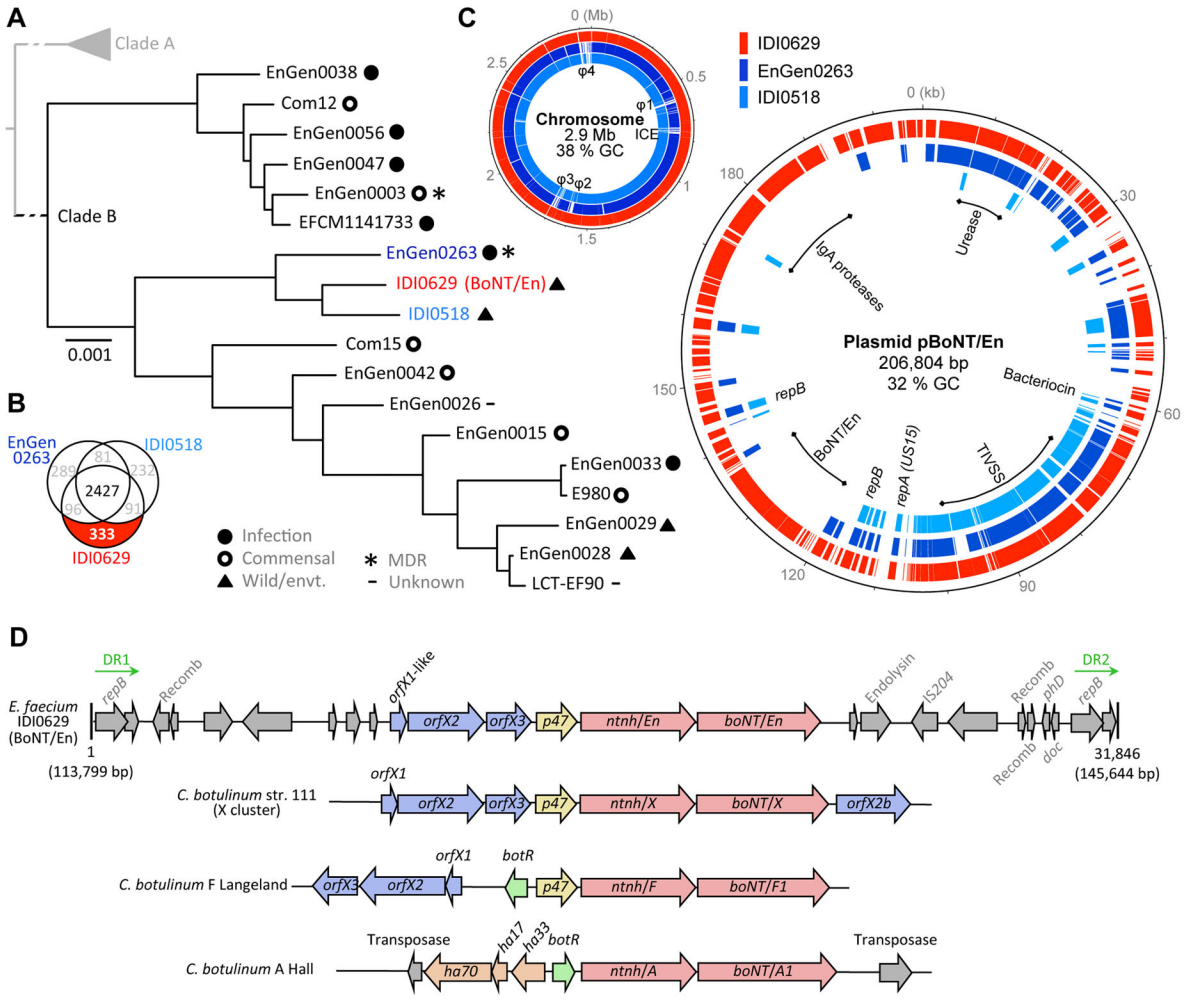


Figure 2. BoNT/En gene cluster is located on a conjugative plasmid in *E. faecium*

(A) RAxML SNP-based tree based on the concatenated alignments of DNA sequences of 1,824 single-copy core genes in 18 Clade B *E. faecium* genomes. Bootstrapping was performed with 1,000 replicates. Branch lengths reflect the number of substitution per site. The origins and characteristics of the strains are indicated. Additional strain metadata are detailed in Table S1. Strain IDI0629 (red) and its closest neighbors EnGen0263 (dark blue) and IDI0518 (light blue) are highlighted.

(B) Venn diagram illustrating the gene content comparison between the genomes of strains IDI0629, EnGen0263 and IDI0518. The 333 genes found uniquely in strain IDI0629 are highlighted.

(C) *Left panel:* comparisons of the draft chromosome of IDI0629 with the strains EnGen0263 and IDI0518. Scaffolds of IDI0629 are ordered arbitrary from the largest to the smallest. Predicted phages (φ1–4) and integrated conjugative element (ICE) in the genome of IDI0629 are indicated. *Right panel:* comparisons of pBoNT/En with the *rep*US15 plasmids found in EnGen0263 and IDI0518. Outer circle, genes of pBoNT/En are painted in red; Middle circle, homologous genes of EnGen0263 are painted in dark blue; Inner circle, homologous genes of IDI0518 are painted in light blue. Origins of replication (*rep* genes)

and other features of interest (BoNT/En; TIVSS, Type IV secretion system and pilus assembly; IgA, putative immunoglobulin A protease; Urease, urea catabolism operon) are indicated.

(D) Schematic representation of the BoNT/En gene cluster (31.8 kb), which is flanked by two direct repeats (DR1 and DR2), in comparison with other typical BoNT gene clusters. The BoNT/En cluster contains a neighboring *ntnh*-like gene as well as upstream *orfX*-like genes. Length and reference plasmid pBoNT/En coordinates are indicated. Recomb.: putative recombinase; IS204: insertion element 204; PhD-doc: a putative toxin antitoxin system. Genes with no labels represent open reading frames coding for proteins of unknown function.