Recombinant derivatives of botulinum neurotoxin A engineered for trafficking studies and neuronal delivery

Philip A. Band a, Steven Blais b, Thomas A. Neubert a,b, Timothy J. Cardozo a, Konstantin Ichtchenko a,*

a Department of Pharmacology, New York University School of Medicine, New York, NY, United States
b Kimmel Center for Biology and Medicine at the Skirball Institute, New York University School of Medicine, New York, NY, United States

A R T I C L E   I N F O
Article history:
Received 6 October 2009
and in revised form 4 December 2009
Available online 4 January 2010

Keywords:
Botulinum neurotoxin A
Clostridium botulinum
Baculovirus
Protease
Sfp phosphopantetheinyl transferase

A B S T R A C T
Work from multiple laboratories has clarified how the structural domains of botulinum neurotoxin A (BoNT/A) disable neuronal exocytosis, but important questions remain unanswered. Because BoNT/A intoxication disables its own uptake, light chain (LC) does not accumulate in neurons at detectable levels. We have therefore designed, expressed and purified a series of BoNT/A atoxic derivatives (ad) that retain the wild type features required for native trafficking. BoNT/A1ad and BoNT/A1adf are full length derivatives rendered atoxic through double point mutations in the LC protease (E224 > A; Y366 > A). ΔLC-peptide-BoNT/Atev and ΔLC-GFP-BoNT/Atev are derivatives wherein the catalytic portion of the LC is replaced with a short peptide or with GFP plus the peptide. In all four derivatives, we have fused the S6 peptide sequence GDSLWLLRLLN to the N-terminus of the proteins to enable site-specific attachment of cargo using Sfp phosphopantetheinyl transferase. Cargo can be attached in a manner that provides a homogeneous derivative population rather than a polydisperse mixture of singly and multiply-labeled molecular species. All four derivatives contain an introduced cleavage site for conversion into disulfide-bonded heterodimers. These constructs were expressed in a baculovirus system and the proteins were secreted into culture medium and purified to homogeneity in yields ranging from 1 to 30 mg per liter. These derivatives provide unique tools to study toxin trafficking in vivo, and to assess how the structure of cargo linked to the heavy chain (HC) influences delivery to the neuronal cytosol. Moreover, they create the potential to engineer BoNT-based molecular vehicles that can target therapeutic agents to the neuronal cytoplasm.

© 2009 Elsevier Inc. All rights reserved.

A B S T R A C T

Introduction

Botulinum neurotoxins (BoNTs) are a family of structurally similar proteins that cause peripheral neuromuscular blockade and respiratory paralysis, with an extremely low LD50 (1–50 ng/kg) [1]. There are 7 major serotypes (A–G) and multiple subtypes [2], but all have common structural features and a similar mechanism of action. BoNTs are synthesized as single chain propeptides (Mr approximately 150,000; approximately 1300 amino acids). The majority are activated by proteolytic cleavage to generate a disulfide-bonded heterodimer consisting of light (approximately 50 kDa) and heavy (approximately 100 kDa) chains (LC and HC). The BoNT/A heterodimer contains three functional domains. Toxicity is associated with metalloprotease activity confined to the LC; neuron binding activity is associated with the C-terminal half of the HC (HC C); and translocation activity responsible for delivering the LC protease to the neuronal cytosol is associated with the N-terminal half of the HC (HC N) [3,4]. Although significant evidence supports a multi-step mechanism culminating in LC delivery into the neuronal cytosol, currently available methodologies have not permitted direct detection of LC in the neuronal cytosol. In the case of wt BoNTs, neuron intoxication disables further toxin uptake, and consequently LC does not accumulate to levels allowing direct visualization. Nonetheless, researchers from many laboratories, using indirect approaches, have described BoNT trafficking pathways, and have deduced how the different domains of BoNT polypeptides contribute to its unique targeting mechanism. The HC and LC of wt BoNTs can be separated, individually radiolabeled, reconstructed into the disulfide-bonded heterodimer, and subsequently used to study intracellular trafficking in neurons [5]. However because the LC–HC separation and reconstitution process results in loss of ~90% of the toxin’s biological activity, it is difficult to conclude with confidence that the tracer localization corresponds to that of the...
biologically active fraction (~10% of the radiolabeled preparation). Investigators attempting to reconstitute HC with recombinant atoxic LC likewise found that the reconstituted BoNT heterodimer had a severely reduced ability to transport LC into the neuronal cytosol [6]. Strategies to take advantage of BoNT trafficking for carrying cargo into neurons have also proven difficult to develop. Isolated wt HC has been chemically coupled to dextran, but the internalized HC adduct remained localized to the endosomal compartment and no fluorescent-labeled dextran was delivered to the neuronal cytosol [7]. These studies illustrate the difficulty of reengineering separated HCs and LCs and reconstituting native configuration including disulfide bonds. Moreover, they illustrate that chemical methods to label or attach cargo to BoNT are insufficiently selective, can produce a heterogeneous population of derivatives, and are generally too harsh to retain native BoNT activity. Such problems limit the utility of chemically labeled BoNTs as probes for definitively demonstrating BoNT trafficking pathways, or as carriers for efficiently delivering therapeutics to the neuronal cytosol. Because of these limitations, we have focused on developing genetic constructs and expression systems that enable production of full-length, disulfide-bonded, atoxic, recombinant BoNT derivatives, that retain the key structural features required for native toxin trafficking. The large size, multi-domain structure, critical disulfide bonding and mechanical sensitivity [8] of the BoNT heterodimer make it challenging to express recombinant full-length BoNT proteins that retain native configuration and trafficking. Several laboratories have reported expression of recombinant, full-length BoNTs in Escherichia coli. Kiyatkin et al. reported the expression of BoNT/C in E. coli, with three inactivating point mutations (H229 > G; E230 > T; H233 > N) in the LC protease. This atoxic BoNT/C derivative was competent for transport across epithelia and was immunogenic when orally administered [9]. Rummel et al. described the expression of full-length single chain BoNT/C, D, B and A in E. coli, either as the wt or with the LC protease inactivated by point mutation [10–12]. Pier et al. described expression of recombinant, full-length BoNT/A holotoxin in the non-toxic Clostridium botulinum strain LNT01. The mutations R664 > A and Y366 > F were introduced into the LC (BoNT/ARY), rendering the protein unable to cleave the substrate SNAP 25 in vitro. Immunization with this holotoxin effectively protected mice against lethal BoNT/A challenge [13]. Webb et al. described expression of synthetic BoNT/A gene constructs optimized for codon bias in Pichia pastoris with the mutations H229 > A; E232 > A; H233 > A to render the toxin inactive (cBoNT/A HP), which also provided excellent protective immunity [14]. Although these latter two reports are encouraging, neither provides data on the utility of these novel recombinant BoNT derivatives for trafficking studies or for neuronal delivery. The efforts to express recombinant BoNTs have succeeded in producing effective immunogens, which in some cases are competent for epithelial transcytosis, but none have reported the production of recombinant proteins with the structural features required for targeting the neuronal cytosol with the efficiency of wt toxins. In the work reported here, we describe the design, expression and purification of recombinant, full length, atoxic BoNT/A heterodimers that retain key structural elements required for native BoNT trafficking. Moreover, our constructs are designed to also contain a short peptide sequence that enables site selective attachment of cargo molecules under physiological conditions. All of the BoNT/A derivatives reported here were designed to be atoxic, in that they are either expressed with the LC inactivated by a double mutation or with the entire catalytic region of the LC removed, enabling their uptake and accumulation into neurons at levels allowing direct detection, trafficking studies and cargo delivery.

Methods

Creation of the construct encoding BoNT/Aad<sup>α</sup> in plitmus vector

Modification of plitmus28i by replacement of the existent polylinker with a custom polylinker

A vector with a custom polylinker derived from Litmus 28i (New England Biolabs, Cat # N3528S, 2823 bp) was used for subcloning the full-length BoNT/Aad<sup>α</sup>. This derivative of Litmus 28i (plitmus28C1) was created by restriction digestion of Litmus 28i with Bgl II and Aat II followed by dephosphorylation. In all subcloning procedures described below, vectors were digested with the same set of restriction endonucleases as were the DNA fragments for subcloning, followed by vector dephosphorylation, unless indicated otherwise. Annealed phosphorylated oligonucleotides C1-1S and C1-1A (Table 1) were ligated into digested Litmus 28i, resulting in the intermediate vector, plitmus28C1INT (2842 bp). Vector plitmus28C1INT was digested with Kst I and Stu I and dephosphorylated. Annealed phosphorylated oligonucleotides C1-2S and C1-2A were ligated into digested plitmus28C1INT, resulting in the vector plitmus28C1 (2890 bp). The sequence of this vector was deposited in Genbank under accession number GQ855199.

Cloning of BoNT/Aad<sup>α</sup> into plitmus28C1

In the first cloning step, plitmus28C1 was digested with Hind III and Not I and dephosphorylated. Annealed phosphorylated oligonucleotides CP15 and CP1A were ligated into digested plitmus28C1, resulting in vector plitSB3A1 (2968 bp). Vector plitSB3A1 carries several unique S– restriction sites, including Bgl II, Bbs I, Rsr II and Bst B1 which provide considerable flexibility for subsequent subcloning of the full-length construct into expression vectors. This polylinker was followed by the enhancer sequence AAACCTCTAAAAACCCGCACCC, followed by the honeybee melittin signal peptide sequence.

In the next step, plitSB3A1 was digested with Not I and Mlu I and dephosphorylated. Annealed phosphorylated oligonucleotides CP25 and CP2A were ligated into digested plitSB3A1, resulting in vector plitSB3A2 (2987 bp), which carried a 10-His affinity tag downstream of the honeybee melittin signal peptide sequence. Next, plitSB3A2 was digested with Mlu I and Nhe I and dephosphorylated. Annealed phosphorylated oligonucleotides CP35 and CP3A were ligated into digested plitSB3A2 to generate plitSB3A3 (3056 bp), which carries the tobacco etch virus (TEV) protease cleavage site, followed by the 56 peptide tag sequence downstream of the 10-His affinity tag sequence.

In the next three steps a full-length BoNT/Aad<sup>α</sup> vector optimized for expression in two hosts, Spodoptera frugiperda and E. coli, was progressively constructed in the plitSB3A3 vector. DNA fragments for constructing the full-length BoNT/Aad<sup>α</sup> gene were provided by a commercial supplier (Genscript) in the form of three contiguous DNA segments, as subclones in pUC57. Numbers in the text below correspond to amino acid numbering in wt BoNT/A [15]. First, a 1347 bp DNA fragment encoding the sequence P<sub>2</sub>–L<sub>442</sub> was isolated from pUC57–Baa11 by restriction digestion with BssH II and Adh I and subcloned into plitSB3A4 to generate the 4355 b.p. vector, plitSB3A4. Then, a 1233 bp DNA fragment encoding the sequence D<sub>443</sub>–L<sub>549</sub> was isolated from pUC57–Baa12 by digest with Xba I and EcoR V and subcloned in plitSB3A4 to generate the 5570 b.p. vector, plitSB3A5. Finally, a 1401 bp DNA fragment encoding the sequence P<sub>560</sub>−L<sub>1296</sub> followed by a short linker, the Streptag II sequence, and a triplet of termination codons in three reading frames was isolated from pUC57–Baa13 by digest with EcoR V and Xho I, and subcloned into plitSB3A5 to generate the 6925 b.p. vector, plitSB3A5. The
sequence of pLitSB3A has been deposited in GenBank under accession number GQ855200.

Construct for expression of BoNT/Aad<sup>ES</sup>

The plasmid pLitSB3A was digested with the restriction endonucleases Rsrl II and Xho I. The 4190 bp DNA fragment encoding full-length BoNT/Aad<sup>ES</sup> was isolated from an agarose gel, purified and ligated into the baculovirus transposition vector, pFastBac<sup>TM</sup> (Invitrogen, Cat. # 10360-014, 4776 bp) to generate the 8968 bp vector, pFB1SB3AEGK. The DNA sequence of pFB1SB3AEGK with annotations has been deposited in GenBank under accession number GQ855201. Features of the 1372 aa protein, translated from the open reading frame of the insert DNA (BoNT/Aad<sup>ES</sup>), are shown in Table 1 (this figure represent an alignment of all recombinant proteins expressed and described in the current work in comparison to wt BoNT/A).

Construct for expression of BoNT/Aad<sup>EV</sup>

In this construct DNA encoding the enterokinase cleavage site positioned between LC and HC was replaced with a sequence encoding a second TEV cleavage site. The 348 bp fragment containing the enterokinase cleavage site was excised from pFB1SB3AEGK with annotations has been deposited in GenBank under accession number GQ855202. Features of the 1372 aa protein along with annotations has been deposited in GenBank under accession number GQ855202.

Construct for expression of ALC-Peptide-BoNT/A<sup>EV</sup>

A DNA fragment containing the sequence encoding the ALC-peptide was obtained from the vendor (GenScript) in the form of synthetic DNA cloned into pUC57 (pUC57-Pept), pUC57-Pept was digested with BssH II and Xba I and the 200 bp DNA fragment was isolated from an agarose gel, purified and ligated into digested pFB1SB3AEGK to generate the 8971 bp vector, pFB1SB3AETEV. The DNA sequence of pFB1SB3AETEV, encoding a 1373 aa protein along with annotations has been deposited in GenBank under accession number GQ855203.

Construct for expression of ALC-GFP-BoNT/A<sup>EV</sup>

DNA encoding the GFP sequence was obtained by PCR amplification from plasmid pAcGFPI-C2 (Clontech, Cat. # 632481, 4722 bp) using primers CP7S and CP7A (Table 1) and PrimeSTAR HS DNA polymerase (Takara, Cat. # TAK R010A) in a GeneAmp PCR system 9700 (PE/Applied Biosystems). The reaction buffer and conditions were set according to the protocol provided by Takara. The PCR product was digested with restriction endonucleases Rsrl I and Xho I. 7822 bp DNA fragment was isolated from an agarose gel and purified. The vector, pFB1SBPepB2ATEV, was digested with Xho I and partially digested (short incubation time and small amount of the restriction endonuclease) with Acc65 I. A 7822 bp DNA fragment was isolated from an agarose gel, purified and dephosphorylated. The purified GFP PCR product and linearized pFB1SBPepB2ATEV were ligated, resulting in the 7822 bp vector, pFB1SGFPB2ATEV. The DNA sequence of pFB1SGFPB2ATEV encoding a 1230 aa protein along with annotations has been deposited in GenBank under accession number GQ855204.

Custom-synthetic oligonucleotides were obtained from Sigma-Aldrich. E. coli strain TOP10, used for plasmid transformation and amplification, was purchased in the form of electrocompetent cells from Invitrogen (Cat. # C404052), except as specified otherwise. Restriction endonucleases, T4 DNA ligase, and arctic shrimp alkaline phosphatase were purchased from New England Biolabs. All DNA fragments isolated from agarose gels after enzymatic treatment were purified with a Qiagen II DNA extraction kit (Qiagen, Cat. # 20051). DNA sequences of all final constructs were obtained using overlapping sets of primers at the DNA sequencing facility at NYU’s Skirball Institute of Biomolecular Medicine. The sequencing data obtained proved the constructs to be free of unexpected mutations.

Generating recombinant baculovirus

pFastBac<sup>TM</sup> constructs were used for transposition of the cloned DNA into the shuttle vector (bacmid) for use in the Bac-to-Bac<sup>®</sup> baculovirus expression system (Invitrogen, Cat. # 10361-012), according to the manufacturer’s recommendations. The recombinant bacmids were isolated and purified from harvested cells using a Qiagen Large-Construct Kit (Cat. # 12462), according to manufacturer’s protocol. Transposition of the cloned genes into recombinant bacmids was confirmed by PCR using the bacmids as templates and sets of primers specific for each cloned construct. Transfection of insect cells with the bacmid DNA and collection of P1 baculoviral stock were performed according to the manufacturer’s protocol (Invitrogen). To generate P2 high-titer stock, 1 ml of P1 stock was added to 25 ml of S9 cells grown in SF-900 II medium in a shaker flask. Cell density at the time of infection was approximately 1.5 × 10<sup>6</sup> cells/ml. Cells were incu-
The disulfide bridges between residues of the light and heavy chains are indicated by horizontal brackets. Cargo to the expressed proteins. *GFP* on a green background represents a portion of green fluorescent protein. The five proteins are aligned to illustrate homology between Amino acids on an orange background represent enterokinase recognition sequence. *S6* on a pink background identifies a peptide tag used for site-specific attachment of indicated by *StrepTag II* on a purple background. *TEV* and amino acid sequence on a yellow background represent tobacco etch virus protease recognition sequence. indicated by *SP* on blue background. Tags used for affinity chromatography are indicated: polyhistidine tag is indicated by *10 His* on a gray background; *StrepTag II* is black. Introduced mutations are shown in red. Added amino acids are shown in blue. Signal peptide required for insect cell secretion of the expressed derivatives into medium sequence identity omitted for simplicity. Numbers in the upper row correspond to amino acid residues in Cat. # CDUF002LT) and dialyzed against Talon approximately 3-fold on a Prep/Scale TFF cartridge (Millipore, GF/F filter (Schleicher&Schuell, Cat. # 1825090), concentrated filtration. Dialyzed supernatant was loaded on Talon resin chromatography loading buffer (500 mM NaCl, 25 mM TrisHCl, 20 mM imidazole, pH 8.0) and protein was eluted with 5 volumes of elution buffer (500 mM NaCl, 25 mM TrisHCl, 200 mM imidazole, pH 8.0). For *ΔC-Peptide-BoNT/Avec* and *ΔC-GFP-BoNT/Avec* derivatives, chromatography was slightly modified by inclusion of 1% Triton X-100 in the loading buffer. For these derivatives after loading and column wash with 10 volumes of loading buffer, the column was washed with 10 volumes of 20 mM imidazole wash buffer with 1% Triton X-100 followed by washing with 10 volumes of 20 mM imidazole wash buffer without detergent. Protein was eluted with 5 volumes of elution buffer containing 0.2% n-octyl-[β-D-glucopyranoside (USB, Cat. # 29836-26-8). Eluted proteins were dialyzed against 100 mM NaCl, 25 mM TrisHCl buffer, pH 8.0 either without (BoNT/Avec) and BoNT/Avec), or with *ΔC-Peptide-BoNT/Avec* and *ΔC-GFP-BoNT/Avec* 0.2% n-octyl-[β-D-glucopyranoside, and concentrated to a final total protein concentration of approximately 10 mg/ml using Millipore Amicon Ultra 30,000 MWCO centrifugal filter units (Cat. # UFC803024). Dialyzed and concentrated fractions were loaded on StrepTactin superfusion agarose (Novagen, Cat. # 71592) pre-equilibrated with 100 mM NaCl, 25 mM TrisHCl buffer, pH 8.0 with or without 0.2% n-octyl-[β-D-glucopyranoside, and washed with the same buffer. Bound protein was eluted with 3 mM D-desthiobiotin (Novagen, Cat. # 71610). Fractions containing pure protein were combined, dialyzed and concentrated using Amicon Ultra centrifugal filter units.

bated for 72 h at 28 °C in a humidified incubator. Cells were removed from high titer P2 stock by centrifugation at 100g. The titer of the P2 stock was measured by viral plaque assay. P2 titers for the constructs were approximately 1−2 × 10⁸ pfu/ml. P2 stock was used to infect Sf9 cells for expression of the proteins of interest.

**Protein expression, purification and processing**

Sf9 cells grown to a density of approximately 1.5 × 10⁶ cells/ml in a shaker flask in SF-900 II serum-free medium were infected with recombinant P2 baculovirus stock. For each recombinant protein-expression vector, the optimal multiplicity of infection (MOI) and time after infection for harvesting were determined empirically. For BoNT/Avec and BoNT/Avec optimal conditions were an MOI of approximately 0.1 and an incubation time of 60 h; for *ΔC-Peptide-BoNT/Avec* and *ΔC-GFP-BoNT/Avec* optimal conditions were an MOI of approximately 0.5 and an incubation time of 72 h. At the time of harvest, medium was separated from the cells by centrifugation at 100g, filtered through a Whatman D-glucopyranoside (USB, Cat. # 29836-26-8). Eluted proteins were dialyzed against 100 mM NaCl, 25 mM TrisHCl buffer, pH 8.0 either without (BoNT/Avec) and BoNT/Avec), or with *ΔC-Peptide-BoNT/Avec* and *ΔC-GFP-BoNT/Avec* 0.2% n-octyl-[β-D-glucopyranoside, and concentrated to a final total protein concentration of approximately 10 mg/ml using Millipore Amicon Ultra 30,000 MWCO centrifugal filter units (Cat. # UFC803024). Dialyzed and concentrated fractions were loaded on StrepTactin superfusion agarose (Novagen, Cat. # 71592) pre-equilibrated with 100 mM NaCl, 25 mM TrisHCl buffer, pH 8.0 with or without 0.2% n-octyl-[β-D-glucopyranoside, and washed with the same buffer. Bound protein was eluted with 3 mM D-desthiobiotin (Novagen, Cat. # 71610). Fractions containing pure protein were combined, dialyzed and concentrated using Amicon Ultra centrifugal filter units.

**Fig. 1.** Amino acid sequence and features of the expressed BoNT/A derivatives in comparison with wt. Spaces between arrowheads and arrow tails represent regions of sequence identity omitted for simplicity. Numbers in the upper row correspond to amino acid residues in wt BoNT/A. Residues that are identical in all proteins are shown in black. Introduced mutations are shown in red. Added amino acids are shown in blue. Signal peptide required for insect cell secretion of the expressed derivatives into medium indicated by *SP* on blue background. Tags used for affinity chromatography are indicated: polyhistidine tag is indicated by "10 His" on a gray background; *StrepTag II* is indicated by "StrepTag II" on a purple background. "TEV" and amino acid sequence on a yellow background represent tobacco etch virus protease recognition sequence. Amino acids on an orange background represent enterokinase recognition sequence. *S6* on a pink background identifies a peptide tag used for site-specific attachment of cargo to the expressed proteins. "GFP" on a green background represents a portion of green fluorescent protein. The five proteins are aligned to illustrate homology between respective structural domains. Gaps are used to facilitate the alignment. Spaces between rectangle-enclosed sequences represent sites of proteolytic cleavage. The disulfide bridges between residues of the light and heavy chains are indicated by horizontal brackets.
**Processing of BoNT/A**

The 10-His tag was removed from recombinant protein using AcTEV protease (Invitrogen, Cat. # 12575-015). The protein was incubated with AcTEV protease (1 U of enzyme per μg of protein) at 30 °C for 6 h in digestion buffer (50 mM NaCl, 0.5 mM EDTA, 3 mM glutathione, 0.3 mM oxidized glutathione, pH 8.0). Completion of the cleavage was confirmed by Western blotting with anti-His monoclonal antibodies (Santa Cruz, H-3 His probe, Cat # sc-8036 HRP). The AcTEV and potentially under-digested peptides were removed from the reaction mixture by Talon™ resin chromatography. De-tagged protein was collected in the flow-through fractions, which were analyzed, combined, dialyzed against low-salt buffer and concentrated using Amicon Ultra centrifugal filter units. Pro-peptide was converted to the heterodimer by treatment with recombinant enterokinase (rEK, Novagen, Cat. # 71537-3). The protein was incubated with Tag-off high activity rEK (0.007 U of enzyme per 1 μg of protein) at 16 °C for 16 h in digestion buffer (50 mM NaCl, 20 mM TrisHCl, 2 mM CaCl₂, pH 7.8). Approximately 96% of the protein was processed to heterodimer by this treatment as assessed by Coomassie-stained SDS–PAGE under reducing conditions. The rEK was removed from the reaction mixture by Tag-off cleavage capture kit (Novagen, Cat. # 71540-3), according to the manufacturer’s protocol. Protein was dialyzed against buffer (50 mM NaCl, 40 mM sodium phosphate, 40% glycerol, pH 7.2) and concentrated using Amicon Ultra filtration units.

**Processing of BoNT/A**

Removal of the N-terminal 10-His tag and propeptide cleavage were performed simultaneously, according to the protocol described for BoNT/Aadek™ above. ΔLC-Peptide-BoNT/A**<sup>ΔN</sup>** and ΔLC-GFP-BoNT/A**<sup>ΔN</sup>** required the presence of 0.2% n-octyl-β-D-glucopyranoside in the digestion. The detergent was removed after cleavage by dialysis and concentration on Amicon Ultra filtration units.

**Enzymatic labeling of BoNT/A**

All expressed proteins incorporated the S6 peptide GDSLS-RLLN [16] at the N-terminus of the smaller subunit of the heterodimer. Recombinant Sfp phosphopantetheinyl transferase (kindly provided as the pET29 vector with cloned C-terminally His6-tagged Sfp phosphopantetheinyl transferase from Bacillus subtilis by Dr. Jun Yin, University of Chicago) was expressed in *E. coli* strain BI21 (DE3) (Novagen, Cat. # 69450) and purified according to the protocol described in [17]. CoA 547 was purchased from Covalsy (Cat. # SV124) or from New England Biolabs (Cat. # S9349S). Prior to enzymatic labeling, the fluorescent substrate was dissolved in DMSO to a final concentration of 1 mM. The labeling reaction was optimized and performed according to the following protocol using a molar ratio of labeled protein: Sfp phosphopantetheinyl transferase: CoA 547 of 4:1:5. A pilot reaction was performed by assembling approximately 18 μg of processed BoNT/A**<sup>ΔN</sup>** in a 30 μl reaction mixture containing 1× un-supplemented Grace’s insect cell culture medium (prepared from the dry medium, Invitrogen, Cat. # 11300-043), 5 mM MgCl₂, a mixture of short peptides used to suppress the background labeling (and consisting of 150 μM Neurokinin A (Peninsula Laboratories, Cat. # 7359), 75 mM Substance P (Peninsula Laboratories, Cat. # 7451), 2 mM poly-L-lysine hydrobromide, MW 500–2000 (Sigma, Cat. # P8954)), 1 μM Sfp phosphopantetheinyl transferase and 5 μM CoA 547, pH 7.0. Fluorescent substrate was added as the final component and the reaction mixture was incubated for 15 min at 30 °C. For visualization after labeling, the reaction was stopped by the addition of 2× Laemmli SDS loading buffer followed by SDS–PAGE, transfer of the proteins to a nitrocellulose membrane, and scanning on a Typhoon scanner. For the preparative isolation of labeled proteins, after the 15 min incubation, 10 volumes of Talon™ resin loading buffer (500 mM NaCl, 25 mM TrisHCl, pH 8.0) was added to the reaction mix and the resulting solution was passed through Talon™ chelating resin. Flow-through fractions were combined and immediately concentrated using Amicon ultrafiltration units. Concentrated protein was dialyzed against glyceral-phosphate buffer (100 mM NaCl, 40 mM sodium phosphate, 40% glycerol, pH 7.2).

**Structural analysis of BoNT/A**

**In-gel tryptic digest**

Gel bands were digested in gel with sequencing grade trypsin (Promega) using a modified method of Shevchenko et al. [18] that included an additional step of dehydration and rehydration of gel pieces in 25 μl of 0.1% RapiGest SF (Waters, Milford, MA) at 37 °C for 10 min before dehydration and rehydration in buffer containing trypsin.

**MALDI-TOF analysis of BoNT/A**

Samples were desalted using C18 ZipTips (Millipore, Billerica, MA) and eluted with 50% acetonitrile, 0.1% formic acid after which 1 μl of sample was mixed with 1 μl of alpha-cyano-4-hydroxyquinoline (Agilent Technologies, Santa Clara, CA) and 1 μl of the mixture was then spotted on a steel target plate and allowed to air dry at room temperature. Analysis was performed on a Bruker Autoflex MALDI-TOF mass spectrometer (Billerica, MA) in positive ion reflectron mode using standard operating conditions.

**HPLC-Q-TOF MS/MS analysis**

A Q-TOF Premier mass spectrometer (Waters, Milford, MA) equipped with a Waters nano-ESI source coupled directly to a Nano-Acquity UPLC system (Waters) with a 100 μm x 15 cm reverse phase column (BEH C18, Waters) was used for all LC-MS/MS analyses. Mascot software (version 2.2.1, Matrix Science, London, U.K.) was used for database searching and spectral interpretation.

**Results**

The full-length BoNT/A ad (atoxic derivatives) DNA and proteins were generated under biosafety level 2 containment (project approved by CDC on 02.07.2006 for the registered entity C20060207-0419).

To improve the yield of recombinant proteins, the DNA sequence encoding the full-length construct was synthesized *de novo*, and optimized for expression in both *S. flexens* insect cells and *E. coli* as explained in Methods. Our attempts to express BoNT/A**<sup>ΔN</sup>** in *E. coli* were not successful despite repeated attempts, as described in the Supplemental materials to this report.

**Expression, purification and processing of BoNT/A**

A four amino acid enterokinase cleavage site, DDDD, was introduced between amino acid residues N<sub>447</sub> and K<sub>448</sub> in the first construct, BoNT/A**<sup>ΔN</sup>** (Fig. 1). The 21 bp cis-dna sequence, AACCTCTAAAAACCGCCACC, was inserted to increase exogenous gene expression in baculovirus-infected insect cells [19]. The signal peptide, MKFLVNLAVFMVVVISYYA, positioned in front of the first methionine, was inserted to direct transport of the expressed protein into the culture medium, and was removed by processing during intracellular trafficking and secretion [20]. For the purpose of protein purification, an N-terminal 10-His tag and C-terminal Streptag II were also present in the construct. The expressed
propeptide was detected in the secreted medium with polyclonal antibodies raised against BoNT/A holotoxin. The mobility of the protein band was similar to the mobility of the unprocessed form of wt BoNT/A (data not shown). After optimization of expression, the BoNT/Aad* propeptide was purified to virtual homogeneity from SF-900 II medium in two steps, metal chelate affinity resin (Fig. 2, Panel A), followed by StrepTactin affinity chromatography (Fig. 2, Panel B).

The purified protein was then processed to the heterodimer by cleavage with recombinant enterokinase. In the pilot reaction (Fig. 3, Panel A1 and A2) the optimal enzyme/protein ratio for cleavage was determined. It was noted that an excess of enterokinase led to non-specific protein degradation (Fig. 3, Panels A1 and A2, lanes 4–6). This degradation could be attributed either to secondary activity of the enterokinase, or to contaminants in the commercially available enterokinase preparations. The bulk of the expressed BoNT/Aad* propeptide was processed with 0.007 U of enzyme per microgram of protein for 12 h at 16 °C. This resulted in approximately 95% completion of cleavage without visible degradation of light and heavy chains in the processed BoNT/Aad* heterodimer. Recombinant enterokinase was removed from the reaction mixture by incubation with Tag-off cleavage capture kit (Novagen).

To facilitate removal of the 10-His tag from the BoNT/Aad* propeptide, a TEV protease recognition sequence was introduced downstream of the metal chelate affinity tag in the expressed protein. Due to the small size (19 aa) of the peptide released as a result of the cleavage, the shift in BoNT/Aad* propeptide mobility and the degree of the enzymatic cleavage were not evident on Coomassie-stained gels (data not shown). To examine and optimize conditions for TEV digest, a time course pilot reaction was conducted. Samples of BoNT/Aad* without addition of the enzyme were used as a control. Aliquots from the reaction mixture were taken at times ranging from 1 to 6 h, separated by SDS–PAGE, transferred to nitrocellulose membrane, and probed with anti-His-tag monoclonal antibody. The results are shown in Fig. 3, Panel B. Incubation of one microgram of the BoNT/A1ad* with one unit of the AcTEV protease at 30 °C for 6 h led to almost complete removal of the His tag from the propeptide (Fig. 3, Panel B, lane 10). The mobility

Fig. 2. BoNT/Aad* propeptide purification. Reduced 12% SDS–PAGE stained with Coomassie BB R-250. Panel A: Talon® chromatography purification: lane 1, wt BoNT/A, control; lane 2, unfractionated sample of concentrated and dialyzed SF-900 II medium containing secreted propeptide BoNT/Aad* prior to loading on column; lane 3, column flow through; lane 4, wash 1, loading buffer; lane 5, wash 2, loading buffer with 20 mM imidazole; lane 6, elution, loading buffer with 200 mM imidazole. Panel B: StrepTactin agarose chromatography: lane 1, sample of concentrated and dialyzed fraction from Panel A, lane 6 prior to loading on column; lane 2, flow through; lanes 3–7, washes with loading buffer; lanes 8–12, elutions with loading buffer and 3 mM desthiobiotin; lane 13, wt BoNT/A, control.

Fig. 3. Processing of BoNT/Aad* propeptide. Panels A1, A2: Processing to heterodimer by proteolytic cleavage with recombinant enterokinase (rEK). One microgram of BoNT/Aad* propeptide per lane was treated with 0–5 U rEK at 16 °C for 12 h, separated by 12% SDS–PAGE, and stained with Coomassie BB R-250. Panel A1: non-reduced samples. Panel A2: samples reduced by addition of β-mercaptoethanol. Lane 1, no rEK; lane 2, 0.001 U rEK; lane 3, 0.01 U rEK; lane 4, 0.1 U rEK; lane 5, 1 U rEK; lane 6, 5 U rEK; lane 7, wt BoNT/A, control. Panel B: Removal of 10-His tag from BoNT/Aad* propeptide by treatment with AcTEV. BoNT/Aad* was either treated with buffer (odd lane numbers), or treated with AcTEV protease (even lane numbers; 1 U per microgram, 30 °C; see Methods for details) for the times indicated: lanes 1 and 2, 1 h; lanes 3 and 4, 2 hours; lanes 5 and 6, 3 h; lanes 7 and 8, 4 h; lanes 9 and 10, 6 h. Samples were loaded on a 12% SDS–PAGE in the presence of β-mercaptoethanol, separated and transferred to nitrocellulose. Western blot was probed with HRP-coupled anti-His MAb (Santa Cruz, H-3 His probe, Cat # sc-8036 HRP). Note: The low MW band (approximately 30 kDa) in even lanes represents AcTEV protease which was supplied by Invitrogen as a 6-His tagged recombinant enzyme.
of the propeptide band on SDS–PAGE before and after cleavage did not indicate any apparent non-specific proteolytic activity associated with AcTEV. The AcTEV protease was removed from the reaction mixture by affinity chromatography on Talon® resin. Purification and processing of BoNT/A derivative was summarized and shown in Table 2 and yielded approximately 30 mg BoNT/A derivative per liter of insect cell culture. To rule out the presence of post-translational glycosylation in our protein, we compared the mobility of the expressed, processed and denatured BoNT/A derivative after treatment with Endo-α-N-acetylgalactosaminidase and PNGaseF (New England Biolabs, Cat # P0733S, P0704S) with untreated samples. No difference in mobility of treated vs. untreated samples were detected by SDS–PAGE (data not shown).

Due to secondary activity in enterokinase samples detected during the propeptide processing we confirmed the structural integrity of expressed and processed protein. N-terminal sequencing (Edman degradation) of separated light and heavy chain, performed at the Molecular Structure Facility at UC Davis, identified the first six amino acids as GAGDSL for LC, and ALNDLC from the BoNT/A derivative. The sequence of the peptide was confirmed by the MS/MS spectrum (Fig. 4, Panel B3).

### Site-selective cargo attachment to BoNT/A derivative

A 12 aa S6 peptide tag placed downstream of the 10-His sequence and upstream of the N-terminus of the LC was incorporated in BoNT/A derivative and the other derivatives described in this work as a site for selective cargo attachment. Sfp phosphopantetheinyl transferase catalyzes incorporation of small-molecule-CoA-based cargo to a specific serine residue within the S6 tag [16]. As a prototypic cargo molecule, and to create a molecular probe to study BoNT/A trafficking, a commercially available fluorescent conjugate of CoA (CoA547, NEB) suitable for standard TAMRA and Cy3 microscopy emission filter sets was used for BoNT/A derivative labeling. The original report and New England Biolabs protocols described conditions used for in vitro labeling of recombinant proteins expressed on the cell surface, i.e., when complex mixtures of various biopolymers are present in the reaction. It was noted that under the conditions tested, an increase of the CoA substrate concentration usually results in a higher background and does not necessarily increase the signal to background ratio. The addition of fetal calf serum or 0.5% BSA to the reaction mixture reduces the background staining. During optimization of this site-specific BoNT/A derivative labeling in vitro, background staining was also noticed. To minimize the background and to avoid introducing contaminating proteins such as BSA into the labeling reaction, BSA was replaced by a mixture of short peptides that were subsequently removed by dialysis and ultrafiltration. The results of the enzymatic labeling are shown in Fig. 5. In this figure labeled and unlabeled samples of BoNT/A derivative were separated on SDS–PAGE and either stained with Coomassie, or transferred to a nitrocellulose membrane and scanned using a 532/580 nm excitation/emission filter set. The recombinant Sfp phosphopantetheinyl transferase was removed from the reaction mixture by affinity chromatography on Talon® resin; the excess of CoA 547 and other low molecular weight components were removed by dialysis and ultrafiltration.

After enzymatic labeling of BoNT/A derivative light chain by Sfp phosphopantetheinyl transferase with CoA 547, three separate in-gel
Fig. 4. Structural analysis of tryptic peptides isolated from BoNT/Aad. Panel A: ESI Q-TOF MS/MS spectrum of the C-terminal tryptic peptide isolated from enterokinase-processed BoNT/Aad LC under reducing conditions. The b and y ion series have been included at the top of the figure to identify the peptide fragment peaks found in the spectrum, where the position of vertical bars separating each amino acid(s) correspond to the position of the respective ion m/z within the spectrum. For clarity, only the most intense peaks from y series have been labeled.

Panels B1-B3: MALDI-TOF mass spectra and ESI Q-TOF MS/MS spectrum of the tryptic dipeptide with an internal disulfide bridge linking light and heavy chains of enterokinase-processed BoNT/Aad LC separated and isolated from a reduced SDS–PAGE. Panel B1: MALDI-TOF mass spectrum of an in-gel tryptic digest of the enterokinase-processed BoNT/Aad LC separated and isolated from a reduced SDS–PAGE; Panel B2: MALDI-TOF mass spectrum of an in-gel tryptic digest of the enterokinase-processed BoNT/Aad separated and isolated from an unreduced SDS–PAGE. A peak at m/z 1489.84, matched the predicted m/z of the dipeptide with internal disulfide bridge; Panel B3: ESI Q-TOF MS/MS spectrum of the dipeptide with m/z 1489.8 shown in Panel B2, confirming presence of the S-S bond in the dipeptide. The b and y ion series have been included at the top of the panel to identify the peptide fragment peaks found in the spectrum, where the position of vertical bars separating each amino acid(s) correspond to position of the respective ion m/z within the spectrum. Only the most intense peaks have been labeled for clarity.
tryptic digests were analyzed by LC-MS in duplicate, and compared to the LC-MS spectra of digests of protein incubated with enzyme without CoA 547. In the absence of exogenous CoA derivative, the N-terminal peptide was not labeled with endogenous CoA, demonstrating that, as intended, the inserted signal peptide directs the expressed protein for secretion. Efficiency of cargo attachment could therefore be estimated by measuring the reduction of unmodified N-terminal peptide peak in MS analyses of tryptic digests. Mean ion intensity of the unmodified N-terminal peptide was not labeled with endogenous CoA, 547. Lanes 3, 4, 6, 8: 0.1 µg BoNT/A/AD. Panel A: 10.5–14% Criterion gel (Bio-Rad) stained with Bio-Safe Coomassie (Bio-Rad). Panel B: Western blot of gel shown in panel A scanned on a Typhoon 9500 scanner (GE Healthcare) using 300 V PMT, 532/580 nm excitation/emission filter set (green).

Expression, purification and processing of BoNT/A/ADev

The BoNT/A/ADev construct is very similar to the BoNT/A/AD construct, but enables the heterodimer-forming cleavage step and removal of the 10-His tag to be performed during a single processing step with TEV protease. The design of this construct, the procedure for Sf9 infection and culture growth, and the affinity purification procedure were similar to those used for BoNT/A/AD, as described above and in Methods. Removal of the 10-His tag and processing of the propeptide were performed simultaneously by incubating 1 µg of the BoNT/A/ADev propeptide with 2 U of AcTEV protease at 30 °C for 6 h. Processed peptide was separated from AcTEV protease by affinity chromatography on Talon® resin and collected in flow-through fractions. Fig. 6 shows the purified protein separated by SDS-PAGE under reducing and non-reducing conditions. Purification and processing of BoNT/A/ADev are summarized in Table 2 and yielded approximately 30 mg BoNT/A/ADev per liter of insect cell culture.

Expression, purification and processing of ALC-Peptide-BoNT/Aev and ALC-GFP-BoNT/Aev

The ALC constructs were developed to evaluate the role of the LC peptide in BoNT-mediated delivery to the neuronal cytosol, and the limits on cargo that can be targeted using HC-mediated mechanisms. The design of the constructs and expression-purification procedures were similar to the previous two constructs as shown in Table 2, except for differences in MOI and incubation time after infection, which were approximately 0.5 and 72 h, respectively. With both ALC derivatives, precipitation was noted during the concentration step. To prevent protein precipitation, prior to the concentration step, Triton X-100 was added to filtered medium. Triton X-100 was present throughout the Talon® chromatography procedure, and was replaced with dialyzable n-octyl-D-glucopyranoside for StrepTactin affinity chromatography. Removal of the 10-His tag and processing of the propeptides were performed simultaneously by incubating 1 µg of the propeptides with 2 U of AcTEV protease at 30 °C for 6 h in the presence of 0.2% n-octyl-D-glucopyranoside, which was removed by dialysis and ultrafiltration after the cleavage. Processed peptides were separated from AcTEV protease by affinity chromatography on Talon® resin.
and collected in flow-through fractions. The yield of ΔLC-Peptide-BoNT/Atev was approximately 2 mg per liter and ΔLC-GFP-BoNT/Atev 1 mg per liter of culture medium. Fig. 6 shows the purified proteins separated by SDS–PAGE under reducing and non-reducing conditions.

Discussion

BoNTs are large, multi-domain, disulfide-bonded heterodimers, with mutual stabilization of domains through hydrogen bonds and hydrophobic interactions [21–23]. It is therefore challenging to produce recombinant BoNT derivatives that retain structural features required for native BoNT trafficking. Factors affecting the success of protein expression include the design of expression constructs and the choice of expression system. When domains are expressed separately, they can be denatured, poorly soluble [24–26], unstable [19], unusually sensitive to mild agitation [8] and difficult to reconstitute into native disulfide-bonded heterodimers. This report builds on our prior work describing constructs produced from C. botulinum A1 Hall genomic DNA, expressed using the baculovirus system, and purified as full-length disulfide-bonded heterodimers [27]. The major problems encountered in this earlier work were related to low yield of the protein after purification (approximately 0.35 mg per liter of insect cell culture medium), and our inability to establish a single step purification procedure based on the original 6-His tag incorporated. Here we have utilized fully synthetic gene constructs in order to optimize multiple aspects of protein expression and purification.

We considered it essential for recombinant BoNT to be expressed as a full-length, disulfide-bonded, single chain propeptide. More specifically, E. coli has a strong codon bias against AT-rich chloridial genes, and to be secreted into the host culture medium as soluble protein that could be purified under mild, non-denaturing conditions. For this reason we chose baculovirus-infected SF9 cells as the expression host, and included a signal sequence at the N-terminus of our recombinant proteins. Several attempts in our laboratory to express BoNT/A derivatives in E. coli as the full-length disulfide-bonded heterodimer failed under multiple experimental variations, despite reports by others of successful expression of clostridial neurotoxins in E. coli [9–12,28]. Expressing large disulfide-bonded proteins in E. coli is difficult for multiple reasons, including the reducing environment of the E. coli cytosol, and the tendency of E. coli to segregate unfolded recombinant proteins into inclusion bodies aggregates requiring harsh extraction. More importantly, E. coli has a strong codon bias against AT-rich chloridial genes, and a propensity toward non-specific proteolysis of improperly folded proteins. Prior studies have found that BoNT/A propeptide can be cleaved at K444, and K448, leaving uncertainty regarding the natural C-terminus of the LC [29,30]. We rationalized that our constructs should not shorten the length of the peptide loop excised because this might influence the 3D structural constraints on this region, but should rather be made resistant to trypsin-like proteases. Therefore, mutations K438 > H, K440 > Q and K444 > Q were introduced into our constructs to render our propeptide derivatives resistant to proteolytic cleavage by trypsin-like proteases and to yield uniformly processed LC C-termini in the heterodimers. Trypsin treatment of wt BoNT/A can also lead to cleavage of its receptor-binding HC domain, and complete loss of toxicity [31]. Therefore, a mutation K871 > N at the HC9–HC1 junction was introduced into our constructs, rendering the HC insensitive to this type of non-specific cleavage, as well [31,32].

The four distinct BoNT/A derivatives described here were expressed using the SF9-baculovirus system. All were expressed as soluble, single chain, disulfide bonded propeptides, secreted into the host medium, and purified to near homogeneity under non-denaturing conditions. Two mutations were introduced into the catalytic core responsible for substrate cleavage and toxicity. Amidino acids E224 and Y366 which are conserved among different BoNT serotypes, were mutated to A, rendering the full-length BoNT/A ad unable to cleave SNAP 25 [6,33]. Moreover, mutation of E224 and Y366 has been demonstrated not to change the secondary structure, topography of aromatic amino residues, Zn2+ content, or substrate binding ability of the LC metalloprotease [34]. We suggest that the SF9-baculovirus expression system is advantageous because E. coli and P. pastoris show high bias against clostridial AT-rich DNA, resulting in slow growth [21], and premature termination of protein synthesis or initiation of irrelevant translation from an alternative reading frame [26].

The first derivative, BoNT/AadΔk, replaces the enterokinase site between LC and HC with a TEV protease domain. The structural authenticity of the expressed protein was confirmed by Western blotting with BoNT/A polyclonal antibodies and by Edman degradation of LC and HC. MALDI-TOF and ESI-Q-TOF LC-MS/MS analysis of LC tryptic digest confirmed uniformity of the enterokinase cleavage site at the LC–HC junction through identification of the predicted C-terminal peptide of the LC (Fig. 4, Panel A), and the presence of the disulfide bridge formed between Cys430 and Cys445 (aa numbers are from the sequence of wt BoNT/A [15]) (Fig. 4, Panels B1–B3). The second derivative, BoNT/AadΔev is similar to BoNT/AadΔk but replaces the enterokinase site between LC and HC with a TEV recognition sequence. This avoided the non-specific cleavage we observed with excess enterokinase treatment, and reduced the number of steps required for protein purification. This derivative was also purified to virtual homogeneity, yielding 30 mg per liter of culture medium. The structural authenticity of the expressed heterodimers and formation of the S–S bridge between LC and HC in this derivative were confirmed by reduced and unreduced SDS–PAGE and Western blotting (Fig. 6).

In the third and fourth derivatives, derivative, ΔLC-Peptide-BoNT/AΔev and ΔLC-GFP-BoNT/AΔev, the entire catalytic domain (P2–F390) of the light chain was removed, leaving a 54 aa LC segment (N391–G445) that forms the disulfide bridge and otherwise interacts with the HCN (aa numbers are from the sequence of wt BoNT/A [15]). Two TEV protease cleavage sites allow simultaneous removal of the N-terminal 10-His tag and processing of the precursor to generate the ΔLC–HC heterodimer. These derivatives were likewise expressed as soluble single chain proteins secreted into culture medium, but were found to have a tendency to precipitate during concentration. Moreover, detergents were necessary during concentration and purification to maintain solubility. After processing with ACTEV protease to generate the ΔLC–HC heterodimers and remove the 10-His affinity tags, it became possible to remove detergent because the solubility of the processed heterodimers increased markedly. Yield of the proteins was 2 and 1 mg per liter of culture medium for ΔLC-Peptide-BoNT/AΔev and ΔLC-GFP-BoNT/AΔev, respectively. SDS–PAGE under reducing agent and non-reducing conditions, with Western blotting performed with polyclonal antibodies raised against BoNT/A holotoxin confirmed the structural identity of the ΔLC-Peptide-BoNT/AΔev heavy chain, but did not detect the truncated LC (Fig. 6). Blotting with monoclonal antibody against GFP confirmed the structural identity of ΔLC-GFP-BoNT/AΔev light and heavy chains (Fig. 6).

The relatively low yield and aggregation of the ΔLC derivatives may be attributed to conformational instability and spatial tension in the expressed propeptides. The published X-ray structure of BoNT/A holotoxin [35] suggests that the N-terminal portion of the N391–G445 sequence is a flexible loop and not part of the distinct protein fold. We hypothesize that deletion of the entire LC catalytic domain exposes hydrophobic areas of the propeptide that cannot optimally collapse due to the constraints imposed by the
tethering of the loop on both ends, and that AcTEV cleavage releases this constraint by untethering one end so that the propeptide can collapse into a more soluble conformation.

During early pilot stages of this work we expressed a full-length BoNT/A derivative without any introduced mutations to inactivate the LC metalloprotease. When this non-mutated derivative was tested in mouse phrenic nerve-hemidiaphragm preparations, it blocked neural transmission in 167 ± 17 min. At a concentration of \(1 \times 10^{-11}\) M \((n = 4)\), which is comparable to the potency of pharmaceutical BoNT/A preparations. Pre-incubation of the derivative with polyclonal BoNT/A antiserum completely eliminated its paralytic effect on the hemidiaphragm throughout the duration of the experiment \((n = 3,\) approximately 400 min of tissue monitoring), demonstrating that the blockade was attributable to a BoNT-type action \((L.\ Simpson,\ unpublished\ communication)\). This experiment provided the only data currently available indicating that the methodology described in this report is capable of producing recombinant BoNT derivatives with native biological activity, and by inference, native trafficking properties.

All recombinant BoNT/A derivatives described in this work carry an S6 tag in their N-terminal region. Specific conditions for in vitro CoA-fluorophore labeling were optimized and shown for BoNT/A\textsubscript{ad\textsubscript{k}} \((\text{Fig. 5})\). Prior studies using fluoroscein-labeled BoNTs to evaluate either LC trafficking in neurons or BoNT/A uptake by epithelial or neuronal cells have primarily relied on two methods: (a) transient expression following transfection with LC sequences tagged with a fluorescent marker \((\text{e.g. GFP})\ [36]\), or (b) attachment of fluorescent tags to LC–HC BoNT heterodimers using chemical methods [37,38]. The transient expression system provides important information on LC interactions after the recombinant protein accumulates in neurons, but cannot be used to explore the normal trafficking route that delivers LC to the neuronal cytoplasm. The chemical method for attaching probes is limited by the lack of selective fluorophore attachment \((\text{e.g. both LC and HC will be modified})\), and the inevitable formation of complex mixtures \((unlabeled,\ singly\ and\ multiply\ labeled\ species)\). The latter problem makes it particularly difficult to confidently attribute behavior of fluorescently labeled BoNT molecules as being representative of native trafficking and biological activity; the most intensely fluorescently labeled molecules would produce the most intense signal despite potentially being the least related to native trafficking, while a small population of unmodified molecules could be responsible for any observed biological activity. Several studies comparing chemical labeling to site-selective enzymatic labeling have demonstrated the superiority of the latter for preserving biological function [39].

In conclusion, we have produced a series of atoxic, full-length and truncated BoNT derivatives that preserve important structural features of native BoNT. All of the derivatives can be recovered from culture media as soluble disulfide-bonded heterodimers, and can be purified to homogeneity using two-stage, non-denaturing and highly selective affinity purification. The ability to recover the expressed derivatives as soluble proteins is extremely important, because it obviates the need to recover insoluble expressed derivatives from inclusion bodies using denaturing conditions. Retaining native BoNT structure after harsh solubilization and purification steps is extremely challenging [8,40]. The sensitivity of BoNT/A to isolation and purification conditions is also reflected in the wide batch-to-batch variability observed during pharmacological BoNT/A production from clostralid cultures in terms of BoNT/A specific activity units per mg of therapeutic protein preparation \((E.\ Johnson,\ personal\ communication)\). Because the expression and purification methodology employed in this work does not use denaturing conditions, the expressed BoNT derivatives described here retain native BoNT structure to a greater extent than methods requiring exposure to harsh reagents. Experiments are currently underway to evaluate the systemic and neuronal targeting properties of these BoNT derivatives, and to evaluate their toxicity in vivo.

Acknowledgments

We thank Dr. Eric Johnson for discussion. Support for these studies was provided by NIH-NIAID AI072466 to K.L., NIH-NINDS NS050276, NIH-NCCR RR017990 to T.A.N., and NIH Office of the Director DP2-OD004631 to T.J.C.\textsuperscript{wr} BoNT/A was supplied by Willy Tepp from the laboratory of Dr. Eric Johnson \((\text{University of Wisconsin at Madison})\), pET29 vector carrying C-terminally His\textsubscript{6}-tagged Sfp phosphopantetheinyl transferase from \(B.\ subtilis\) was kindly provided by Dr. Jun Yin \((\text{University of Chicago})\). Authors thank Dr. Eric Milner for reading and careful editing of the manuscript.

Appendix A. Supplementary data


References


