

Published in final edited form as:

*Protein Expr Purif.* 2010 May ; 71(1): 62–73. doi:10.1016/j.pep.2009.12.013.

## Recombinant Derivatives of Botulinum Neurotoxin A Engineered for Trafficking Studies and Neuronal Delivery

Philip A. Band<sup>1</sup>, Steven Blais<sup>2</sup>, Thomas A. Neubert<sup>1,2</sup>, Timothy J. Cardozo<sup>1</sup>, and Konstantin Ichtchenko<sup>1,\*</sup>

<sup>1</sup> Department of Pharmacology, New York University School of Medicine, New York, New York

<sup>2</sup> Kimmel Center for Biology and Medicine at the Skirball Institute, New York University School of Medicine, New York, New York

### Abstract

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/A1*ad*<sup>ek</sup> and BoNT/A1*ad*<sup>tev</sup> are full length derivatives rendered atoxic through double point mutations in the LC protease (E<sub>224</sub>>A; Y<sub>366</sub>>A). ΔLC-peptide-BoNT/A<sup>tev</sup> and ΔLC-GFP-BoNT/A<sup>tev</sup> 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.

### Keywords

Botulinum neurotoxin A; *Clostridium botulinum*; baculovirus; protease; *Sfp* phosphopantetheinyl transferase

---

Botulinum neurotoxins (BoNTs) are a family of structurally similar proteins that cause peripheral neuromuscular blockade and respiratory paralysis, with an extremely low LD<sub>50</sub> (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

---

© 2009 Elsevier Inc. All rights reserved.

\*Corresponding author. Mailing address: Department of Pharmacology, NYU School of Medicine, MSB 320, 550 First Avenue New York NY 10016. Phone (212) 263 2301 Fax (212) 263 7133. ichtck01@med.nyu.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

chain propeptides (Mr approximately 150,000; approximately 1,300 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<sub>C</sub>); and translocation activity responsible for delivering the LC protease to the neuronal cytosol is associated with the N-terminal half of the HC (HC<sub>N</sub>) (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, reconstituted 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 renaturing 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 *E. coli*. Kiyatkin et al. reported the expression of BoNT/C in *E. coli*, with three inactivating point mutations (H<sub>229</sub>>G; E<sub>230</sub>>T; H<sub>233</sub>>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/G, /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 holotoxoid in the non-toxic *Clostridium botulinum* strain LNT01. The mutations R<sub>364</sub>>A and Y<sub>366</sub>>F were introduced into the LC (BoNT/A<sup>RYM</sup>), rendering the protein unable to cleave the substrate SNAP 25 *in vitro*. Immunization with this holotoxoid 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 *P. pastoris*, with the mutations H<sub>223</sub>>A; E<sub>224</sub>>A; H<sub>227</sub>>A to render the toxin inactive (ciBoNT/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>ek</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 b.p.) was used for subcloning the full-length BoNT/Aad<sup>ek</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 b.p.). Vector pLitmus28C1INT was digested with Kas 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 b.p.). The sequence of this vector was deposited in Genbank under accession number [GQ855199](#).

**Cloning of BoNT/Aad<sup>ek</sup> into pLitmus28C1**—In the first cloning step, pLitmus28C1 was digested with Hind III and Not I and dephosphorylated. Annealed phosphorylated oligonucleotides CP1S and CP1A were ligated into digested pLitmus28C1, resulting in vector pLitSB3A1 (2968 b.p.). Vector pLitSB3A1 carries several unique 5'-restriction sites, including Bgl II, Bbs I, Rsr II and BstB I which provide considerable flexibility for subsequent subcloning of the full-length construct into expression vectors. This polylinker was followed by the enhancer sequence AACTCCTAAAAACCGCCACC, followed by the honeybee mellitin signal peptide sequence.

In the next step, pLitSB3A1 was digested with Not I and Mlu I and dephosphorylated. Annealed phosphorylated oligonucleotides CP2S and CP2A were ligated into digested pLitSB3A1, resulting in vector pLitSB3A2 (2987 b.p.), which carried a 10-His affinity tag downstream of the honeybee mellitin signal peptide sequence. Next, pLitSB3A2 was digested with Mlu I and Nhe I and dephosphorylated. Annealed phosphorylated oligonucleotides CP3S and CP3A were ligated into digested pLitSB3A2 to generate pLitSB3A3 (3056 b.p.), which carries the tobacco etch virus (TEV) protease cleavage site, followed by the S6 peptide tag sequence downstream of the 10-His affinity tag sequence.

In the next three steps a full-length BoNT/Aad<sup>ek</sup> vector optimized for expression in two hosts, *S. frugiperda* and *E. coli*, was progressively constructed in the pLitSB3A3 vector. DNA

fragments for constructing the full-length BoNT/Aad<sup>ek</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 b.p. DNA fragment encoding the sequence P<sub>2</sub> - L<sub>442</sub> was isolated from pUC57-BAad-1 by restriction digest with BssH II and Afl II and subcloned into pLitSB3A3 to generate the 4355 b.p. vector, pLitSB3A4. Then, a 1233 b.p. DNA fragment encoding the sequence D<sub>443</sub> - I<sub>849</sub> was isolated from pUC57-BAad-2 by digest with Xba I and EcoR V and subcloned in pLitSB3A4 to generate the 5570 b.p. vector, pLitSB3A5. Finally, a 1401 b.p. DNA fragment encoding the sequence P<sub>850</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-BAad-3 by digest with EcoR V and Xho I, and subcloned into pLitSB3A5 to generate the 6925 b.p. vector, pLitSB3A. The sequence of pLitSB3A has been deposited in GenBank under accession number [GQ855200](#).

**Construct for expression of BoNT/Aad<sup>ek</sup>**—The plasmid pLitSB3A was digested with the restriction endonucleases Rsr II and Xho I. The 4190 b.p. DNA fragment encoding full-length BoNT/Aad<sup>ek</sup> was isolated from an agarose gel, purified and ligated into the baculovirus transposition vector, pFastBac1™ (Invitrogen, Cat. # 10360-014, 4776 b.p.) to generate the 8968 b.p. vector, pFB1SB3AEK. The DNA sequence of pFB1SB3AEK 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>ek</sup>), are shown in Figure 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>tev</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 b.p. fragment containing the enterokinase cleavage site was excised from pFB1SB3AEK by digestion with Xba I and Afl II. A 351 b.p. Xba I/Afl II DNA fragment carrying the TEV sequence was excised from the plasmid pUC57-TEV (GenScript) and ligated into the digested pFB1SB3AEK to generate the 8971 b.p. vector, pFB1SB3ATEV. The DNA sequence of pFB1SB3ATEV, encoding a 1373 aa protein along with annotations has been deposited in GenBank under accession number [GQ855202](#).

**Construct for expression of ΔLC-Peptide-BoNT/A<sup>tev</sup>**—A DNA fragment containing the sequence encoding the ΔLC-peptide was obtained from the vendor (GenScript) in the form of synthetic DNA cloned into pUC57 (pUC57-Pept). pUC 57-Pept was digested with BssH II and Xba I and the 200 b.p. DNA fragment was isolated from an agarose gel, purified and ligated into digested pFB1SB3ATEV, resulting in the 7831 b.p. vector, pFB1SPepB2ATEV. The DNA sequence of pFB1SPepB2ATEV encoding a 993 aa protein along with annotations has been deposited in GenBank under accession number [GQ855203](#).

**Construct for expression of ΔLC-GFP-BoNT/A<sup>tev</sup>**—DNA encoding the GFP sequence was obtained by PCR amplification from plasmid pAcGFP1-C2 (Clontech, Cat. # 632481, 4722 b.p.) 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 Nhe I and Acc65 I, and the 700 b.p. digested DNA was isolated from an agarose gel and purified. The vector, pFB1SPepB2ATEV, was digested with Nhe I and partially digested (short incubation time and small amount of the restriction endonuclease) with Acc65 I. A 7822 b.p. DNA fragment was isolated from an agarose gel, purified and dephosphorylated. The purified GFP PCR product and linearized pFB1SPepB2ATEV were ligated, resulting in the 7822 b.p. 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 Qiaex 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™ constructs were used for transposition of the cloned DNA into the shuttle vector (bacmid) for use in the Bac-to-Bac® 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 *Sf9* cells grown in Sf-900 II medium in a shaker flask. Cell density at the time of infection was approximately  $1.5 \times 10^6$  cells/ml. Cells were incubated for 72 hours at 28°C in a humidified incubator. Cells were removed from high titer P2 stock by centrifugation at 1000 g. The titer of the P2 stock was measured by viral plaque assay. P2 titers for the constructs were approximately  $1-2 \times 10^8$  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  $1.5 \times 10^6$  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/Aad<sup>ek</sup> and BoNT/Aad<sup>tev</sup> optimal conditions were an MOI of approximately 0.1 and an incubation time of 60 hours; for ΔLC-Peptide-BoNT/A<sup>tev</sup> and ΔLC-GFP-BoNT/A<sup>tev</sup> optimal conditions were an MOI of approximately 0.5 and an incubation time of 72 hours. At the time of harvest, medium was separated from the cells by centrifugation at 1000 g, filtered through a Whatman GF/F filter (Schleicher&Schuell, Cat. # 1825090), concentrated approximately 3 fold on a Prep/Scale TFF cartridge (Millipore, Cat. # CDUF002LT) and dialyzed against Talon® resin chromatography loading buffer (500 mM NaCl, 25 mM TrisHCl, 5 mM imidazole, pH 8.0). For ΔLC-Peptide-BoNT/A<sup>tev</sup> and ΔLC-GFP-BoNT/A<sup>tev</sup>, 0.3% Triton X-100 was added to the medium prior to concentration. Dialyzed supernatant was loaded on Talon® resin (Clontech, Cat. # 63506) pre-equilibrated with loading buffer. After loading, resin was washed with 10 volumes of loading buffer, 20 volumes of wash 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 ΔLC-Peptide-BoNT/A<sup>tev</sup> and ΔLC-GFP-BoNT/A<sup>tev</sup> 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 100mM NaCl, 25 mM TrisHCl buffer, pH

8.0 either without (BoNT/Aad<sup>ek</sup> and BoNT/Aad<sup>tev</sup>), or with ( $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> and  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup>) 0.2% n-octyl- $\beta$ -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 superflow agarose (Novagen, Cat. # 71592) pre-equilibrated with 100mM NaCl, 25 mM TrisHCl buffer, pH 8.0 with or without 0.2% n-octyl- $\beta$ -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.

**Processing of BoNT/Aad<sup>ek</sup>**—The 10-His tag was removed from recombinant protein using AcTEV protease (Invitrogen, Cat. # 12575-015). The protein was incubated with AcTEV protease (1 unit of enzyme per  $\mu$ g of protein) at 30°C for 6 hours 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 protein 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. Propeptide 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 units of enzyme per 1  $\mu$ g of protein) at 16°C for 16 hours in digestion buffer (50 mM NaCl, 20 mM TrisHCl, 2 mM CaCl<sub>2</sub>, pH 7.8). Approximately 95% 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 (100 mM NaCl, 40 mM sodium phosphate, 40% glycerol, pH 7.2) and concentrated using Amicon Ultra filtration units.

**Processing of BoNT/A ad<sup>tev</sup>,  $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> and  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup>**—Removal of the N-terminal 10-His tag and propeptide cleavage were performed simultaneously, according to the protocol described for BoNT/Aad<sup>ek</sup> above.  $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> and  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup> required the presence of 0.2% n-octyl- $\beta$ -D-glucopyranoside during the cleavage. The detergent was removed after cleavage by dialysis and concentration on Amicon Ultra filtration units.

**Enzymatic labeling of BoNT/Aad<sup>ek</sup>**—All expressed proteins incorporated the S6 peptide GDSL<sup>S</sup>WLLRLLN (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 *B. subtilis* by Dr. Jun Yin, University of Chicago) was expressed in *E. coli* strain BL21 (DE3) (Novagen, Cat. # 69450) and purified according to the protocol described in (17). CoA 547 was purchased from Covalys (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  $\mu$ g of processed BoNT/Aad<sup>ek</sup> in a 30  $\mu$ l reaction mixture containing 1X unsupplemented Grace's insect cell culture medium (prepared from the dry medium, Invitrogen, Cat. # 11300-043), 5 mM MgCl<sub>2</sub>, a mixture of short peptides used to suppress the background labeling (and consisting of 150  $\mu$ M 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  $\mu$ M *Sfp*

phosphopantetheinyl transferase and 5  $\mu$ M CoA 547, pH 7.0. Fluorescent substrate was added as the final component and the reaction mixture was incubated for 15 minutes at 30°C. For visualization after labeling, the reaction was stopped by the addition of 2x 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 minute 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 glycerol-phosphate buffer (100 mM NaCl, 40 mM sodium phosphate, 40% glycerol, pH 7.2).

### Structural analysis of BoNT/Aad<sup>ek</sup> and fluorescently labeled BoNT/Aad<sup>ek</sup>

**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  $\mu$ l 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/Aad<sup>ek</sup> tryptic digest**—Samples were desalted using C18 ZipTips (Millipore, Billerica, MA) and eluted with 50% acetonitrile, 0.1% formic acid after which 1  $\mu$ l of sample was mixed with 1  $\mu$ l of alpha-cyano-4-hydroxycinnamic acid (Agilent Technologies, Santa Clara, CA) and 1  $\mu$ l 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  $\mu$ m 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 *Sf9* insect cells and *E. coli* as explained in Methods. Our attempts to express BoNT/Aad<sup>ek</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/Aad<sup>ek</sup>

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/Aad<sup>ek</sup> (Fig. 1). The 21 b.p. *cis*-DNA sequence, AACTCCTAAAAAACCGCCACC, was inserted to increase exogenous gene expression in baculovirus-infected insect cells (19) (Table 2). The signal peptide, MKFLVNVALVFMVVYISYIYAA, 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 Strep-tag 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*<sup>ek</sup> 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*<sup>ek</sup> propeptide was processed with 0.007 units of enzyme per microgram of protein for 12 hours at 16°C. This resulted in approximately 95% completion of cleavage without visible degradation of light and heavy chains in the processed BoNT/*Aad*<sup>ek</sup> 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*<sup>ek</sup> 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*<sup>ek</sup> 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*<sup>ek</sup> without addition of the enzyme were used as a control. Aliquots from the reaction mixture were taken at times ranging from 1 to 6 hours, separated by SDS PAGE, transferred to nitrocellulose membrane, and probed with anti His-tag monoclonal antibody. The results are shown in Figure 3, Panel B. Incubation of one microgram of the BoNT/*Aad*<sup>ek</sup> with one unit of the AcTEV protease at 30°C for 6 hours led to almost complete removal of the His tag from the propeptide (Fig. 3, Panel B, lane 10). The mobility 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/*Aad*<sup>ek</sup> is summarized and shown in Table 2 and yielded approximately 30 mg BoNT/*Aad*<sup>ek</sup> 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/*Aad*<sup>ek</sup> after treatment with Endo- $\alpha$ -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 wanted to confirm 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 for HC, which confirmed the predicted sequences of the respective N-termini. In addition, Figure 4, Panel A shows an ESI Q-TOF MS/MS spectrum of the intact tryptic peptide originating from the C-terminus of BoNT/*Aad*<sup>ek</sup> light chain, which confirmed the sequence of the predicted enterokinase cleavage site in the loop between LC and HC. The mobility and behavior of dichain protein under reducing and non-reducing conditions confirmed the presence of a disulfide bridge between subunits; additional structural analysis identified these cysteine residues as Cys<sub>430</sub> and Cys<sub>454</sub> (numbering from *wt* BoNT/A, 15). Figure 4, Panels B1-B3 show MALDI-TOF MS (Fig. 4, Panel B2) and ESI Q-TOF MS/MS (Fig. 4, Panel B3) spectra of a tryptic peptide of sequence LLCVR from the BoNT/*Aad*<sup>ek</sup> light chain linked *via* a disulfide bridge to a tryptic peptide of sequence ALNDLCIK from the BoNT/



*Aad*<sup>ek</sup> heavy chain. A peptide of observed m/z 1489.84 was detected in the MALDI-TOF mass spectrum of the non-reduced toxin (Fig.4, Panel B2), which closely matches the predicted m/z of the singly charged monoisotopic dipeptide ion of 1489.82. The peptides are absent from the MALDI-TOF MS spectrum of a tryptic digest of the reduced BoNT/*Aad*<sup>ek</sup> (Fig. 4, Panel B1). The sequence of the peptide was confirmed by the MS/MS spectrum (Fig. 4, Panel B3).

### Site-selective cargo attachment to BoNT/*Aad*<sup>ek</sup>

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/*Aad*<sup>ek</sup> 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/*Aad*<sup>ek</sup> 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/*Aad*<sup>ek</sup> 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 Figure 5. In this figure labeled and unlabeled samples of BoNT/*Aad*<sup>ek</sup> 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/*Aad*<sup>ek</sup> light chain by *Sfp* phosphopantetheinyl transferase with CoA 547, three separate in-gel 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 unlabeled N-terminal peptide peak in MS analyses of tryptic digests. Mean ion intensity of the unmodified N-terminal peptide, after normalization to the total ion intensity of all observed tryptic peptides from each respective protein, was reduced by 69±6% compared to the intensity of the unmodified peptide.

### Expression, purification and processing of BoNT/*Aad*<sup>tev</sup>

The BoNT/*Aad*<sup>tev</sup> construct is very similar to the BoNT/*Aad*<sup>ek</sup> 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/*Aad*<sup>ek</sup>, 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/*Aad*<sup>tev</sup> propeptide with 2 units of AcTEV protease at 30°C for 6 hours. Processed peptide was separated from AcTEV protease by affinity chromatography on Talon® resin and collected in flow-through fractions. Figure 6 shows the purified protein separated by SDS PAGE under

reducing and non-reducing conditions. Purification and processing of BoNT/Aad<sup>tev</sup> are summarized in Table 2 and yielded approximately 30 mg BoNT/Aad<sup>tev</sup> per liter of insect cell culture.

### Expression, purification and processing of $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> and $\Delta$ LC-GFP-BoNT/A<sup>tev</sup>

The  $\Delta$ LC 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 hours, respectively. With both  $\Delta$ LC 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- $\beta$ -D-glucopyranoside for StrepTactin affinity chromatography. Removal of the 10-His tag and processing of the propeptides were performed simultaneously by incubating 1  $\mu$ g of the propeptides with 2 units of AcTEV protease at 30°C for 6 hours in the presence of 0.2% n-octyl- $\beta$ -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  $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> was approximately 2 mg per liter and  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup> 1 mg per liter of culture medium. Figure 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 *Clostridium 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 here 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, 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 body

aggregates requiring harsh extraction. Moreover, *E. coli* has a strong codon bias against AT-rich clostridial genes, and a propensity toward non-specific proteolysis of improperly folded proteins. Prior studies have found that BoNT/A propeptide can be cleaved at K<sub>444</sub>, and K<sub>448</sub>, 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 K<sub>438</sub>>H, K<sub>440</sub>>Q and K<sub>444</sub>>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<sub>C</sub> domain, and complete loss of toxicity (31). Therefore, a mutation K<sub>871</sub>>N at the HC<sub>N</sub>-HC<sub>C</sub> 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. Amino acids E<sub>224</sub> and Y<sub>366</sub>, 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 E<sub>224</sub> and Y<sub>366</sub> has been demonstrated not to change the secondary structure, topography of aromatic amino residues, Zn<sup>2+</sup> content, or substrate binding ability of the LC metalloprotease (34). We suggest that the *Sf9*-baculovirus expression system is advantageous because *E. coli* and *Pichia 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/Aad<sup>ek</sup>, represents substantially intact BoNT/A with an enterokinase site for processing the expressed propeptide into the heterodimer. 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 digests 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 Cys<sub>430</sub> and Cys<sub>454</sub> (aa numbers are from the sequence of *wt* BoNT/A, **15**) (Fig. 4, Panels B1-B3).

The second derivative, BoNT/Aad<sup>tev</sup> is similar to BoNT/Aad<sup>ek</sup> 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 milligrams per liter of culture medium. The structural authenticity of the expressed heterodimer, 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,  $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> and  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup>, the entire catalytic domain (P<sub>2</sub> - F<sub>390</sub>) of the light chain was removed, leaving a 54 aa LC segment (N<sub>391</sub> - G<sub>445</sub>) that forms the disulfide bridge and otherwise interacts with the HC<sub>N</sub> (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  $\Delta$ LC -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 prior to the first affinity chromatography. Detergents were necessary during concentration and purification to maintain solubility. After processing with AcTEV protease

to generate the  $\Delta$ LC-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  $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> and  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup>, 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  $\Delta$ LC-Peptide-BoNT/A<sup>tev</sup> heavy chain, but did not detect the truncated LC (Figure 6). Blotting with monoclonal antibody against GFP confirmed the structural identity of  $\Delta$ LC-GFP-BoNT/A<sup>tev</sup> light and heavy chains (Figure 6).

The relatively low yield and aggregation of the  $\Delta$ LC 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 N<sub>391</sub>-G<sub>445</sub> 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 relieves 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  $\sim 1 \times 10^{-11}$  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 minutes 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/Aad<sup>ek</sup> (Fig. 5). Prior studies using fluorescent-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 (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 (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 pharmaceutical BoNT/A production from clostridial 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*.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

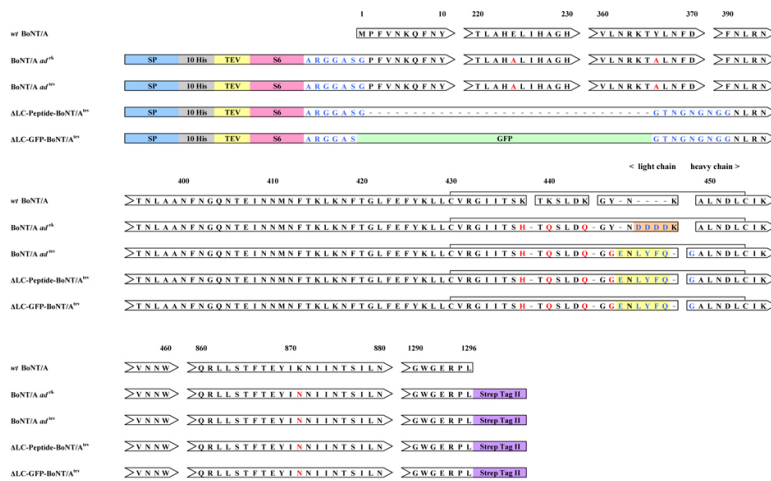
We thank Dr. Eric Johnson for discussion. Support for these studies was provided by NIH-NIAID AI072466 to K.L., NIH-NINDS NS050276, NIH-NCRR RR017990 to T.A.N., and NIH Office of the Director DP2-OD004631 to T.J.C. *wt* BoNT/A was supplied by William Tepp from the laboratory of Dr. Eric Johnson (University of Wisconsin at Madison). pET29 vector carrying C-terminally His6-tagged *Sfp* phosphopantetheinyl transferase from *B. subtilis* was kindly provided by Dr. Jun Yin (University of Chicago). Authors thank Dr. Eric Milner for reading and careful editing of the manuscript.

## References

1. National Institute of Occupational Safety and Health. Registry of Toxic Effects of Chemical Substances (R-TECS). National Institute of Occupational Safety and Health; Cincinnati, Ohio: 1996.
2. Smith TJ, Lou J, Geren IN, Forsyth CM, Tsai R, Laporte SL, Tepp WH, Bradshaw M, Johnson EA, Smith LA, Marks JD. Sequence variation within botulinum neurotoxin serotypes impacts antibody binding and neutralization. *Infect. Immun* 2005;73(9):5450–5457. [PubMed: 16113261]
3. Simpson LL. Identification of the major steps in botulinum toxin action. *Annu. Rev. Pharmacol. Toxicol* 2004;44:167–193. [PubMed: 14744243]
4. Koriazova LK, Montal M. Translocation of botulinum neurotoxin light chain protease through the heavy chain channel. *Nat. Struct. Biol* 2003;10(1):13–18. [PubMed: 12459720]
5. Dekleva ML, DasGupta BR, Sathyamoorthy V. Botulinum neurotoxin type A radiolabeled at either the light or the heavy chain. *Arch. Biochem. Biophys* 1989;274(1):235–240. [PubMed: 2774575]
6. Zhou L, de Paiva A, Liu D, Aoki R, Dolly JO. Expression and purification of the light chain of botulinum neurotoxin A: a single mutation abolishes its cleavage of SNAP-25 and neurotoxicity after reconstitution with the heavy chain. *Biochemistry* 1995;34(46):15175–15181. [PubMed: 7578132]
7. Goodnough MC, Oyler G, Fishman PS, Johnson EA, Neale EA, Keller JE, Tepp WH, Clark M, Hartz S, Adler M. Development of a delivery vehicle for intracellular transport of botulinum neurotoxin antagonists. *FEBS Lett* 2002;513:163–168. [PubMed: 11904143]
8. Toth SI, Smith LA, Ahmed SA. Extreme sensitivity of botulinum neurotoxin domains towards mild agitation. *J. Pharm. Sci* 2009;98(9):3302–3311. [PubMed: 19226630]
9. Kiyatkin N, Maksymowych AB, Simpson LL. Induction of an immune response by oral administration of recombinant botulinum toxin. *Infect. Immun* 1997;65:4586–4591. [PubMed: 9353037]

10. Rummel A, Mahrhold S, Bigalke H, Binz T. The HC<sub>C</sub>-domain of botulinum neurotoxins A and B exhibits a singular ganglioside binding site displaying serotype specific carbohydrate interaction. *Mol. Microbiol* 2004;51(3):631–643. [PubMed: 14731268]
11. Rummel A, Karnath T, Henke T, Bigalke H, Binz T. Synaptotagmins I and II act as nerve cell receptors for botulinum neurotoxin G. *J. Biol. Chem* 2004;279(29):30865–30870. [PubMed: 15123599]
12. Bade S, Rummel A, Reisinger C, Karnath T, Ahnert-Hilger G, Bigalke H, Binz T. Botulinum neurotoxin type D enables cytosolic delivery of enzymatically active cargo proteins to neurons via unfolded translocation intermediates. *J. Neurochem* 2004;91(6):1461–1472. [PubMed: 15584922]
13. Pier CL, Tepp WH, Bradshaw M, Johnson EA, Barbieri JT, Baldwin MR. Recombinant holotoxoid vaccine against botulism. *Infect. Immun* 2008;76(1):437–442. [PubMed: 17967862]
14. Webb RP, Smith TJ, Wright P, Brown J, Smith LA. Production of catalytically inactive BoNT/A1 holoprotein and comparison with BoNT/A1 subunit vaccines against toxin subtypes A1, A2, and A3. *Vaccine* 2009;27(33):4490–4497. [PubMed: 19450643]
15. Thompson DE, Brehm JK, Oultram JD, Swinfield TJ, Shone CC, Atkinson T, Melling J, Minton NP. The complete amino acid sequence of the *Clostridium botulinum* type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene. *Eur. J. Biochem* 1990;189(1):73–81. [PubMed: 2185020]
16. Zhou Z, Cironi P, Lin AJ, Xu Y, Hrvatin S, Golan DE, Silver PA, Walsh CT, Yin J. Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. *ACS Chem. Biol* 2007;2(5):337–346. [PubMed: 17465518]
17. Yin J, Lin AJ, Golan DE, Walsh CT. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat. Protoc* 2006;1(1):280–285. [PubMed: 17406245]
18. Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen P, Shevchenko A, Boucherie H, Mann M. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. U S A* 1996;93(25):14440–14445. [PubMed: 8962070]
19. Sano K, Maeda K, Oki M, Maeda Y. Enhancement of protein expression in insect cells by a lobster tropomyosin cDNA leader sequence. *FEBS Lett* 2002;532(12):143–146. [PubMed: 12459479]
20. von Heijne G. Signals for protein targeting into and across membranes. *Subcell. Biochem* 1994;22:1–19. [PubMed: 8146878]
21. Baldwin MR, Bradshaw M, Johnson EA, Barbieri JT. The C-terminus of botulinum neurotoxin type A light chain contributes to solubility, catalysis, and stability. *Protein Expr. Purif* 2004;37(1):187–195. [PubMed: 15294297]
22. Brunger AT, Breidenbach MA, Jin R, Fischer A, Santos JS, Montal M. Botulinum neurotoxin heavy chain belt as an intramolecular chaperone for the light chain. *PLoS Pathog* 2007;3(9):e113, 1191–1194.
23. Fischer A, Mushrush DJ, Lacy DB, Montal M. Botulinum neurotoxin devoid of receptor binding domain translocates active protease. *PLoS Pathog* 2008;4(12):e1000245, 1–9. [PubMed: 19096517]
24. Ahmed SA, Smith LA. Light chain of botulinum A neurotoxin expressed as an inclusion body from a synthetic gene is catalytically and functionally active. *J. Protein Chem* 2000;19(6):475–487. [PubMed: 11195972]
25. Zhou Y, Singh BR. Cloning, high-level expression, single-step purification, and binding activity of His<sub>6</sub>-tagged recombinant type B botulinum neurotoxin heavy chain transmembrane and binding domain. *Protein Expr. Purif* 2004;34(1):8–16. [PubMed: 14766296]
26. Lacy DB, Stevens RC. Recombinant expression and purification of the botulinum neurotoxin type A translocation domain. *Protein Expr. Purif* 1997;11(2):195–200. [PubMed: 9367816]
27. Ichtchenko K, Band PA. Full-length *Clostridium botulinum* serotype A derivatives with native structure and properties. *Neurotox. Res* 2006;9:234.
28. Li Y, Foran P, Lawrence G, Mohammed N, Chan-Kwo-Chion CK, Lisk G, Aoki R, Dolly O. Recombinant forms of tetanus toxin engineered for examining and exploiting neuronal trafficking pathways. *J. Biol. Chem* 2001;276(33):31394–31401. [PubMed: 11402045]
29. DasGupta BR, Dekleva ML. Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and around the nicking site. *Biochimie* 1990;72(9):661–664. [PubMed: 2126206]

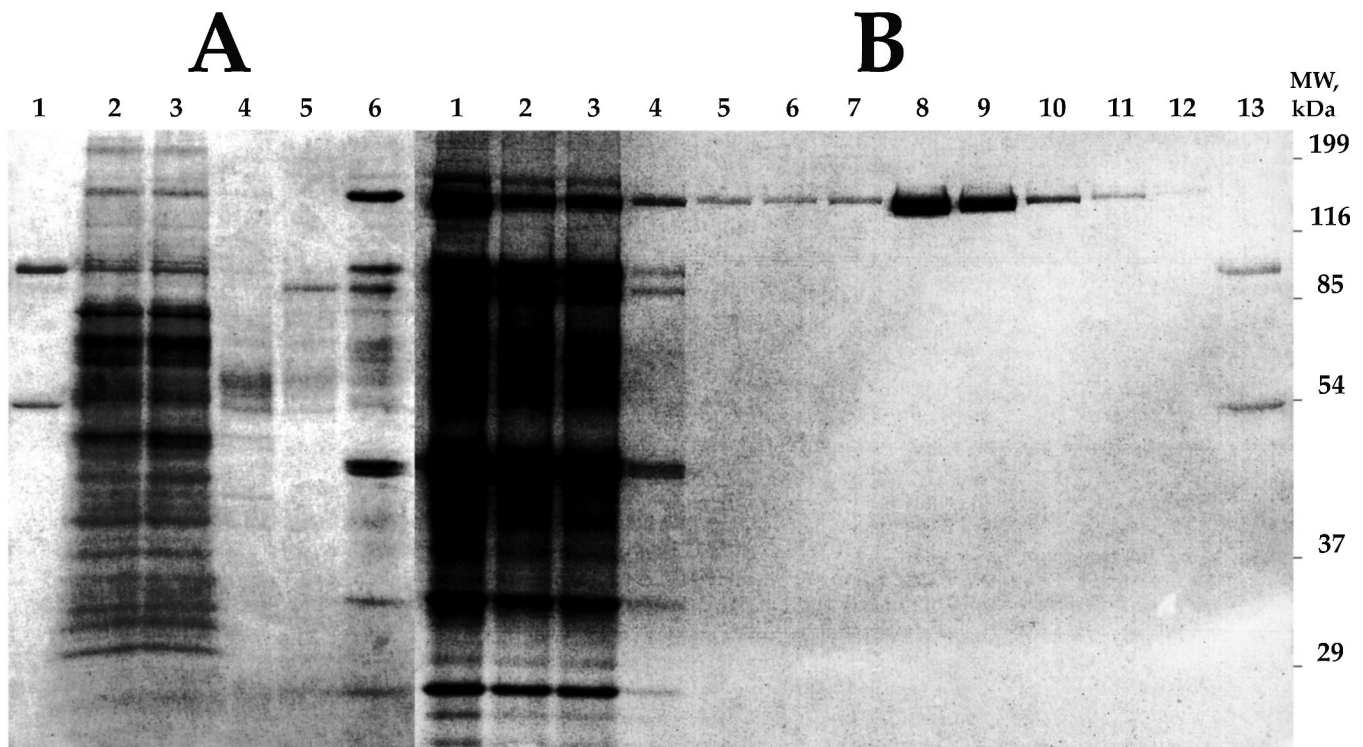
30. Kriegelstein KG, DasGupta BR, Henschen AH. Covalent structure of botulinum neurotoxin type A: location of sulfhydryl groups, and disulfide bridges and identification of C-termini of light and heavy chains. *J. Protein Chem* 1994;13(1):49–57. [PubMed: 8011071]
31. Shone CC, Hambleton P, Melling J. Inactivation of *Clostridium botulinum* type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity. *Eur. J. Biochem* 1985;151(1):75–82. [PubMed: 3896784]
32. 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(2):219–228. [PubMed: 12135553]
33. Rigoni M, Caccin P, Johnson EA, Montecucco C, Rossetto O. Site-directed mutagenesis identifies active-site residues of the light chain of botulinum neurotoxin type A. *Biochem. Biophys. Res. Commun* 2001;288(5):1231–1237. [PubMed: 11700044]
34. Li L, Binz T, Niemann H, Singh BR. Probing the mechanistic role of glutamate residue in the zinc-binding motif of type A botulinum neurotoxin light chain. *Biochemistry* 2000;39(9):2399–2405. [PubMed: 10694409]
35. Lacy DB, Tepp W, Cohen AC, DasGupta BR, Stevens RC. Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nat. Struct. Biol* 1998;5(10):898–902. [PubMed: 9783750]
36. Fernandez-Salas E, Steward LE, Ho H, Garay PE, Sun SW, Gilmore MA, Ordas JV, Wang J, Francis J, Aoki KR. Plasma membrane localization signals in the light chain of botulinum neurotoxin. *Proc Natl Acad Sci U S A* 2004;101(9):3208–3213. [PubMed: 14982988]
37. Grumelli C, Verderio C, Pozzi D, Rossetto O, Montecucco C, Matteoli M. Internalization and mechanism of action of clostridial toxins in neurons. *Neurotoxicology* 2005;26(5):761–767. [PubMed: 15925409]
38. Ahsan CR, Hajnóczky G, Maksymowych AB, Simpson LL. Visualization of binding and transcytosis of botulinum toxin by human intestinal epithelial cells. *J. Pharmacol. Exp. Ther* 2005;315(3):1028–1035. [PubMed: 16144978]
39. Taki M, Shiota M, Taira K. Transglutaminase-mediated N- and C-terminal fluorescein labeling of a protein can support the native activity of the modified protein. *Protein Eng. Des. Sel* 2004;17(2):119–126. [PubMed: 15047907]
40. Bouvier A, Chapline J, Boerner R, Jeyarajah S, Cook S, Acharya PS, Henderson I, Schrimsher JL, Shepard SR. Identifying and modulating disulfide formation in the biopharmaceutical production of a recombinant protein vaccine candidate. *J. Biotechnol* 2003;103(3):257–271. [PubMed: 12890612]



**Figure 1. Amino acid sequence and features of the expressed BoNT/A derivatives in comparison with *wt***

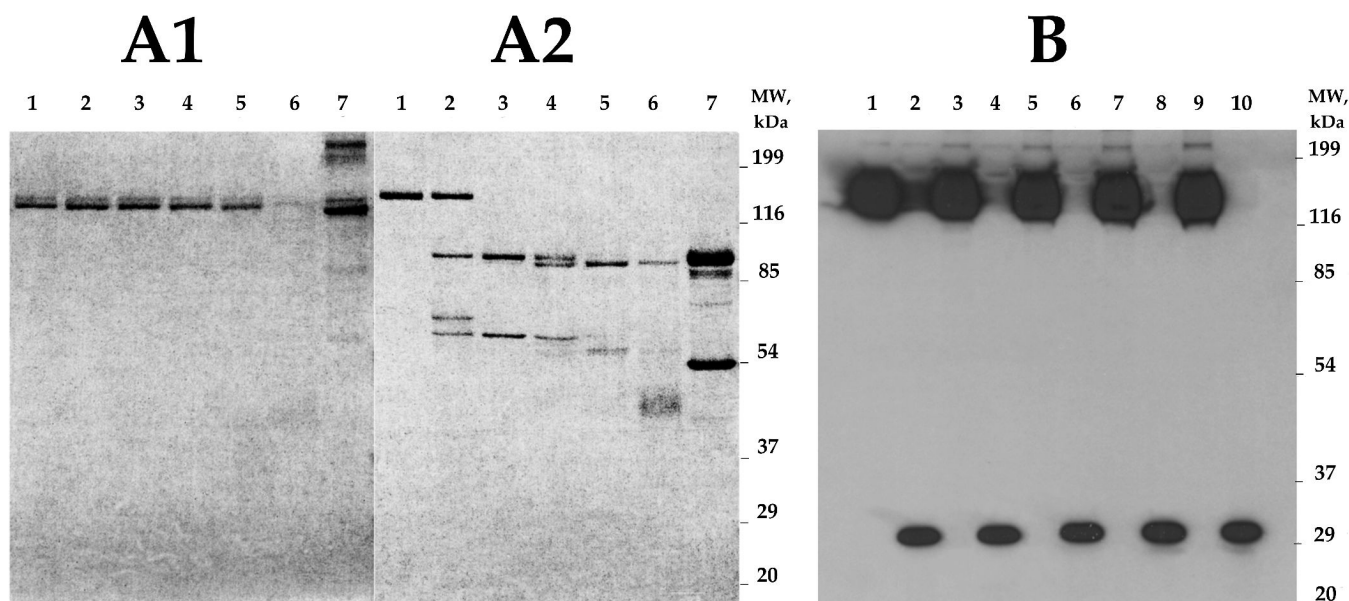
Spaces between arrowheads and arrow tails represent regions of sequence identity omitted for simplicity. Numbers in the upper row correspond to amino acid residues in *wt* BoNT/A. Residues that are identical in all proteins are shown in black. Introduced mutations are shown in red. Added amino acids 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 have been introduced 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.





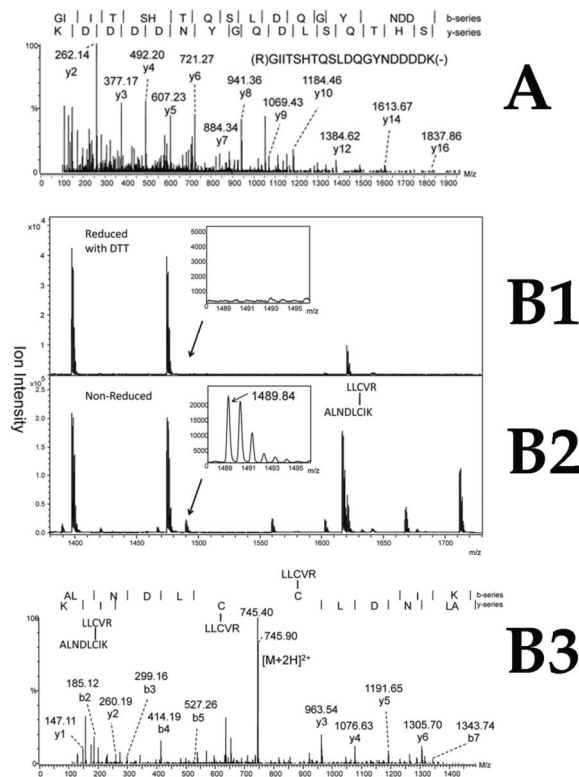
**Figure 2. BoNT/Aad<sup>ek</sup> propeptide purification**

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<sup>ek</sup> 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 3mM desthiobiotin; lane 13, *wt* BoNT/A, control.



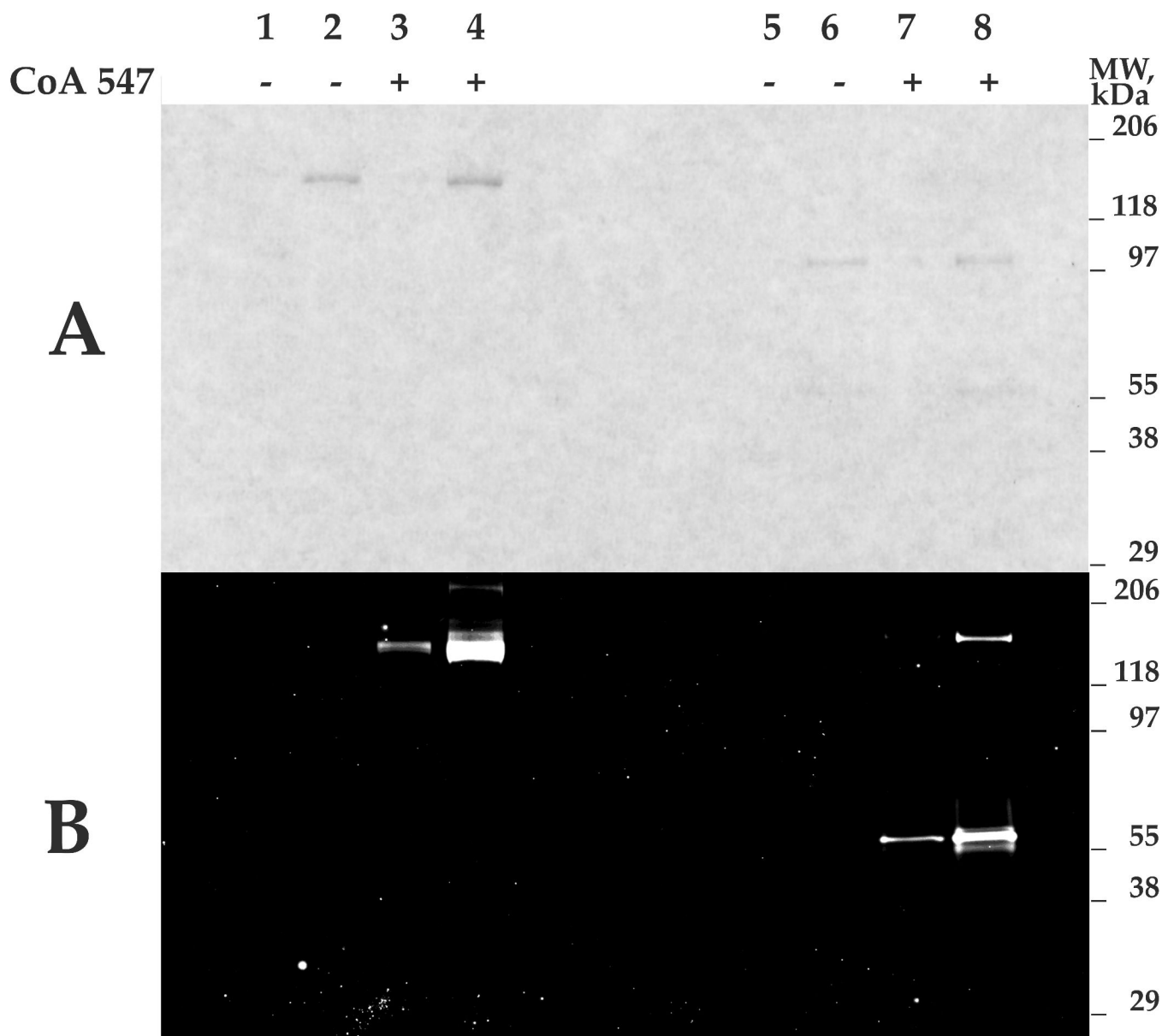
### Figure 3. Processing of BoNT/Aad<sup>ek</sup> propeptide

**Panels A1, A2:** Processing to heterodimer by proteolytic cleavage with recombinant enterokinase (rEK). One microgram of BoNT/Aad<sup>ek</sup> propeptide per lane was treated with 0 - 5 U rEK at 16°C for 12 hours, separated by 12% SDS PAGE, and stained with Coomassie BB R-250. **Panel A1:** non-reduced samples. **Panel A2:** samples reduced by addition of  $\beta$ -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<sup>ek</sup> propeptide by treatment with AcTEV. BoNT/Aad<sup>ek</sup> 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, one hour; lanes 3 and 4, two hours; lanes 5 and 6, three hours; lanes 7 and 8, four hours; lanes 9 and 10, six hours. Samples were loaded on a 12% SDS PAGE in the presence of  $\beta$ -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.



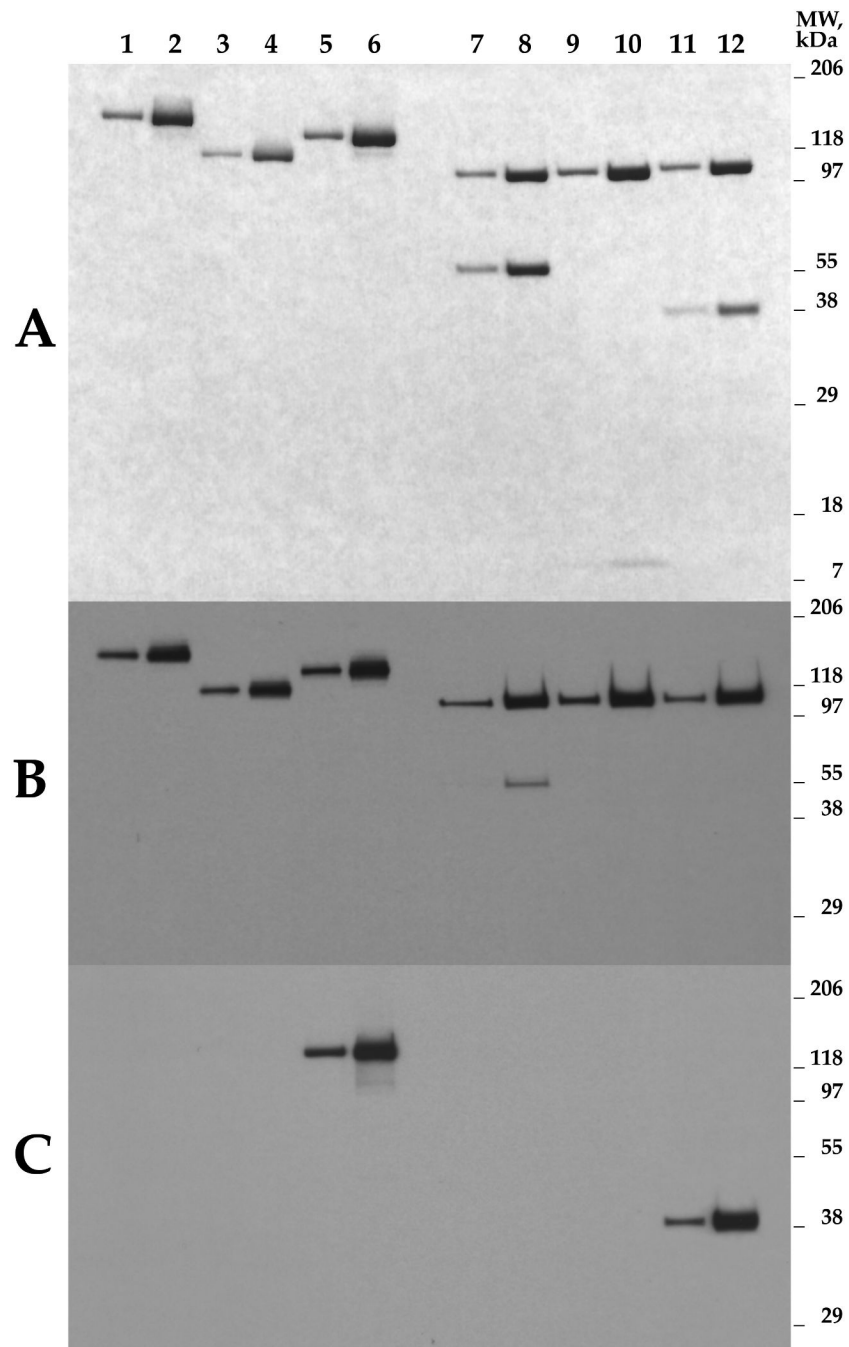
**Figure 4. Structural analysis of tryptic peptides isolated from BoNT/Aad<sup>ek</sup>**

**Panel A:** ESI Q-TOF MS/MS spectrum of the C-terminal tryptic peptide isolated from enterokinase-processed BoNT/Aad<sup>ek</sup> 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<sup>ek</sup>. **Panel B1:** MALDI-TOF mass spectrum of an in-gel tryptic digest of the enterokinase-processed BoNT/Aad<sup>ek</sup> 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<sup>ek</sup> 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.



**Figure 5. *In vitro* fluorescent labeling of AcTEV-treated and rEK-treated BoNT/Aad<sup>ek</sup> with *Sfp* phosphopantetheinyl transferase and CoA 547**

Lanes 1 - 4, unreduced samples; lanes 5 - 8, samples reduced by addition of  $\beta$ -mercaptoethanol. Lanes 1, 3, 5, 7: 0.02  $\mu$ g BoNT/Aad<sup>ek</sup>; lanes 2, 4, 6, 8: 0.1  $\mu$ g BoNT/Aad<sup>ek</sup>. **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 300V PMT, 532/580 nm excitation/emission filter set (green).



**Figure 6.** BoNT/Aad<sup>tev</sup>, ΔLC-Peptide-BoNT/A<sup>tev</sup>, and ΔLC-GFP-BoNT/A<sup>tev</sup> expressed in the baculovirus system, purified by metal chelate and StrepTactin affinity chromatography, and processed with AcTEV

Lanes 1 - 6, unreduced samples; lanes 7 - 12 samples reduced by addition of β-mercaptoethanol. Lanes 1, 2, 7, 8: BoNT/Aad<sup>tev</sup>; lanes 3, 4, 9, 10: ΔLC-Peptide-BoNT/A<sup>tev</sup>; lanes 5, 6, 11, 12: ΔLC-GFP-BoNT/A<sup>tev</sup>. **Panel A:** 10.5 - 14 % Criterion gel (Bio-Rad) stained with Bio-Safe Coomassie (Bio-Rad); odd lanes: 0.3 μg samples; even lanes: 1.0 μg samples. **Panel B:** Western blot probed with polyclonal antibody Pol001 raised against BoNT/A holotoxoid (Staten Serum Institut, Denmark); odd lanes: 3 ng samples; even lanes: 10 ng samples. **Panel C:** Western blot

probed with monoclonal antibody against GFP (Clontech); odd lanes: 3 ng samples; even lanes: 10 ng samples.

**Table 1**

## Oligonucleotides used for cloning experiments

Abbreviation:	Length:	Sequence:
C1-1S	103 b.	5'-pGAT CTG AAG ACA TAA GTC GGT CCG TTC GAA CCA GAA CTC TGG AAG CTT GAC GCG GCC GCT ATC CAT GGC ACA CGC GTT CAG CTA GCT TAG GCG CCT ATG ACG T-3'
C1-1A	95 b.	5'-pCAT AGG CGC CTA AGC TAG CTG AAC GCG TGT GCC ATG GAT AGC GGC CGC GTC AAG CTT CCA GAG TTC TGG TTC GAA CGG ACC GAC TTA TGT CTT CA-3'
C1-2S	108 b.	5'-pGCG CCT ATG CGC GCT AAC CGC GGT CAC TTA AGT ATG ATA TCT CTC TGC AGT TAC CCG GGC ATG ACG TCT ATA TGC ATA TTC TCG AGG CAT GCG AGC TCC CTC AGG AGG-3'
C1-2A	104 b.	5'-pCCT CCT GAG GGA GCT CGC ATG CCT CGA GAA TAT GCA TAT AGA CGT CAT GCC CGG GTA ACT GCA GAG AGA TAT CAT ACT TAA GTG ACC CGG GTT AGC GCG CAT AG-3'
CP1S	88 b.	5'-pAGC TTA ACT CCT AAA AAA CCG CCA CCA TGA AAT TCT TAG TCA ACG TTG CCC TTG TTT TTA TGG TCG TAT ACA TTT CTT ACA TCT ATG C-3'
CP1A	88 b.	5'-pGGC CGC ATA GAT GTA AGA AAT GTA TAC GAC CAT AAA AAC AAG GGC AAC GTT GAC TAA GAA TTT CAT GGT GGC GGT TTT TTA GGA GTT A-3'
CP2S	38 b.	5'-pGGC CGC TCA TCA CCA CCA TCA TCA CCA TCA CCA CCA CA-3'
CP2A	38 b.	5'-pCGC GTG TGG TGG TGA TGG TGA TGA TGG TGG TGA TGA GC-3'
CP3S	78 b.	5'-pCGC GTG AAA ACC TGT ATT TTC AGG GCG CCG GTG ACT CCC TGT CTT GGC TGC TCC GTC TGC TCA ACG CGC GCG GTG GCG-3'
CP3A	78 b.	5'-pCTA GCG CCA CCG CGC GCG TTG AGC AGA CGG AGC AGC CAA GAC AGG GAG TCA CCG GCG CCC TGA AAA TAC AGG TTT TCA-3'
CP7S	31 b.	5'-TCA TGC TAG CGT GAG CAA GGG CGC CGA GCT G-3'
CP7A	32 b.	5'-TAT AGG TAC CCT TGT ACA GCT CAT CCA TGC CG-3'

**Table 2**

## Purification and processing of BoNT/A derivatives

<b>BoNT/Aad<sup>ek</sup>:</b>				
<b>Step</b>		<b>Volume (ml)</b>	<b>Total protein (mg)</b>	<b>%Purity</b>
1.	Filtered <i>S</i> / <i>9</i> culture supernatant	1075	N/A	N/A
2.	Concentrated and dialyzed supernatant	310	N/A	~1.5
3.	Pooled and concentrated elution fractions from Talon® resin	20	195	~18
4.	Pooled and concentrated elution fractions from StrepTactin agarose	7.5	32.5	>98
5.	Concentrated fraction after AcTEV treatment and enzyme removal on Talon® resin	7	31.8	>98
6.	Concentrated fraction after rEK treatment and enzyme removal with Tag-off kit	5	30.4	>98

<b>BoNT/Aad<sup>lev</sup>:</b>				
<b>Step</b>		<b>Volume (ml)</b>	<b>Total protein (mg)</b>	<b>%Purity</b>
1.	Filtered <i>S</i> / <i>9</i> culture supernatant	1100	N/A	N/A
2.	Concentrated and dialyzed supernatant	310	N/A	~1.7
3.	Pooled and concentrated elution fractions from Talon® resin	18	175	~20
4.	Pooled and concentrated elution fractions from StrepTactin agarose	7	33.7	>98
5.	Concentrated fraction after AcTEV treatment and enzyme removal on Talon® resin	7	32.3	>98

<b>ΔLC-Peptide-BoNT/A<sup>lev</sup>:</b>				
<b>Step</b>		<b>Volume (ml)</b>	<b>Total protein (mg)</b>	<b>%Purity</b>
1.	Filtered <i>S</i> / <i>9</i> culture supernatant	1150	N/A	N/A
2.	Concentrated and dialyzed supernatant	310	N/A	~0.1
3.	Pooled and concentrated elution	2.5	22	~10.5



<b>ALC-Peptide-BoNT/A<sup>tev</sup>:</b>				
<b>Step</b>		<b>Volume (ml)</b>	<b>Total protein (mg)</b>	<b>%Purity</b>
	fractions from Talon® resin			
4.	Pooled and concentrated elution fractions from StrepTactin agarose	5	2.3	>98
5.	Concentrated fraction after AcTEV treatment and enzyme removal on Talon® resin	2	2	>98
<b>ALC-GFP-BoNT/A<sup>tev</sup>:</b>				
<b>Step</b>		<b>Volume (ml)</b>	<b>Total protein (mg)</b>	<b>%Purity</b>
1.	Filtered S <sub>9</sub> culture supernatant	2200	N/A	N/A
2.	Concentrated and dialyzed supernatant	330	N/A	~0.1
3.	Pooled and concentrated elution fractions from Talon® resin	3.5	35	~5.7
4.	Pooled and concentrated elution fractions from StrepTactin agarose	5	2.0	>98
5.	Concentrated fraction after AcTEV treatment and enzyme removal on Talon® resin	2	1.8	>98