RECOMBINANT DERIVATIVES OF BOTULINUM NEUROTOXINS ENGINEERED FOR TRAFFICKING STUDIES AND NEURONAL DELIVERY

Inventors: Konstantin Ichtchenko, Brooklyn, NY (US); Phillip A. Band, West Orange, NJ (US)

Assignee: New York University, New York, NY (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 783 days.

Appl. No.: 13/013,518
Filed: Jan. 25, 2011

Prior Publication Data
US 2011/0206616 A1 Aug. 25, 2011

Related U.S. Application Data
Provisional application No. 61/298,078, filed on Jan. 25, 2010.

Int. Cl.
C07K 1/00 (2006.01)
C07K 14/00 (2006.01)
C07K 17/00 (2006.01)
A61K 39/08 (2006.01)
A61K 39/02 (2006.01)
C12N 9/52 (2006.01)
C07K 14/33 (2006.01)
C01N 33/50 (2006.01)
A61K 38/00 (2006.01)

U.S. Cl.
CPC: C12N 9/52 (2013.01); C07K 14/33 (2013.01); G01N 33/5058 (2013.01); A61K 38/00 (2013.01); C07K 23/19/02 (2013.01); C07K 21/9/06 (2013.01); C07K 23/19/50 (2013.01); G01N 2333/33 (2013.01)
USPC: ................ 424/247.1; 424/234.1; 424/239.1; 530/350

Field of Classification Search
None
See application file for complete search history.

References Cited
U.S. PATENT DOCUMENTS
5,846,929 A 12/1998 Johnson et al.
6,001,806 A 12/1999 Hilbert et al.
6,022,950 A 2/2000 Murphy
6,461,617 B1 10/2002 Shone et al.
6,967,098 B1 11/2005 Williams et al.
7,172,764 B2 2/2007 Li et al.
7,785,606 B2 8/2010 Ichtchenko et al. ....... 424/234.1
8,187,834 B2 * 5/2012 Foster et al. ............ 435/69.1
8,655,186 B2 * 10/2014 Ichtchenko et al. ....... 424/234.1
2006/0024794 A1 5/2006 Li et al.
2010/0196421 A1 * 8/2010 Ichtchenko et al. ....... 424/239.1
2012/021002 A1 * 1/2012 Ichtchenko et al. ....... 424/239.1
2012/0277335 A1 * 8/2012 Foster et al. ............ 424/94.3
2012/0276132 A1 * 11/2012 Feng et al. ............ 424/239.1

FOREIGN PATENT DOCUMENTS
EP 0209281 1/1987

OTHER PUBLICATIONS
Vazquez-Cintron et al, Abstracts Toxins 2011 / Toxicon 68 (2013): 77
Abstract only.*

Primary Examiner — Nita M Minnifield
(74) Attorney, Agent, or Firm — LeClairRyan, a Professional Corporation
(57) ABSTRACT
This invention relates to isolated Clostridium botulinum proteptides and neurotoxins, isolated nucleic acid molecules encoding Clostridium botulinum proteptides and neurotoxins, methods of expression, treatment methods, and methods of detecting neurotoxin trafficking. The isolated Clostridium botulinum proteptides have a light chain region; a heavy chain region, where the light and heavy chain regions are linked by a disulfide bond; an intermediate region connecting the light and heavy chain regions and comprising a highly specific protease cleavage site; and an S6 peptide sequence according to SEQ ID NO:2 positioned upstream from, but not attached directly to, the N-terminus of the neurotoxin proteptide at the light chain region to enable site specific attachment of cargo.

34 Claims, 7 Drawing Sheets


References Cited

OTHER PUBLICATIONS


References Cited

OTHER PUBLICATIONS


* cited by examiner
RECOMBINANT DERIVATIVES OF BOTULINUM NEUROTOXINS ENGINEERED FOR TRAFFICKING STUDIES AND NEURONAL DELIVERY

This application claims benefit of U.S. Provisional Patent Application Ser. No. 61/298,078, filed Jan. 25, 2010, which is hereby incorporated by reference in its entirety.

This invention was made with government support under NIH-NIAID AI072466, NIH NINDS NS050276, NIH-NCCR RR017709, and NIH Office of the Director DP2-OD004631. The government has certain rights in the invention.

FIELD OF THE INVENTION

This invention relates to isolated Clostridium botulinum preproteases and neurotoxins, isolated nucleic acid molecules encoding Clostridium botulinum preproteases and neurotoxins, methods of expression, treatment methods, methods of detecting neurotoxin trafficking, and methods of detecting levels of neuronal activity.

BACKGROUND OF THE INVENTION

Botulinum neurotoxins ("BoNTs") are a family of structurally similar proteins that cause peripheral neuromuscular blockade and respiratory paralysis. BoNTs are exceedingly toxic, possessing an extremely low LD₅₀ (1-50 mg/kg) (National Institute of Occupational Safety and Health, Registry of Toxic Effects of Chemical Substances (R-TECS), Cincinnati, Ohio: National Institute of Occupational Safety and Health, 1996). There are seven major BoNT serotypes (BoNT A-G) and multiple subtypes (Smith et al., "Sequence Variation Within Botulinum Neurotoxin Serotypes Impacts Antibody Binding and Neutralization," Infect. Immun. 73(9):5450-5457 (2005)), but all have common structural features and a similar mechanism of action. BoNTs are synthesized as single chain preproteases with intramolecular disulfide bonds (Mr approximately 150,000; approximately 1,300 amino acids) with extensive areas of sequence homology. The majority are activated by proteolytic cleavage to generate a disulfide-bonded heterodimer containing light (approximately 50 kDa) and heavy (approximately 100 kDa) chains ("LC" and "HC" respectively).

Botulinum neurotoxin serotype A ("BoNT/A") heterodimer has been extensively studied and has been found to contain three functional domains. Toxicity is associated with metallopeptase activity confined to the LC. Neuron binding activity is associated with the C-terminal half of the HC (HC₀), and translocation activity responsible for delivering the LC protease to the neuronal cytosol is associated with the N-terminal half of the HC (HC₃) (Johnson, "Cistroidal Toxins as Therapeutic Agents: Benefits of Nature's Most Toxic Proteins," Annu. Rev. Microbiol. 53:551-575 (1999); Montecucco et al., "Structure and Function of Tetanus and Botulinum Neurotoxins," Q. Rev. Biophys. 28(4):423-472 (1995)).

The toxicity of BoNTs is a consequence of a multi-step mechanism culminating in a LC-mediated proteolytic event that disrupts the neuronal machinery for synaptic vesicle exocytosis. During BoNT poisoning, BoNTs must first cross epithelial barriers by transcytosis (Simpson, "Identification of the Major Steps in Botulinum Toxin Action," Annu. Rev. Pharmacol. Toxicol. 44:167-193 (2004)). The BoNT then passes into the circulation by an unknown pathway, from which it selectively targets the presynaptic membrane of motor neurons at neuromuscular junctions. Toxicity of the neuromuscular junction involves (i) binding to the plasma membrane, (ii) internalization into endocytic vesicles, (iii) activation within an acidic endosomal compartment that enables HC-mediated translocation of the LC into the neuronal cytoplasm, and (iv) catalytic cleavage by the LC zinc-endopeptidase of protein components in the neuronal machinery required for synaptic vesicle exocytosis.


Several laboratories have reported expressing recombinant, full-length BoNTs in E. coli. Rummel et al., “Two Carbohydrate Binding Sites in the HC, Domain of Tetanus Neurotoxin Are Required for Toxicity,” J. Mol. Biol. 326(3): 835-847 (2003); Rummel et al., “The HC Domain of Botulinum Neurotoxins A and B Exhibits a Singular Ganglioside Binding Site Displaying Serotype Specific Carbohydrate Interaction,” Mol. Microbiol. 51(3): 631-643 (2004); Rummel et al., “Synaptotagmins I and II Act as Nerve Cell Receptors for Botulinum Neurotoxin G,” J. Biol. Chem. 279(29): 30865-30870 (2004); and Bade et al., “Botulinum Neurotoxin Type D Enables Cytosolic Delivery of Enzymatically Active Cargo Proteins to Neurons Via Unfolded Translocation Intermediates,” J. Neurochem. 91(6): 1461-1472 (2004), described the expression of full-length single chain BoNT/G, D, B, and A in E. coli, either as the wt, or with a thrombin-specific cleavage site inserted between the HC and LC, or with the LC protease inactivated by a point mutation. Kiyatkin et al., “Induction of an Immune Response by Oral Administration of Recombinant Botulinum Toxin,” Infect. Immun. 65:4586-4591 (1997), reported the expression of BoNT/C in E. coli, with three inactivating point mutations (H237=D; E238=L; S239=N) in the LC protease, without the insertion of any specific proteolytic cleavage site between the LC and HC. There was no evidence that this BoNT/C single chain was processed into a disulfide-bonded heterodimer in vivo, but it was effective as an immunogen when orally administered. In all reports, the single chain holotoxin expressed in E. coli was not secreted into the culture medium or periplasm and had to be recovered from whole cell lysates. Expression problems in E. coli are associated with improper protein folding stemming from the reducing environment in the E. coli cytosol, the tendency of E. coli to segregate unfolded recombinant proteins within aggregates of inclusion bodies, proteolytic degradation, and a strong codon bias against AT-rich clostridial genes.

Recently, a recombinant, atoxic BoNT/A holotoxin was expressed in the non-toxic strain of Clostridium botulinum, LNT01, with a yield of approximately 1 mg/L (Pier et al., “Recombinant Holotoxin Vaccine Against Botulism,” Infect. Immun. 76(1): 437-442 (2008)). This recombinant holotoxin had the mutations R380A and Y496F introduced into the LC (BoNT/A), and lacked the ability to cleave the substrate SNAP 25 in vitro. Mice were challenged with up to 1 µg of this derivative (approximately 3.3 x 10^5 mouse LD50) and monitored for 96 hours. All mice survived challenge with 1 µg of single-chain or trypsin-nicked dichain of BoNT/ A1± in vivo. Immunization with this holotoxin effectively protected mice against lethal BoNT/A challenge. Although this report is encouraging, no information has yet been provided regarding the physiological trafficking of BoNT/A in vivo with BoNT/A1±.

The most recent report describes the production of catalytically inactive BoNT/A holoprotein (H237=D; E238=L; S239=A) in P. pastoris (cibBoNT/A HP) (Webb et al., “Production of Catalytically Inactive BoNT/A1 Holoprotein and Comparison With BoNT/A1 Subunit Vaccines Against Toxin Subtypes A1, A2, and A3,” Vaccine 27(33): 4490-4497 (2009)). The protein expressed from the synthetic gene, which was optimized for codon bias in the host, accumulated intracellularly. There was no introduction of an artificial cleavage site into the loop between LC and HC in the propeptide. The protein was purified in several steps with conventional ion exchange chromatographies. The yield of highly purified product was reported to be approximately 1 milligram from four grams of the frozen methlylthrotic yeast. cibBoNT/A HP provided excellent protective immunity, not only against the homologous toxin, but also against two distinct toxin subtypes with significant amino acid divergence. Mice challenged with 50 µg of this derivative (approximately 1.7 x 10^10 mouse LD50) and monitored for 240 hours did not display discernable signs of BoNT intoxication.

The selectivity of BoNT targeting to neurons has led several laboratories to consider using BoNT-based molecular vehicles for delivering therapeutic agents. Early work reported that the HC and LC of wt BoNTs could be separated, and that the wt HC could be reconstituted in vitro with either wt LC, or with recombinant LC which could carry point mutations, such as H237=D, which rendered the LC atoxic...
(Zhou et al., “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* 34(46):15175-15181 (1995); Maisey et al., “Involvement of the Constituent Chains of Botulinum Neurotoxins A and B in the Blockade of Neurotransmitter Release,” *Eur. J. Biochem.* 177(3):683-691 (1988); Suthyamourthi et al., “Separation, Purification, Partial Characterization and Comparison of the Heavy and Light Chains of Botulinum Neurotoxin Types A, B, and E,” *J. Biol. Chem.* 260(19):10461-10466 (1985)). The reconstituted BoNT holotoxin derivatives had a severely reduced ability to transport LC into the neuronal cytosol, probably resulting from the harsh conditions required for HC-LC separation and the difficulty of renaturing the protein and reconstituting native disulfide bonds. Attempts have also been made to use isolated HC for targeted delivery, by chemically coupling dextran to the HC to provide sites for attaching fluorescent markers or therapeutic agents (Goodnough et al., “Development of a Delivery Vehicle for Intracellular Transport of Botulinum Neurotoxin Antagonists,” FEBS Lett. 513:163-168 (2002)). Although this “semi-synthetic” BoNT derivative was internalized by neurons, the dextran remained localized to the endosomal compartment and the specificity of the uptake was uncertain. Direct chemical or biochemical attachment of cargo molecules to the HC of BoNTs may not be sufficient for achieving cytosolic delivery, because structural features associated with the toxin LC are required for translocation to the cytosol (Baldwin et al., “The C-Terminus of Botulinum Neurotoxin Type A Light Chain Contributes to Solubility, Catalysis, and Stability,” *Protein Expr. Purif.* 37(1):187-195 (2004); Brunger et al., “Botulinum Neurotoxin Heavy Chain Belt as an Intramolecular Chaperone for the Light Chain,” *PLoS Pathog.* 3(2):e113 (2007)). Moreover, when chemical methods are used to attach cargo to BoNT toxoids, cargo attachment is not sufficiently selective and, consequently, produces a heterogeneous population of derivatives. These problems limit the utility of chemically labeled BoNTs as probes for definitive demonstration of BoNT trafficking pathways.

Bade et al., “Botulinum Neurotoxin Type D Enables Cytosolic Delivery of Enzymatically Active Cargo Proteins to Neurons Via Unfolded Translocation Intermediates,” *J. Neurochem.* 91(6):1461-1472 (2004), described recombinant full-length derivatives of BoNT/D as effective delivery vehicles which were expressed in *E. coli* with or without an inactivating mutation (E230C-A) to the LC protease. To evaluate the delivery of prototypic cargo proteins in neuronal cultures, green fluorescent protein (“GFP”), dihydrofolate reductase, firefly luciferase, or BoNT/A LC were fused to the amino terminus of the recombinant BoNT/D holotoxin. Delivery to the cytosol was evaluated by measuring cleavage of the BoNT/D cytoplasmic substrate, synaptotagmin, dihydrofolate reductase and BoNT/A LC were reported to be effectively delivered. When luciferase or GFP were the cargo, delivery of the corresponding BoNT/D LC catalytic activity to the cytosol was significantly reduced, presumably due to the large size of the cargo (luciferase) or its rigidity (GFP) (Breje et al., “Structural Basis for Dual Excitation and Photosomerization of the Aequorea victoria Green Fluorescent Protein,” *Proc. Natl. Acad. Sci. (USA)* 94(6):2306-2313 (1997); Palm et al., “The Structural Basis for Spectral Variations in Green Fluorescent Protein,” *Nat. Struct. Biol.* 4(5):361-365 (1997)).

It has proven particularly difficult to successfully engineer translocation of recombinant toxin LCs from an endosomal compartment to the cytosol. This translocation requires acidification of the luminal milieu, either to trigger a conformational change in the BoNT heterodimer or to enable its interaction with a translocation mediator (Brunger et al., “Botulinum Neurotoxin Heavy Chain Belt as an Intramolecular Chaperone for the Light Chain,” *PLoS Pathog.* 3(2):e113 (2007); Kama et al., “Involvement of Phospholipids in the Intoxication Mechanism of Botulinum Neurotoxin,” *Biochim. Biophys. Acta.* 1199(1):65-68 (1994); Tortorella et al., “Immunochemoanalytical Analysis of the Structure of Diphtheria Toxin Shows all Three Domains Undergo Structural Changes at Low pH,” *J. Biol. Chem.* 270(46):27439-27445 (1995); Tortorella et al., “Immunochemoanalytical Analysis Shows All Three Domains of Diphtheria Toxin Penetrate Across Model Membranes,” *J. Biol. Chem.* 270(46):27446-27452 (1995)). A requirement for cooperation between the BoNT LC and the translocation domain of the HC is supported by evidence demonstrating that a decapeptide motif, common to the HCα of several BoNT serotypes as well as to diphtheria and anthrax toxins, is required for successful translocation of the LC to the cytosol (Raits et al., “A Conserved Motif in Transmembrane Helix 1 of Diphtheria Toxin Mediates Catalytic Domain Delivery to the Cytosol,” *Proc. Natl. Acad. Sci. (USA)* 102(43):15635-15640 (2005)). Future development of BoNTs as carrier vehicles will require a deeper understanding of how the LC itself, and its interaction with HCα, contributes to this mechanism.

Although efforts to express recombinant BoNTs have succeeded in producing effective immunogens, which in some cases are competent for epithelial transcytosis, these efforts have not produced recombinant proteins with the structural features required for targeting the neuronal cytosol with the efficiency of wt toxins. These limitations emphasize the importance of selecting an expression system capable of producing full-length BoNT derivatives that retain native toxin structure, disulfide bonding, and physiological trafficking.

Also, work from multiple laboratories has clarified how the structural domains of wt botulinum neurotoxin A (BoNT/A) disable neuronal exocytosis, but important questions remain unanswered. Because BoNT/A intoxication disables its own uptake, wt light chain does not accumulate in neurons at detectable levels.

The present invention is directed to overcoming these and other limitations in the art.

**SUMMARY OF THE INVENTION**

One aspect of the present invention relates to an isolated *Clostridium botulinum* neurotoxin propeptide. The propeptide has a light chain region; a heavy chain region, where the light and heavy chain regions are linked by a disulfide bond; an intermediate region connecting the light and heavy chain regions and comprising a highly specific protease cleavage site which has three or more specific adjacent amino acid residues that are recognized by the highly specific protease in order to enable cleavage; and a peptide sequence to enable site-specific attachment of cargo, where the peptide sequence is positioned upstream of the light chain region and is separated from the N-terminus of the light chain region by an amino acid spacer sequence.

Another aspect of the present invention relates to an isolated nucleic acid molecule encoding the *Clostridium botulinum* neurotoxin propeptide of the present invention as well as expression systems and host cells containing this nucleic acid molecule.

A further aspect of the present invention relates to an isolated, physiologically active *Clostridium botulinum* neurotoxin produced by cleaving the *Clostridium botulinum* neurotoxin propeptide of the present invention. The propeptide is
cleaved at the highly specific protease cleavage site. The light and heavy chain regions are linked by a disulfide bond.

Yet another aspect of the present invention relates to a method of expressing a recombinant physiologically active *Clostridium botulinum* neurotoxin. This method involves providing a nucleic acid construct having a nucleic acid molecule encoding a *Clostridium botulinum* neurotoxin propeptide of the present invention. The nucleic acid construct has a heterologous promoter operably linked to the nucleic acid molecule and a 3' regulatory region operably linked to the nucleic acid molecule. The nucleic acid construct is introduced into a host cell under conditions effective to express the physiologically active *Clostridium botulinum* neurotoxin.

A further aspect of the present invention relates to a treatment method. This method involves providing the isolated *Clostridium botulinum* neurotoxin of the present invention, where the cargo comprises a therapeutic agent and administering the isolated *Clostridium botulinum* neurotoxin to an individual in need of treatment under conditions effective to provide treatment to the individual.

Another aspect of the present invention relates to a method of detecting *Clostridium botulinum* neurotoxin trafficking. This method involves expressing a recombinant physiologically active *Clostridium botulinum* neurotoxin as described herein. A fluorophore is coupled to the neurotoxin. Trafficking of the neurotoxin is then detected by detecting one or more locations of the fluorophore.

A further aspect of the present invention relates to a method of detecting levels of neuronal activity. This method involves providing the isolated *Clostridium botulinum* neurotoxin of the present invention and administering the neurotoxin to an individual or a tissue sample. The method further involves detecting location of the neurotoxin, where detection of the neurotoxin at a specific site in the individual or tissue sample indicates an increased level of activity of neurons at that site.

The invention described herein relates to the design, expression, and purification of recombinant, full-length, BoNT heterodimers that retain all key structural elements required for native BoNT trafficking. Moreover, the expression constructs have been designed to contain a short peptide sequence that enables site selective attachment of cargo molecules using mild enzymatic conditions that do not contribute to protein denaturation. Because the BoNT derivatives of the present invention contain point mutations that inactivate the LC protease, they will be capable of accumulating in neurons at higher levels than wt BoNT, which will improve their efficiency for trafficking studies, cargo delivery, detection of neuronal activity, and therapeutic use. Using small molecule fluorophores as proteotypic cargo, these derivatives provide unique molecular tools for studying the sequential steps in BoNT translocation and targeting events, and for defining the limits of the potential cargo that can be delivered to the neuronal cytoplasm with the use of this system for site-specific cargo attachment.

A series of BoNT/A derivatives have been designed, expressed, and purified that retain the wild type features required for native trafficking. For example, BoNT/A11a" and BoNT/A11a" are full length derivatives rendered atoxic through double point mutations in the LC protease (E224A and Y350G/A). ΔLc-peptide -BoNT/A11a" and ΔLc-GFP-BoNT/A11a" are derivatives where the catalytic portion of the LC is replaced with a short peptide or with GFP plus the peptide. In all of these derivatives, the S6 peptide sequence GDSLSEWLLRLLN (SEQ ID NO:2) has been fused to the N-terminus of the proteins to enable site-specific attachment of cargo using Sf 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 of these exemplary 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. Derivatives of the present invention provide unique tools to study toxin trafficking in vivo, and to assess how the structure of cargo linked to the heavy chain influences delivery to the neuronal cytoplasm. Moreover, they enable engineering of BoNT-based molecular vehicles that can target therapeutic agents to the neuronal cytoplasm.

**BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 is an amino acid sequence alignment showing features of expressed BoNT/A derivatives, including BoNT/ Aad" (SEQ ID NO:3), BoNT/Aad" (SEQ ID NO:4), ΔLc-Peptide-BoNT/Aad" (SEQ ID NO:5), and ΔLc-GFP-BoNT/Aad" (SEQ ID NO:6), according to several embodiments of the present invention, in comparison with wild type BoNT A (GenBank Accession No. ABP48106 (SEQ ID NO:7)). Space 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 regular type. Introduced mutations are shown in italics. Added amino acids are shown in bold. A signal peptide for insect cell secretion of the expressed derivatives into medium is indicated by “SP”. Tags used for affinity chromatography are indicated: polyhistidine tag is indicated by “His”; Strep II tag is indicated by “Strep II”. “TEV” and the amino acid sequence on light shading represent tobacco etch virus protease recognition (TEV) sequence. Amino acids on dark shading represent an enterokinase recognition sequence. “S6” identifies a peptide tag used for site-specific attachment of cargo to the expressed proteins. “GFP” 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 bridge between residues of the light and heavy chains are shown as long horizontal brackets.

FIGS. 2A-B show BoNT/Aad" propo peptide purification. A reduced 12% SDS PAGE was stained with Coomassie BB R-250. FIG. 2A: TAILON® chromatography purification: lane 1, wt BoNT/A, control; lane 2, unfractionated sample of concentrated and dialyzed Sf-900 II medium containing secreted propo 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. FIG. 2B: Strep II antigen agarose chromatography: lane 1, sample of concentrated and dialyzed fraction from FIG. 2A, 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. FIGS. 3A-B show processing of BoNT/Aad" propo peptide to heterodimer by proteolytic cleavage with recombinant enterokinase (rEK). One microgram of BoNT/Aad" propo 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. FIG. 3A: non-reduced samples. FIG. 3B: samples reduced by addition of β-mercaptoethanol. Lane 1, no EK;
lanes: 3 ng samples; even lanes: 10 ng samples. FIG. 8C, Western blot probed with monoclonal antibody against GFP (Clontech); odd lanes: 3 ng samples; even lanes: 10 ng samples.

FIG. 9 is an illustration of the chemical synthesis of cargo attachment to Clostridium botulinum proteases using the S6 peptide of SEQ ID NO-2 (i.e., GDSLSWLRLLN).

FIG. 10 is a Western blot illustrating that BoNT/A ad competes with wt BoNT/A for binding to receptors and protects SNAP 25 from cleavage. Rat spinal cord cells were exposed to the indicated concentration of BoNT/A ad in ice-cold culture medium supplemented with 56 mM KCl and 0.5 mM CaCl2 for 15 min. wt BoNT/A in the same ice-cold culture medium was then added and incubation was continued on ice for an additional 15 min. Cells were then washed twice with ice-cold culture medium, fresh culture medium was added, and cells were incubated at 37°C for 3 hours. Cells were harvested in SDS sample buffer, and lysates were analyzed by Western blot using anti-SNAP 25 antibody (Synaptic Systems).

FIGS. 11A-C, are images of immunostaining that show symptoms of BoNT/A ad poisoning are concurrent with BoNT/A ad LC accumulation at the neuromuscular junction in vivo. 6 week old mice were injected ip with 1 μg of BoNT/A ad. The first signs of physiological effect of BoNT/A ad were observed —4 hours after injection (low movement activity, ruffled fur, heavy breathing), and were pronounced at the time of animal euthanasia 12 hours after injection. Triangularis sterni nerve-muscle preparations were stained with primary antibody and probed with AlexaFluor 555-conjugated secondary antibody and Alexa-Fluor 488-conjugated α-bungarotoxin. FIG. 11A illustrates the postsynaptic acetylcholine receptors; FIG. 11B is the BoNT/A LC (Mur F;1-40); FIG. 11C is a merge of images shown in FIGS. 11A and B.

The scale bar is 5 μm.

FIG. 12 is a series of images of immunostaining showing continuous uptake of BoNT/A ad in the primary culture of rat hippocampal neurons. Row 1 shows results after culture was incubated with 15 nM BoNT/A ad for 30 mins. Row 2 shows results after culture was incubated with the same concentration of BoNT/A ad for 90 minutes, followed by 30 min chase with the BoNT/A ad-free medium. Row 3 shows results after culture was incubated with the same concentration of BoNT/A ad for 90 minutes, followed by 90 minute chase with BoNT/A ad-free medium. At the end of incubation time, cells were washed and processed for immunocytochemical staining. The staining was as follows: Column A: Primary—tau mouse monoclonal IgG2b; secondary—AlexaFluor 488 conjugated donkey-anti-IgG2b-mouse; Column B: Primary—mouse-anti-BoNT/A LC monoclonal IgG1 (F;1-40), secondary—AlexaFluor 555 conjugated goat-anti-lgG1-mouse; Column C: Primary rabbit-anti-SNAP 25 polyclonal antibodies, secondary—AlexaFluor 647 conjugated goatanti-rabbit IgG. Column D is a merged image of the staining from three (Row 1 and 2), or two (Row 3) channels. Tau from Row 3 is omitted to better visualize the co-staining pattern with internalized LC ad and SNAP 25. The scale bar is equal to 10 μm.

FIGS. 13A-C are images of Western blots showing that LC ad binds SNAP 25 when BoNT/A ad is internalized by rat spinal cord cells. Cultured E18 rat spinal cord cells were either (1) untreated; (2) treated with 30 nM BoNT/A ad; or (3) treated with cross-linked/inactivated 30 nM BoNT/A ad. Following 8 hr treatment, cells were washed and fractionated according to Bernasconi et al., “Sequential Detergent Fractionation of Primary Neurons for Proteomics Studies,” Proteomics 8 (5):930-938 (2008), which is hereby incorporated by
reference in its entirety. Cytosolic extracts were immunoprecipitated with anti-SNP 25 antibodies, and protein A and Protein G magnetic beads and separated by reduced SDS PAGE. FIG. 13A shows results from input proteins probed with anti-BoNT/A polyclonal antibodies (Statens Serum Institut, Denmark). FIGS. 13B and C show the Western blot of immunoprecipitates. FIG. 13B shows results from probe with mouse monoclonal anti-LC. FIG. 13C shows results from probe with rabbit anti-SNP 25.

FIGS. 14A and B are images showing internalization of palmitoylated ΔLC-GFP-BoNT/A in COS7 cells. Palmitoylated derivative (25 nM) was added to COS7 cells and cells were incubated for 5 minutes (FIG. 14A), or for 1 hour (FIG. 14B). Image scanning was performed on a Nikon LSM 510 confocal microscope equipped with argon laser, producing an excitation line of 488 nm. The scale bar is equal to 25 μm.

DETAILED DESCRIPTION OF THE INVENTION

One aspect of the present invention relates to an isolated Clostridium botulinum neurotoxin propeptide. The propeptide has a light chain region; a heavy chain region, where the light and heavy chain regions are linked by a disulfide bond; an intermediate region connecting the light and heavy chain regions and comprising a highly specific protease cleavage site which has three or more specific adjacent amino acid residues that are recognized by the highly specific protease in order to enable cleavage; and a peptide sequence to enable site-specific attachment of cargo, where the peptide sequence is positioned upstream of the light chain region and is separated from the N-terminus of the light chain region by an amino acid spacer sequence.

BoNT/A propeptide has two chains, a light chain of Mr ~50,000 and a heavy chain of Mr ~100,000, linked by a disulfide bond between Cys240 and Cys245. All seven BoNT serotype propeptides have a light chain region and a heavy chain region linked by a disulfide bond. Two essential Cys residues, one adjacent to the C-terminus of the light chain, and a second adjacent to the N-terminus of the heavy chain are present in all seven BoNT serotypes. These two Cys residues form the single disulfide bond holding the HC and LC polypeptides together in the mature neurotoxin. This disulfide bond enables the mature neurotoxin to accomplish its native physiological activities by permitting the HC and LC to carry out their respective biological roles in concert. The intermediate region (i.e., Lys35-Arg46 of BoNT/A) identifies the amino acids eliminated during maturation of wild-type BoNT/A, and believed to be excised by a protease endogenous to the host microorganism. This cleavage event generates the biologically active BoNT HC-LC dimer.

All seven BoNT serotypes contain Lys or Arg residues in the intermediate region, which make the propeptides susceptible to activation by trypsin. Native BoNT/A propeptide recovered from young bacterial cultures can be activated by trypsin-lysis, with production of intact, S—S bound light and heavy chain. Though multiple additional trypsin-susceptible sites are present in the propeptides, they are resistant to proteolysis due to their spatial positions within the native toxin molecule (Dekleva et al., “Nicking of Single Chain Clostridium botulinum Type A Neurotoxin by an Endogenous Protease,” Biochem. Biophys. Res. Commun. 162:767-772 (1989); Lacy et al., “Crystal Structure of Botulinum Neurotoxin Type A and Implications for Toxicity,” Nat. Struct. Biol. 5:898-902 (1998), which are hereby incorporated by reference in their entirety). A second site in the native propeptide of several BoNT serotypes can be susceptible to trypsin cleavage when subjected to higher enzyme concentrations or incubation times (Chaddock et al., “Expression and Purification of Catalytically Active, Non-toxic Endopeptidase Derivatives of Clostridium botulinum Toxin Type A,” Protein Expr. Purif. 25:219-228 (2002), which is hereby incorporated by reference in its entirety). This trypsin-susceptible site is located in the region adjacent to the toxin receptor binding domain. This region of the HC peptide is found to be exposed to solvent in BoNT serotypes for which information is available on their 3-D crystal structure (Lacy et al., “Crystal Structure of Botulinum Neurotoxin Type A and Implications for Toxicity,” Nat. Struct. Biol. 5:898-902 (1998); Swaminathan et al., “Structural Analysis of the Catalytic and Binding Sites of Clostridium botulinum Neurotoxin B,” Nat. Struct. Biol. 7:693-699 (2000), which are hereby incorporated by reference in their entirety).

Propeptides of the present invention have an intermediate region connecting the light and heavy chain regions which has a highly specific protease cleavage site and no low-specificity protease cleavage sites. For purposes of the present invention, a highly specific protease cleavage site has three or more specific adjacent amino acid residues that are recognized by the highly specific protease in order to permit cleavage (e.g., an enterokinase cleavage site or a TEV recognition sequence). In contrast, a low-specificity protease cleavage site has two or less adjacent amino acid residues that are recognized by a protease in order to enable cleavage (e.g., a trypsin cleavage site).

In all seven BoNT serotypes, the amino acid preceding the N-terminus of the heavy chain is a Lys or Arg residue which is susceptible to proteolysis with trypsin. This trypsin-susceptible site can be replaced with a five amino acid enterokinase cleavage site (i.e., DDDDK (SEQ ID NO:8)) upstream of the heavy chain's N-terminus, as illustrated by the amino acids on the dark shading in FIG. 1. Alternatively, the trypsin-susceptible site can be replaced with a TEV recognition sequence, as shown by the amino acid sequence on light shading in FIG. 1. Either of these modifications enables standardization activation with specific enzymes. In serotypes A and C, additional Lys residues within this region may be mutated to either Gln or His, thereby eliminating additional trypsin-susceptible sites which might result in undesirable non-specific activation of the toxin. Trypsin-susceptible recognition sequences also occur upstream of the heavy chain's receptor-binding domain in serotypes A, E, and F. This region's susceptibility to proteolysis is consistent with its exposure to solvent in the toxin's 3-D structure, as shown by X-ray crystallography analysis. Therefore, in serotypes A, E, and F, the susceptible residues are modified to Asn.

Propeptides of the present invention also include a cargo attachment peptide sequence to enable site-specific attachment of cargo (i.e., a cargo attachment peptide sequence or cargo attachment peptide). Examples of cargo attachment peptides include an S6 sequence having a sequence according to SEQ ID NO:2. The S6 sequence enables site specific attachment of cargo using Spf phosphopantetheinyl transferase from B. subtilis, which targets the S6 amino acid of SEQ ID NO:2 as a substrate. Zhou et al., “Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Spf and AcpS Phosphopantetheinyl Transferases,” ACS Chem. Biol. 2(5): 337-346 (2007), which is hereby incorporated by reference in its entirety. In addition, an N-terminally placed 12 aa sequence GSIDLMLEWSLM ("A1") (SEQ ID NO:45) enables site specific attachment of cargo using AcpS phosphopantetheinyl transferase from E. coli, which targets the S6 amino acid of SEQ ID NO:45. Zhou et al., “Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Spf and AcpS Phosphopantetheinyl Transferases,” ACS
Chem. Biol. 2(5): 337-346 (2007), which is hereby incorporated by reference in its entirety. Another example of a cargo attachment peptide sequence is the N-terminally placed 8 amino acid sequence DSLDMLEW ("A4") (SEQ ID NO:46) that enables site specific attachment of cargo using AcpS phosphopantetheinylation transferase from E. coli, which targets the S$_{0}$ amino acid sequence (SEQ ID NO:46) as a substrate. Zhou et al., "An Eight Residue Fragment of an Acyl Carrier Protein Suffix for Post-translational Introduction of Fluorescent Panthetheiny1 Arms in Protein Modification in vitro and in vivo," J. Am. Chem. Soc. 130(30): 9925-9930 (2008), which is hereby incorporated by reference in its entirety. Yet another example is an N-terminally placed amino acid sequence (e.g., MSGLVDFEEAEKIEWH (SEQ ID NO:47)) that enables site specific attachment of cargo using biotin ligase, which targets the K$_{12}$ amino acid of SEQ ID NO:47 as a substrate. The N-terminally placed 7 amino acid sequence PKPQQFM ("Q tag") (SEQ ID NO:48) enables site specific attachment of cargo using transglutaminase, which targets the QQ amino acids of SEQ ID NO:48 as a substrate. Lin et al., "Transglutaminase-catalyzed Site-specific Conjugation of Small-molecule Probes to Proteins in vitro and on the Surface of Living Cells," J Am Chem Soc. 128 (14): 4542-4543 (2006), which is hereby incorporated by reference in its entirety. The N-terminally placed 5 amino acid sequence GGGGG (SEQ ID NO:49) enables site specific attachment of cargo using sortase A from S. aureus, which targets the Nterminus of the G$_{1}$ amino acid of SEQ ID NO:49 as a substrate. Antos et al., "Site-specific N- and C-terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity," J. Am. Chem. Soc. 131(31): 10800-10801 (2009), which is hereby incorporated by reference in its entirety.

Propeptides of the present invention may include a signal peptide coupled to the cargo attachment peptide sequence, suitable to permit secretion of the neurotoxin propeptide from a eukaryotic cell to a medium. Coupling of a signal peptide sequence to the S6 peptide sequence is illustrated in the propeptides shown in FIG. 1 by the "SP" designation.

Propeptides of the present invention may additionally include a 10-His affinity tag positioned between and connecting the signal peptide to the peptide sequence. The designation "10 His" in FIG. 1 illustrates the placement of the polyhistidine tag.

Propeptides of the present invention may additionally include a TEV recognition sequence positioned between and connecting the 10-His affinity tag to the cargo attachment peptide. For example, as shown in FIG. 1 by the designation "TEV", a TEV sequence is positioned upstream of the S6 peptide sequence and downstream from the signal peptide.

Propeptides of the present invention may additionally include an 8 amino acid Streptag II connected to the propeptide at the C-terminus (see FIG. 1, "Streptag II").

In one embodiment, the entire catalytic domain of the light chain region has been removed from the isolated propeptide. According to this embodiment, the propeptide may have a fluorophore connected to the N-terminus of the light chain region.

Attachment of cargo to the propeptides of the present invention can be facilitated by the cargo attachment peptide sequence, an amino acid sequence that allows site-selective enzyme-specific attachment of cargo to the propeptide. For example, the chemistry involved in attaching cargo to the propeptide via the S6 peptide is illustrated in FIG. 9. Specifically, CoA derivatives, having the structure

CoA Derivatives

where R (i.e., the cargo) is any prosthetic group, can be biochemically coupled to CoA, with the resulting CoA adduct recognized as a substrate for Sfp phosphopantetheinylation transferase used for enzymatic labeling, as described in Zhou et al., "Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Sfp and AcpS Phosphopantetheinylation Transferases," ACS Chemical Biology 2(5): 337-346 (2007), which is hereby incorporated by reference in its entirety.

Suitable cargo for attachment to propeptides of the present invention may include, without limitation, lipid moieties, therapeutic agents, marker molecules, and targeting agents.

Exemplary lipid moieties include fatty acids (e.g., saturated, unsaturated, greater than four carbon chain length, prostanooids, leukotienes, ecosanoids, etc.), neutral lipids (e.g., cholesterol and esters thereof, triglycerides, sterols, spermacetin (ethyl palmitate), waxes, fatty alcohols, etc.), phospholipids (e.g., phosphatidyl choline, phosphatidyl serine, ethanolamine, phosphatidyl inositol, platelet activating factor, fatty acid glycerol ethers, cardiolipins, etc.), and complex lipids (e.g., sphingolipids, ceramides, glycolipids, gangliosides, sulfolipids, etc.). In one particular embodiment, the lipid is selected from a group consisting of palmitoyl-CoA, C-22 aliphatic CoA, or cholesterol CoA. Incorporation of lipid moieties into the propeptide of the present invention anchors the propeptides to plasma membranes at the injection site and restricts its diffusion away from the site of intended action.

Exemplary therapeutic agents may include, without limitation, the peptide described in Zuniga et al., "A Potent Peptidomimetic Inhibitor of Botulinum Neurotoxin Serotype A
Has a Very Different Conformation than SNAP-25 Substrate," *Structure* 16:1588-1597 (2008) (which is hereby incorporated by reference in its entirety), which is an effective BoNT inhibitor. Other therapeutic agents include any agent with a therapeutic target in, e.g., the neural cytosol including, without limitation, agents for treating neuropathic pain, Alzheimer's Disease, and virus inhibitors (e.g., HSV2 inhibitors).

Exemplary marker molecules include, without limitation, fluorophores having a photoluminescent property that can be detected and easily identified with appropriate detection equipment to permit neurotoxin trafficking studies. Exemplary fluorescent labels include, without limitation, fluorescent dyes, semiconductor quantum dots, lanthanide atom-containing complexes, and fluorescent proteins. The fluorophore used in the present invention is characterized by a fluorescent emission maxima that is detectable either visually or using optical detectors of the type known in the art.

Exemplary dyes include, without limitation, Cy2TM, YO-PROTM-1, YOYO1M-1, Calcine, FITC, FluorXTM, AlexaTM, Rhodamine 110, 5-FAM, Oregon GreenTM 500, Oregon GreenTM 488, RiboGreenTM, Rhodamine GreenTM, Rhodamine 123, Magnesium Green, Cy5TM, TO-PROTM-1, TO-TO®-1, JOE, BODIPY® 530/550, DII, BODIPY® TMR, BODIPY® 558/568, BODIPY® 564/570, Cy3TM, Cy564, TRITC, Magnesium Orange, Phycoerythrin R&B, Rhodamine Phalloidin, Calcium OrangeTM, Pyronin Y, Rhodamine B, TAMRA, Rhodamine RedTM, Cy3.5TM, ROX, Calcium翠绿TM, AlexaTM 594, TEXAS RED®, Nile Red, YO-PRO-3TM, YOETM-3, R-phycocyanin, C-Phycocyanin, TO-PROTM-3, TO-TO®-3, DDDIC(5), CySTM, ThioflavineС, and Cy5.5TM. Other dyes now known or hereafter developed can similarly be used as long as their excitation and emission characteristics are compatible with a light source and non-interfering with other fluorophores that may be present.

Exemplary proteins include, without limitation, both naturally occurring and modified (i.e., mutant) green fluorescent proteins (Prasher et al., Gene 111:229-233 (1992); PCT Application WO 95/07463, which are hereby incorporated by reference in their entirety) from various sources such as *Aequorea* and *Renilla*; both naturally occurring and modified blue fluorescent proteins (Karantani et al., *Photochem. Photobiol.* 55(2):293-299 (1992); Lee et al., *Methods Enzymol. (Biolum. Chemilum.)* 57:226-234 (1978); Gask et al., *Biochem. Biophys. Res. Commun.* 80(1):14-21 (1978), which are hereby incorporated by reference in their entirety) from various sources such as *Vibrio* and *Photobacterium*; and phycho- bilioproteins of the type derived from cyanobacteria and eukaryotic algae (Apt et al., *J. Mol. Biol.* 238:79-96 (1995); Glazer, *Ann. Rev. Microbiol.* 36:173-198 (1982); Fairchild et al., *J. Biol. Chem.* 269:8686-8694 (1994); Pilot et al., *Proc. Natl. Acad. Sci. USA* 81:6983-6987 (1984); Liu et al., *Plant Physiol.* 103:293-294 (1993); Hounard et al., *J. Bacteriol.* 170:5512-5521 (1988), which are hereby incorporated by reference in their entirety), several of which are commercially available from ProZyme, Inc. (San Leandro, Calif.). Other fluorescent proteins now known or hereafter developed can similarly be used as long as their excitation and emission characteristics are compatible with the light source and non-interfering with other fluorophores that may be present.

Exemplary lanthanide atoms include, without limitation, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, and Lv. Of these, Nd, Er, and Tb are preferred because they are commonly used in imaging applications.

Exemplary targeting agents may include, without limitation, agents that direct trafficking of *Botulinum* neurotoxins to specific neurons and other cell types. Ligands for cell receptors (e.g., NGF, EGF, and others) and antibodies against receptors may also be used as targeting agents. Selectively incorporating lipid moieties into recombinant BoNT/A derivatives by enzymatic coupling to the S6 peptide tag, in order to restrict the diffusion of the protein adduct from the site of injection, is described herein in the Examples.

Propptides of the present invention may include a disabling mutation in an active metalloprotease site of the propeptide.

As noted supra, propptides of the present invention may have a cargo attachment peptide sequence to enable site-specific attachment of cargo. The cargo attachment peptide is positioned upstream of the light chain region and is separated from the N-terminus of the light chain region by an amino acid spacer sequence. The amino acid spacer (or linker) sequence may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21-25, 26-30, 31-35, or 36-40, or more, amino acid residues. The amino acid spacer (or linker) sequence serves to preserve and protect conformational independence of the cargo attachment peptide and the *botulinum* neurotoxin. An exemplary amino acid spacer (or linker) sequence is shown in FIG. 1 (bold type) as the 7 amino acid spacer ARGGGASG (SEQ ID NO:9).

Propptides of the present invention may further include a neuron-specific protease cleavage site. In one embodiment, the cargo is attached at the N-terminus of the propeptide's light chain via a peptide linker carrying a neuron-specific protease cleavage site (e.g., a BACE1 cleavage site). In one embodiment, the neuron-specific cleavage site may be positioned between the cargo attachment peptide and the linker or spacer sequence. In one embodiment in which the cargo is a lipid moiety, the lipided proppeptide will remain anchored to the membrane until it is released by endogenous neuron-specific protease and will be internalized by interaction with endogenous neuronal receptors, leading to release of the active light chain into the neuronal cytoplasm. In non-neuronal cells, which are neuron-specific protease deficient, lipided proppeptide can only be internalized through a non-specific mechanism and directed into a degradation pathway, thereby contributing to removal of excess proppeptide from the circulation.

Propptides of the present invention may also possess a non-native motif in the light chain region that is capable of inactivating light chain metalloprotease activity in a toxic *Botulinum* neurotoxin, as described in U.S. Patent Application Publication No. 2006/0204524 to Ichichochno et al., which is hereby incorporated by reference in its entirety.

In one embodiment, propptides of the present invention have light and heavy chains that are not truncated. In another embodiment, the entire catalytic domain of the light chain has been removed.

Another aspect of the present invention relates to isolated nucleic acid molecules encoding propptides of the present invention.

Wildtype BoNT/A has an amino acid sequence as set forth in GenBank Accession No. ABP48106 (SEQ ID NO:7), as follows:

```
Of these, Nd, Er, and Tb are preferred because they are commonly used in imaging applications.

Exemplary targeting agents may include, without limitation, agents that direct trafficking of *Botulinum* neurotoxins to specific neurons and other cell types. Ligands for cell receptors (e.g., NGF, EGF, and others) and antibodies against receptors may also be used as targeting agents. Selectively incorporating lipid moieties into recombinant BoNT/A derivatives by enzymatic coupling to the S6 peptide tag, in order to restrict the diffusion of the protein adduct from the site of injection, is described herein in the Examples.

Propptides of the present invention may include a disabling mutation in an active metalloprotease site of the propeptide.

As noted supra, propptides of the present invention may have a cargo attachment peptide sequence to enable site-specific attachment of cargo. The cargo attachment peptide is positioned upstream of the light chain region and is separated from the N-terminus of the light chain region by an amino acid spacer sequence. The amino acid spacer (or linker) sequence may comprise at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21-25, 26-30, 31-35, or 36-40, or more, amino acid residues. The amino acid spacer (or linker) sequence serves to preserve and protect conformational independence of the cargo attachment peptide and the *botulinum* neurotoxin. An exemplary amino acid spacer (or linker) sequence is shown in FIG. 1 (bold type) as the 7 amino acid spacer ARGGGASG (SEQ ID NO:9).

Propptides of the present invention may further include a neuron-specific protease cleavage site. In one embodiment, the cargo is attached at the N-terminus of the propeptide's light chain via a peptide linker carrying a neuron-specific protease cleavage site (e.g., a BACE1 cleavage site). In one embodiment, the neuron-specific cleavage site may be positioned between the cargo attachment peptide and the linker or spacer sequence. In one embodiment in which the cargo is a lipid moiety, the lipided proppeptide will remain anchored to the membrane until it is released by endogenous neuron-specific protease and will be internalized by interaction with endogenous neuronal receptors, leading to release of the active light chain into the neuronal cytoplasm. In non-neuronal cells, which are neuron-specific protease deficient, lipided proppeptide can only be internalized through a non-specific mechanism and directed into a degradation pathway, thereby contributing to removal of excess proppeptide from the circulation.

Propptides of the present invention may also possess a non-native motif in the light chain region that is capable of inactivating light chain metalloprotease activity in a toxic *Botulinum* neurotoxin, as described in U.S. Patent Application Publication No. 2006/0204524 to Ichichochno et al., which is hereby incorporated by reference in its entirety.

In one embodiment, propptides of the present invention have light and heavy chains that are not truncated. In another embodiment, the entire catalytic domain of the light chain has been removed.

Another aspect of the present invention relates to isolated nucleic acid molecules encoding propptides of the present invention.

Wildtype BoNT/A has an amino acid sequence as set forth in GenBank Accession No. ABP48106 (SEQ ID NO:7), as follows:```
An exemplary nucleic acid molecule of the present invention is set forth in GenBank Accession No. GQ855201 (SEQ ID NO:15), as follows:

gacgcgcctcgtacggcgcattaagcgcggcggggttggcgttcggggtgctgtacgcggcgcggcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctgcttctg
tggaccaccgc gagtaaccgc attgtgcttt ggagcgcggca ggcatagacc ttaaaa 3420
acacatccaa caagcctgta aaacgcctca accgccggtta caaaccgtgc ggattgctca 3480
agtctgccag cagttgctgg agccgcatcg atcactgctat taaactttaa gaaaccgat 3540
gggttatgtc aagttgggttg cttgacctca cagtccttcac ggtgttgctgt accgcggcact 3600
cgccggcag acgcaagcg ccagtatttct tgcctggctgt gcagaacggac gcaggaatgc 3660
ggtctccagc catgctccag ctattgccccgc ccattggact ttctacgagc aagttgctgt 3720
cacagctcg cctgtggccttc agagagccgg aagaacctcg ccgctgccgg ccttgccgctt 3780
ggigtgctgac ccgtatagta cgggtctgcttc ctctgtgctctt ctggaagccg acgcatgcttt 3840
ggccggccgc gacttctag tcagttgctcg tgcctggcctc gtaatacttc ggaattatgaa 3900
tagatcattg agataataaa aagatataacc acctgccgaa taataaatgta tttaattgctt 3960
tcgattaca cttgtagtta aaacataacta taatatatcc ggaattttca tacgctccca 4020
catccggggc ggagcgtccgg tcgcattcag caaacaactt gcagaagctaa ccttctagaaa 4080
acccgccacg tggaaaaccgt ataagctgctc gctgtgctgctt ctctgtgctttt ctctgctgcttt 4140	tagtctagtc gaaccagtc taggctgctctt ctctgtgctctt ctctgctgctttt ctctgctgcttt 4200
tcgccttttc ggcgtgcttcc ctctgtgctctt ctctgtgctctt ctctgctgctttt ctctgctgcttt 4260
agtctgatgata gcggcggttc tggcttctag tggcttctag tggcttctag tggcttctag 4320
gtggacatttc tgttttctggta ggtgttgttc gtcgatcct gcgtatcctgc gaaatcgtgc 4380
aacatactcgc aacaataaattt cgggcggtg acaatggctac ccctgcccgact 4440
gggtcagctgc acacgtgctta cagttgctgt gccgttgctctt gcgtgtctgtc ggtgtgtctg 4500
tctgtttgccccg gcataggtcaga ggaaccggac taccgcaggt gcgtgtgtgtgt gcgtgtgtgtgt 4560
agatattacct ctattgacacac gcctgctgcct ctctgtgctctt ctctgctgctttt ctctgctgcttt 4620	tcctgctgctctt ctctgtgctctt ctctgctgctttt ctctgctgctttt ctctgctgctttt ctctgctgcttt 4680
tgagcgtgcag cagattgacct gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 4740
tgattgctgttt gcgtgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc 4800
tccggcagcttttgcgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc 4860
ggccgtgttc gagcgtgctgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 4920
tccggcagcttttgcgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc 4980
atcagctggca aacggagtgtc gaaagaatct ctaataactgc gagggtctgtgc gagggtctgtgc 5040
gtcgtttctct gcggctcagac ggtgctgcc ctgctgctgcc ctgctgctgcc ctgctgctgcc 5100
tgctgggaag ttggtgtctgc cggcgtgcttt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5160
acacgtgcag ccacctggctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5220
cccgtggtgtgc tggcggtgtgc acatgggtgctc gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5280
aaatctccgca acatctgctgc cttgatgtcg taccgcaggt gcgtgtgctctt gcgtgtgctctt 5340
tccgtgctgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5400
ataattcttc ttctgctgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5460
tcccggacta ataatcttctt gcgtgtgtgt cggcgtgcttt gcgtgtgctctt gcgtgtgctctt 5520
aatctcaact tctgctgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5580
tccggccgctc accggtgggtt attaataactgc ggaatattgc cggcgtgcttt gcgtgtgctctt 5640
tcctgatata gcgttttctc tgcggtgttct gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5700
agcgtgctgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt gcgtgtgctctt 5760
-continued

taatctcct tggatctat ccagcaaatat taacctgctct ttaacctgata taacagcggc

gaaaaacctct ccctcgggaa cctccagcctt gacactctgg tgcagttgga gctgaagtccaa

aacttgccact tgcctccttt cccacagggc tggagcctggag cctcctcgg gctgatcgc

tactataagag cgcggggaat ggacactcct ggctcctgctt cactccttcct ctaacctgcctg

aatggaacgtc ttgatccttct gcatatcttct tctatacgta taactctctc tctctgtggtc

aactggactgc acctgtcctgt gccttctctt cctctcttctt ctctctctctt ctctctctc

aatctgcctgc gctgatcctt cctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

aatctgagctc acctgtcctgt gccttctctt cctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc

cagacaatttc cctctctctt ctctctctctt ctctctctctt ctctctctctt ctctctctc
This nucleotide sequence encodes the neurotoxin BoNT/A

Another exemplary nucleic acid molecule of the present invention is set forth in GenBank Accession No. GQ855202 (SEQ ID NO:10), as follows:

gacgogcct tcgacgogcctgtagtggtgtagtggttacgc cagcogcctgcacc 60
gtccacgcttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 120
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 180
agtcgcttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 240
ccagcttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 300
gacgcttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 360
aatccgttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 420
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 480
atccgttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 540
ccagcttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 600
aatccgttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 660
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 720
atccgttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 780
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 840
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 900
aatccgttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 960
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1020
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1080
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1140
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1200
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1260
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1320
aagcogccttc gcgacgccttcgtagtggtgtagtggttacgc cagcogcctgcacc 1380
cagggcaacta tggtgagaacg aatagacag atcggctgaga tagttgctctc actgattaag 1440
cattgtgacg tgcgagcaca agttaacctc tatatacttt agatgatttt aaaacctcat 1500
ttttatattta aagggctata ggtgaagatt catctttttgata atctoatgac caaaaatcctt 1560
tacgccgatagt tctgcttcct cacacgcgtg gaacccctgag aaaagatcaca agattcttctt 1620
tgagatcttt tttttctgcg ctaaaccctgc tgctggccaa caaaaaacac accgctacca 1680
gcgggtggttt gttgctgccga tcaagagccta ccaactcttt ttcgcaagtt aaactggttcc 1740
agcagagccg acgataccaa tactgtctctt tctggtgagc cgtagttaag ccacacactc 1800
agaaactcttg tgcacccgcgct cataactcct gctgtaagtaa tcctgttacc cgttgccgtg 1860
gctcggtcg gaatggcagtct tcttacccggg tgggaactcaa gacgatatgt accggataag 1920
cgacacgcttg ccggggctgagg gggggttcg ccggagaacgccg ccaaacgaccc 1980
tacaccggag tgcagatactc agacggctgat cctggaagaa gcgcctccgct tcggaggccc 2040
agaaagcgccg aacgctatccg gtaaccgcgg agaggctgcaacc aagggcaggg cagacgcagc 2100
ttcctcgggg gcacagctgct tgcctttcatt atgcctgctg ggtgctgcac gctctgtagct 2160
gacctgtcat ttttctcgtg ctgcctgaggg gggggagccc taaaagaaaa cgggacgaac 2220
gggtcccttt tacctgccttc gcgctcttgg gcctggttcttg tctacagcctt tttctcctgcg 2280
ctttacccct acctgctgctg tacccctatt accccttttg agtgacgcatt tacccctgctg 2340
cgactcgcca gcacggagcgc cagcggtcga ggtggcaggg aagggagaag cgggctgatg 2400
cggatattttc tctctacgcg tctggtgctgt atcccaacc gcacacccgc cgcggtaacctc 2460
ggcaaaactt tgcgttctgg gtaaataata ggtgcccttg ctaggctgggt tggggtggcg 2520
caataagatct taaaactggg aaaaaatgct taaaactagct aacataaag tctaaaactag 2580
acagaactag tggcataactt gcctgcgctt gttgagtttc gaaaaacgact ggtgtctctgct 2640
tggttatagcc aagggacact ttatcttttt tggataccca attgcctcct gtattaaaag 2700
gggggtgcgc ccggggtctg gtaagacact tattccgctgc gttgtgacaaa ctctacggcag 2760
aacctcgccg ccggagaggcc gagtccctgcg gccagaacc tggctgggtgc ggggtcgcttgg 2820
tggtatatccg tctcttcctct atgcctacaact tggattgttag agccatctgg cgggctgctc 2880
gggtggctgaga cgttaacgtcc ttgcaagtag ttttcatactc ccaacacgagc gttggcctca 2940
tggtggaggg gaatggctgac gcggggtgca atgcctgctc tccggctgctg gcgggctgac 3000
gcaggatactt atctaccccg ggtgctgctag aacctgccttg gctctctacca aacctaaatcg 3060
ggagagccg cacaacgctgg ttctctgctgtt aaggggacggt ggcgggctgctg cggactggcggt 3120
gcggcagatg ccgaaaggtc taacgagctgc gcggctgctgtt ccgggatgagtt ggtcgtcatcc 3180
ccggcctcaac gcggacaaac gtccggagca cggggacagt agggtgctttgtg cgggctggcgg 3240
agctacatag tggcaagttg gcaccataca ctggctggatc tcttggttcgctt gggggagctg 3300
ttcgctgtccg gctaaaaagctt gcctggtcgtt ctctggagtc acacttcacctt gatgctcttg 3360
tggacacagaag ggtgctgctgg atctgctacaagaac gttggctgctg ctgctgctgctg 3420
acagtctacat ctggccagttc aaaaagctgac gcccgctggtgc gccggctgctg ctgctgctgctg 3480
-gggtggaggt cagtgctgtg atgggatacag cgctggtcgtt ggggtggcgg ggtggcggcgg 3540
-gggtggaggt cagtgctgtg atgggatacag cgctggtcgtt ggggtggcgg ggtggcggcgg 3600
ttcgctgtccg gctaaaaagctt gcctggtcgtt ctctggagtc acacttcacctt gatgctcttg 3660
gggtggaggt cagtgctgtg atgggatacag cgctggtcgtt ggggtggcgg ggtggcggcgg 3720
ccgggactcc gccgggtctg gggggagctg cggggactcc gccgggtctg gggggagctg 3780
-continued

gggtctgacc cggacagtcg tggctgggct cctctgct ggacggtgct acgctgttct 3940
gtccgccag cgcgtcctg catcttctcg ggtgcttgtt cttgatcatt tggatggtcct 3900
tagtaactg gtaaaaata gggtaaatgt taaccgtaga tcagatagtg 3960
tcctgactgc ccctgacctc tacttctcg cctcttgtta cctctgcact 3920
tgaatggggg ggaagaacgcttg tggctgctcc gggcagcgttc gggcagcgttc 4020
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4080
tggagctcttc ctcagccttt gaaataaatg gggcagcgttc gggcagcgttc 4140
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4200
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4260
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4320
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4380
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4440
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4500
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4560
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4620
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4680
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4740
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4800
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4860
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4920
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 4980
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5040
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5100
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5160
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5220
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5280
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5340
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5400
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5460
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5520
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5580
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5640
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5700
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5760
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5820
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5880
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 5940
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 6000
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 6060
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 6120
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 6180
tcacttgca gggcagcgttc tggctgctcc gggcagcgttc gggcagcgttc 6240
This nucleotide sequence encodes the neurotoxin BoNT/Aad."n
Another exemplary nucleic acid molecule of the present invention is set forth in GenBank Accession No. GQ855203 (SEQ ID NO:11), as follows:

gagcgcgcc ttgacggcgc attaagcgcg ggggtggtgg tggtaacgac cagcgtgacc 60
getacaacttg ccaagccgtct agcgcgccct ctttctgctt ttcttctctc ttttcttgcc 120
agtgcgcggt gtcttccgct tcaagactaa aacgagcgcc cgcttttgag ggctcagatt 180
agtgcctttcc ggcaccctcc ccaaaaaaaa ctggatattg tgatgccgct ctcgtagttg 240
cctgcgcctt gatagacggt tttgccgctt ttcagctgtgg agtcacgtct cttaataagt 300
ggactcttgtg cccaaacgctt attacacttct cggct cacgttcctttttgttattttgta 360
tagggaattt tgcgatattc ggcctattgg taaaaaatgtc gtacatattg aaacaattt 420
aacgccgaatt tsacaaatgg ataatcgtt accattcccc gggccggatt tggggaaat 480
gtggagggaa cccctatattttatatgg taaaaatgtc gatagattg aacgacgacg 540
agaccaatac tctgatatg gccacatgaa tattgaaagg gcagccacgtag tgcgatgg 600
cattctcgtt tgcgacctt cttcttcttt gccggtactt ttcctgctca 660
ccagaacacgg ttgtagaattt aaaaagatct gaaatcagt ggcggctcgg ctgggtcgacag 720
atctcaagct atccaaacag ggcacctatt gtcgacgag ttgagcggag aacaacgtttt 780
ccatcaagct gcacacatc cagctgtgg ctttaaatcag cctcagatg gctcagatcag 840
ggagaagcc ctggccctcg cccgtccctt ctagaagct ggaagagctt cggcgggttca 900
cactcagca ccaacgtacg ttgctgatttc gatcataatc acaatcagaga attacagcag 960
atatccctca tctgacgagc ccgcgacctt cattacctag gacagagaaa ggcacgagct 1020
gagcgtcatgct ctttttttctt cccatggtgct tttcctgctg aacgactattt 1080
cggcagctat agaagacatc acacaaggac cagcgtgaccc ccaagctgctgg tggggtcatt 1140
gcaacactg tgcgcaacact atataagctct gcgtctacct tatactcttt aaccagagtt 1200
atatagact ggtagagcgc gataaagtt ggagaccccc tttattcgcg ggcctctcgg 1260
getgtagtgc tttaggtct taaattcag ggcgctgacg tggagtcatg tggggtcatt 1320
gcgcctcctg ccgcgactc ggggctctc tttctctgtc ggcacggggtt 1380
cgcagacactg tgcgcatgg aactagccgc atgctgatgct gccggtgtata 1440
catatgctac tgcacagct gagttctcat actttactct ctttaaatcg cagacgagtt 1500
attttataa aacgatctca ggtgagatct ttttttctttt gacatctcat ccaagcctttt 1560
atatagact ggtgtagct tttccgcc gatcagctctct ctggattgtc aagacgacgct 1620
tgacggtgct ctttggtggc ctgaaaatcc ggggctctc aaccacgctattttgt 1680
gccggtggttt gtttgcgctg tcaagagctg ctcaggtcttt ttcgaagcat gcagttttct 1740
agcagagcgc gtagatccac tactttctcc cttgtaggc cggcgggttt 1800
-continued

gttggagcta ggcgggtcac caaagccac ggttaacggtg gtaatctgag aaacacaac 4320
tggtgccca acctcaacgg ccaacatacc gcacatttt tactacatcct tac 4380
aagaacctta ctaggctgtg cacgacacac agctgctctc gctggcagctg gttatcacc 4440
tccattactag aaagagttcccc cacagccacac ggttactttc ctagaacttc 4500
acagttgctg gttttgagttt gtttctttat gtaagataggc gtaaaccattt 4560
aagaaacagtcc taacagcccc agaagatgagc gtcggtggtaa cttttt 4620
gagaaatcct cttttggcttc caatacaggta ctttttaacct tt 4680
cacgaacctta tactcattgag ggtacctctc caaacctttac ctacattaac 4740
cagccacagc gaaacatctt gtaacactgc ggacaggtcct tagatgtttt 4800
cactattcctag aagagatttca ctttttatgta cattttc 4860
gttacactagt cccggtgactt gcggtactggt ggtttttttttc ttctacactc 4920
acagttgctg gttttgagttt gtttctttat gtaagataggc gtaaaccattt 4980
taccagcattc ccaagagagc cgacctaggtg gggaggtttg 5040
atgatagcctc aatagatgcctt cttactgggtg 5100
tggtgatgtggt tagggtagttt ggttaaaccgg 5160
gctgggtagttg tggcagccaga aatcctagtttcc caataacttc 5220
tgtagagctt ccaccccttctttaaagt gtcagtcggac aata 5280
tacagttgctg gttttgagttt gtttctttat gtaagataggc gtaaaccattt 5340
agtactgcatt gcaggggttt gcagaggttag gagagttttttttc ttctacactc 5400
aaccataactgc aagttttagc ctttcctacagtt cttttatggct ggaagccagtct 5460
agaaatcctc ggcggtgactt gcggtactggt ggtttttttttc ttctacactc 5520
taccagcattc ccaagagagc cgacctaggtg gggaggtttg 5580
agaaatcctc ggcggtgactt gcggtactggt ggtttttttttc ttctacactc 5640
agaactagctagc aatagatgcctt cttactgggtg 5700
agtactgcatt gcaggggttt gcagaggttag gagagttttttttc ttctacactc 5760
aaccataactgc aagttttagc ctttcctacagtt cttttatggct ggaagccagtct 5820
agaaatcctc ggcggtgactt gcggtactggt ggtttttttttc ttctacactc 5880
agaactagctagc aatagatgcctt cttactgggtg 5940
agaaatcctc ggcggtgactt gcggtactggt ggtttttttttc ttctacactc 6000
agaactagctagc aatagatgcctt cttactgggtg 6060
agaaatcctc ggcggtgactt gcggtactggt ggtttttttttc ttctacactc 6120
agaactagctagc aatagatgcctt cttactgggtg 6180
agaaatcctc ggcggtgactt gcggtactggt ggtttttttttc ttctacactc 6240
agaactagctagc aatagatgcctt cttactgggtg 6300
agaactagctagc aatagatgcctt cttactgggtg 6360
agaactagctagc aatagatgcctt cttactgggtg 6420
agaactagctagc aatagatgcctt cttactgggtg 6480
agaactagctagc aatagatgcctt cttactgggtg 6540
agaactagctagc aatagatgcctt cttactgggtg 6600
agaactagctagc aatagatgcctt cttactgggtg 6660
This nucleotide sequence encodes the neurotoxin ΔLC-Peptide-BoNT/A™.

Another exemplary nucleic acid molecule of the present invention is set forth in GenBank Accession No. GQ855204 (SEQ ID NO: 12), as follows:

gacgogogctg ttagcgggagct ataagagcgc cgggctgtgg ctagttacgc cagctgugcc  60
gcgtcactgg ccacgctctg cccttcttgct cctcccttctt cctctgtgg gctttttactg  120
agctgcctctg ccctctcttct ttctctctttct ttctttctttct ttctttctttct  180
agctgcctctg ccctctcttct ttctctctttct ttctttctttct ttctttctttct  240
ccatgcctctg cctctctcttct ttctctctttct ttctttctttct ttctttctttct  300
ggctctctct ttctctctttct ttctctctttct ttctttctttct ttctttctttct  360
taagggattg tgcctgcgttg gcttactggt cctttaagttg cccctcccttct ggttgctgtgg  420
aagggcctatt ttaaaggtttc ccctctctttct ttctttctttct ttctttctttct  480
gtgcgcctgg ccctctctttct ttctttctttct ttctttctttct ttctttctttct  540
agccactttc ctgtagatcc atctcgatgg cagcagcttg gcttctcttg ccttctttctttct  600
ccatgctcttg gcttactggt cctttaagttg cccctcccttct ggttgctgtgg  660
acacccagtc ctgtagatcc atctcgatgg cagcagcttg gcttctcttg ccttctttctttct  720
atgacatgg gcttactggt cctttaagttg cccctcccttct ggttgctgtgg  780
ccacacaccc gcttactggt cctttaagttg cccctcccttct ggttgctgtgg  840
acacccagtc ctgtagatcc atctcgatgg cagcagcttg gcttctcttg ccttctttctttct  900
ccacacaccc gcttactggt cctttaagttg cccctcccttct ggttgctgtgg  960
atactgatga cgcgttttta agtcggttta cggggttgg ggctggtgt 1020
gagtcattg ggtggtggtt cgggttgg ccaggtgttg ccaggttgct 1080
tccgggtgctt aacacatcct tgcctccct cccttctc 1140
tgtcctctt gcaggtgttg ccggttgg ccaggttgg 1200
tgcctccct tgcctctc ctctctt cccttctc 1260
tgctctctt tgcctctc cccttctc 1320
tgcctctc cccttctc cccttctc 1380
tgcctctc cccttctc cccttctc 1440
tgcctctc cccttctc cccttctc 1500
tgcctctc cccttctc cccttctc 1560
tgcctctc cccttctc cccttctc 1620
tgcctctc cccttctc cccttctc 1680
tgcctctc cccttctc cccttctc 1740
tgcctctc cccttctc cccttctc 1800
tgcctctc cccttctc cccttctc 1860
tgcctctc cccttctc cccttctc 1920
tgcctctc cccttctc cccttctc 1980
tgcctctc cccttctc cccttctc 2040
tgcctctc cccttctc cccttctc 2100
tgcctctc cccttctc cccttctc 2160
tgcctctc cccttctc cccttctc 2220
tgcctctc cccttctc cccttctc 2280
tgcctctc cccttctc cccttctc 2340
tgcctctc cccttctc cccttctc 2400
tgcctctc cccttctc cccttctc 2460
tgcctctc cccttctc cccttctc 2520
tgcctctc cccttctc cccttctc 2580
tgcctctc cccttctc cccttctc 2640
tgcctctc cccttctc cccttctc 2700
tgcctctc cccttctc cccttctc 2760
tgcctctc cccttctc cccttctc 2820
tgcctctc cccttctc cccttctc 2880
tgcctctc cccttctc cccttctc 2940
tgcctctc cccttctc cccttctc 3000
tgcctctc cccttctc cccttctc 3060
tgcctctc cccttctc cccttctc 3120
tgcctctc cccttctc cccttctc 3180
tgcctctc cccttctc cccttctc 3240
tgcctctc cccttctc cccttctc 3300
tgcctctc cccttctc cccttctc 3360
tgcctctc cccttctc cccttctc 3420
acagtcataa caagcattga aacccgcaac tcgcgctgta ccacgctgcg gttggtcga 3480
ggttccgac cagttgctg acgtgccataa cctctgcgtc taacagttc gaaacgcaaca 3540
ggttttgtc aacagtgcctg ggcgttcgcgt ggtcctgtcgt gctcgtgacg aacccgcaac 3600
cctgggtgac aggcgacgc ggctaaattt gcctctgccgt gcggcgcggcc gcgcgggtcg 3660
ggtgcttccag gctcgtcgcc tgcgctgctg ttcacgcagca ggttgctgcg 3720
cgcgtgcttg ctcgtcgtcgcc ggtggttcggt ggtgcgtgcg gtcgctggtgcg 3780
ggtgcttgac ccggtgtaac gttcgtgcgt gctcgtgctt gcgcggtgatg gctcgtggtg 3840
gttcggcctg gctcgtgctg ttcacgccg gttggtccag caagcagttc gggcggctga 3900
tagatcgctct ggcgatattaa aaagtataacct acgcgtcggata aaaaaatgtt ggtgtgcttc 3960
ttcgtaaacggatcttgataaaaaacttaa taataataattgc gggcgtattac ccgttacca 4020
cagcggcctg cgcggccgtg tcggcgcgctg ccggacactt ggcgagttcct ctccctataa 4080
cgcggacacagtgcgatctc gctacggtgg ctggtggttaa ggtggttggat ggtggtgtgc 4140
tacagcataag gggtgcggcgc gccgctgcgc gcgtcgtgcct gcgggtgtgcg 4200
tgtctttttt tgggtgcgctg gtcttggttcggt ggtggttggtgcgtc gcgcgtggtgcg 4260
ggtggtggcg gccggcggcc ggtggttcggt tcgcgctgtcgt ccggtcgtgcg 4320
cacagcttcctt ggcggttgac ggctggtgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 4380
cacagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 4440
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4500
cacagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 4560
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 4620
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4680
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4740
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4800
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4860
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4920
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 4980
tggttaacag acttgctgatg cgtctgatg ggcgtgccgg gguagtcgac ccgtcctgac ggcgtggtgcg 5040
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5100
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5160
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5220
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5280
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5340
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5400
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5460
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5520
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5580
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5640
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5700
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5760
acagcttcctt ggcggttgac ccgtcctgac ggtggttggtc gcgcgtggtgcg 5820
-continued

gcctgtacct ctacggyttgc cggtatctgg cttgagctca ttctgaaat ccgcatccot 5880
gtgcctggca cgctctgctt ggttcctcatc atagcgaata aagctctgac ccgtgacaga 5940
atcagaatacg ctctctccaa acctaaaccaaat aatagacgag cgcttcatcag aatcagctg 6000
accaactggc tggacaaggt taacccacccaa attagcttgca tcggagaaag aatgaaggag 6060
gtctcctgga aaccaagttg acgctactaa gcctattacca atacacgtag taaccaagtct 6120
aacacaaagg aaaacaattta cataatcttc acatcagtag atctgctctc aataaatgaa 6180
gatgcctcata caacagcgtat cctacacatc caacaaatcc tgaatcagctg cctcgctcct 6240
tacctctaga actcctagat ccctacacggt gttgaagcgcgcc cctgaagaccc ctctgcagcag 6300
cgggaagg ccagtctcagc atacacattc gataactgeg gcggcttgat tgtgccaggt 6360
gaatcgtcag agaacaggtt taacaatattcct tgtgacaccag atatatccttta tgaatgttcc 6420
aaatcaggtg ataaccacagcct tgtgctctct actttacacg aataccattaa ccaacattttc 6480
aacacatccaa ttcttacccct ggcgatatgc tctacatcata tgatgatcct gttccagggct 6540
gccacagaag aacacatgagg ccagccagagtt atccttcatgc cgcctgaaa gccaacatcc 6600
cagttgcttc acctccgaaag ctccaaatcc aagctctgacc taaggaatacg catgctcttc 6660
aactccagt tctgaaatgtt cctacacatcct tttgctgatta gcacccggt aaatctcacc 6720
taatctctcg tgtgaaagttt ataacacgtata aattaactgtg tggaaagataat cctgggttg 6780
aagctctctct tgaactatag tggactctgcg aagatcggcg agagatcggaa 6840
caacgttggag tgttttaaat gacgccagatgc atacaattcat ctgactcatac caaacagattg 6900
atctttgtcaca cctataacaac caactcttcttg aataactctc aaccatctac caacgtctgt 6960
tgatcagcag aagctacatc tccaaactcct ggcagacctt atgctctccaa taacattool 7020
ctgaaagctg aagctttcagc ctgaaacacg atcagctgatctg gaaagtggtcctt gacaccagct 7080
aggtaagg tgtagttggccg ctagggcccc gccagccccct tttggtcttg 7140
gccctctcgga aacatcctctg ggggccacac cttccgctacg ataagccata ttaatctctg 7200
aacatcttatg accttccacaa atatcgntgc gttgaancatg cctgtctcag tggccatctg 7260
tacctcaag aacaggctggg tagcgttttgt gaaacaccaca tttacgctagaa tagctctcttg 7320	
tacggcgtg aagatcctctt tatacaagag tagctcctcag gcaccaaggg caaacatgtg 7380
aggaacattg atcgctgtgtata cctacaacggc ctgggagacg aataagggca cegttgggcg 7440
acacacgtct ctcacggcttg aagttgcagag atctcgagcc ccttgggtatg ccacagcctt 7500
gccacactgga gccaagctgct tgtgatgaaaaga agaagagatt ccagagctat ccaacaaaaa 7560
tgcacacata gacccactagc cagccactgcg aacagcatct ctttcacagt ttctccacag 7620
ctacacatgta tgcagcagcct ccaggttctc aacagcgcat gctgaggtca 7680
tccgctacct cctggtcact gcggagaaaatt ataaccatcctt ggggagaaattg 7740
cccctgtggtgg ccagctaggct caacccttcatct tgtgagaaagct atataattaa agataataa 7800
agctggagct atcgctgtct caacccttcatct ccagctagcct cagctcgact gctcgggcct 7860
getttcgcttg aatagctgggtcagcc ccgctgcctgctctggagtt ctgaggtgcc 7920
tagctgacta taataactgtcta caataaaa acttgcatcag aatcctcactt ccataaaaga 8000
ctttcttctc tggctctctct gcctctcctcact gtatcatttct ttaagctctgc 8060
tggatgctgta cctgtttgta gccctaggata ccgctgcctgctctggagtt ctgaggtgcc 8120
aatcttttatt cttcttttat gctgtaagcttt gctgctgctt gcccacaattt 8180
This nucleotide sequence encodes the neurotoxin ΔLC-GFP-BoNT/A**.

Nucleic acid molecules of the present invention also include nucleic acid sequences that are at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to the nucleic acid molecules of SEQ ID NOs: 10-12 and 15.

Nucleic acid molecules of the present invention may encode the amino acid sequences of FIG. 1. In particular, the nucleic acid molecules of the present invention are modified from the wild type BoNT serotype sequences to have one or more characteristics shown in FIG. 1. Other modifications may include, without limitation, a mutation which renders the encoded protease resistant to low-specificity proteolysis, one or more silent mutations that inactivate putative internal DNA regulatory elements, and/or one or more unique restriction sites. Mature neurotoxin stability and yield may be optimized by site-directed mutation of residues within the intermediate region of the protease, thereby reducing the protepese's susceptibility to non-specific proteolysis and poisoning of the host organism used for expression of the mature neurotoxin. Also, silent mutations are introduced into DNA regulatory elements that can affect RNA transcription or expression of the proteptides in the expression system of choice.

A nucleic acid molecule of the present invention may have a disabling mutation in a region encoding an active metalloprotease site of the protease.

A nucleic acid molecule of the present invention may also have a mutation in a region encoding the light chain region, such that the nucleic acid molecule encodes, in the light chain region, a non-native motif capable of inactivating light chain metalloproteinase activity in a neuron intoxicated by wt Clostridium botulinum neurotoxin.

A nucleic acid molecule of the present invention may have a mutation encoding one or more of the following mutations in the neurotoxin: Δ124–A, ΔY260–A, ΔK43–H, ΔK48–Q, ΔK44–Q, and ΔK67–N.

A further aspect of the present invention relates to an expression system having a nucleic acid molecule encoding an isolated Clostridium botulinum neurotoxin protease of the present invention in a heterologous vector.

Yet another aspect of the present invention relates to a host cell having a heterologous nucleic acid molecule encoding an isolated Clostridium botulinum neurotoxin protease of the present invention.

Still another aspect of the present invention relates to a method of expressing a recombinant physiologically active Clostridium botulinum neurotoxin of the present invention. This method involves providing a nucleic acid construct having a nucleic acid molecule encoding an isolated Clostridium botulinum neurotoxin protease of the present invention. The nucleic acid construct has a heterologous promoter operably linked to the nucleic acid molecule and a 3' regulatory region operably linked to the nucleic acid molecule. The nucleic acid construct is then introduced into a host cell under conditions effective to express the physiologically active Clostridium botulinum neurotoxin.

In a preferred embodiment, the expressed neurotoxin is contacted with a highly specific protease under conditions effective to effect cleavage at the intermediate region. Preferably, the intermediate region of the Clostridium botulinum neurotoxin protease is not cleaved by proteases endogenous to the expression system or the host cell.

Expression of a Botulinum neurotoxin of the present invention can be carried out by introducing a nucleic acid molecule encoding a Botulinum neurotoxin protease into an expression system of choice using conventional recombinant technology. Generally, this involves inserting the nucleic acid molecule into an expression system to which the molecule is heterologous (i.e., not normally present). The introduction of a particular foreign or native gene into a mammalian host is facilitated by first introducing the gene sequence into a suitable nucleic acid vector. “Vector” is used herein to mean any genetic element, such as a plasmid, plage, transposon, cosmid, chromosome, virus, virion, etc., which is capable of replication when associated with the proper control elements and which is capable of transferring gene sequences between cells. Thus, the term includes cloning and expression vectors, as well as viral vectors. The heterologous nucleic acid molecule is inserted into the expression system or vector in proper sense (5'→3') orientation and correct reading frame. The vector contains the necessary elements for the transcription and translation of the inserted Botulinum neurotoxin proteptide-coding sequences.

U.S. Pat. No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference in its entirety, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including prokaryotic organisms and eukaryotic cells grown in tissue culture.

Recombinant genes may also be introduced into viruses, including vaccinia virus, adenovirus, and retroviruses, including lentivirus. Recombinant viruses can be generated by transfection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt W35.t.b.2, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pAC184, pUC8, pUC18, pUC19, pLG339, pHR20, pKc37, pKc101, SV 40, pBluescript IIISK +/- or KS +/- (see “Stratagene Cloning Systems” Catalog (1993) from Stratagene, La Jolla, Calif., which is hereby incorporated by reference in its entirety), pQG, pHR21, pGEX, pFastBac series (Invitrogen), pET series (Studier et al., “Use of T7 RNA Polymerase to Direct Expression of Cloned Genes,” Gene Expression Technology Vol. 185 (1990), which is hereby incorporated by reference in its entirety), and any derivatives thereof. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, mobilization, or elec-
troporation. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, N.Y. (1989), which is hereby incorporated by reference in its entirety.

A variety of host-vector systems may be utilized to express the *Botulinum* neurotoxin prepeptide-encoding sequence in a cell. Primarily, the vector system must be compatible with the host cell used. Host-vector systems include, but are not limited to, the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmids DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria. The expression elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

Different genetic signals and processing events control many levels of gene expression (e.g., DNA transcription and messenger RNA ("mRNA") translation).

Transcription of DNA is dependent upon the presence of a promoter which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eukaryotic promoters differ from those of prokaryotic promoters. Furthermore, eukaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a prokaryotic system and, further, prokaryotic promoters are not recognized and do not function in eukaryotic cells.

Similarly, translation of mRNA in prokaryotes depends upon the presence of the proper prokaryotic signals which differ from those of eukaryotes. Efficient translation of mRNA in prokaryotes requires a ribosome binding site called the Shine-Dalgarno ("SD") sequence on the mRNA. This sequence is a short nucleotide sequence of RNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression see Roberts and Lauer, Methods in Enzymology 68:473 (1979), which is hereby incorporated by reference in its entirety.

Promoters vary in their "strength" (i.e., their ability to promote transcription). For the purposes of expressing a cloned gene, it is desirable to use strong promoters in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promoters such as the P1 promoter, T7 phage promoter, lac promoter, trp promoter, recA promoter, ribosomal RNA promoter, the P<sub>r</sub> and P<sub>L</sub> promoters of coliphage lambda and others, including but not limited to, lacUV5, ompF, bla, lpp, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid trp-lacUV5 (tac) promoter or other *E. coli* promoters produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promoter unless specifically induced. In certain operons, the addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the lac operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as trp, pro, etc., are under different controls.

Specific initiation signals are also required for efficient gene transcription and translation in prokaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promoter, may also contain any combination of various "strong" transcription and/or translation initiation signals. For instance, efficient translation in *E. coli* requires a Shine-Dalgarno ("SD") sequence about 7-9 bases 5' to the initiation codon (ATG) to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the cro gene or the N gene of coliphage lambda, or from the *E. coli* tryptophan E, D, C, B, or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

Depending on the vector system and host utilized, any number of suitable transcription and/or translation elements, including constitutive, inducible, and repressible promoters, as well as minimal 5' promoter elements may be used.

The *Botulinum* neurotoxin-encoding nucleic acid, a promoter molecule of choice, a suitable 3' regulatory region, and if desired, a reporter gene, are incorporated into a vector-expression system of choice to prepare a nucleic acid construct using standard cloning procedures known in the art, such as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Third Edition, Cold Spring Harbor: Cold Spring Harbor Laboratory Press, New York (2001), which is hereby incorporated by reference in its entirety.

The nucleic acid molecule encoding a *Botulinum* neurotoxin is inserted into a vector in the sense (i.e., 5'→3') direction, such that the open reading frame is properly oriented for the expression of the encoded *Botulinum* neurotoxin prepeptide under the control of a promoter of choice. Single or multiple nucleic acids may be ligated into an appropriate vector in this way, under the control of a suitable promoter, to prepare a nucleic acid construct.

Once the isolated nucleic acid molecule encoding the *Botulinum* neurotoxin prepeptide has been inserted into an expression vector, it is ready to be incorporated into a host cell. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, lipofection, protoplast fusion, mobilization, particle bombardment, or electroporation. The DNA sequences are incorporated into the host cell using standard cloning procedures known in the art, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Springs Laboratory, Cold Springs Harbor, N.Y. (1989), which is hereby incorporated by reference in its entirety. Suitable hosts include, but are not limited to, bacteria, virus, yeast, fungi, mammalian cells, insect cells, plant cells, and the like. Preferable host cells of the present invention include, but are not limited to, *Escherichia coli*, insect cells, and *Pichia pastoris* cells.

Typically, an antibiotic or other compound useful for selective growth of the transformed cells only is added as a supplement to the media. The compound to be used will be dictated by the selectable marker element present in the plasmid with which the host cell was transformed. Suitable genes are those which confer resistance to gentamycin, G418, hygromycin, puromycin, streptomycin, spectinomycin, tetracycline, chloramphenicol, and the like. Similarly, "reporter genes"
which encode enzymes providing for production of an identifiable compound, or other markers which indicate relevant information regarding the outcome of gene delivery, are suitable. For example, various luminescent or phosphorescent reporter genes are also appropriate, such that the presence of the heterologous gene may be ascertained visually.

In one embodiment of the present invention, the expressed neurotoxin propeptide is contacted with a highly specific protease (e.g., enterokinase or TEV sequence) under conditions effective to enable cleavage at the intermediate region of the propeptide of the present invention. By this means, the intermediate region is not cleaved by proteases endogenous to the host cell. The expressed neurotoxin propeptide has one or more disulfide bridges.

Another aspect of the present invention relates to an isolated, physiologically active Clostridium botulinum neurotoxin produced by cleaving an isolated Clostridium botulinum neurotoxin propeptide of the present invention. The propeptide is cleaved at the highly specific protease cleavage site. The light and heavy chain regions are linked by a disulfide bond.

The Clostridium botulinum neurotoxin of the present invention can be isolated at a yield or concentration of at least about 0.1 mg/L, at least about 0.5 mg/L, at least about 1 mg/L, at least about 5 mg/L, at least about 10 mg/L, about 10-20 mg/L, about 20-30 mg/L, or at least about 30 mg/L. One of the particular advantages of the propeptides of the present invention and the method of their expression described herein is that BoNT neurotoxins can be purified to a homogeneity using a two-stage, non-denaturing, and highly selective affinity purification, as described in greater detail infra.

As discussed supra, Botulinum neurotoxins are synthesized as single chain propeptides which are later activated by a specific proteolysis cleavage event, generating a dimer joined by a disulfide bond. These structural features can be illustrated using BoNT/A as an example, and are generally applicable to all Clostridium botulinum serotypes. The mature BoNT/A is composed of three functional domains of Mr ~50,000, where the catalytic function responsible for toxicity is confined to the light chain (residues 1-437), the translocation activity is associated with the N-terminal half of the heavy chain (residues 448-872), and cell binding is associated with its C-terminal half (residues 873-1,295) (Johnson, “Clostridial Toxins as Therapeutic Agents: Benefits of Nature’s Most Toxic Proteins,” Annu. Rev. Microbiol. 53:551-575 (1999); Montecucco et al., “Structure and Function of Tetanus and Botulinum Neurotoxins,” Q. Rev. Biophys. 28:423-472 (1995), which are hereby incorporated by reference in their entirety).

Optimized expression and recovery of recombinant neurotoxins for BoNT serotypes in a native and physiologically active state is achieved by the introduction of one or more alterations to the nucleotide sequences encoding the BoNT propeptides, as discussed supra. These mutations are designed to maximize yield of recombinant Botulinum neurotoxin, while retaining the native toxins’ structure and biological activity.

Isolated, full-length Clostridium botulinum neurotoxins of the present invention are physiologically active. This physiological activity includes, but is not limited to, toxin immunogenicity, trans- and intracellular trafficking, and cell recognition. The mechanism of cellular binding and internalization of Clostridial toxins is still poorly understood. No specific receptor has been unambiguously identified, and the binding constants have not been characterized. The C-terminal portion of the heavy chain of all Botulinum neurotoxins binds to gangliosides (sialic acid-containing glycolipids), with a preference for gangliosides of the Gs series (Montecucco et al., “Structure and Function of Tetanus and Botulinum Neurotoxins,” Q. Rev. Biophys. 28:423-472 (1995); Montecucco, “How Do Tetanus and Botulinum Toxins Bind to Neuronal Membranes?” TIBS 11:314-317 (1986); and Van Heyningen et al., “The Fixation of Tetanus Toxin by Ganglioside,” J. Gen. Microbiol. 24:107-119 (1961), which are hereby incorporated by reference in their entirety). The sequence responsible for ganglioside binding has been identified for the structurally similar TeNT molecule, and is located within the 34 C-terminal amino acid residues of its heavy chain. BoNT/A, /B, /C, /E, and /F share a high degree of homology with TeNT in this region (Shapiro et al., “Identification of a Ganglioside Recognition Domain of Tetanus Toxin Using a Novel Ganglioside Photoaffinity Ligand,” J. Biol. Chem. 272:30380-30386 (1997), which is hereby incorporated by reference in its entirety). Multiple types of evidence suggest the existence of at least one additional component involved in the binding of Botulinum neurotoxins to neuronal membranes (Montecucco et al., “Structure and Function of Tetanus and Botulinum Neurotoxins,” Q. Rev. Biophys. 28:423-472 (1995); Montecucco, “How Do Tetanus and Botulinum Toxins Bind to Neuronal Membranes?” TIBS 11:314-317 (1986), which are hereby incorporated by reference in their entirety). In two reports (Nishikai et al., “The High-Affinity Binding of Clostridium Botulinum Type B Neurotoxin to Synaptotagmin II Associated with Gangliosides G_{1,3},G_{1,4},G_{1,5},FEBLS Lett. 378: 253-257 (1996); Dong et al., “Synaptotagmins I and II Mediate Entry of Botulinum Neurotoxin B into Cells,” J. Cell Biol. 162:1293-1303 (2003), which are hereby incorporated by reference in their entirety), synaptotagmins were identified as possible candidates for the auxiliary BoNT/B receptor, and synaptotagmins I and II were implicated as neuronal receptors for BoNT/B (Rommel et al., “Synaptotagmins I and II Act as Nerve Cell Receptors for Botulinum Neurotoxin G,” J. Biol. Chem. 279:30865-30870 (2004), which is hereby incorporated by reference in its entirety). However, despite the structural similarity in the putative receptor-binding domain of Botulinum neurotoxins, other toxin subtypes show no affinity for synaptotagmins or synaptotagmin-related molecules. Lipid rafts (Herrero et al., “Lipid Rafts Act as Specialized Domains for Tetanus Toxin Binding and Internalization Into Neurons,” Mol. Biol. Cell 12:2947-2960 (2001), which is hereby incorporated by reference in its entirety) have been implicated as a specialized domain involved in TeNT binding and internalization into neurons, but these domains are widely distributed on multiple cell types, and therefore cannot simply explain the high specificity of the toxins for neurons.


The toxicity of Botulinum neurotoxins is a result of a multistep mechanism. From the circulation, BoNT targets the presynaptic membrane of neuromuscular junctions, where it is internalized to directly exert its toxic effect on the peripheral nervous system (Dolly et al., “Acceptors for Botulinum Neurotoxin Reside on Motor Nerve Terminals and Mediate Internalization,” Nature 307:457-460 (1984), which is hereby incorporated by reference in its entirety). Toxicity at the neuromuscular junction involves neuron binding: internalization into endocytic vesicles, similar to those involved in synaptic vesicle recycling; activation within an acidic compartment to the proteolytically active toxin which then penetrates into the neuronal cytoplasm; and target recognition and catalytic cleavage of substrates in the neuronal machinery for synaptic vesicle exocytosis.

Clostridium botulinum neurotoxins of the present invention are physiologically active, and may be either toxic or atoxic. Atoxic neurotoxins have a toxicity that is reduced from the wt BoNT by at least about 1000-fold. In certain exemplary embodiments, the LD_{50} of an atoxic Clostridium botulinum neurotoxin of the present invention is between 1,000 and 150,000; between 50,000 and 150,000; between 75,000 and 150,000; between 100,000 and 150,000; between 1,000 and 100,000; between 50,000 and 100,000; between 75,000 and 100,000; 1,000; 25,000; 50,000; 75,000; 100,000; or 150,000-fold higher than the LD_{50} of the wt Clostridium botulinum.

The endopeptidase activity responsible for Botulinum neurotoxin toxicity is believed to be associated with the presence of a HExxH box H (SEQ ID NO:1) motif in the light chain, characteristic of metalloproteases. Mutagenesis of BoNT/A light chain, followed by microinjection of the corresponding mRNA into presynaptic cholinergic neurons of Apisidae californica, allowed the minimal essential domain responsible for toxicity to be identified (Kurazono et al., “Minimal Essential Domains Specifying Toxicity of the Light Chains of Tetanus Toxin and Botulinum Neurotoxin Type A,” J. Biol. Chem. 267:14721-14729 (1992), which is hereby incorporated by reference in its entirety). Site-directed mutagenesis of BoNT/A light chain pinpointed the amino acid residues involved in Zn2+ coordination, and formation of the active metalloendopeptidase core which cleaves SNAP-25 (Rigon et al., “Site-Directed Mutagenesis Identifies Active-Site Residues of the Light Chain of Botulinum Neurotoxin Type A,” Biochem. Biophys. Res. Commun. 288:1231-1237 (2001), which is hereby incorporated by reference in its entirety). The three-dimensional structures of Botulinum neurotoxins and their derivatives confirmed the mutagenesis results, and detailed the spatial organization of the protein domains. For the BoNT/A holotoxin, crystal structure was obtained to a resolution of 3.3 Å (Lacy et al., “Crystal Structure of Botulinum Neurotoxin Type A and Implications for Toxicity,” Nat. Struct. Biol. 5:898-902 (1998), which is hereby incorporated by reference in its entirety). The BoNT/B holotoxin crystal structure was determined at 1.8 and 2.6 Å resolution (Swaminathan et al., “Structural Analysis of the Catalytic and Binding Sites of Clostridium botulinum Neurotoxin B,” Nat. Struct. Biol. 7:693-699 (2000), which is hereby incorporated by reference in its entirety). Recently, a crystal structure for BoNT/E catalytic domain was determined to 2.1 Å resolution (Agarwal et al., “Structural Analysis of Botulinum Neurotoxin Type E Catalytic Domain and Its Mutant Glu_{232} to Glu Reveals the Pivotal Role of the Glu_{232} Carbonylate in the Catalytic Pathway,” Biochemistry 43:6637-6644 (2004), which is hereby incorporated by reference in its entirety). The later study provided multiple interesting structural details, and helps explain the complete loss of metalloendoproteolytic activity in the BoNT/E L.C.E_{232} Q mutant. The availability of this detailed information on the relationship between the amino acid sequence and biological activities of Clostridial toxins enables the design of modified toxin genes with properties specifically altered for therapeutic goals.

Thus, the physiologically active and atoxic Botulinum neurotoxin of the present invention may have a disabling mutation in an active metalloprotease site. The physiologically active Botulinum neurotoxins of the present invention may also have a non-native motif (e.g., a SNARE motif) in the light chain region that is capable of inactivating light chain metalloprotease activity in a toxic Botulinum neurotoxin, as described in U.S. Patent Application Publication No. 2006/0204524 to Lehtchenko et al., which is hereby incorporated by reference in its entirety.

Exemplary Clostridium botulinum neurotoxins of the present invention have structures as shown in Fig. 1. Specifically, the neurotoxin may have a structure selected from BoNT/Aa (\textsuperscript{5}), BoNT/A\textsuperscript{a}, \textDelta L.C-Peptide-BoNT/A\textsuperscript{a}, or \textDelta L.C-GFP-BoNT/A\textsuperscript{a}, or other such similar derivatives. Alternatively, neurotoxins of the present invention have an amino acid sequence that is at least 90%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:3-6 set forth below.

The full-length BoNT/Aa (\textsuperscript{5}) neurotoxin has an amino acid sequence of SEQ ID NO:3, as follows:
The full-length BoNT/Ad neurotoxin has an amino acid sequence of SEQ ID NO.4, as follows:

```
MKFLVNYALVFMVYISYYAANHHEBHHHHNHHTRRENLYFQOADSLSLWLRLLNNARGAGA60
SGPFBVKQFNYKDPVQVVEATYKPIQAGQMQPVKAFKIHNKIWIPERDFTPNPGDDL120
NPPPRAKVPVSTTVQSYLDTDNKQFVGTLVKPERYSTDLMRMNLLTIVSGGIPFWG180
GSTIDTNEVADTMCINVIQPGOSYRESNELNLVIGPASAIQFPCBSCPGHEVNLNLTRWG240
YGSTQYIIFSDDPDPFEEPEVDVTNPLAGAVPQATPDPAVTLAHLINAGHRLLGIAINFP300
NRPVFVNTVYMENMVLEFHLTFQDKISPQHEFPRFLVYNGQESDIATNLNK360
AKSIVGOTTALQINMNYFKEKYYLSEFDSEKIFPSVDELETLYTDHNFVTKFK420
VLRKSTANFIDKAVPHKINPVNKKVYIDPGVNRTHNLAAFNQPVQTEEINNNSFRKLKHF480
TGFLPEYKLKCVRRGIITSHQQSLQVYSSDDDKAMNLCLIKNMWDLLPFSPEDNPTIDL540
NKHEEITSDNIEASERNIQLDQVYUYLTQFNQHPENLIIQNLESLDDIQLEMPNHIE600
RPPNHMKKLYDLDTYFMHLKASQFPEHKSRIALTNQVHBAHNLPSRVTTFPSYNYKIVV660
KATEAMFLGQVQVNDPFTDESEYSTTDKIAITIIPYIQAPNYNMLYKCEOFTVOGA720
LIPSQAVILLIEFEIAFVLPFTALPSYYAQNKLVYQTIQNDALSNKNEKMDEVYKIVT780
NNLAVKQTQDLIREKNMELQAMARATKQYINQYQIVTKEKNQNQNNIDGSLKSNH840
SINKAMININQKPLMQCQVYYNLMSHIPQYYKRLHDFDLASLKEALQYIDNGSLIQYIQ900
RLEQDNYRNLSTDIPQQLNLQVNDQRLSTPEYMIININLTSLNRYESNHLDLSRTY960
SKINIGSKVQPDIDQKQIQLPMLLESSKELVKNAYIVMTNENPSSTSFRIPIYKPS1020
ISLNAVIRRIPQCNDSWENQVSVQSYIIEIHLQLDTQHEIQKVVRKQSGMIYSLVKIRI1080
FVTITNHNLINIKYIKNLQIDQPSHNLGNIHASNINMFVLDQCRETHRIVIKYFHLF1140
EKELNREDIKLTVQDQNGSTAYKDFLQGYKQPKVYLMHLYDMMKSYVDVDNVRGQYMV1200
LAEQPFRGVITNTLWSSLRTKFFIKKLYASQKGIENVRMNDVFYYVVENKERYLAT1260
NAQQAVQEKELSGEIDVYNLQQVYMKERQDQIYNECQLNQDQNNIQGIGPQHQF1320
BNNAKLJNANVNYRQISBRSTRTLCSWEPFVPDVQNGGERPGLAQGWSHPQFQEK1380
```

The full-length BoNT/Ad neurotoxin has an amino acid sequence of SEQ ID NO.4, as follows:
The full-length ΔLC-Neurotoxin-BoNT/D<sup>α</sup> neurotoxin has an amino acid sequence of SEQ ID NO:5, as follows:

```
MKFLVSLVAL FMVYTIYIV AAAMHHHHHH HHHTRENLHF QAGDSLSWL LLNLARGGA  60
SQHTLGKGG NHNRTMLAA NHPQCEKHL IHPFTFHEF TQLPEPHYLL CVRGITSTH  120
QSLOQGQIL YPQGALNDIC IQNNUDFGEP SSDSPANPQW LQKEHTSD THIESAAAERN 180
SLIDQGQQYTL TPNFPRPNEE IISSTLSLED SIQLEPMFIH EREPPHGKYE LDKYMFHIL 240
RAQPEFPHIES RIALTHSVNKL ALLMPREYTVF PFSSDYVKEV HEATEAAMFM GWVEVLVYDP 300
TDETSVEVST DKN dives PIGFPAHNI HMLYKDQDPGL ALIPSQAVIL LEPFPEIAIP 360
VLFPLQVSVY IANQVLVQTV IDNAEQHKE KQKQYKIVY TMNLAVHTQ IDLNEKKMK 420
ALENQASEKT ALNQVQNYQ TEEENNEMF HHQDLSKLN ESINQMMI NKPLAQCSVS 480
YLMQMSIPYG VERLPSFPEA LDKALLYII YRQGLGIPQ DNLRTQKLL RTMIDPQLSL 540
KVDRQRLLS TPTERNINII NTSLHLRAVY SNRLSIDR YSKINIESK NFPODFIMQI 600
QPLNLESSK ETVLGAITYY YAMRHENPSLSK IRPDIRFKN FYLSPHPQNL INCNEMNSGQ 660
KVSNYGEII WTQGTPQEK QRVFKYQSM IINSYDIENP IVTPITNRL NSKIIYNGR 720
LIDQPKPNL GNHASNNLM FKLDIGRTH YRINIFPKHL FDEKNEKEE KELDYNQCSS 780
GILKFPGQEy LQYDKPEYLL NLYDPHHKVD VNHQIGORM YLQGPRGSM TTTIIYNSSL 840
YRGTPTIKE YASQKKNIV RNDRVYINV VVQKEVYLA TIAQAOVEK ILASEIPDEV 900
GHLQQVVYMK SNDQGQTHN CMNLQCGEG MDQFIGFHPQ FHNIAKLVAS NINRQIERIS 960
SRTLOCSWEF IPVDQGGER PLGAQHSPFQ PEK
```

The full-length ΔLC-GFP-BoNT/D<sup>α</sup> neurotoxin has an amino acid sequence of SEQ ID NO:6, as follows:

```
MKFLVSLVAL FMVYTIYIV AAAMHHHHHH HHHTRENLHF QAGDSLSWL LLNLARGGA  60
SVSKEAELEF GYFIVILLEI QGNNHKKQF SVMEQEDRAT QKTLKIPCT TQKLPWFQPT 120
LVITLSTGQQ CFSPFDMNKM QHEDFQESAMP RQYIQERTIF PEDOQYNKSR AEQFPFGDTL 180
VNRLLETDQ FEDCNLQHC IQTMNHASNY VYIMDKKAN GOKVNKHKH NIEDGSVQA 240
DHQVMTPQIP DQOVLQPPHNY YLTSQSALS DMRKRENMN YQPVTAAAI THQMDLKYI 300
THQMGNLH RQILANTFQH QMTEINNNNN NPTKHAFQ LTDYKLCVR GIITSHQSL 360
DDQGENLYQF GALNDDCSK KHNWLPFFPS EDNTHNLQ GEEITGTTH AEADNISLDO 420
LQCDQYITPM FQQNEPENI INLSSDIIQQ LELMPFQER FMQKVLQDD KTMHylFDAQ 480
EFKXQKRIA LTSSWQHALL NPSKVTTFPS SDYWKDQK EAAMPGLGQV EQLVYDPTDR 540
TSEVSTDVKT ADITIIIPYI GPRNIGINML YKDQFQGALY FGAVILYEF IPEIAIPVGL 600
```
Still another aspect of the present invention relates to a treatment method. This method involves providing an isolated Clostridium botulinum neurotoxin according to the present invention, where the light chain region and the heavy chain region are linked by a disulfide bond and the cargo comprises a therapeutic agent. The isolated Clostridium botulinum neurotoxin is administered to an individual in need of treatment under conditions effective to provide treatment to the individual.

Administration can be carried out orally, parenterally, for example, subcutaneously, intravenously, intramuscularly, intrathecally, intraperitoneally, by intranasal instillation, or by application to mucous membranes, such as, that of the nose, throat, and bronchial tubes. The neurotoxin may be administered alone or with suitable pharmaceutical carriers, and can be in solid or liquid form such as, tablets, capsules, powders, solutions, suspensions, or emulsions.

The neurotoxin of the present invention may be orally administered, for example, with an inert diluent, or with an assimilable edible carrier, or may be enclosed in hard or soft shell capsules, or may be compressed into tablets, or may be incorporated directly with the food of the diet. For oral therapeutic administration, the neurotoxin (along with any cargo) may be incorporated with excipients and used in the form of tablets, capsules, elixirs, suspensions, syrups, and the like. Such compositions and preparations should contain at least 0.1% of active compound. The percentage of the compound in these compositions may, of course, vary and may conveniently be between about 2% to about 60% of the weight of the unit. The amount of active compound in such therapeutically useful compositions is such that a suitable dosage will be obtained. Preferred compositions according to the present invention are prepared so that an oral dosage unit contains between about 1 and 250 mg of active compound.

The tablets, capsules, and the like may also contain a binder such as gum tragacanth, acacia, corn starch, or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, lactose, or saccharin. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier, such as a fatty oil.

Various other materials may be present as coatings or to modify the physical form of the dosage unit. For instance, tablets may be coated with shellac, sugar, or both. A syrup may contain, in addition to active ingredient, sucrose as a sweetening agent, methyl and propylparabens as preservatives, a dye, and flavoring such as cherry or orange flavor.

The neurotoxin may also be administered parenterally. Solutions or suspensions can be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof in oils. Illustrative oils are those of petroleum, animal, vegetable, or synthetic origin, for example, peanut oil, soybean oil, or mineral oil. In general, water, saline, aqueous dextrose and related sugar solution, and glycols such as, propylene glycol, hyaluronan and its derivatives, or polyethylene glycol, are preferred liquid carriers, particularly for injectable solutions. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In all cases, the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol), and liquid polyethylene glycol), suitable mixtures thereof, and vegetable oils.

The neurotoxin of the present invention may also be administered directly to the airways in the form of an aerosol. For use as aerosols, the neurotoxin of the present invention in solution or suspension may be packaged in a pressurized aerosol container together with suitable propellants, for example, hydrocarbon propellants like propane, butane, or isobutane with conventional adjuvants. The neurotoxin of the present invention also may be administered in a non-aerosolized form such as in a nebulizer or atomizer.

In one embodiment, the neurotoxin of the present invention has a therapeutic agent as cargo, and the method involves treatment by the therapeutic agent. According to this embodiment, the Botulinum neurotoxin serves as a delivery vehicle for the therapeutic agent.

In another embodiment, the neurotoxin of the present invention has cargo other than a therapeutic agent and is administered to treat disorders amenable to treatment by Botulinum neurotoxins (either toxic or atoxic) including, for example, neuronal pathologies, such as Tourette syndrome.

a further aspect of the present invention is directed to a method of detecting _Clostridium botulinum_ neurotoxin trafficking either for mechanistic studies or for imaging specific sites of high neuronal activity. This method involves expressing a recombinant physiologically active _Clostridium botulinum_ neurotoxin according to the methods of the present invention, where a fluorophore is coupled to the neurotoxin, and detecting trafficking of the neurotoxin by detecting one or more locations of the fluorophore.

This method of the present invention can be carried out in vivo or in vitro. Detecting location of a fluorophore or other imaging modality can be carried out using methods well-known in the art. For example, detection devices known are that capable of receiving fluorescent emissions and generating a response to be examined by an operator. Suitable detectors include, without limitation, charge coupled devices (CCDs), photomultiplier tubes (PMTs), avalanche photodiodes (APDs), and photodiodes that contain a semiconductor material such as Si, InGaAs, extended InGaAs, Ge, HgCdTe, PbS, PbSe, or GaAs to convert optical photons to electrical current. CCD can produce an image in extremely dim light, and its resolution (i.e., sharpness or data density) does not degrade in low light. Other imaging procedures suitable for detection may include CAF scan, PET scan, and MRI.

Another aspect of the present invention relates to a method of detecting levels of neuronal activity. This method involves providing the isolated _Clostridium botulinum_ neurotoxin of the present invention and administering the neurotoxin to an individual or a tissue sample. The method further involves detecting location of the neurotoxin, where detection of the neurotoxin at a specific site in the individual or tissue sample indicates an increased level of activity of neurons at that site.

**EXAMPLES**

The following examples are provided to illustrate embodiments of the present invention but are by no means intended to limit its scope.

**Example 1**

Creation of the Construct Encoding BoNT/Aad<sup>AE</sup> in pLmutus Vector

Modification of pLmutus28i by Replacement of the Existent Polylinker with a Custom-Built Polylinker

A vector with a custom polylinker derived from Litusus 28i (New England Biolabs, Cat. #N3528S, 2823 b.p.) was used for subcloning the full-length BoNT/Aad<sup>AE</sup>. This derivative of Litusus 28i (pLmutus28C1) was created by restriction digestion of Litusus 28i with BglII and AatII followed by dephosphorylation. Annealed phosphorylated oligonucleotides C1-1S and C1-1A (Table 1) were ligated into digested Litusus 28i, resulting in the intermediate vector, pLmutus28C1INT (2824 b.p.). Vector pLmutus28C1INT was digested with Kas I and Stru I and dephosphorylated. Annealed phosphorylated oligonucleotides C1-2S and C1-2A were ligated into digested pLmutus28C1INT, resulting in the vector pLmutus28C1 (2890 b.p.). The sequence of this vector was deposited in GenBank under accession number GQ855199, and has a sequence of SEQ ID NO:13, as follows:

```
gttactacct tcaggtggca ctttctgggg aaactgctgg ggaacccctta ttgttttat    60
ttcttaata ctcctaaata tgatctcgtg ctactgctga taccctgtg aatgcttca    120
ataatattga aaaaaggaag gtargagta tcaacatttc cgttgctgccc ttaccttctt   180
cttctgccga ctgttttgcttc ctatcctgggg gaaatcgtcgt gaaatgctgg    240
tgctgaagat cagctttggttg caagatggtgg ttcctgagaa cttctgactga acagctgatg    300
gatcctctag atgtagctttc gcgagagtca tccttccagac atgagccttag ttaaaggct    360
gctactgggc gcgcttttttc ccctgctgatg ggcgggagcc agaagaactcg ggtgcgctat    420
acactttata ccagactgct tttgctacgg ctttctccggt acagagaggg ctcactcagc    480
tgctctgaca ttaagagag atgtgacttc cggcactgaac cgacagagcc ggaagagttcg ttgttttgc    540
cacactactc ctctgacgcgg tcgggagacc gaagagttgg actgtttttt ccctctaccc    600
gggagtctac gattacgccc gattcgctgg ggaacagagg cgtaagtaag cacaatatct    660
cggcaggt gcaggcagcagtg ccgtgtgact gtaaagggac ccctttaactc    720
tgagcgaga ctatctatcg ctctccgggg gcaatatata gtaatgtagg gggaggattaaa    780
agtttcggaa gcactttctcg gtggctgctc cgcttgcttg ctgagttctcg ccctccgga    840
tggcagcggct tgtgcgctgg ggtcgcggtat caatttcggga cggggtccag aggtaaagcc    900
ctctcttttgc gcaggtcctgc gcctcgcgctc gccgcgccacc ggcgctgact gtaaagct    960
acagagctg gtagagcttg ggcagcactg taaaagacctg tctaactgc gtaaagct    1020
tctatata ttttagatgcc attttgctca ctttttggctg gttgaatacc ggaaagcggc cgaagggagcc    1080
aagaggtat aagagagatg ttaaaatcag taaagatg gttgtttggct gttttgtaaa atctggctt    1140
aattttttct aacctagctc ttttttttaa ctaatgctgg aataggggaa aatcctctatt    1200
```
TABLE 1

Oligonucleotides Used for Cloning Experiments

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Length</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1-1S</td>
<td>103 b</td>
<td>5'-PGAT CTG AAC TAA GTC GCT CCG TTC GAA CCA GAA CTC TGG AAG CTG CCC GCC GGC GCT ATC CAT GGC ACA CCC GTT CCA CTA GAT GTG GCG CCA CCT AGT AGC T-3' (SEQ ID NO: 16)</td>
</tr>
<tr>
<td>C1-1A</td>
<td>95 b</td>
<td>5'-PGAT AGG GCG CTA AGC TAG CTG AAG GCG TGT GCC ATG GAT AGC GCC CCC GTC AGG CTT CCA GAG TTC TGG TTC GAA CGG ACC GAC TTA TGT CCT CA-3' (SEQ ID NO: 17)</td>
</tr>
<tr>
<td>C1-2S</td>
<td>108 b</td>
<td>5'-PGCG CCT ATG CCC GCT AAC CCC GGT CAC TTA AGT AGG ATA TCT GTC GTC AGT TAC CCC CCC GGC AGT AGC TGT ATA TGC ATA TTC TCG AGG CAT CGC AGC TTC CTC AGG AGG-3' (SEQ ID NO: 18)</td>
</tr>
</tbody>
</table>
### Table 1—continued

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Length</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl-2A</td>
<td>104 b</td>
<td>5'-pCTT CCT GAG GAA GCT CCC ATC CAG GAA TAT GCA TAT AGA CTA CAT GCG GTA ACT GCA GAG AGA TAT CAT ACT TAA GTC ACC GCG GTC ATT ACG CCG CAT AG-3' (SEQ ID NO: 19)</td>
</tr>
<tr>
<td>CP1S</td>
<td>88 b</td>
<td>5'-pGCG TTA ACT CCT AAA AAA CCG CCA CCA TGA AAT TCT TAG TCA ACG TTC TCC TTA TTA TTA TGG TGA TAT ACA TTT CTT ACT TCT AGT C-3' (SEQ ID NO: 20)</td>
</tr>
<tr>
<td>CP1A</td>
<td>88 b</td>
<td>5'-pGCG CCG GAT GAT GAG ATT GAT TAC GAG CAT AAA AAC AAG GGC AAC GCT GAC TAA GAA TTT CAT GGT GCC GCT TTT TTA GGA GAT A-3' (SEQ ID NO: 21)</td>
</tr>
<tr>
<td>CP2S</td>
<td>38 b</td>
<td>5'-pGCG CCA TCA TCA CCA CCA TCA TCA TCA TCA CCA CCA CA-3' (SEQ ID NO: 22)</td>
</tr>
<tr>
<td>CP2A</td>
<td>38 b</td>
<td>5'-pGCG GTG TGG TGG TGA TGA TGA TGG TGA TGA GC-3' (SEQ ID NO: 23)</td>
</tr>
<tr>
<td>CP3S</td>
<td>78 b</td>
<td>5'-pGCG GTA AAA ACC TCT ATT TTC TCT AGG GCC CCG GTG ACT CCC TGT CTT GCC TGC TTC TGC TCA AGC CCG GCC GTG GCC-3' (SEQ ID NO: 24)</td>
</tr>
<tr>
<td>CP3A</td>
<td>78 b</td>
<td>5'-pCTA GCC CCC CCC CCG TCG AGC AGC AGC AAC GCA GAC AGG GAG TCA GCG GCC GCC TGA AAA TAC AGG TTT TCA-3' (SEQ ID NO: 25)</td>
</tr>
<tr>
<td>CP4S</td>
<td>45 b</td>
<td>5'-pGCG TTA CCA TGG GTG ATC ACC ACC ACC ATC ACC ACC ACC ACG ACC ACC-3' (SEQ ID NO: 26)</td>
</tr>
<tr>
<td>CP4A</td>
<td>45 b</td>
<td>5'-pGCG GTG TGG TGG TGA TGG TGA TGG TGA TGA TCA ACG GTA AGT-3' (SEQ ID NO: 27)</td>
</tr>
<tr>
<td>CP5S</td>
<td>45 b</td>
<td>5'-pGCG TTA CCG TGC TCC ATC ACC ACC ATC ACC ACC ACC ACC ACC-3' (SEQ ID NO: 28)</td>
</tr>
<tr>
<td>CP5A</td>
<td>45 b</td>
<td>5'-pGCG GTG TGG TGG TGA TGG TGA TGG TGA TGG AGT GTA-3' (SEQ ID NO: 29)</td>
</tr>
<tr>
<td>CP6</td>
<td>10 b</td>
<td>5'-pTAC CCT AGG G-3' (SEQ ID NO: 30)</td>
</tr>
<tr>
<td>CP7S</td>
<td>31 b</td>
<td>5'-TCA TGG TAG CAT GAG CAA GGG CCG CGA GCT G-3' (SEQ ID NO: 31)</td>
</tr>
<tr>
<td>CP7A</td>
<td>32 b</td>
<td>5'-TAT AGG TAC CCT TGG TAT GCA GCT CAT CCA TCC CG-3' (SEQ ID NO: 32)</td>
</tr>
</tbody>
</table>

Cloning of BoNT/Aad<sup>4</sup> into pLittus28C1

In the first cloning step, pLittus28C1 was digested with Hind III and Not I and dephosphorylated. Annealed phosphorylated oligonucleotides CP1S and CP1A were ligated into digested pLittus28C1, resulting in vector pLitSB3A1 (2968 b.p.). Vector pLitSB3A1 carried several unique 5'-restriction sites, including Bgl II, Bts I, Rsr II, and BstB I, to be used for the subsequent subcloning of the full-length construct into expression vectors. This polylinker was followed by the enhancer sequence AACCTGTTAAACACCGCGAC (SEQ ID NO:33), 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 CP2S and CP2A were ligated into digested pLitSB3A1, resulting in vector pLitSB3A2 (2987 b.p.), which carried the sequence encoding a 10-lysine affinity tag downstream of the sequence encoding the honeybee melittin signal peptide. Next, pLitSB3A2 was digested with Mlu I and Nhe I and dephosphorylated. Annealed phosphorylated oligonucleotides CP3S and CP3A were ligated into digested pLitSB3A2, resulting in vector pLitSB3A3 (3056 b.p.), which carried the sequence encoding the tobacco etch virus (TEV) protease cleavage site, followed by the sequence encoding S6 peptide tag downstream of the sequence encoding the 10-His affinity tag.

In the next three steps, the synthetic DNA of the full-length BoNT/Aad<sup>4</sup> with the sequence optimized for expression in two hosts, S. frugiperda and E. coli, was consecutively built into the pLitSB3A3 vector. The synthetic gene for full-length BoNT/Aad<sup>4</sup> was synthesized in the form of three contiguous DNA segments, each supplied by the vendor (Genescript) as DNA subcloned in the pUC57 vector. Numbers in the text below correspond to amino acid numbering in wt BoNT/A.

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 (which had been digested with the same set of restriction enzymes and dephosphorylated) to generate the 4355 b.p. vector, pLitSB3A4. Then, a 1233 b.p. DNA fragment encoding the sequence D<sub>357</sub>-S<sub>589</sub> was isolated from pUC57-BAad-2 by digest with Xba I and EcoR V and subcloned in pLitSB3A4 (digested with the same set of restriction enzymes and dephosphorylated) to generate the 5570 b.p. vector, pLitSB3A5. Finally, a 1401 b.p. DNA fragment
encoding the sequence P_{650}^{L1206} followed by a sequence encoding a short linker, the StrepTag II sequence, WSH-PQFEK (SEQ ID NO:34), and a triplet of termination codons in three reading frames was isolated from PUC37-Baad-3 by
digest with EcoRI and XhoI, and subcloned into pLtiS33A5
(digested with the same set of restriction enzymes and
dephosphorylated) to generate the 6925 b.p. vector,
pliS33B. The sequence of pLtiS33A has been deposited
into GenBank under accession number GQ855200, and has a
sequence of SEQ ID NO:14, as follows:

gttactcag tcagttggca cttttcgggg aatacgtgcc ggaaccctaa ttgtttatat  60
tttctataa attcactata tgtatgctgt ctaggagcaaa taaccaattg gccacgctggt  120
ataatagta aaaaagaga gttatgagct tcaactcattct cgctgcgccc tttttcctt  180
tgtgtgggca tttttcctt ttcttttttc ttcacacagaa gcgctgcgtg aagttaaaaa  240
tgctgaagat cagttgggtg cagctgagtt ttaatcgaa aagcgacctg ca 300
gactcctag gttttttgct cggagagcct ttcttcaaat atgtgacact ttaagttctc  360
gccagctgg cccgttttat ccctgtgtag aagccggcag gcaccaattg gcggtcgtct  420
acactattct ccaatgactg tggctgtgat cttcaccact gcagaaacc acgctttactg  480
tggtcagaca gtaaagaat tagctgcttg gcagcataac atgatgctaa acagcggcgc  540
dacatcatct cttcaccact gcagccagac ggaagagct ttccgcttttt cgcacagtct  600
ggggaacat gtaacctgct tgtatgcttg ggcagcgagc tgaataagag ccataaaaac  660
cggagctcg gatgcacgcgc tgtctgtgct gctggctggc atgctctattc cttctgtcagc  720
tggcgtcag actataactg tcttctctgg gcacactattg gactgtgagt gcggacattg  780
agttcgagca ccacacttctg gcctggccct ctcggcttcg tggctctattg ctgataaatc  840
tggacgcttg gactgcggttg ttctcgtcag cttgagctgccttg ggttcttactg tcacgccgct  900
tctccgctg ctatgctagg cagcagcgg cgcctgctgg acctgtctgt ggcgactagg  960
aacagagct gtagatcggc aataccttg ccactgtgat taactcgctg acaacgtaatt 1020
tctctatata cttattatgg attaacctgc gttgtataatc agaataaatgc ccaagcccgt 1080
aatgttgat aagcaaatat ttaaatgtag ttaaagtttaa aagcagttta ttttgattaa attcgagttt 1140
aatgttctgc atcaagcgcgt attttttata ccaatagcgcg aagctgcaaa atcctcttatc 1200
aatcagatag aaaaagaga gttatgagct tcaactcattct cgctgcgccc tttttcctt 1260
cttctatata ccaatgactg tggctgtgat cttcaccact gcagaaacc acgctttactg 1320
ocactattct ccaatgactg tggctgtgat cttcaccact gcagaaacc acgctttactg 1380
aatcagatag aaaaagaga gttatgagct tcaactcattct cgctgcgccc tttttcctt 1440
aacagagct gtagatcggc aataccttg ccactgtgat taactcgctg acaacgtaatt 1500
ocactattct ccaatgactg tggctgtgat cttcaccact gcagaaacc acgctttactg 1560
acactattct ccaatgactg tggctgtgat cttcaccact gcagaaacc acgctttactg 1620
ccacagctg gctcagcctgc acctagacaatctc agccttctgtgcttcttg 1680
tgctgtctac gttgctgtta ggtgctgtta ggtgctgtta ggtgctgtta 1740
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 1800
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 1860
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 1920
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 1980
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 2040
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 2100
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 2160
acagacagtc acagacagtc acagacagtc acagacagtc acagacagtc 2220
agacgatttc ttggtgtgcc tgtctctcag cgggtgcttg atctctgtgg agttccattc 4740
tgaaatgocg atctcttgctgc ttggtgcttg ctgcgtgtgc tcatacattg cgataaaggt 4800
cttgacgttg cagacaaacg ataagcgcttt ctcacaacgt aagcgacta tgggctaggt 4860
ctacaacaac acogcggctgc aacagggtaa acacaaattc atctgacatg 4920
taagaaacgt aaggaggtt ttgagaacca ggcgtgagct acataaagcct taccaacta 4980
ccagttacaa cagttatacg aagagggaaa gataacactc aattctcaaca ttcgagcttc 5040
gtccccaaag cagacaaacg catacataac cgcattagact cacatcaaca aatctctgaa 5100
tacgtcttcc ctgtgcattc ctgatgacct ttcagtccaa ttcggtgtga agcgcgtcga 5160
ggacctgtag gtcgagcgtg ttcggtgtaac acacacagct gaatcagttta ttgccagt 5220
tttagcttggc caagttgacc gctgtaagga caagttcaac atatcctgt gtaaccgata 5280
cctcctcctga ctgctcaatc acogttagaa cccgcatcgtg ttcttccttc atcaccata 5340
cattaacac atatacaaat ctcatacttc ctcctgctgc tcaatgctga atctatcgtg 5400
cgatctgtcc gtcctacgca gcaagataac cctgtgtgctg aatgcctgatg ctgctcgtga 5460
tgacaagcc caaatcgcatt tggctcaacct gcaagctcct aatcctgagtt ttcgctgag 5520
gaatgcacct catcattcag aatacttctc acatctactc ttcggtgtag aagatcagtaa 5580
cctccttcatc ttcctgtctca ttcacatgaaa aacgcatcatt acogtagtaa 5640
gataactctg ggtggtgagg tttctctgaa cattgagaaa atacatggaata ctcgctcaga 5700
taugcacaaggg caaactacgc gttggtgcct taaaataaggt gatgatcatt cgcacacagt 5760
cattaacctc atagtacttcc gctgttcctc ttacacaaat ctgcgctcata ctgcgtcataa 5820
cattcataac ggtcggtgta ctcacacgaaa acacacttca cgtccgctga ccaactcagc 5880
ttcacaatac atacgatttc gttcggagctg ttcgctcgtg acacactgct tcgtgctgatg 5940
catttatttc acatctgccg acaaagactc cttgctgctt ctcgctgcct 6000
tatcctcaaa atctcggcct ttcgctaaga cttggtgctgc gattacttcc actacgataa 6060
gcatactac gacggtgtac gcacatgcac caatctcactc cctggagttt ccaacctgct 6120
tacggttggc tacactgacc tcaacaggcc acogtgggtgc gttgagctcct gtcgctgatg 6180
cotgagtagt ttctcagaga ggtgtctgtgc cgttcctata gagaagtactg cttcagggca 6240
cagcacaact atcgagttgc aacaattgc gtcgactcct gtcgtcggag tcaagatctc 6300
ggataacgcc tggcgccttc aacgtccttc gcgtgcggag gcgacagttc cttgcggttct 6360
gggagactaca gagtctgtcc acaacaaggcc aacaagtctg acaacagagtt 6420
gggagactaca aacaatctgc aatgacaccc gcctgagtatt gcgtttgttt 6480
catcgttttc caacagagtt ctcacatgta gctgagagtgt ggtcgcgcttt 6540
gcaggttggc agttccttgg gattccttgg cttgtaacctt cgggtgtcgg 6600
tggtggtgctc cggacagccct tcgtggtcgtc ccctcgtttc agaagttataa 6660
gttaatagat atataagct gtagctagtc gcgtgctcctc aggagactgt gtagctgtat 6720
tacggaacct ccgcgtttt tcaacgctgtg gcgtggtgggt aacgtgtcctg taccacactc 6780
atctcgtttcg cagcacttcc cctcctgcctg gtggtggtgctg ctcgtccttc taccacactc 6840
gatcggcaatt tggcagctttg aagccagactg ggcgcttcgc ttcggtgtag aagatcagtaa 6900
agcgcgcttt gcgcgccttt ttttt
Creation of Vectors for Protein Expression in E. coli

**pET-19b-Based Construct**

Vector pLtiSB3A was linearized by digestion with Hind III and Mlu I and dephosphorylated. The 6799 bp DNA fragment was isolated from an agarose gel and purified. Synthetic phosphorlated oligonucleotides CP4S and CP4A were annealed and ligated into the isolated vector DNA to generate the vector pLti83A1 (6844 b.p.). Vector pLti83A1 was digested with Nco I and Xho I and the 4078 bp DNA fragment was isolated and purified from an agarose gel. This fragment was ligated into expression vector pET-19b (Novagen, Cat. #69677, 5717 b.p.), that had been digested with Nco I and Xho I and dephosphorylated, generating the 9721 b.p. vector pET19B3A.

**pETCoco2-Based Construct**

Vector pLtiSB3A was linearized by digestion with Hind III and Mlu I and dephosphorylated. The 6799 bp DNA fragment was isolated from an agarose gel and purified. Synthetic phosphorlated oligonucleotides CPSS and CP5A were annealed and ligated into the isolated vector DNA to generate the vector pLti83A2 (6844 b.p.). Vector pLti83A2 was digested with Xho I and treated with CpG methyltransferase (M. Ssrl, New England Biolabs, Cat. #M0226). After completion of methylation, the enzyme was inactivated and removed by phenol/chloroform extraction, followed by DNA precipitation with ethanol. DNA was treated with Klenow fragment of DNA polymerase I (New England Biolabs, Cat. #M021) to fill in 5' overhangs generated by Xho I cleavage, and ligated. Ligated DNA was digested with Pvu I and dephosphorylated. Annealed phosphorlated synthetic oligonucleotide CP 6 was ligated into Pvu I digested vector to generate pLti83A3 (6858 b.p.). This vector was transformed, amplified, and propagated in NEB 10-beta E. coli strain (New England Biolabs, Cat. #C3020K). Following propagation, plasmid DNA was isolated and digested with restriction endonucleases Sph I and Aat II. The 4081 b.p. DNA fragment was isolated from an agarose gel, purified, and ligated into the vector pETCoco2 (Novagen, Cat. #71148-3, 12417 b.p.) (that had been treated with the same set of restriction endonucleases and dephosphorylated) to generate the expression vector pETCoco3A (16263 b.p.).

Creation of Vectors for Protein Expression in Baculovirus

**Construct for Expression of BoNT/Aa1**

The plasmid pLtiSB3A was digested with the restriction endonucleases Rsr II and Xho I. The 4190 b.p. DNA fragment, encoding full-length BoNT/Aa1, was isolated from an agarose gel, purified, and ligated into the baculovirus transposition vector, pFastBacTm (Invitrogen, Cat. #10360-014, 4776 b.p.) (which had been digested with Rsr II and XhoI and dephosphorylated) to generate the 8968 b.p. vector pFB1SB3A-AE1. The DNA sequence of pFB1SB3A-AE1 with annotations has been deposited to GenBank under accession number GQ855201 (SEQ ID NO:15). Features of the 1372 aa protein, translated from the open reading frame of the insert DNA (BoNT/Aa1), are shown in FIG. 1.

**Construct for Expression of BoNT/Aa2**

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 pFB1SB3A-AE1 by digestion with Xba I and Afl II. The linearized vector was dephosphorylated and a 351 b.p. Xba I/Afl II DNA fragment excised from the plasmid pUC57-TEV (GenScript) was ligated into the vector to generate the 8971 b.p. vector pFB1SB3A-TEV. The DNA sequence of pFB1SB3A-TEV, along with annotations, has been deposited to GenBank under accession number GQ855202 (SEQ ID NO:10). Features of the 1373 aa protein, translated from the open reading frame of the insert DNA (BoNT/Aa2), are shown in FIG. 1.

**Construct for Expression of δLC-Peptide-BoNT/A**

Vector pFB1SB3A-TEV was linearized by digestion with BssH II and Xba I and dephosphorylated. The 7631 b.p. dephosphorylated vector was isolated from an agarose gel and purified. A DNA fragment containing the sequence encoding the δLC-pEptide was added to the form of synthetic DNA cloned into pUC57 (pUC57-PEpt), supplied by the manufacturer (GenScript). 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 the linearized pFB1SB3A-TEV, resulting in the 7831 b.p. vector, pFB1SPepB2ATEV. The DNA sequence of pFB1SPepB2ATEV, along with annotations, have been deposited in GenBank under accession number GQ855203 (SEQ ID NO:11). Features of the 993 aa protein, translated from the open reading frame of the insert DNA (δLC-Peptide-BoNT/Aa2), are shown in FIG. 1.

**Construct for Expression of δLC-GFP-BoNT/A**

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 50 µl reaction mixture 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. Vector pFB1SPepB2ATEV was digested with Nhe I and under digested (short incubation time and small amount of the restriction endonuclease) with Acc65 I. The 7822 b.p. DNA fragment was isolated from an agarose gel, purified, and dephosphorylated. The purified GFP PCR product and the linearized pFB1SPepB2ATEV were ligated, resulting in the 7822 b.p. vector pFB1SGFPB2ATEV. The DNA sequence of pFB1SGFPB2ATEV, along with annotations, has been deposited to GenBank under accession number GQ855204 (SEQ ID NO:12). Features of the 1230 aa protein, translated from the open reading frame of the insert DNA (δLC-GFP-BoNT/Aa2), are shown in FIG. 1.

Custom synthetic oligonucleotides were obtained from Sigma-Aldrich. E. coli strain TOP10, used for plasmid transformation and amplification, was purchased in a form of electrocompetent cells from Invitrogen (Cat. #C404052), except as specifically otherwise. Restriction endonucleases, T4 DNA ligase, and arctic shrimp alkaline phosphatase were purchased from New England Biolabs. All DNA fragments isolated from agarase gels after enzymatic treatment were purified with 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 Biomedical Medicine. The sequencing data obtained proved constructs to be free of unexpected mutations.
Example 4

Protein Expression and Purification—E. coli Experiments

Electrocompetent E. coli strain JM109 (DE3) (Promega, Cat. #98901) was transformed with pET10B3A plasmid according to an established protocol. A single colony isolated from a Luria-Bertani (LB) plate containing carbenicillin (100 µg/ml) was inoculated into 10 ml of LB medium containing carbenicillin and grown overnight with agitation (250 rpm) at 37°C. An aliquot (2 ml) of the culture was transferred to 200 ml of fresh LB medium with the addition of carbenicillin (100 µg/ml), and the cells were grown further at 37°C with agitation to reach an OD₆₀₀ of approximately 0.5 (about 4 hours). The suspension was rapidly cooled to approximately 25°C in an ice bath, and IPTG (100 µl of a 10 mM stock) was added to a final concentration of 0.5 mM. The cultures were incubated with agitation for 12 hours at 25°C. Aliquots were collected at 3, 6, and 12 hours after IPTG induction and cells were harvested by centrifugation at 6000 x g for 15 min at 4°C. The cells were resuspended at 4°C in BugBuster protein extraction reagent (Novagen, Cat. #78584) and lysed with the addition of Lysonase™ bioprocessing reagent (lysozyme and benzamidine, Novagen, Cat. #71230-4) in the presence of protease inhibitor cocktail without EDTA (Pierce, Cat. #78430) by trituration the suspension until signs of high viscosity disappeared (approximately 30 minutes). The BugBuster lysate was cleared by centrifugation at 35,000 g for 15 minutes, 4°C, distributed in 0.5 ml aliquots and stored at −80°C. The residual pellet was resuspended at room temperature in 6 M urea, 25 mM Tris-HCl, pH 8.0. Cleared urea solubilize was collected by centrifugation at 35,000 g for 15 minutes at room temperature, distributed in 0.5 ml aliquots, and stored at −80°C. Aliquots of BugBuster and urea lysates were batch incubated with TALON® affinity resin (Clontech, Cat. #63506) at 4°C for 30 minutes. TALON™ beads were sequentially washed with 100 mM NaCl, 25 mM Tris HCl, pH 8.0 and protein was eluted by resuspension of the beads in 1× SDS PAGE loading buffer with β-mercaptoethanol (Bio-Rad, Cat. #161-0710). BugBuster and urea lysates, as well as eluates of these lysates obtained from Talon® resin, were loaded on a 10.5-14% Criterion SDS PAGE gel (Bio-Rad, Cat. #345-0106) and separated. Proteins were transferred to nitrocellulose membranes (Bio-Rad, Cat. #162-0117) and probed with polyclonal antibodies against BoNT/A holotoxin (Polio, raised against BoNT/A holotoxin (Sten Serum Institut, Denmark)). As a positive control, m BoNT/A was loaded on the gel, and JM109 (DE3) transformed with the empty vector was used as the negative control. When the results were analyzed by electrophoresis and Western blotting, it was unexpectedly found that toxin of the proper molecular weight and immunoreactivity could not be produced.

pET30c3A was transformed into competent E. coli Rosetta-gami B (DE3) cells (Novagen, Cat. #71136-3) according to the manufacturer’s protocol. A single bacterial colony, picked from an LB agar plate, was grown in (i) LB medium with 100 µg/ml carbenicillin, 15 µg/ml kanamycin, 12 µg/ml tetracycline, and 34 µg/ml chloramphenicol, and (ii) LB medium with 0.01% L-arabinose, 100 µg/ml carbenicillin, 15 µg/ml kanamycin, 12 µg/ml tetracycline, and 34 µg/ml chloramphenicol for 16 hours at 37°C. An aliquot of overnight cultures (1:100 v/v) was transferred to fresh LB media supplemented with antibiotics, with and without addition of L-arabinose. The cells were grown at 37°C with agitation to reach OD₆₀₀ of approximately 0.4 (about 9 hours). The suspensions were rapidly cooled to approximately 25°C in an ice bath, and IPTG was added to a final concentration of 0.5 mM. The cultures were incubated with agitation for 12 hours at 25°C. Aliquots were collected at 3, 6, and 12 hours after IPTG induction and cells were harvested by centrifugation at 6000 x g for 15 min at 4°C. Protein extraction, sample preparation, SDS PAGE, and Western blotting were performed as described above. When the results were analyzed by electrophoresis and Western blotting, it was unexpectedly found that toxin of the proper molecular weight and immunoreactivity could not be produced.

Example 5

Baculovirus Experiments

Generating Recombinant Bacmids

pFastBac™ constructs were used for transposition of the cloned DNA into the shuttle vector (bacmid). Approximately 1 ng of plasmid DNA was used to transform MAX Efficiency® DH10Bac™ E. coli (Invitrogen, Cat. #10361-012) using a heat-shock method according to the manufacturer’s protocol (Invitrogen, BactoBac™ baculovirus expression system). Colonies were grown on LB agar plates supplemented with 50 µg/ml kanamycin, 7 µg/ml gentamycin, 10 µg/ml tetracycline, 100 µg/ml Bluo-Gal, and 40 µg/ml IPTG, for 48 hours for full color development. For each construct, two white colonies were picked and re-grown for 24 hours at 37°C in 50 ml LB medium supplemented with antibiotics. The recombinant bacmids were isolated and purified from harvested cells using a QIAGEN Large-Construct Kit (Cat. #12462), according to the manufacturer’s protocol. The DNA concentration was measured with a NanoDrop spectrophotometer and adjusted to 0.5 mg/ml with TE buffer. The yield of the bacmid DNA was approximately 1 µg/ml of starting LB culture. Transposition of the cloned genes into recombinant bacmids was confirmed by PCR with bacmid as a template and sets of primers specific for each cloned construct.

Example 6

Transfecting Insect Cells and Collecting Baculoviral Stock for Protein Expression

Sf9 insect cells, grown in shaker flasks in serum-free SF900 II medium (Invitrogen, Cat. #10902-096) were plated in 6-well (35 mm) culture plates at a density of 1x10⁶ cells per well. Plates were incubated at 28°C in a humidified incubator until cells adhered to the surface (approximately 1 hour). Two (2) µg of bacmid DNA (0.5 mg/ml) was diluted with 100 µl unsupplemented Grace’s insect cell medium (Invitrogen, Cat. #11595-030) in a 15-ml sterile polystyrene tube. In a separate tube, 7 µl of CELLfectin® reagent (Invitrogen, Cat. #10362-010) was also diluted with 100 µl of unsupplemented Grace’s medium. Diluted bacmid DNA and CELLfectin® were combined and incubated at room temperature for 40 minutes. While DNA:lipid complexes were incubating, the medium was aspirated from the adherent Sf9 cells, cells were washed with unsupplemented Grace’s medium, and 2 ml of fresh Grace’s medium was added to each well. At the end of the 40 minute incubation, 800 µl of unsupplemented Grace’s medium was added to the DNA-CELLfectin® mixture, the medium covering the cells was aspirated, and 1 ml of the DNA-CELLfectin® mixture was gently added to the adherent Sf9 cells. Cells were incubated at 28°C for 5 hours, after which the DNA-lipid complexes were aspirated and 2 ml of fresh SF900 II serum-free medium was added to
the cells and incubation was continued. P1 low-titer baculoviral stock (2 ml of growth medium) was collected 72 hours after transfection.

To generate P2 high-titer stock, 1 ml of P1 stock was added to 25 ml of SF9 grown in SF-900 II medium in a shaker flask. Cell density at the time of infection was approximately 1.5x10^6 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 (Invitrogen, BAC-TO-BAC® baculovirus expression system). P2 titers for the constructs were approximately 1-2x10^8 pfu/ml. P2 stock was used to infect SF9 cells for expression of the proteins of interest.

Example 7

Protein Expression, Purification, and Processing

SF9 cells grown to a density of approximately 1.5x10^6 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 was determined empirically. For BoNT/Aad^\* and BoNT/Aad^\**, these conditions were MOI approximately 0.1 and 60 hours; for \( \Delta L C - \text{Peptide-BoNT/A}^{\text{rev}} \) and \( \Delta L C - \text{GFP-BoNT/A}^{\text{rev}} \), the conditions were MOI approximately 0.5 and 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. #1825000), concentrated approximately 3 fold on a Millipore Prep/Scale TFF cartridge (Cat. #CDUF002LT), and dialyzed against TALON® resin chromatography loading buffer (500 mM NaCl, 25 mM Tris-HCl, 5 mM imidazole, pH 8.0). For \( \Delta L C - \text{Peptide-BoNT/A}^{\text{rev}} \) and \( \Delta L C - \text{GFP-BoNT/A}^{\text{rev}} \), 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 the loading buffer. After loading, resin was washed with 10 volumes of loading buffer, 20 volumes of wash buffer (500 mM NaCl, 25 mM Tris-HCl, 20 mM imidazole, pH 8.0), and protein was eluted with 5 volumes of elution buffer (500 mM NaCl, 25 mM Tris-HCl, 200 mM imidazole, pH 8.0). For \( \Delta L C - \text{Peptide-BoNT/A}^{\text{rev}} \) and \( \Delta L C - \text{GFP-BoNT/A}^{\text{rev}} \) 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 wash buffer of column wash 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-\( \beta \)-D-glucopyranoside (USB, Cat. #29836-26-8). Eluted proteins were dialyzed against 100 mM NaCl, 25 mM Tris-HCl buffer, pH 8.0, either without (BoNT/Aad^\* and BoNT/Aad^\**) or with (\( \Delta L C - \text{Peptide-BoNT/A}^{\text{rev}} \) and \( \Delta L C - \text{GFP-BoNT/A}^{\text{rev}} \)) 0.2% n-octyl-\( \beta \)-D-glucopyranoside, and concentrated using Millipore Amicon Ultra 30,000 MWCO centrifugal filter units (Cat. #UFC803024) to a final total protein concentration of approximately 10 mg/ml. Dia-lyzed and concentrated fractions were loaded on Streptactin superfowl agarose (Novagen, Cat. #71592) pre-equilibrated with 100 mM NaCl, 25 mM Tris-HCl 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-desethylfibrin (Novagen, Cat. #71610). Fractions containing pure protein were combined, dialyzed, and concentrated using Amicon Ultra centrifugal filter units.

Processing of BoNT/Aad^\*

The 10-His tag was removed from recombinant protein using ActEVE protease (Invitrogen, Cat. #12575-015). The protein was incubated with ActEVE protease (1 unit of enzyme per 1 µg of protein) at 30°C for 6 hours in digestion buffer (50 mM NaCl, 0.5 mM EDTA, 3 mM glutathione, 0.5 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 ActEVE 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 µg of the protein) at 16°C for 16 hours in digestion buffer (50 mM NaCl, 20 mM Tris-HCl, 2 mM CaCl₂, 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^\**, \( \Delta L C - \text{Peptide-BoNT/A}^{\text{rev}} \), and \( \Delta L C - \text{GFP-BoNT/A}^{\text{rev}} \)

N-termially placed 10-His tag removal and propeptide cleavage were performed simultaneously, according to the protocol described for BoNT/Aac^\* above. \( \Delta L C - \text{Peptide-BoNT/A}^{\text{rev}} \) and \( \Delta L C - \text{GFP-BoNT/A}^{\text{rev}} \) 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.

Example 8

Enzymatic Labeling of BoNT/Aad^\*

All expressed proteins incorporated S6 peptide GDLSLWILLRILN (SEQ ID NO:2) (Zhou et al., “Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Sfp and AcpP Phosphopantetheiny1 Transfereases,” ACS Chem. Biol. 2(5):337-346 (2007), which is hereby incorporated by reference in its entirety) at the N-terminus of the smaller subunit of the heterodimer The pET29 vector with cloned C-terminally His6-tagged Sfp phosphopantetheinyl transferase from B. subtilis was kindly provided by Dr. Jun Yin (University of Chicago). Recombinant Sfp phosphopantetheinyl transferase was expressed in E. coli strain BL21 (DE3) (Novagen, Cat. #69450) and purified according to the protocol. CoA 547 was purchased from Covalsy (Cat. #8V124) 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. Aliquots of this solution were used as a stock for further dilution in aqueous buffer-based labeling preparations. The labeling reaction was optimized and performed according to the following protocol: The molar ratio of labeled protein: Sfp phosphopantetheinyl transferase: CoA 547 was 4:1:5. A pilot reaction was performed by assembling approximately 18 µg of processed BoNT/Aad^\* in a 30 µl reaction mixture containing 1x unsupplemented Grace’s insect cell culture medium (prepared from the dry medium, Invitrogen, Cat. #11300-
043), 5 mM MgCl₂, 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. #P8595), 1 µM SiF₆₃ 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 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 TALON¹ resin loading buffer (500 mM NaCl, 25 mM Tris·HCl, pH 8.0) was added to the reaction mix and the resulting slurry was passed through TALON¹ chelating resin. Flow through fractions and three additional washes (each corresponding to the column volume) 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). All procedures that involved fluorescent substrate or labeled protein were performed in the dark.

Example 9

Structural Analysis of BoNT/Aad³⁴ and Fluorescently Labeled BoNT/Aad³⁴

Edman Degradation

BoNT/Aad³⁴ LC and HC, separated under reducing conditions on 10-14% Criterion SDS PAGE (Bio-Rad), were transferred to a PVDF membrane (Bio-Rad, Cat. #162-01258) and stained with Bio-Safe Coomassie BB (C-250) (Bio-Rad, Cat. #161-0786). Proteins were subjected to sequencing on an ABI 494-HT Procise Edman Sequencer at the Molecular Structure Facility at UC Davis.

In-Gel Tryptic Digestion

Gel bands were destained until clear using 30% acetonitrile in 25 mM ammonium bicarbonate. The gel slices were then cut into smaller pieces, approximately 1 mm² in size, and placed in a Speed Vac vacuum centrifuge for complete dehydration. After dehydration, 25 µL of 0.1% RapiGest SF (Waters, Milford, Mass.) was added and the samples were incubated at 37°C for 10 min. All remaining supernatant fluid was discarded and the samples were placed in a Speed Vac vacuum centrifuge for complete dehydration. Then 10 ng/µL sequencing grade trypsin (Promega Corporation, Madison, Wis.) in 50 mM ammonium bicarbonate was added to each sample and they were left to digest overnight at 37°C. Samples were extracted stepwise using acetonitrile, 1% formic acid, and acetonitrile again. Samples were vortexed at 800 rpm between steps and supernatant fluids were transferred to 0.5 mL Eppendorf tubes. Then 10 µL of 10% TFA was added to each extract followed by incubation at 37°C for 10 min. Samples were then dried in the vacuum centrifuge until approximately 3 µL of extract remained.

MALDI-TOF Analysis of BoNT/Aad³⁴ Tryptic Digest

Samples were desalted using C18 ZipTips (Millipore, Billerica, Mass.) and eluted with 50% acetonitrile, 0.1% formic acid after which 1 µL of sample was combined with 1 µL of alpha-cyano-4-hydroxy cinnamic acid (Agilent Technologies, Santa Clara, Calif.). Sample (1 µL) was then spotted on a Bruker 384 steel target frame (Billerica, Mass.) and allowed to air dry at room temperature. Samples were analyzed using a Bruker Autoflex MALDI-TOF mass spectrometer in positive ion reflectron mode under standard operating conditions. HPLC-Q-TOF MS/MS Analysis

A Q-TOF Premier mass spectrometer (Waters, Milford, Mass.) equipped with a Waters nano-ESI source coupled directly to a Nano-Acquity UPLC system (Waters) equipped with a 100 µm x 15 cm reverse phase column (BEH C18, Waters) was used for all LC-MS/MS analyses. Samples were directly loaded on the column using 0.1% formic acid at a flow rate of 0.8 µL/min for 20 min and eluted by a gradient of 1-40% acetonitrile in 0.1% formic acid over 40 min.

Samples for quantitation were analyzed by the mass spectrometer in TOF-MS mode while samples for identification by MS/MS were analyzed in DDA (data-dependent acquisition) mode. The mass range for all survey scans was 300-1400 m/z. Mascot software (version 2.2.1, Matrix Science, London, U.K.) was used for database searching and spectral interpretation.

Example 10

Expression, Purification, and Processing of BoNT/Aad³⁴

The full-length BoNT/A ad (atoxic derivatives) DNA and proteins were generated under biosafety level 2 containment (project approved by CDC on Feb. 7, 2006 for the registered entity C20060207-0419).

To improve the yield of the recombinant protein, the DNA sequence encoding the full-length construct was optimized for expression in both Sf9 insect cells and E. coli and synthesized de novo, as explained in supra.

A four amino acid insert, DDDE (SEQ ID NO:44), was introduced between amino acid residues N₄₄₅ and K₄₄₆ in the first construct, BoNT/Aad³⁴, to create a site recognized and cleaved by enterokinase (FIG. 11).

E. coli Experiments

The BoNT/Aad³⁴ insert was initially created in a non-expression vector derived from the Litmus 281 plasmid and named pLitSB3A. The insert DNA contains a 5' enhancer sequence placed upstream of the initiation codon and signal peptide, both of which are used exclusively for protein expression in the baculovirus system. To generate a vector suitable for E. coli expression, the DNA sequence encoding full-length BoNT/Aad³⁴ was isolated from pLitSB3A and subcloned downstream of the T7lac promoter into expression vector PET19. The resulting vector was transformed into E. coli strain JM109 (DE3), followed by growth of the transformed culture, IPTG induction, harvesting, lysing, SDS PAGE, and Western blotting with BoNT/A polyclonal antibodies. Under the conditions tested, BoNT/A-specific immunoreactivity in the processed samples was unable to be detected.

Next, expression of BoNT/Aad³⁴ was tested in a different expression vector and different E. coli strain. DNA encoding full-length BoNT/Aad³⁴ was isolated from pLitSB3A and subcloned downstream of T7lac promoter in the expression vector pETcoco2. The pETcoco system combines the advantages of T7 promoter-driven protein expression with the ability to control plasmid copy number. The pETcoco vectors are normally maintained at one copy per cell. In the single-copy state pETcoco clones are extremely stable, which is especially important for target genes toxic to the host. Copy number can be amplified to 20-50 copies per cell by the addition of L-arabinose to the culture medium. In ADE3 lysogenic hosts carrying pETcoco vectors expression of the target gene can be increased by as much as 2,500-fold over background
when IPTG is added to the culture media. The expression construct was transformed into E. coli Rosetta-gami B (DE3) competent cells and was grown in LB media with carbencillin, kanamycin, tetracycline, and chloramphenicol. Carbenicillin was added to maintain propagation of the cells carrying the pET10c-based bla marker, kanamycin and tetracycline were added to select for thioredoxin (trxD) and glutathione reductase (gor) mutations, thus improving the chances for proper disulfide bond formation in the E. coli cytoplasm (Derman et al., “Mutations that Allow Disulfide Bond Formation In the Cytoplasm of Escherichia coli.” Science 262 (5140):1744-1747 (1993); Prinz et al., “The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds In the Escherichia coli Cytoplasm,” J. Biol. Chem. 272(25):15661-15667 (1997), which are hereby incorporated by reference in its entirety). Chloramphenicol was added to the medium to maintain helper plasmids that provide tRNAs for rare codons, thereby increasing the expression of proteins encoded by DNA with codons non-cannonical for E. coli. Multiple conditions were tested to optimize expression of the BoNT/Aad⁎⁴. Cultures were grown with and without L-arabinose in the media, and different IPTG concentrations were evaluated for induction. Incubation temperatures and times were also optimized for expression. Optimal conditions were as follows: E. coli cultures were incubated overnight in the presence of L-arabinose at 37°C until reaching an OD₅₀₀ of approximately 0.4. The temperature of the bacterial suspension was then lowered to 25°C over 30 minutes, and IPTG was added to a final concentration of 0.5 mM. After induction, culture growth was allowed to continue on a shaker incubator at 25°C for twelve more hours. The bacterial pellet was harvested by centrifugation, lysed with BugBuster lysis reagent in the presence of benzamidine, lysozyme, and a cocktail of protease inhibitors. The lysate was cleared by centrifugation and purified by incubation with a TALON® affinity resin. The supernatant and eluate from the TALON® beads were run on SDS PAGE and analyzed by Western blot with polyclonal antibodies raised against the full-length BoNT/A holotoxin. Rosetta-gami B (DE3), transformed with the empty vector was used as the negative control. The expressed protein was partially soluble, immunoreactive to the polyclonal BoNT/A antibodies, and could be purified using the metal affinity tag. However, the molecular weight of the recombinant propeptide expressed was significantly lower than that of the native full-length BoNT/A propeptide. Extensive proteolysis was observed with all purification and expression protocols tested, even when the derivative was expressed with the single-copy plasmid. This instability may be related to the systems available in E. coli for post-translational processing of proteins, with improper folding and disulfide bonding making the recombinant protein susceptible to degradation.

Baculovirus Expression

The insert for expression of BoNT/Aad⁎⁴ was isolated from pLtSI3BA and subcloned into the donor vector pFast-Boel under control of the polyhedrin promoter. A 21 bp. cis-DNA sequence AACTCCCTAAAAACCGCCACC (SEQ ID NO:35) that was shown to increase the expression levels of exogenous genes in baculovirus-infected insect cells (Sano et al., “Enhancement of Protein Expression In Insect Cells by a Lobster Tropomyosin cDNA Leader Sequence,” FEBS Lett. 532(1-2):143-146 (2002), which is hereby incorporated by reference in its entirety) was positioned in front of the first methionine, upstream of DNA encoding the honeybee mellitin signal peptide KMTLNVNVALTMVVSVYIYAA (SEQ ID NO:36). The signal peptide is needed for transport of the expressed protein into the culture medium and is removed by processing during intracellular trafficking and secretion (von Hejne, “Signals for Protein Targeting Into and Across Membranes,” Subcell. Biochem. 22:1-19 (1994), which is hereby incorporated by reference in its entirety). For the purpose of protein purification, an N-terminal His tag and C-terminal Strep-tag II were also present in the construct. The generation of recombinant baculovirus and insect cell procedures are detailed supra. 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. After optimization of expression, the BoNT/ Aad⁎⁴ propeptide was highly enriched to virtual homogeneity from SF-900 II medium in two steps: metal chelate affinity resin (FIG. 2A), followed by Strep-Tactin affinity chromatography (FIG. 2B).

The purified protein was then processed to the heterodimer by cleavage with recombinant enterokinase. In the pilot reaction shown in FIGS. 3A-B, the optimal enzyme/protein ratio for the cleavage was determined. However, it was also noted that an excess of enterokinase led to non-specific protein degradation (FIG. 3B, lanes 4-7). 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⁎⁴ propeptide was processed with 0.007 units of enzyme per microgram of protein for 12 hours at 16°C, and resulted in ~95% completion of cleavage without visible degradation of light and heavy chains in the processed BoNT/Aad⁎⁴ 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⁎⁴ 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⁎⁴ propeptide mobility and the degree of the enzymatic cleavage were not evident on the Coomassie-stained gels. To examine and optimize conditions for TEV digest, a time course pilot reaction was conducted. Samples of BoNT/Aad⁎⁴ 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 FIG. 4. Incubation of one microgram of the BoNT/AAd⁎⁴ 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. 4, lane 10). The mobility of the propeptide band on SDS PAGE before and after the 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.

Some recombinant proteins expressed in a secreted form in the baculovirus system can be excessively glycosylated (Sydow et al., “Overexpression of a Functional NMDA Receptor Subunit (NMDAR1) In Baculovirus-Infected Trichoplusia Insect Cells,” Brain Res. Mol. Brain. Res. 41(1-2):228-240 (1996); Pechan et al., “Heterologous Expression of Maize (Zea mays L.) M1r1 Cysteine Proteinase in Fukuyotic and Prokaryotic Expression Systems;” Protein Expr. Purif. 34(1): 134-141 (2004), which are hereby incorporated by reference in its entirety). To rule out the presence of this post-translational modification, the mobility of the expressed, processed, and denatured BoNT/Aad⁎⁴ after treatment with Endo-o-acetylgalactosaminidase and PNGaseF (New
England Biolabs, Cat #P07335, P07045) was compared with untreated samples. No difference in mobility of treated versus untreated samples were detected by SDS PAGE.

The 12 aa S6 peptide tag placed downstream of the 10-His sequence and upstream of the N-terminus of the LC was incorporated as a target for Sfp phosphopantetheinyl transferase in BoNT/Aad⁶ and the other derivatives described herein. Sfp phosphopantetheinyl transferase catalyzes incorporation of small -molecule-CoA-based cargo to a specific serine residue within the S6 tag (Zhou et al., “Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Sfp and AcpS Phosphopantetheinyl Transferases,” ACS Chem. Biol. 2(5):337-346 (2007), which is hereby incorporated by reference in its entirety). As a prototypic cargo molecule, and to create a molecular probe to study BoNT/A trafficking, a commercially available fluorescent conjugate of CoA (CoA547, New England Biolabs) suitable for standard TAMRA and Cy3 microscopy emission filter sets was used for BoNT/Aad⁶ labeling.

A vector expressing recombinant C-terminally 6-His-tagged Sfp phosphopantetheinyl transferase from B. subtilis was expressed in E. coli, purified, and concentrated according to the cited procedure (Yin et al., “Site-Specific Protein Labeling by Sfp Phosphopantetheinyl Transferase,” Nat. Protoc. 1(1):280-285 (2006), which is hereby incorporated by reference in its entirety). 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 conditions tested, an increase of the CoA substrate concentration usually resulted in a higher background and did not necessarily increase the signal to background ratio. The addition of fetal calf serum or 0.5% BSA to the reaction mixture reduced the background staining.

During optimization of this site-specific BoNT/Aad⁶ 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 fluorescent derivative of the BoNT/Aad⁶ heterodimer was synthesized under physiological conditions, as described supra. The results of the enzymatic labeling are shown in Figs 5A-B. In Figs. 5A-B, labeled and unlabeled samples of BoNT/Aad⁶ 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. Under the tested conditions, 7% of the fluorescently labeled BoNT/Aad⁶ light chain were visualized, an amount otherwise undetectable on Coomassie-stained gels. 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.

N-Terminal Sequencing of BoNT/Aad⁶

BoNT/Aad⁶ LC and HC, separated under reducing conditions on 10-14% Criterion SDS PAGE (Bio-Rad), were transferred to PVDF membrane (Bio-Rad, Cat. #162-0182) and stained with Bio-Safe Coomassie BB G-250 (Bio-Rad, Cat. #161-0786). The proteins were subject to sequencing on an ABI 494-FIT Precise Edman Sequencer at the Molecular Structure Facility at UC Davis. N-terminal sequencing of LC identified the first six amino acids as GAGDSL (SEQ ID NO:37), and for the HC as ALNDLC (SEQ ID NO:38), which confirmed the predicted sequences of the protein N-termini.

Mass Spectrometric Analysis of BoNT/Aad⁶ and CoA547-BoNT/Aad⁶

FIG. 6 shows an ESI-Q-TOF MS/MS spectrum confirming the identity of the intact tryptic peptide from the C-terminus of BoNT/Aad⁶ light chain, and confirming the sequence of the predicted enterokinase cleavage site in the loop between LC and HC.

FIGS. 7A-C show MALDI-TOF MS (FIG. 7B) and ESI Q-TOF MS/MS (FIG. 7C) spectra of a tryptic peptide of sequence L1CVR (SEQ ID NO:39) from the BoNT/Aad⁶ light chain linked via a disulfide bridge to a tryptic peptide of sequence ALNDLC (SEQ ID NO:40) from the BoNT/Aad⁶ heavy chain. The peptide of observed m/z 1489.84 was detected in the MALDI-TOF mass spectrum of the non-reduced toxin (FIG. 7B), 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/Aad⁶ light chain (FIG. 7A). The sequence of the peptide was confirmed by the MS/MS spectrum (FIG. 7C), and the peptide representing the heavy chain was identified by Mascot database searching when including the mass of the disulfide bonded light chain peptide.

After enzymatic labeling of BoNT/Aad⁶ 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 to estimate the labeling efficiency of the reaction. Because the labeled N-terminal peptide was not observed by mass spectrometry, reduction of the amount of unlabeled peptide was used to estimate labeling efficiency. 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.

Example 11

Expression, Purification, and Processing of BoNT/Aad⁶

The BoNT/Aad⁶ construct is very similar to the BoNT/ Aad⁶ 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 for expression in the baculovirus system is provided in detail supra. The generation of baculovirus stock and the procedure for Sf9 infection and culture growth was similar to BoNT/Aad⁶, as described supra. The propeptide was purified using the 2-step affinity chromatography procedure, and yielded approximately 30 mg per liter of insect cell culture. Removal of the 10-His tag and processing of the propeptide were performed simultaneously by incubating 1 µg of the BoNT/Aad⁶ propeptide with 2 units of AcTEV protease at 30⁰C for 6 hours. Processed peptide was separated from AcTEV protease by affinity chromatography on TALON® resin and collected in flow-through fractions. Loss of protein during this procedure was ~5%. FIGS. 8A-C show the purified protein separated by SDS PAGE under reducing and non-reducing conditions.

Example 12

Expression, Purification, and Processing of ΔLC-Peptide -BoNT/Aad⁶ and ΔLC-GFP-BoNT/Aad⁶

The ΔLC 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 constructs for expression in the baculovirus system is provided in detail supra. The generation of baculovirus stock and the procedure for Sf9 infection and culture growth was similar to the previous two constructs, except for differences in MOI and incubation time after infection, which were approximately 0.5 and 72 hours, respectively. The propeptides were purified using 2-step affinity chromatography, as described supra. With both ΔLc 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-acyctyl-D-glucopranoside for Step I 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 units of AcETE protease at 30°C for 6 hours in the presence of 0.2% n-octyl-

β-D-glucopyranoside, which was removed by dialysis and ultrafiltration after the cleavage. Processed peptides were separated from AcETE protease by affinity chromatography on TALON® resin and collected in flow-through fractions. The yield of ALC-Peptide:BoNT/A" was approximately 2 mg per liter and ΔLc-GFP-BoNT/A" was 1 mg per liter of insect cell culture medium. FIGS. 8A-C show the purified proteins separated by SDS PAGE under reducing and non-

reducing conditions.

Example 13

Discussion of BoNT/A Derivatives that Retain Wild Type Features Required for Native Trafficking

Botulinum neurotoxin type A (BoNT/A) is the most toxic protein known, with LD50 values for mice of ~1 μg/g. The consequent high risk associated with handling large amounts of this toxin has hindered the study of BoNT/A absorption and trafficking in vitro and in vivo. High toxicity has also prevented use of quantities of toxin that can be detected using standard protein protocols. Potential solutions to this challenge include studying the protein domains expressed as separate entities, or studying atactic versions of the full-length toxin, as, for example, variants carrying point mutations that eliminate protein toxicity associated with light chain metalloprotease. Because BoNTs are large (150 kDa), multi-domain, disulfide-bonded heterodimers, with mutual stabilization of the protein domains through multiple hydrogen bonds and hydrophobic interactions, it is challenging to produce recombinant BoNTs that retain all of the structural features and trafficking properties of native BoNTs. Factors affecting the success of protein expression include the design of the expression construct and the choice of expression system. When domains are expressed separately, they can be denatured, poorly soluble (Ahmed et al., “Light Chain of Botulinum A Neurotoxin Expressed As an Inclusion Body From a Synthetic Gene Is Catalytically and Functionally Active,” J. Protein Chem. 19(6):475-487 (2000); Zhou et al., “Cloning, High-Level Expression, Single-Step Purification, and Binding Activity of His-6-Tagged Recombinant Type B Botulinum Neurotoxin Heavy Chain Transmembrane and Binding Domain,” Protein Expr. Purif. 34(1):8-16 (2004); Lacy et al., “Recombinant Expression and Purification of the Botulinum Neurotoxin Type A Translocation Domain,” Protein Expr. Purif. 11(2):195-200 (1997), which are hereby incorporated by reference in their entirety), or unstable (Baldwin et al., “The C-Terminus of Botulinum Neurotoxin Type A Light Chain Contributes to Solubility, Catalysis, and Stability,” Protein Expr. Purif. 37(1):187-195 (2004), which is hereby incorporated by reference in its entirety). BoNTs are also unusually sensitive to mechanical factors, and mild agitation has been reported to denature the active toxin structure (Toth et al., “Extreme Sensitivity of Botulinum Neurotoxin Domains Towards Mild Agitation,” J. Pharm. Sci. 98(9): 3302-3311 (2009), which is hereby incorporated by reference in its entirety). Even when the isolated domains are soluble and properly folded, their functionality in comparison with full-length native toxin is limited.

For these reasons, expression of mutated, atoxic forms of the full-length botulinum neurotoxin A was pursued. The mature wt BoNT/A is a heterodimer formed between light (LC) and heavy (HC) chains, connected via a disulfide bridge; a second, intra-chain disulfide bridge is also present at the C-terminus of the toxin’s receptor-binding domain (Kriegstein et al., “Covalent Structure of Botulinum Neurotoxin Type A: Location of Sulfhydryl Groups, and Disulfide Bridges and Identification of C-Terminal of Light and Heavy Chains,” J. Protein Chem. 13(1):49-57 (1994), which is hereby incorporated by reference in its entirety). An endogenous clostridial protease is involved in processing of wt BoNT/A propeptide to the mature form. The cleavage occurs at a basic amino acid residue in the loop between LC and HC. The recombinant BoNT/A derivatives described herein were expressed as single chain propeptides and were subsequently enzymatically processed to form LC-HC heterodimers, as in wt BoNT/A. However, the specificity of the cleavage was increased by introducing a more complex sequence (~5 aa) into the loop between LC and HC as a protease recognition site.

In prior work (U.S. Patent Application Publication No. 2006/0204524 to Ichchenko et al.; Ichchenko et al., “Full-Length Clostridium botulinum Serotype A Derivatives with Native Structure and Properties,” Neurotox. Res. 9:234 (2006), which are hereby incorporated by reference in their entirety) expression of several full-length BoNT/A derivatives was described. The DNA construct encoding the first of these was generated by consecutive subcloning of two phosphorylated linkers and five DNA fragments obtained from PCR reactions performed on genomic DNA isolated from Clostridium botulinum A1 Hall strain. Although this derivative was not able to be expressed in E. coli, the baculovirus system enabled its expression in a secreted form that could be readily purified from the culture medium without harsh processing. The major problems encountered in earlier baculovirus work were related to low yield of the protein after purification (approximately 0.35 mg per liter of insect cell culture medium) and the inability to establish a single step purification procedure based on the original 6-His tag incorporated at the N-terminus of the construct for this purpose. Either because of insufficient length, or partial burial of the tag in the globule of the light chain, protein was eluted from the Ni2+ affinity resin in 40 mM imidazole along with multiple major contaminants. As a solution, a 4-step conventional/affinity chromatography protocol was established. The complexity of the protocol contributed to the low yield of the final product. These results were considered when designing and synthesizing the constructs described herein.

The BoNT/A ad constructs reported here represent full length botulinum neurotoxin type A carrying two mutations, E234A and Y360A, introduced in the light chain. Both original amino acids, E234 and Y360, are conserved among different botulinum neurotoxin serotypes and are part of the catalytic core of the light chain metalloprotease responsible for cleavage of the substrate. The light chain in mutated full-
length BoNT/A ad lacks the ability to cleave SNAP 25 (Rigoni et al., "Site-Directed Mutagenesis Identifies Active Site Residues of the Light Chain of Botulinum Neurotoxin Type A." Biochem. Biophys. Res. Commun. 288(5):1251-1257 (2001); Zhou et al., "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 34(46):15175-15181 (1995); Li et al., "Probing the Mechanistic Role of Glutamate Residue In the Zinc-Binding Motif of Type A Botulinum Neurotoxin Light Chain," Biochemistry 39(9):2399-2405 (2000), which are hereby incorporated by reference in its entirety. It was also shown that mutation of E224 and Y366 did not change the secondary structure, topography of aromatic amino residues, Zn²⁺ content, or substrate binding ability of the LC metalloprotease (Li et al., "Probing the Mechanistic Role of Glutamate Residue In the Zinc-Binding Motif of Type A Botulinum Neurotoxin Light Chain," Biochemistry 39(9):2399-2405 (2000), which is hereby incorporated by reference in its entirety). Several popular expression systems, including E. coli and Pichia pastoris, show high bias against clostridial AT-rich DNA, resulting in slow growth (Balduini et al., "The C-Terminus of Botulinum Neurotoxin Type A Light Chain Contributes to Solubility, Catalysis, and Stability," Protein Expr. Purif. 37(1):187-195 (2004), which is hereby incorporated by reference in its entirety), premature termination of protein synthesis, or initiation of irrelevant translation from an alternative reading frame (Lacy et al., "Recombinant Expression and Purification of the Botulinum Neurotoxin Type A Translocation Domain," Protein Expr. Purif. 11(2):195-200 (1997), which is hereby incorporated by reference in its entirety). Similar problems were observed in attempts to express protein in E. coli, and the initial attempts in the baculovirus expression system resulted in low protein yields. Typically, the use of E. coli strains supplemented with rare tRNAs, silent mutagenesis of the native clostridial DNA, or codon-optimized synthetic DNA are used in attempts to resolve this problem (Ahmed et al., "Light Chain of Botulinum A Neurotoxin Expressed As An Inclusion Body From A Synthetic Gene Is Catatytically and Functionally Active," J. Protein Chem. 19(6):475-487 (2000); Sutton et al., "Preparation of Specifically Activatable Endopeptidase Derivatives of Clostridium botulinum Toxins Type A, B, and C and Their Applications," Protein Expr. Purif. 40(1):31-41 (2005), which are hereby incorporated by reference in their entirety). Here, a fully synthetic DNA sequence was created de novo, encoding full-length constructs that were optimized for expression in both S. coelic and E. coli, as explained supra. A 7 aa spacer sequence was introduced, to separate the N-terminus of the light chain domain from the upstream length. In the constructs described here the length of the metal affinity tag was increased from 6 to 10 histidine residues, which significantly improved the intended single step affinity purification using non-denaturing conditions.

The constructs described herein additionally contain an 8 aa StrepTag II added at their C-termini to improve the selectivity of purification. While the addition of extra amino acids to the C-terminus of the clostridial neurotoxins causes some concern for interference with the protein binding to its cognate receptors, recent data obtained from X-ray studies of a StrepTag II C-terminally fused BoNT/B in complex with its receptors showed that the amino acid sequence of this tag is spatially separated from the receptor-binding epitopes of BoNT/B (Jia et al., "Botulinum Neurotoxin B Recognizes Its Protein Receptor With High Affinity and Specificity," Nature 444(7122):1092-1095 (2006), which is hereby incorporated by reference in its entirety). Therefore, it was decided to use this tag in the design of the neurotoxins described herein. The benefit of having both N- and C-terminally tagged protein is to allow separation of the full-length protein from N- or C-terminally truncated products generated adventitiously in the host expression system.

In early structural studies of BoNT/A isolated from Clostridium botulinum, it was shown that an endogenous protease, or trypsin, can cleave BoNT/A propeptide at K344, followed by cleavage at K344, which generates the tetrapeptide G443YNK446 (SEQ ID NO:41) which is released as a cleavage product (DasGupta et al., "Botulinum Neurotoxin Type A: Sequence of Amino Acids at the N-Terminus and Around the Nicking Site," Biochemistry 72(9):661-664 (1990), which is hereby incorporated by reference in its entirety). In subsequent work, K348 was identified as the C-terminus of the mature LC, generated by the excision of the 10 amino acid peptide, T350KSLDKYNK446 (SEQ ID NO:42) from BoNT/A precursor (Kriegel et al., "Cova lent Structure of Botulinum Neurotoxin Type A: Location of Sulhydryl Groups, and Disulfide Bridges and Identification of C-Term in of Light and Heavy Chains," J. Protein Chem. 13(1):49-57 (1994), which is hereby incorporated by reference in its entirety). The most C-terminal amino acid of the light chain visible in the published X-ray structure of the BoNT/A holotoxin is R332 (Lacy et al., "Crystal Structure of Botulinum Neurotoxin Type A and Implications for Toxicity," Nat. Struct. Biol. 5(10):898-902 (1998), which is hereby incorporated by reference in its entirety). There is thus a discrepancy between the native form of the light chain, extending to position 438, and the form seen in the crystal structure, extending only to position 432. This discrepancy, which could influence BoNT properties, could arise because the C-terminal residues beyond R332 of the light chain in crystal are present, but are too flexible to generate distinct electron density on X-ray. Alternatively, a trypsin-like protease may have cleaved the LC-HC propeptide at R332, so that this fragment is not actually present in the crystal structure. It was rationalized that the constructs described herein should not be shortened by the length of the peptide loop excised from wt BoNT/A precursor because this might influence the 3D structural constraints on this region, but should rather be made resistant to trypsin-like proteases. Therefore, mutations K348H, K444Q, and K444Q were introduced into the constructs described herein to render the 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 HC₃-domain, and complete loss of toxicity (Shone et al., "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. 151(1):75-82 (1985), which is hereby incorporated by reference in its entirety). Therefore, the mutation I117N at the HC₃-HC₄ junction was introduced into the constructs described herein, rendering the HC insensitive to this type of non-specific cleavage as well (Chaddock et al., "Expression and Purification of Catalytically Active, Non-Toxic Endopeptidase Derivatives of Clostridium botulinum Toxin Type A," Protein Expr. Purif. 25(2):219-228 (2002); Shone et al., "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. 151(1):75-82 (1985), which are hereby incorporated by reference in their entirety).
The biological activity of botulinum neurotoxins requires proper folding and disulfide bonding of the protein during post-translational processing. There are 9 cysteine residues in wt BoNT/A, five of which carry free sulhydryl groups and 4 of which are involved in formation of two disulfide bridges (Kriegstein et al., “Covalent Structure of Botulinum Neurotoxin Type A: Location of Sulphydryl Groups, and Disulfide Bridges and Identification of C-Termini of Light and Heavy Chains,” J. Protein Chem. 13(1):49-57 (1994), which is hereby incorporated by reference in its entirety). Expression of individual BoNT/A domains with endogenous cysteines in the reduced state has been reported to contribute to formation of protein aggregates (Baldwin et al., “The C-Terminal of Botulinum Neurotoxin Type A Light Chain Contributes to Solubility, Catalysis, and Stability,” Protein Expr. Purif. 37(1):187-195 (2004); Lacy et al., “Recombinant Expression and Purification of the Botulinum Neurotoxin Type A Translocation Domain,” Protein Expr. Purif. 11(2):195-200 (1997), which are hereby incorporated by reference in their entirety). The use of maltose-binding protein as a fusion partner for expression of LC-HC, clostridial polypeptides in E. coli has been beneficial for formation of five S-S bonds in E. coli-based expression systems, albeit the nature of the bond(s) formed was not confirmed (Sutton et al., “Preparation of Specifically Activatable Endoproteinase Derivatives of Clostridium botulinum Toxins Type A, B, and C and Their Applications,” Protein Expr. Purif. 40(1):31-41 (2005), which is hereby incorporated by reference in its entirety). Maltose-binding protein probably not only contributes to solubility of the fusion polypeptides, but rather acts as a chaperone, promoting partial export of the fusion protein to the oxidizing environment of the periplasm. A recent review describes a variety of methods to optimize expression of cysteine-containing proteins in E. coli (de Marco, “Strategies for Successful Recombinant Expression of Disulfide Bond-Dependent Proteins In Escherichia coli,” Microbiol. Cell. Fact. 8.26 (2009), which is hereby incorporated by reference in its entirety). The currently available reports of soluble, properly folded, and biologically active, full-length clostridial polypeptides expressed in E. coli did not utilize strains that contain any of these specialized features required for proper post-translational modifications (Rummel et al., “Two Carbohydrate Binding Sites in the HC, Domain of Tetanus Neurotoxin Are Required for Toxicity,” J. Mol. Biol. 326(3):835-847 (2003); Rummel et al., “The HC, Domain of Botulinum Neurotoxins A and B Exhibits a Singular Ganglioside Binding Site Displaying Serotype Specific Carbohydrate Interaction,” Mol. Microbiol. 51(3):631-643 (2004); Rummel et al., “Synaptotagmins I and II Act as Nerve Cell Receptors for Botulinum Neurotoxin G,” J. Biol. Chem. 279(29):30865-30870 (2004); Bude et al., “Botulinum Neurotoxin Type E Enables Cytosolic Delivery of Enzymatically Active Cargo Proteins to Neurons Via Unfolded Translocation Intermediates,” J. Neurochem. 91(6):1461-1472 (2004); Li et al., “Recombinant Forms of Tetanus Toxin Engineered for Examining and Exploiting Neuronal Trafficking Pathways,” J. Biol. Chem. 276(33):31394-31401 (2001), which are hereby incorporated by reference in their entirety). Attempts to obtain full-length BoNT/A (104 kDa) in two different strains of E. coli have not been successful. However, efforts with the baculovirus expression system have been successful on multiple levels, as described herein.

Four distinct BoNT/A derivatives were expressed and characterized using the baculovirus system. The first derivative, BoNT/Aad, represents the substantially intact BoNT/A, with two mutations, F 3254->A and Y 3607->A, introduced to inactivate the toxin protease, and thereby render the derivative toxoid, and with an enterokinase site for processing the expressed propeptide into the heterodimer. This derivative was expressed as a soluble single chain protein secreted into insect cell culture medium, and was purified to homogeneity by chromatography on metal affinity resin and StrepTactin column, yielding 30 milligrams per liter of culture medium. The propeptide was treated with AcTEV protease to remove the N-terminal 10-His tag, and was treated with enterokinase to produce the LC-HC 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 at the LC-HC junction through identification of the predicted C-terminal peptide of the LC (FIGS. 7A-C). The presence of the disulfide bridge formed between Cys 533 and Cys 542 (as numbers are from the sequence of wt BoNT/A in the expressed protein was confirmed by MALDI-TOF and ESI-Q-TOF LC-MS/MS analysis through identification of the S-S linked dipeptide in tryptic digests of processed BoNT/A (104 kDa) heterodimer (FIG. 8).

The second derivative, BoNT/Aad 7A-C is similar to BoNT/Aad, but replaces the enterokinase site between LC and HC with a TEV recognition sequence. This avoids non-specific cleavage associated with excess enterokinase, and reduces the number of steps required for protein processing. This derivative was also expressed as a soluble single chain protein secreted into insect cell culture medium, and was purified to virtual homogeneity by chromatography on metal affinity resin and StrepTactin column, yielding 30 milligrams per liter of culture medium. The propeptide was further processed with AcTEV protease to simultaneously remove the N-terminal 10-His tag and to generate the LC-HC heterodimer. 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 (FIGS. 8A-C).

In the third derivative, LC-Peptide-BoNT/A, the entire catalytic domain (P 33-390) of the light chain was removed, leaving a 54 aa LC segment (N 53-G 193) that forms the disulfide bridge and otherwise interacts with the HC. Two TEV protease cleavage sites in this derivative allowed simultaneous removal of the N-terminal 10-His tag and processing of the precursor to generate the LC-HC heterodimer. This derivative was expressed as a soluble single chain protein secreted into insect cell culture medium, but was found to have a tendency to precipitate during concentration prior to the first affinity chromatography. To prevent precipitation, TRITON® X-100 was added to the medium during concentration and was maintained throughout processing. The protein was eluted from TALON® resin and purified further on a StrepTactin column in the presence of n-octyl-β-D-glucopyranoside. After processing with AcTEV protease to generate the LC-HC heterodimer and remove the 10-His affinity tag, it became possible to remove detergent because the solubility of the processed heterodimer increased markedly. Yield of the protein was 2 milligrams per liter of culture medium. The purified AcTEV-processed protein migrated as ~110 kDa band on SDS PAGE in the absence of reducing agent and, as expected in the presence of β-mercaptoethanol, separated as LC-remnant and HC with apparent mobilities of 10 and 100 kDa, respectively. Western blotting performed with polyclonal antibodies raised against BoNT/A holotoxin confirmed the structural identity of the LC-Peptide-BoNT/A heavy chain, but did not detect the truncated LC (FIGS. 8A-C).
The fourth derivative, δLC-GFP-BoNT/A", originates from the third derivative by insertion of green fluorescent protein (GFP) into the sequence of the short LC remnant peptide at its N-terminus. It likewise includes the 54 aa LC segment (N39,G44) that forms the disulfide bridge and otherwise interacts with the HC. Two TEV protease cleavage sites in this derivative allow simultaneous removal of the N-terminal 10-His tag and processing of the precursor to the LC-HC heterodimer. Aggregation of this protein during concentration prior to chromatography also required use of a combination of detergents and the purification scheme described supra for δLC-Peptide-BoNT/A". Yield of the protein was 1 milligram per liter of culture medium. The purified AcTeV-processed protein migrated as ~140 kDa band on SDS PAGE in the absence of reducing agent and, as expected in the presence of β-mercaptoethanol, migrated as inseparable Biocytin-LC-GFP and HC with apparent mobilities of 40 and 100 kDa, respectively. Western blotting performed with polyclonal antibodies raised against BoNT/A holotoxin and monoclonal antibody against GFP confirmed the structural identity of δLC-GFP-BoNT/A" light and heavy chains (Figs. 8A-C).

The relatively low yield and aggregation of derivatives three and four, where the majority of the LC has been removed, is attributed to conformational instability and spatial tension in the expressed proteptides. The published X-ray structure of BoNT/A holotoxin (Lacy et al., “Crystal Structure of Botulinum Neurotoxin Type A and Implications for Toxicity,” Nat. Struct. Biol. 5(10):898-902 (1998), which is hereby incorporated by reference in its entirety) suggests that the N-terminal portion of the N39,G44 sequence is a flexible loop and not part of the distinct protein fold. It was hypothesized that deletion of the entire LC catalytic domain exposes hydrophobic areas of the proteptide 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 untying one end so that the proteptide can collapse into a more soluble conformation.

During early pilot stages using natural genes instead of the synthetic genes reported here, to evaluate the utility of the baculovirus expression system and the physiological activity of the designed constructs, a full-length BoNT/A derivative without any introduced mutations that would inactivate the LC metalloprotease was expressed. When the non-mutated derivative was tested in vitro and in vivo, it demonstrated toxicity and absorption kinetics comparable to that of the native toxin. The expressed toxic derivative was tested in mouse phrenic nerve-hemidiaphragm preparations. Exocytosis was evoked by stimulation of the nerve trunk (0.2 Hz), and the muscle twitch was monitored. When added at a concentration of ~1×10^{-11} M, the toxic derivative produced blockage of transmission in 167±17 min (n=4), comparable to pharmaceutical preparations of BoNT/A. To ensure that the blockade was attributed to a botulinum toxin-type action, the derivative was pre-incubated (room temperature, 60 min) with rabbit antiserum raised against the carboxy-terminal half of the native BoNT/A heavy chain (i.e., receptor-binding domain). When these experiments (n=3) were performed in the presence of this polyclonal BoNT/A antiserum, there was no onset of paralysis throughout the duration of the experiment (approximately 400 minutes of tissue monitoring) (U.S. Patent Application Publication No. 2006/0204524 to Ichtech-enko et al., Ichtechenko et al., “Full-Length Clostridium botulinum Serotype A Derivatives with Native Structure and Properties,” Neurotox. Res. 9:234 (2006), which are hereby incorporated by reference in their entirety).

According to one embodiment, the recombinant BoNT/A derivatives described herein carry an S6 tag in their N-terminal region. Specific conditions for in vitro CoA-fluorophore labeling were optimized and shown for BoNT/Aad". 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: (i) transient expression following transfection with LC sequences tagged with a fluorescent marker (e.g., GFP) (Oyler et al., “Trafficking and Post-Translational Modifications of BoNT Light Chains Within Cells,” Abstracts of the 5th International Conference on Basic and Therapeutic Aspects of Botulinum and Tetanus Toxins, Denver, Colo. 24 (2005); Fernandez-Salas et al., “Plasma Membrane Localization Signals In the Light Chain of Botulinum Neurotoxin,” Proc. Natl. Acad. Sci. (USA) 101(9):3208-3213 (2004); Fernandez-Salas et al., “Is the Light Chain Subcellular Localization an Important Factor In Botulinum Toxin Duration of Action?” Mov. Disord. 19(Suppl 8):S23-S34 (2004), which are hereby incorporated by reference in their entirety), or (ii) attachment of fluorescent tags to LC-HC BoNT heterodimers using chemical methods (Verderio et al., “Internalization and Proteolytic Action of Botulinum Toxins in CNS Neurons and Astrocytes,” J. Neurochem. 73(1):372-379 (1999); Lalli et al., “Functional Characterization of Tetanus and Botulinum Neurotoxins Binding Domains,” J. Cell. Sci. 112(Pt 16):2715-2724 (1999); Grumelli et al., “Internalization and Mechanism of Action of Clostridial Toxins In Neurons,” Neurotoxicology 26(5):761-767 (2005); Alhassan et al., “Visualization of Binding and Transcytosis of Botulinum Toxin by Human Intestinal Epithelial Cells,” J. Pharmacol. Exp. Ther. 315(3):1028-1035 (2005), which are hereby incorporated by reference in their entirety). 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 could 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. One study showed that a protein with relatively simple spatial organization, such as ~50 kDa glutathione S-transferase (GST), can lose 90% of its activity as a result of chemical labeling, while targeted enzymatic incorporation of a fluorescent probe into GST completely preserved its biological function (Taki et al., “Transglutaminase-Mediated N- and C-Terminal Fluorescein Labeling of a Protein Can Support the Native Activity of the Modified Protein,” Protein Eng. Des. Sel. 17(2):119-126 (2004), which is hereby incorporated by reference in its entirety). In the original work describing the design and selection of the S6 tag for site-selective enzymatic attachment of fluorophores to recombinant proteins, the ligand-binding properties of two enzymatically labeled proteins, EGF and transferrin receptors, were tested. The data confirmed that site-specific labeling of the expressed receptors with CoA-fluorophore did not interfere with physiological binding or recycling of either EGF or transferrin (Zhou et al., “Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Sfp and Acp5 Phosphopantetheinylation Trans-


One aspect of the present invention provides experimental evidence demonstrating enzymatic site-specific fluorescent labeling of a recombinant full-length BoNT/A derivative, with the fluorophore attached to a specific site located on the N-terminal peptide, designed as such to minimally interfere with native BoNT structure. Interestingly, the ability to incorporate CoA 547 into recombinant protein also indirectly confirms the sorting mechanism through which these derivatives are secreted into insect culture medium. Sfp phosphopantetheinylation transferase is a ubiquitous enzyme that is present in the cytoplasm of insect cells, and would therefore be expected to attach endogenous CoA to the BoNT derivatives described herein if they were not secreted into the culture medium and thereby separated from the endogenous phosphopantetheinylation transferase in the insect cell cytoplasm. The unmodified serine residue within the S6 tag identified by Edman degradations confirms that the recombinant protein undergoes a sorting mechanism that protects it from enzymatic action of the endogenous cytoplasmic Sfp phosphopantetheinylation transferase in the Sf9 cell expression host. According to mass spectrometric analysis of CoAS547-modified versus non-modified derivative, the fluorophore was attached to approximately 69.6±6% of the BoNT/Aad46 added to the enzymatic labeling reaction mixture. With the existing commercially available derivatives, such as CoA-biotin (New England Biolabs, Cat. #S9351S), it is possible to achieve 100% incorporation via streptavidin affinity enrichment. The chemistry behind coupling of various moieties to CoA has been described (Yin et al., “Site-Specific Protein Labeling by Sfp Phosphopantetheinylation Transferase,” *Nat. Protoc.* 1(1):280-285 (2006), which is hereby incorporated by reference in its entirety) and can provide almost unlimited opportunities in terms of coupling small molecules, bioactive peptides, and peptidomimetics to the toxic BoNT/A derivatives described herein, exemplifying a versatile platform for targeting diverse therapeutic agents to the neuronal cytoplasm.

Three previously mentioned reports describe expression and purification of recombinant, atoxic, full-length BoNT derivatives, with point mutations to the LC active site (Kiyatkin et al., “Induction of an Immune Response by Oral Administration of Recombinant Botulinum Toxin,” *Infect. Immun.* 65:4586-4591 (1997); Pier et al., “Recombinant Holotoxoid Vaccine Against Botulism,” *Infect. Immun.* 76(1):437-442 (2008); Webb et al., “Production of Catalytically Inactive BoNT/A1 Holoprotein and Comparison With BoNT/A1 Subunit Vaccines Against Toxin Subtypes A1, A2, and A3,” *Vaccine* 27(33):4490-4497 (2009), which are hereby incorporated by reference in their entirety). All three of these recombinant holotoxoids were developed with the specific intention of producing a recombinant BoNT vaccine, rather than with the intention of developing probes for BoNT trafficking studies, or delivery vehicles that can target the neuronal cytosol. The recombinant metalloprotease-inactivated BoNT/C reported by Kiyatkin et al., “Induction of an Immune Response by Oral Administration of Recombinant Botulinum Toxin,” *Infect. Immun.* 65:4586-4591 (1997), which is hereby incorporated by reference in its entirety, was non-toxic at doses up to 10 μg per mouse. BoNT/A46 described by Pier et al., “Recombinant Holotoxoid Vaccine Against Botulism,” *Infect. Immun.* 76(1):437-442 (2008), which is hereby incorporated by reference in its entirety showed no apparent toxicity up to doses of 1 μg per mouse, and cibro/A HP described by Webb et al., “Production of Catalytically Inactive BoNT/A1 Holoprotein and Comparison With BoNT/A1 Subunit Vaccines Against Toxin Subtypes A1, A2, and A3,” *Vaccine* 27(33):4490-4497 (2009), which is hereby incorporated by reference in its entirety, was non-toxic at doses up to 50 μg per mouse, indicating that the respective holotoxoids can elicit immunity at the non-toxic doses mentioned supra. Although data on toxicity at higher doses was not included in these reports, the relative absence of toxicity in different vaccine candidates might be considered somewhat surprising if in fact the recombinant proteins follow the native BoNT targeting pathway. BoNT/A LC mutated in its active site has been shown in crystallographic studies to still be capable of binding its substrate SNAP-25 (Breidenbach et al., “Substrate Recognition Strategy for Botulinum Neurotoxin Serotype A,” *Nature* 432:925-929 (2004), which is hereby incorporated by reference in its entirety) and if delivered to neurons at high doses might be expected to bind and potentially interfere with the excytotic machinery at high doses. Second, HCs delivered to the endosomal compartment of motor neurons might be expected to retain their pore-forming capability in the acidic endosomal environment, and likewise potentially disrupt physiologic neuronal activity (Shone et al., “A 50-KDa Fragment From the NH2-Terminus of the Heavy Subunit of Closstridium botulinum Type A Neurotoxin Forms Channels In Lipid Vesicles,” *Eur. J. Biochem.* 167(1):175-180 (1987), which is hereby incorporated by reference in its entirety). Since the mutated LC does not disable the apparatus for its own endocytosis, these recombinant holotoxoids should accumulate in neurons when administered at high concentrations, and might be expected to have some effect on neuronal function. It should be noted in this context that immunogenic activity can be retained in the absence of structural features required for native BoNT trafficking. Thus, although it is clear that these recombinant holotoxoids are effective vaccine candidates and that their low toxicity may provide the acceptable therapeutic index required, the data presented do not address the question of whether the recombinant holotoxoids follow the native BoNT trafficking pathway, and therefore do not currently support their potential utility for BoNT trafficking studies, or as delivery vehicles that can target the neuronal cytosol.

The above data demonstrate that the baculovirus expression system, used in conjunction with synthetic gene constructs described herein, enables production of a series of atoxic, full-length, and truncated BoNT derivatives that preserve important structural features of native BoNT. 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 obviates the need to recover insoluble expressed derivatives from inclusion bodies using denaturing conditions. Retaining native BoNT structure during harsh solubilization and purification steps is challenging. Renaturation is equally challenging, because of the large, complex, disulfide-bonded structure of BoNTs. Attempts to restore disulfide bonding in recombinant BoNT/A1 HC, domain expressed in P. pastoris, which is only ½ of the full-length molecule, illustrate the difficulty of achieving physiologically relevant protein folding during recombinant protein expression (Bouvier et al., “Identifying and Modulating Disulfide Formation In the Biopharmaceutical Production
of a Recombinant Protein Vaccine Candidate,” J. Biotechnol. 103(3):257-271 (2003), which is hereby incorporated by reference in its entirety). Another recent publication highlights vulnerability of botulinum neurotoxins to rapid and irreversible denaturation during handling, as even mild agitation was shown to alter the secondary structure of HCs and LCs from diverse BoNT serotypes (Toth et al., “Extreme Sensitivity of Botulinum Neurotoxin Domains Towards Mild Agitation,” J. Pharm. Sci. 98(9):3302-3311 (2009), which is hereby incorporated by reference in its entirety). 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. Because the expression and purification methodology employed in the present invention circumvents all types of denaturing conditions, the expressed BoNT derivatives described herein have been designed to retain native BoNT structure to a greater extent than methods requiring exposure to harsh reagents.

Example 14

BoNT/Aad<sup>e</sup> is Non-Toxic to Neuronal Cultures and Can Effectively Compete with wt BoNT/A

To confirm that BoNT/A ad is non-toxic, but otherwise has characteristics similar to wt BoNT/A, primary rat spinal cord cells were exposed to as much as 500 nM BoNT/A ad. In particular, these primary rat spinal cord neurons were used to determine whether BoNT/Aad<sup>e</sup> itself has any observable effects on cultured neuronal cells, and whether it has the ability to compete with, or otherwise antagonize, the effects of wt BoNT/A on the neuronal cultures. This assay can detect toxicity of femtomolar amounts of wt BoNT/A (measured by SNAP 25 cleavage) (Pellett et al., “A Neuronal Cell-Based Botulinum Neurotoxin Assay for Highly Sensitive and Specific Detection of Neuronal Secretion Antibodies,” FEBS Lett. 581(25): 4803-08 (2007); Pellett et al., “Comparison of the Primary Rat Spinal Cord Cell (RSC) Assay and the Mouse Bioassay for Botulinum Neurotoxin Type A Potency Determination,” J. Pharmacol. Toxicol. Methods 61(3):304-310 (2010), which are hereby incorporated by reference in their entirety), yet incubation of cells with 500 nM BoNT/A ad resulted in no detectable BoNT-specific proteolytic activity (Fig. 10). In other words, exposure of cells to as much as 500 nM BoNT/Aad<sup>e</sup> did not result in detectable cleavage of intracellular SNAP 25, the target of the wt toxin.

Structural and partial functional identity of wt BoNT/A and BoNT/A ad was further confirmed by the dose-dependent ability of BoNT/A ad to block SNAP 25 cleavage by wt BoNT/A (Fig. 10). In this competition assay, pre-incubation of the primary rat spinal cord cells with 500 nM BoNT/Aad<sup>e</sup> blocked intracellular SNAP 25 cleavage induced by 0.5 nM wt BoNT/A.

Example 15

Toxicity Studies of BoNT/Aad<sup>e</sup>

To examine toxicity in vivo, the LD<sub>50</sub> of BoNT/A ad was determined by mouse bioassay, and was approximately 50 µg/kg intraperitoneally, which is about 100,000-fold higher than the LD<sub>50</sub> of wt BoNT/A.

The wt BoNT/A is targeted to the neuromuscular junction where it cleaves SNAP 25 and causes neuromuscular paralysis by disabling the machinery of regulated exocytosis. Specificity of BoNT/A ad binding to the presynaptic sites at the neuromuscular junction (“NMJ”) was confirmed by immunocytochemical analysis of triangularis sterni nerve-muscle preparations after in vivo injection of mice (FIGS. 11A-C). These results indicate that residual toxicity of BoNT/A ad is associated with NMJ-specific accumulation.

Next, it was determined if the light chain of BoNT/A ad undergoes translocation to the neuronal cytosol. Rat hippocampal neuronal cultures (Vicario, “Long-term Culture of Hippocampal Neurons,” in Current Protocols in Neuroscience Suppl. 26: 3.2.1-3.2.13 (John Wiley & Sons, Inc. 2004), which is hereby incorporated by reference in its entirety) were exposed to BoNT/A ad, and analyzed by immunostaining using an antibody, MAb F1-40, which is a well characterized BoNT/A L-chain specific antibody (Stunker et al., “Development and Partial Characterization of High-affinity Monoclonal Antibodies for Botulinum Toxin Type A and Their Use in Analysis of Milk by Sandwich ELISA,” J. Immunol. Methods 336 (1):1-8 (2008); Scoetri et al., “Epitope Characterization and Variable Region Sequence of F1-40, a High-Affinity Monoclonal Antibody to Botulinum Neurotoxin Type a (Holl strain)” PLoS One 4 (3):e4924 (2009), which are hereby incorporated by reference in their entirety) with high specificity, sensitivity, and reproducibility in immunocytochemistry.

Immunostaining of neurons with three different treatment and chase regimens of BoNT/A ad is shown in FIG. 12. BoNT/A ad uptake could be visualized when neuronal cultures were exposed to BoNT/A ad for 30-90 minutes at 37°C as an extracellular punctate pattern (FIG. 12, Columns B and D, Row 1), representing the active synaptic contacts and points of BoNT/A ad entry at the axon termini. This staining gradually disappears during the chase (FIG. 12, Column D, Rows 2 and 3), due to continuous intracellular uptake of the extracellular BoNT/A ad and absence of additional recombinant protein in the medium. When cells are chased for 90 min after incubation with BoNT/A ad (FIG. 12, Row 3), most of the LC ad staining is concentrated intracellularly, adjacent to the plasma membrane and co-localized with SNAP 25. This type of staining has never been shown for wt BoNT/A because the amount of wt LC delivered to the neuronal cytoplasm is exceedingly low (below the level for visualization).

The immunostaining of the neurons shown in Row 3 of FIG. 12 is consistent with earlier reports where LCA/GFP constructs were expressed in neuronal and nonneuronal cultures. The pattern of the LCA/GFP distribution after transfection shows that intrinsic properties of LCA contributed to accumulation of LCA/GFP on the inner leaflet of the neuronal plasma membrane after expression (Fernández-Salas et al., “Is the Light Chain Subcellular Localization an Important Factor in Botulinum Toxin Duration of Action?” Mol. Disorder., 19, Suppl8: 23-34 (2004); Wang et al., “Novel Chimeras of Botulinum Neurotoxins A and E Unveil Contributions From the Binding, Translocation, and Protease Domains to Their Functional Characteristics” J. Biol. Chem. 283(25): 16993-17002 (2008); Tsai et al., “Targeting Botulinum Neurotoxin Persistence by the Ubiquitin-proteasome System,” Proc. Natl. Acad. Sci. USA., 107 (38):16554-16559 (2010), which are hereby incorporated by reference in their entirety).

It was then directly examined whether LC ad interacted with SNAP 25 in the cytosol of neuronal cells exposed to BoNT/A ad. As shown in FIGS. 13A and 13B, the ~52 kDa band corresponding to LC ad was co-precipitated with an anti-SNAP 25 antibody. Neither SNAP 25 nor LC ad were detected in a control experiment with BHK fibroblasts treated with BoNT/A ad under the same conditions. If LC ad binds SNAP 25 in the neuronal cytoplasm without cleaving it, then
binding could lead to sequestration of SNAP 25 from the tripartite complex and disruption of neuronal activity. This would suggest that the toxicity of BoNT/A ad at high doses was a result of increased LC ad accumulation in the neuronal cytosol due to nitroimidazole inactivation—a feature beneficial for neuronal delivery of a wide range of therapeutic moieties. Therefore, it was examined whether exocytosis was affected by exposure of neuronal cells to BoNT/A ad.

Rat hippocampal neurons were pre-loaded under depolarizing conditions with FM 143 dye. When 100-500 nM BoNT/A ad was added to the depolarizing medium, an inhibition of exocytosis (fusion of FM-143 labeled synaptic vesicles with plasma membrane) occurred in a dose-dependent manner. Similar results were obtained for wt BoNT/A, albeit the testing was performed with picomolar concentrations of the toxin. Finally, the absorption profile of a BoNT derivative modified through enzymatic lipidation was assessed (Figs. 14A-B). Lipidation should result in rapid binding to cell membranes, thereby reducing diffusion. The ΔLC-GFPBoNT/A derivative is readily taken up by neuronal cells due to the presence of BoNT/A heavy chain which binds to the neuron-specific SV2 receptor. In non-neuronal cells, however, which are deficient in SV2 (such as COS7 or BHK), there is no specific binding. For lipidation, the ΔLC-GFP-BoNT/A derivative was labeled enzymatically with palmitoyl-CoA. Non-neuronal COS7 cells were used to assess lipid-mediated binding in the absence of receptor-mediated binding. Incubation of COS7 cells for various times with ΔLC-GFP-BoNT/A did not produce any visible pattern of protein uptake. In comparison, when the palmitoylated derivative of ΔLC-GFPBoNT/A was added to the cells, within 5 minutes the staining of plasma membrane was evident. Longer incubation with palmitoylated ΔLC-GFPBoNT/A resulted in intracellular uptake of the protein consistent with endosome/lysosome internalization similar to other studies with lipided GFP derivatives (Antos et al., "Lipid Modification of Proteins Through Sortase-Catalyzed Transpeptidation," J. Am. Chem. Soc. 130 (48):16338-16343 (2008), which is hereby incorporated by reference in its entirety).

Prophetic Example 16

Formation of BoNT/Aad Derivatives to Which Lipid Moieties Have Been Selectively Incorporated at the S6 Cargo Site, with the Purpose of Improved Targeting of BoNT Action, and Minimizing Unintended Diffusion and Pharmacologic Action at Unintended Sites

BoNT/Aad, atoxic derivative retains the essential wild type toxin features required for native trafficking, but has been rendered atoxic through the introduction of metalloprotease-inactivating mutations in the light chain of BoNT/A. To produce a pharmacologically active BoNT/A derivative, the inactivating point mutations can be restored to their native sequence. Both such toxic and atoxic derivatives carry an S6 peptide tag upstream of the linker sequence adjacent to the N-terminus of the light chain, which allows site-specific enzymatic attachment of various molecules for potential use in therapeutic intervention.

Here, a method is described to selectively incorporate lipid moieties into recombinant BoNT/A derivatives by enzymatic coupling to a cargo attachment peptide (e.g., the S6 peptide tag), in order to restrict the diffusion of the protein adduct from the site of injection. The goal of this project is to test a novel approach for precisely localizing the pharmacological action of lipided BoNT/Aad within the targeted neuromuscular junction at the site of injection, thereby preventing effects associated with unintended dispersal and spread of the toxin beyond the site of application.

To accomplish this, the derivative will be additionally modified to contain a sequence specifically cleaved by the BACE-1 enzyme (B-secretase) specifically localized to the surface of neurons, inserted between the spacer sequence and S6 cargo attachment site. This enables the BoNT/A derivative to be specifically released from its attachment at the external surface of the plasma membrane of neurons, and thereby to improve localization of its pharmacologic action to neurons at the site of application. As an example, lipided ΔLC-GFP-BoNT/A ad, described above, can be used to study the association of the protein with cells through the lipid tail and/or through receptor-mediated endocytosis via the HC of BoNT/A ad. To study the fate of the internalized protein (light chain) after translocation from an endocytic compartment to the cytoplasmic compartment of neurons, see infra, lipided BoNT/A ad derivatives will be used, because the ΔLC-GFP-BoNT/A ad derivative will remain stuck in the endocytic compartment due to the rigidity of the GFP portion of ΔLC-GFP-BoNT/A ad.

Expression and purification of atoxic BoNT/A derivative with cleavage recognition sequence positioned between the S6 tag and the spacer sequence upstream of the N-terminus of BoNT/Aad will be performed as described supra. Removal of the 10-His purification tag and processing of BoNT/Aad single chain prepropeptide to heterodimer will be performed with TEV protease, as described supra.

Labeling of BoNT/Aad with palmitoyl-CoA will be performed with recombinant Sfp phosphopantetheinylation transferase from B. subtilis as described supra. Yield of palmitoylated BoNT/Aad will be evaluated by mass-spectral analysis. Purification of palmitoylated protein from unmodified BoNT/Aad will be performed by fractionation with Triton X-114 with modification: Instead of ion exchange chromatography, either affinity chromatography on Streptavidin agarose or hydrophobic chromatography on octyl-sepharose will be used.

The length and type of lipid moiety attached to the protein not only contributes to protein diffusibility and kinetics of absorption, but to the route of internalization and fate of internalized proteins in vitro (Antos et al., "Lipid Modification of Proteins Through Sortase-Catalyzed Transpeptidation," J. Am. Chem. Soc. 130 (48):16338-16343 (2008), which is hereby incorporated by reference in its entirety). Therefore, in addition to generating palmitoylated, aliphatic (C-16) aduct of BoNT/A ad derivative, an aduct will be created with a longer aliphatic chain (C-22) and an aduct will be created where cholesterol will be used as a lipid moiety. These aducts will be obtained by modifying purified BoNT/A ad protein (Band et al., "Recombinant Derivatives of Botulinum Neurotoxin A Engineered for Trafficking Studies and Neuroan Delivery," Protein Expr. Purif. 71(1):62-73 (2010), which is hereby incorporated by reference in its entirety) through Sfp phosphopantetheinylation transferase labelling (Zhou et al., "Genetically Encoded Short Peptide Tags for Orthogonal Protein Labelling by Sfp and AcpS Phosphopantetheinylation Transferases," ACS Chem. Biol. 2 (5):337-346 (2007), which is hereby incorporated by reference in its entirety) with lipided CoA precursors. The precursors will include commercially available (palmitoyl-CoA, C-16, Sigma-Aldrich) and synthesized (C-22 aliphatic -CoA or cholesterol-CoA, Irvine Chemistry Laboratory) compounds.

A test will be done to compare the diffusion pattern and effects of local and systemic distribution of BoNT/Aad and
palmitoylated BoNT/A ad (or other lipiddated BoNT/A ad noted above) with targeted solubility after injection into mouse hindlimb muscle.

It is expected that limited diffusion of palmitoylated BoNT/A ad from the site of injection will be observed, in comparison with unpalmitoylated derivative, as has been described for other in vitro lipiddated proteins.

Mice will be used as an animal model, because their small muscle size increases the sensitivity of immunossays. Swiss-Webster adult male CD1 mice weighing ~30 g will be used for injection. Mice will be housed in groups of six and food and water will be provided ad libitum. Mice will be maintained on a 12-h light/dark photoperiod for 4 days before the start of experiments. All work with animals will be performed by personnel trained in the safe and humane use of laboratory animals according to existing and pending animal protocols.

Animals (time-pregnant rats) are used as a source for producing adherent cultures of primary rat hippocampal neurons, embryonic spinal cord cells (BACE1-positive), and fetal myoblast cultures (BACE1-negative) as a model for the type of cells encountered after administration of the derivatives in vivo. Live mice will also be used to test toxicity, diffusion, and systemic distribution of pairs of nonlipiddated/lipiddated BoNT/A ad variants in order to identify a lead candidate for future development of the pharmaceutically active recombinant version of BoNT/A, obtained by reversion of the inactivating mutations. Primary cell cultures: Rat embryonic hippocampal neurons and spinal cord cells represent the widely in vitro used system to dissect physiological mechanism of BoNT/A trafficking and internalization. Live time-pregnant rats will be used as a source of the primary neuronal cultures. Though established immortalized cell lines are available with some neuron-like features, their exclusive use for studies of BoNT/A ad derivatives trafficking is undesirable due to the up/down regulation and mutations in multiple gene products involved in the BoNT pathway. The requirement for primary neuronal culture for these types of studies is widely recognized as the standard approach. All work with animals will be performed by personnel trained in the safe and humane use of laboratory animals according to existing and pending animal protocols.

In vitro studies, while adherent cultures of primary rat hippocampal neurons (Vicario-Abejon, “Long-term Culture of Hippocampal Neurons,” in Current Protocols in Neuroscience, Suppl. 26, 3.2.1.-3.2.13 (John Wiley & Sons, Inc., 2004), which is hereby incorporated by reference in its entirety), embryonic spinal cord cells (Pellett et al., “A Neuronal Cell-based Botulinum Neurotoxin Assay for Highly Sensitive and Specific Detection of Neutralizing Serum Antibodies,” FEBS Lett., 581 (25):4803–4808 (2007); Pellett et al. “Comparison of the Primary Rat Spinal Cord Cell (RSC) Assay and the Mouse Bioassay for Botulinum Neurotoxin Type A Potency Determination,” J. Pharmacol. Toxicol. Methods 61(3):304-310 (2010), which are hereby incorporated by reference in their entirety) (BACE1-positive), and fetal myoblast cultures (Pin et al., “Embryonic and Fetal Rat Myoblasts Express Different Phenotypes Following Differentiation in vitro,” Dev. Genet. 14 (5):356-368 (1993); Pin et al., “Embryonic and Fetal Rat Myoblasts Form Different Muscle Fiber Types in an Ectopic in vivo Environment,” Dev. Dyn. 224 (3):253-266 (2002), which are hereby incorporated by reference in their entirety) (BACE1-negative) will be used as models of the most abundant nerve and muscle tissue at the site of injection in vivo, some assays will rely on use of neuronal cell lines (such as PC 12 or Neuro 2A, BACE1-positive) or non-neuronal cell lines (such as COS7 or BHK, BACE1-negative) in suspension cultures.

The first step will be to measure protein association with the cell membrane and to assess resistance to diffusion. To measure relative ability to associate with cells, lipiddated and nonlipiddated ΔLC-GFP-BoNT/A derivatives will be added to cells suspended in serum-free medium, and the extent to which the presence of the lipid tail and differences in the lipid moiety contribute to binding of the proteins to cells will be quantified. Membrane binding kinetics will be measured by comparing fluorescence of the medium vs. fluorescence of the cells as described (Antos et al., “Lipid Modification of Proteins Through Sortase-Catalyzed Transpeptidation,” J. Am. Chem. Soc. 130 (48):16338-16343 (2008), which is hereby incorporated by reference in its entirety). It is expected that the majority of lipiddated ΔLC-GFP-BoNT/A derivatives will be rapidly and effectively bound by all cells through insertion of the lipid moiety into the plasma membrane.

Diffusibility will be measured in monolayers of cells grown on a nitrocellulose matrix. ΔLC-GFP-BoNT/A or/and BoNT/A ad derivatives will be added in milliliter inserts (Millipore) positioned in the center of the growth circle. The diffusibility of the proteins will be directly measured by visualizing and quantifying the diameter of green fluorescence and its intensity relative to the diameter of the insert (0% diffusibility) and the diameter of the plate (100% diffusibility) as described (Flaumenhaft et al., “Heparin and Heparan Sulfate Increase the Radius of Diffusion and Action of Basic Fibroblast Growth Factor,” J. Cell Biol., 111 (4):1651-1659 (1994), which is hereby incorporated by reference in its entirety). Those lipiddated BoNT derivatives that exhibit a ≥90% reduction in diffusibility compared to their nonlipiddated counterparts will be selected for further characterization.

Possible cytotoxic effects of the BoNT/A ad and lipiddated BoNT/A ad will be determined, as described in Francis et al., “Enhancement of Diphtheria Toxin Potency by Replacement of the Receptor Binding Domain with Tetanus Toxin C-fragment: a Potential Vector for Delivering Heterologous Proteins to Neurons,” J. Neurochem. 74 (6):2528-2536 (2000), which is hereby incorporated by reference in its entirety. The cytoxicity will be evaluated over a concentration range from 1 to 500 nM. Cultures will also be observed microscopically for signs of toxicity up to 96 hours post-exposure. Cultures exposed to wt BoNT/A and untreated cells will be included as controls.

Comparison of protein uptake in neuronal and non-neuronal cultures will also be conducted. An attractive feature of the lipiddated BoNT derivatives is the possibility of their efficient uptake and degradation by non-neuronal cells, which would abet removal of excess toxin. It is known that the type of lipid tail affects protein internalization. Antos et al., “Lipid Modification of Proteins Through Sortase-Catalyzed Transpeptidation,” J. Am. Chem. Soc. 130 (48):16338-16343 (2008), which is hereby incorporated by reference in its entirety. To ensure lysosomal targeting of the lipiddated BoNT/A ad in non-neuronal cells, internalization patterns of different lipiddated BoNT/A derivatives will be compared using (a) immunocytochemical staining and (b) subcellular fractionation of the treated cells, followed by Western blotting. The rate of uptake relative to nonlipiddated BoNT/A ad will be calculated. The internalization of the proteins by cells in culture will be assessed at different time points by two methods: first, by comparison of the intensity of LC ad signal on Western blot with antibodies against LC, and second, by LC ad association with specific cellular markers, such as SNAP 25 (cytoplasmic target for LC ad binding), Rab5/EFA1...

Additional evidence of LC-ad presence in the neuronal cytoplasm will be performed by co-immunoprecipitation from cultures with anti-SNAP 25 antibodies, as described in Example 15 (FIGS. 13A-C). As many as 3 pairs of nonlabeled/lipidated BoNT constructs will be tested. Those that exhibit, relative to BoNT/A ad, a 90% reduction in diffusibility and a 10-fold increased rate of absorption by cells in culture, and which are targeted to a lysosomal degradation pathway in non-neuronal cells but to physiological LC-ad translocation to the cytoplasm in neuronal cells, will be tested further.

The experimental outcome of the studies related to BoNT/A diffusion from the site of injection is not expected to be in agreement with common perceptions based on the principles that larger proteins diffuse more slowly through an identical aqueous medium compared with smaller proteins. According to this principle, it would be predicted that BoNT’s of greater size or molecular weight will be less likely to diffuse outside the target tissue compared with those of smaller size. Thus, in this view, BOTOX® (unpurified form of 900 kDa complexes), should be less likely to diffuse outside the target tissue compared with DYSPORT® (a heterogeneous mixture of 500-900 kDa complex sizes) or XEOMIN® (pure 150 kDa toxin). In multiple studies it was found that BoNT/A injected intramuscularly exhibited some diffusion to muscles adjacent to the site of injection, and when the same amount of active neurotoxin was used, regardless of manufacturer or average MW, the observed effect was indistinguishable. However, the volume of injection and the total amount of injected protein are factors affecting both local and systemic spread of the toxin. Therefore, for the planned experiments, the same amount of either palmitoylated or non-palmitoylated BoNT/Aad will be used for injection.

The BoNT/Aad and palmitoylated BoNT/Aad protein concentration will be determined by the BCA-protein assay kit (Pierce), equalized and reconstituted with 0.9% sodium chloride to a final concentration 0.4 mg/ml. Before injection, mice will be anesthetized. The injection volume will be 25 μl. The protein will be injected in the tibialis anterior muscle in one hindlimb while carrier alone (control) will be injected in the contralateral muscle.

The extent of diffusion of BoNT/Aad versus palmitoylated BoNT/Aad will be evaluated by examining the direct pattern of protein immunostaining and by the effect of BoNT on muscles located at different distances from the site of injection. In particular, sections of the soleus muscle, which is close to the injected tibialis anterior muscle, the gastrocnemius muscle, which is next to the soleus, and the even more distant quadriceps femoris muscle, located in a rostral position, will be evaluated.

Lethal doses of BoNT/A released to circulation cause death from respiratory arrest resulting from neuromuscular paralysis produced by accumulation of toxin in the phrenic nerve. Therefore, potential dispersal of the toxin through the circulation following injection into tibialis anterior muscle will be assessed by immunostaining of the diaphragm with antibodies against BoNT/A holotoxin.

According to a previous study, the most notable changes in diffusion of the toxin from the site of injection occurred within 1-48 hours after injection. Therefore, the immunostaining pattern of BoNT/Aad diffusion will be studied after injection within this time frame.

The expression of N-CAM in mouse hindlimb muscles at different times following injection will be used as a readout of the effect generated by BoNT/Aad. N-CAM can be detected with high sensitivity and spatial resolution by histochemical and Western blot analyses, N-CAM is present on the surface of embryonic myotubes, but it is lost as development proceeds. N-CAM is nearly absent from adult muscle, but muscle denervation induces the reappearance of N-CAM. Paralysis of skeletal muscle by BoNT/A is sufficient to activate N-CAM expression, however, the action of the toxin, stemming from accumulation of its light chain in the cytosol of affected neurons, causes significantly delayed reappearance of N-CAM relative to a rapidly changed BoNT/A diffusion pattern. N-CAM is usually detected within 5-30 days after injection. This same time frame will be used.

Immunostaining will be performed 1, 2, 6, 12, and hours, and one and two days after injection (for direct BoNT/A immunodetection) and 2, 7, 14, 30, and 60 days after injection (for N-CAM immunostaining). Tibialis anterior, soleus, gastrocnemius, and quadriceps femoris muscles will be dissected from both hindlimbs of the mice and frozen in isopentane precooled in liquid nitrogen. Preparation of mouse diaphragm will be performed in situ. Muscle cryosections (10 μm) will be probed with polyclonal antibodies against BoNT/A holotoxin (Staten Serum Institut, Denmark) diluted 1:1000, or polyclonal antibodies to N-CAM (Chemicon) diluted 1:500 in blocking solution. Goat anti-rabbit IgG conjugated to Alexa-555 (Invitrogen) diluted 1:200 in blocking solution will be used for detection. Alexa-488 alpha-bungarotoxin will be used for counterstaining of NMJ. Images will be collected using a confocal microscope.

The presence of BoNT/Aad and the expression level of N-CAM in different muscles after injection with either form of BoNT/Aad will also be assessed quantitatively by Western blot analysis. The resulting data will be normalized using tubulin as a loading standard and will be plotted as intensity units relative to the basal value of the non-injected muscle.

After dissection, tissue lysates will be prepared for analysis by Western blot. Samples will be loaded (10 μg/lane) on 4%-12% Tris-HCl polyacrylamide gels (Bio-Rad) and subjected to electrophoresis. The proteins will be transferred to nitrocellulose membranes (Bio-Rad), and probed either with polyclonal antibodies against BoNT/A holotoxin (Staten Serum Institut, Denmark) diluted 1:5000 or polyclonal antibodies to N-CAM (Chemicon) diluted 1:300 in TBS/T supplemented with 5% nonfat milk. Membranes will also be probed with anti-tubulin monoclonal antibody (1:1,000, Sigma Aldrich), which will serve as an internal standard for protein quantification. Densitometric analysis will be performed using Typhoon image analysis software (GE Healthcare).

The immunostaining panels and bars on quantification plots will be representative of several sections/protein extract preparations (n=4-6) made from each of the different injected animals (n=4). Descriptive statistics of means and standard deviations (±SD) will be calculated. The t-test will be used post-hoc to determine significance of differences.

Another preferred way to analyze the diffusion of pairs of nonlipidated/lipidated BoNT/Aad variants is to examine the presence of the protein by ELISA. The double sandwich
ELISA for wt BoNT/A is more sensitive that direct immuno- 
staining and is preferable. The local diffusion of BoNT/A ad 
will be evaluated by assessing presence of BoNT by 
immunoreactivity in three distal muscle groups: the soleus muscle, 
which is close to the injected tibialis anterior muscle, the 
gastrocnemius muscle, which is distal to the soleus, and the 
more distal quadriceps femoris muscle, located in a rostral 
position. The percent reduction in diffusibility will be calcu-
lated both by distance (muscle group) and by quantity 
(ELISA signal).

For evaluation of systemic spread of BoNT, the immunore-
activity will be tested by ELISA in phrenic nerve preparations 
(the phrenic nerve is known to be a target of wt BoNT/A in 
circulation), the tibialis anterior muscle of the other (nonin-
jected) limb, and in serum. Samples for ELISA will be pre-
pared as described in Whelchel et al., “Molecular Targets of 
Botulinum Toxin at the Mammalian Neuromuscular Junc-
tion,” Mov. Disord. 19, Suppl 8: S7-S16 (2004), which is 
hereby incorporated in its entirety. The most notable changes 
in diffusion of wt toxin from the site of injection occur within 
1-48 hours after injection (Tang-Liu et al., “Intramuscular 
Injection of 125I-botulinum Neurotoxin-complex Versus 125I-
Botulinum-free Neurotoxin: Time Course of Tissue Distribu-
tion,” Toxicon 42 (5): 461-469 (2003), which is hereby incor-
porated by reference in its entirety). Therefore, tissue will 
be collected for analysis at six time points 1-48 hours after 
injection. The results will be normalized against tissue from 
animals injected with vehicle only.

Percent reduction in systemic exposure will be calculated 
from ELISA data as rate of appearance and quantity in 
phrenic nerve and serum. 5 mice are expected to be used per 
time point, as was used previously (Carl et al., “Assay of 
Diffusion of Different Botulinum Neurotoxin Type A Formu-
lations Injected in the Mouse Leg,” Muscle Nerve 40 (3): 
374-380 (2009), which is hereby incorporated by reference in 
it entirety.

With the control group, the projected number of animals 
for the pair of lipidated/nonlipidated BoNT/A ad derivatives 
will be 120. Descriptive statistics of means and standard 
deviations (±SD) will be calculated. The t-test will be used 
post-hoc to determine significance of differences. Limited 
diffusion of lipidated BoNT/A is expected from the site of 
injection in comparison with nonlipidated derivatives, as has 
been described for other in vitro lipidated proteins (Antos et 
al., “Lipid Modification of Proteins Through Sortase-Cata-
lyzed Transpeptidation,” J. Am. Chem. Soc. 130 (48):16358-
16343 (2008); Grogan et al., “Synthesis of Lipidated Green 
Fluorescent Protein and its Incorporation in Supporting Lipid 
Bilayers,” J. Am. Chem. Soc. 127 (41), 14383-14387 (2005), 
which are hereby incorporated by reference in their entirety).

This experiment will investigate a novel approach for pre-
cisely localizing the pharmacological action of botulinum 
neurotoxin A (BoNT/A) (e.g., BOTOX®, DYSPORIN®) within 
the targeted neuromuscular junction at the site of injection, 
thereby preventing effects associated with unintended dis-
seal and spread of the toxin beyond the site of entry. 
BoNT/A therapy has a good safety record, which depends 
partly on the toxin’s ability to remain relatively localized at 
the site of injection. The ability of the toxin to spread from 
the site of the injection to distant sites is a consequence of the fact 
that the toxin is soluble in aqueous solution and can be trans-
ported from the site of injection through the circulation or 
diffused locally to nearby tissue. Safety concerns are likely to 
become more important as increasing BoNT/A doses are used 
to treat conditions such as cerebral palsy or spasticity. In this 
regard, symptoms of generalized weakness have been described 
in BoNT-treated patients. These concerns are 
clearly valid, since therapeutic use of BoNT resulted in 28 
deaths between 1989 and 2003, while the number of therapeu-
tic applications of the toxin is rapidly increasing.

The low toxicity of BoNT/Aad makes it an ideal candidate 
for trafficking studies, because the amount of protein used 
in vivo can provide reliable and direct immunodetection, a goal 
which has never been accomplished using wt BoNT/A 
because of its extremely high toxicity. In addition, the ability 
to selectively incorporate lipidated moieties into recombinant 
BoNT/Aad by enzymatic coupling can restrict the diffusion 
of the protein adduct from the site of injection. Both of these 
properties—low toxicity and the ability to enzymatically 
modify the recombinant protein will be used to advantage.

It is anticipated that palmitoylated BoNT/Aad will be tar-
geted for insertion into the plasma membrane of cells proxim-
al to the injection site. Thus, palmitoylated BoNT/Aad will 
be excluded from the circulation immediately after injection, 
and its capacity for systemic dispersal will be minimal. An 
additional important modification of the expressed BoNT/
Aad will be a 13 amino acid sequence, SEISYJVEEFLRWKK 
(SEQ ID NO:43), which will be positioned between the cargo 
attachment peptide sequence and the spacer sequence 
upstream of the N-terminus of the recombinant protein. This 
peptide provides the specific substrate sequence for BACE1, 
an aspartic acid membrane-bound protease involved in the 
pathology of Alzheimer’s disease, and which is predomin-
antly expressed in cells of neural origin. Among non-neu-
ronal cells at the site of injection, the internalization of palmi-
toylated BoNT/Aad should result in non-specific uptake of 
the recombinant palmitoylated protein. In neurons, in con-
trast, because of the presence of BACE1 on the cell surface, 
recombinant BoNT/Aad should be cleaved, releasing soluble 
dichain BoNT/Aad in the proximity of synaptic contact. 
Released recombinant BoNT/Aad should enter the cell via 
the normal, heavy chain-mediated, double receptor mech-
anism of internalization followed by translocation of the light 
chain into the neuronal cytosol.

Prior to incorporation of the canonical BACE1 recognition 
sequence into existing BoNT/A ad constructs (as mentioned 
above), a cell-based FRET assay will be used along with a 
variety of commercially available FRET-based BACE1 sub-
strates (Sigma, Invitrogen, Siena Biotech, etc.) (Gruninger-
Leitch et al., “Substrate and Inhibitor Profile of BACE (beta-
secretase) and Comparison with Other Mammalian Aspartic 
is hereby incorporated by reference in its entirety) to optimize 
the BACE1 recognition sequence for the greatest cleavage 
efficiency in neuronal cells and minimal/absent cleavage in 
non-neuronal cells. The amino acid sequence of the substrate 
that best fits these criteria will be incorporated in BoNT/A ad 
as an alternative to the canonical sequence. To estimate the 
efficiency of BACE1-mediated cleavage after internalization 
of lipidated BoNT/A ad adducts will also be used in which 
the lipid moiety is C13-labeled. Three BoNT/A ad constructs 
will be used to compare the relative yield of LC ad internal-
ized through BoNT/A-receptor mediated endocytosis, which 
should result in LC ad translocation into the cytoplasm: non-
lipidated BoNT/A ad, lipidated BoNT/A ad, and lipidated 
BACE1-BoNT/A ad. Cytosolic fractions from neuronal cul-
tures exposed to all three types of proteins will be prepared by 
digitonin solubilization, as described in Example 15 (FIGS. 
13A-C). LC ad from these fractions will be co-immunopreci-
pitation with anti-SNAP 25 antibodies, separated on SDS 
PAGE and the resulting Western blot will be probed with 
anti-LC MAB F1-40 as described (FIGS. 13A-C). It is 
expected that there will be comparable signals from LC ad 
from the cultures treated with BoNT/A ad and lipidated
BACE1-BoNT/A ad, but minimal or no signal from lipidated BoNT/A ad, which should remain anchored to the membrane. It is also expected that LC ad from the fraction obtained after treatment with lipidated BACE1-BoNT/A ad will lack the radioactive lipid moiety as a consequence of BACE1-mediated cleavage. This project is intended to complete preliminary steps in the pre-clinical development of a non-diffusible botulinum neurotoxin A formulation with targeted solubility, which is intended as a therapeutic for neurological disorders.

Although the invention has been described in detail for the purposes of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 4

<210> SEQ ID NO 1
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Botulinum Neurotoxin Light Chain Motif
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (3) .. (6)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (6) .. (7)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid

<400> SEQUENCE: 1
His Glu Xaa Xaa His Xaa Xaa His
  1  5

<210> SEQ ID NO 2
<211> LENGTH: 12
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: S6 Peptide Sequence
<400> SEQUENCE: 2
Gly Asp Ser Ser Trp Leu Leu Arg Leu Leu Aasn
  1  5  10

<210> SEQ ID NO 3
<211> LENGTH: 1372
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aadex Neurotoxin
<400> SEQUENCE: 3
Met Lys Phe Leu Val Aasn Val Ala Leu Val Phe Met Val Val Tyr Ile
  1  5  10  15
Ser Tyr Ile Tyr Ala Ala His His His His His His His
  20  25  30
His Thr Arg Glu Aasn Leu Tyr Phe Glu Ala Gly Asp Ser Leu Ser
  35  40  45
Trp Leu Leu Arg Leu Leu Aasn Ala Arg Gly Gly Ala Ser Gly Pro Phe
  50  55  60
Val Aasn Lys Glu Phe Aasn Tyr Lys Asp Pro Val Aasn Gly Val Asp Ile
  65  70  75  80
 Ala Tyr Ile Lys Ile Pro Aasn Ala Gly Glu Met Glu Pro Val Lys Ala
  85  90  95
Phe Lys Ile His Aasn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe
 100 105
Thr Aasn Pro Glu Glu Gly Asp Leu Aasn Pro Pro Pro Glu Ala Lys Glu
Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr Thr Asp Asn Glu
130 135 140
Lys Asp Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr
145 150 155 160
Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile
165 170 175
Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp
180 185 190
Thr Asn Cys Ile Asn Val Ile Glu Pro Asp Gly Ser Tyr Arg Ser Glu
195 200 205
Glu Leu Asn Leu Val Ile Gly Pro Ser Ala Asp Ile Ile Glu Phe
210 215 220
Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr Arg Asn Gly
225 230 235 240
Tyr Gly Ser Thr Gin Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly
245 250 255
Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly
260 265 270
Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Ala Leu Ile His
275 280 285
Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe
290 295 300
Lys Val Asn Thr Asn Ala Tyr Gly Met Ser Gly Leu Glu Val Ser
305 310 315 320
Phe Glu Glu Leu Arg Thr Phe Gly Gin His Asp Ala Lys Phe Ile Asp
325 330 335
Ser Leu Gin Glu Asn Glu Phe Arg Leu Tyr Tyr Asn Lys Phe Lys
340 345 350
Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gin Thr Thr
355 360 365
Ala Ser Leu Gin Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu
370 375 380
Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu Lys Phe Asp
395 395 400
Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val
405 410 415
Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Ala Leu Asn Phe Asp Lys
420 425 430
Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr
435 440 445
Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn Phe Asn Gly
450 455 460
Gln Asn Thr Glu Ile Asn Met Asn Phe Thr Lys Leu Lys Asn Phe
465 470 475 480
Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile
485 490 495
Thr Ser His Thr Gin Ser Leu Asp Gin Gly Tyr Asn Asp Asp Asp
500 505 510
Lys Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Thr Trp Asp Leu Phe
515 520 525
Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu
530 535 540
Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser 545 550 555 560
Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu 565 570 575
Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Ser Asp Ile Ile Gly Gln 580 585 590
Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr 595 600 605 610
Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Glu Glu Phe 615 620
Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala 625 630 635 640
Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val 645 650 655
Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val 660 665 670 675
Glu Gln Leu Val Tyr Asp Phe Thr Glu Thr Ser Glu Val Ser Thr 680 685
Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro 690 695 700
Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala 705 710 715 720
Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile 725 730 735
Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn 740 745 750
Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn 755 760 765
Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala 770 775 780
Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala 785 790 795 800
Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Asn Tyr Glu Tyr 805 810 815
Asn Glu Tyr Thr Glu Glu Gly Lys Asn Ile Asn Phe Asn Ile Asp 820 825 830 835
Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn 840 845
Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser 850 855 860
Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu 865 870 875 880
Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile 885 890 895
Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr 890 900 905 910
Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu 915 920 925
Ser Thr Phe Thr Glu Tyr Ile Asn Asn Ile Ile Asn Thr Ser Ile Leu 925 930 935 940
Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu Ser Arg Tyr Ala 945 950 955 960
Ser Lys Ile Asn Ile Gly Ser Lys Val Asn Phe Asp Pro Ile Asp Lys
965 970 975
Asn Glu Ile Glu Leu Phe Asn Leu Glu Ser Ser Lys Ile Glu Val Ile
980 985 990
Leu Lys Asn Ala Ile Val Tyr Asn Ser Met Tyr Glu Asn Phe Ser Thr
995 1000 1005
Ser Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn Ser Ile Ser Leu
1010 1015 1020
Asn Asn Glu Tyr Thr Ile Ile Asn Cys Met Glu Asn Asn Ser Gly
1025 1030 1035
Trp Lys Val Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr Leu Gln
1040 1045 1050
Asp Thr Glu Glu Ile Lys Glu Arg Val Val Phe Lys Tyr Ser Gln
1055 1060 1065
Asp Thr Glu Glu Ile Lys Glu Arg Val Val Phe Lys Tyr Ser Gln
1070 1075 1080
Ile Thr Asn Arg Leu Asn Asn Ser Lys Ile Tyr Ile Asn Gly
1085 1090 1095
Arg Leu Ile Asp Glu Lys Pro Ile Ser Asn Leu Gly Asn Ile His
1100 1105 1110
Asn Ser Asn Ile Met Phe Lys Leu Asp Gly Cys Arg Asp Thr
1115 1120 1125
His Arg Tyr Ile Trp Ile Lys Tyr Phe Asn Leu Phe Asp Lys Glu
1130 1135 1140
Leu Asn Glu Lys Glu Ile Lys Asp Leu Tyr Asp Asn Gln Ser Asn
1145 1150 1155
Ser Gly Ile Leu Lys Asp Phe Trp Gly Asp Tyr Leu Gln Tyr Asp
1160 1165 1170
Lys Pro Tyr Tyr Met Leu Asn Leu Tyr Asp Pro Asn Lys Tyr Val
1175 1180 1185
Asp Val Asn Asn Val Gly Ile Arg Gly Tyr Met Tyr Leu Lys Gly
1190 1195 1200
Pro Arg Gly Ser Val Met Thr Thr Asn Ile Tyr Leu Asn Ser Ser
1205 1210 1215
Leu Tyr Arg Gly Thr Lys Phe Ile Ile Lys Lys Tyr Ala Ser Gly
1220 1225 1230
Asn Lys Asp Asn Ile Val Arg Asn Asp Asp Arg Val Tyr Ile Asn
1235 1240 1245
Val Val Val Lys Asn Lys Gly Tyr Arg Leu Ala Thr Asn Ala Ser
1250 1255 1260
Gln Ala Gly Val Glu Lys Ile Leu Ser Ala Leu Glu Ile Pro Asp
1265 1270 1275
Val Gly Asn Leu Ser Gln Val Val Val Met Lys Ser Lys Asn Asp
1280 1285 1290
Gln Gly Ile Thr Asn Lys Cys Lys Met Asn Leu Gln Asp Asn Asn
1295 1300 1305
Gly Asn Asp Ile Gly Phe Ile Gly Phe His Gln Phe Asn Asn Ile
1310 1315 1320
Ala Lys Leu Val Ala Ser Asn Trp Tyr Asn Arg Gln Ile Glu Arg
1325 1330 1335
Ser Ser Arg Thr Leu Gly Cys Ser Thr Glu Phe Ile Pro Val Asp
1340 1345 1350
Asp Gly Trp Gly Glu Arg Pro Leu Gly Ala Gly Trp Ser His Pro
Gln Phe Glu Lys

<210> SEQ ID NO 4
<211> LENGTH: 1370
<212> TYPE: PRO
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aadetv Neurotoxin

<400> SEQUENCE: 4

Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
1  5     10    15
Ser Tyr Ile Tyr Ala Ala Ala His His His His His His
20   25   30
His Thr Arg Glu Asn Leu Tyr Phe Glu Gly Ala Gly Asp Ser Leu Ser
35   40   45
Trp Leu Leu Arg Leu Asn Ala Arg Gly Gly Ala Ser Gly Pro Phe
50   55   60
Val Asn Lys Glu Phe Asn Tyr Lys Asp Pro Val Asn Gly Val Asp Ile
65   70   75   80
Ala Tyr Ile Lys Ile Pro Asn Ala Gly Glu Met Glu Pro Val Lys Ala
85   90   95
Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe
100  105  110
Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu Ala Lys Glu
115  120  125
Val Pro Val Ser Tyr Tyr Asp Ser Ser Tyr Leu Ser Thr Asp Asn Glu
130  135  140
Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr
145  150  155  160
Ser Thr Asp Leu Gly Arg Met Leu Thr Ser Ile Val Arg Gly Ile
165  170  175
Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Lys Val Ile Asp
180  185  190
Thr Asn Cys Ile Asn Val Ile Glu Glu Gly Ser Tyr Arg Ser Glu
195  200  205
Glu Leu Asn Leu Val Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe
210  215  220
Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Thr Arg Asn Gly
225  230  235  240
Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly
245  250  255
Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly
260  265  270
Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Ala Leu Ile His
275  280  285
Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe
290  295  300
Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser
305  310  315  320
Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp
325  330  335
Ser Leu Glu Asn Glu Phe Arg Leu Tyr Tyr Asn Lys Phe Lys
Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr
340 345 350
 Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu
355 360 365
 Ser Glu Asp Thr Ser Gly Lys Ser Val Asp Lys Leu Lys Phe Asp
370 375 380
 Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp Asn Phe Val
385 390 395 400
 Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Ala Leu Asn Phe Asp Lys
405 410 415
 Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr Thr Ile Tyr
420 425 430
 Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Asn Phe Asn Gly
435 440 445
 Gln Asn Thr Glu Ile Asn Met Asn Phe Thr Lys Leu Lys Asn Phe
450 455 460
 Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg Gly Ile Ile
465 470 475 480
 Thr Ser His Thr Glu Ser Leu Asp Gln Gly Gly Gly Asn Leu Tyr Phe
485 490 495
 Gln Gly Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu
500 505 510
 Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly
515 520 525
 Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile
530 535 540
 Ser Leu Asp Leu Ile Gln Gln Tyr Thr Leu Thr Phe Asn Phe Asp Asn
545 550 555 560
 Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly
565 570 575
 Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys
580 585 590
 Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Glu Glu
595 600 605
 Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu
610 615 620
 Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr
625 630 635 640
 Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp
645 650 655
 Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser
660 665 670
 Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly
675 680 685
 Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly
690 695 700
 Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu
705 710 715 720
 Ala Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala
725 730 735
 Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala
740 745 750
 Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg
755 760 765
| Aam Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile ValThr Asn Trp Leu | 770 | 775 | 780 |
| Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Met Lys Glu | 785 | 790 | 795 | 800 |
| Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Asn Tyr Gln | 805 | 810 | 815 |
| Tyr Asn Gln Tyr Thr Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile | 820 | 825 | 830 |
| Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile | 835 | 840 | 845 |
| Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn | 850 | 855 | 860 |
| Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser | 865 | 870 | 875 | 880 |
| Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu | 885 | 890 | 895 |
| Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser | 900 | 905 | 910 |
| Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu | 915 | 920 | 925 |
| Leu Ser Thr Phe Thr Glu Tyr Ile Asn Asn Ile Ile Asn Thr Ser Ile | 930 | 935 | 940 |
| Leu Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu Ser Arg Tyr | 945 | 950 | 955 | 960 |
| Ala Ser Lys Ile Asn Ile Gly Ser Lys Val Asn Phe Asp Pro Ile Asp | 965 | 970 | 975 |
| Lys Asn Gln Ile Glu Leu Phe Asn Leu Glu Ser Ser Lys Ile Glu Val | 980 | 985 | 990 |
| Ile Leu Lys Asn Ala Ile Val Tyr Asn Ser Met Tyr Glu Asn Phe Ser | 995 | 1000 | 1005 |
| Thr Ser Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn Ser Ile Ser | 1010 | 1015 | 1020 |
| Leu Asn Aam Glu Tyr Thr Ile Ile Asn Cys Met Glu Aam Asn Ser | 1025 | 1030 | 1035 |
| Gly Trp Lys Val Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr Leu | 1040 | 1045 | 1050 |
| Gln Asp Thr Gln Glu Ile Lys Glu Arg Val Val Phe Lys Tyr Ser | 1055 | 1060 | 1065 |
| Gln Met Ile Asn Ile Ser Asp Tyr Ile Asn Arg Trp Ile Phe Val | 1070 | 1075 | 1080 |
| Thr Ile Thr Asn Asn Arg Leu Asn Asn Ser Lys Ile Tyr Ile Asn | 1085 | 1090 | 1095 |
| Gly Arg Leu Ile Asp Gln Lys Pro Ile Ser Asn Leu Gly Asn Ile | 1100 | 1105 | 1110 |
| His Ala Ser Asn Asn Ile Met Phe Lys Leu Asp Gly Cys Arg Asp | 1115 | 1120 | 1125 |
| Thr His Arg Tyr Ile Trp Ile Lys Tyr Phe Asn Leu Phe Asp Lys | 1130 | 1135 | 1140 |
| Glu Leu Asn Glu Lys Glu Ile Lys Asp Leu Tyr Asp Asn Gln Ser | 1145 | 1150 | 1155 |
| Asn Ser Gly Ile Leu Lys Asp Phe Trp Gly Asp Tyr Leu Glu Tyr | 1160 | 1165 | 1170 |
Asp Lys Pro Tyr Tyr Met Leu Asn Leu Tyr Asp Pro Asn Lys Tyr 1176 1180 1185
Val Asp Val Asn Asn Val Gly Ile Arg Gly Tyr Met Tyr Leu Lys 1190 1195 1200
Gly Pro Arg Gly Ser Val Met Thr Thr Asn Ile Tyr Leu Asn Ser 1205 1210 1215
Ser Leu Tyr Arg Gly Thr Lys Phe Ile Ile Lys Lys Tyr Ala Ser 1220 1225 1230
Gly Asn Lys Asp Asn Ile Val Arg Asn Asn Asp Arg Val Tyr Ile 1235 1240 1245
Asn Val Val Val Lys Asn Lys Glu Tyr Arg Leu Ala Thr Asn Ala 1250 1255 1260
Ser Glu Ala Gly Val Glu Lys Ile Leu Ser Ala Leu Glu Ile Pro 1265 1270 1275
Asp Val Gly Asn Leu Ser Gln Val Val Val Met Lys Ser Lys Asn 1280 1285 1290
Asp Gln Gly Ile Thr Asn Lys Cys Lys Met Asn Leu Gln Asp Asn 1295 1300 1305
Asn Gly Asn Asp Ile Gly Phe Ile Gly Phe His Gln Phe Asn Asn 1310 1315 1320
Ile Ala Lys Leu Val Ala Ser Asn Trp Tyr Asn Arg Gln Ile Glu 1325 1330 1335
Arg Ser Ser Arg Thr Leu Gly Cys Ser Trp Glu Phe Ile Pro Val 1340 1345 1350
Asp Asp Gly Trp Gly Glu Arg Pro Leu Gly Ala Gly Trp Ser His 1355 1360 1365
Pro Gln Phe Glu Lys 1370

<210> SEQ ID NO 5
<211> LENGTH: 993
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Delta LC-Peptide-BoNT/Aev Neurotoxin
<400> SEQUENCE: 5
Met Lys Phe Leu Val Asn Val Val Leu Val Phe Met Val Val Tyr Ile 1 5 10 15
Ser Tyr Ile Tyr Ala Ala Ala His His His His His His 20 25 30
His Thr Arg Glu Asn Leu Tyr Phe Gln Gly Ala Gly Asp Ser Leu Ser 35 40 45
Trp Leu Leu Arg Leu Leu Asn Ala Arg Gly Ala Gly Asp Ser Gly Thr 50 55 60
Asn Gly Asn Gly Asn Gly Leu Arg Asn Thr Aan Leu Ala Ala 65 70 75 80
Asn Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys 85 90 95
Leu Lys Asn Phe Thr Gly Leu Phe Gly Phe Tyr Lys Leu Leu Cys Val 100 105 110
Arg Gly Ile Ile Thr Ser His Thr Gln Ser Leu Asp Gln Gly Gly Glu 115 120 125
Ile Thr Thr Ala Leu Asn Asp Leu Cys Ile Lys Val Asn 130 135 140
<table>
<thead>
<tr>
<th>125</th>
<th>126</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>-continued</strong></td>
<td></td>
</tr>
<tr>
<td>Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp</td>
<td></td>
</tr>
<tr>
<td>Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala</td>
<td></td>
</tr>
<tr>
<td>Glu Glu Asn Ile Ser Leu Asp Leu Ile Gin Gin Tyr Tyr Leu Thr Phe</td>
<td></td>
</tr>
<tr>
<td>Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser</td>
<td></td>
</tr>
<tr>
<td>Asp Ile Ile Gly Gin Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro</td>
<td></td>
</tr>
<tr>
<td>Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu</td>
<td></td>
</tr>
<tr>
<td>Arg Ala Gln Glu Phe Glu His Gly Lys Tyr Ser Arg Ile Ala Leu Thr Asn</td>
<td></td>
</tr>
<tr>
<td>Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe</td>
<td></td>
</tr>
<tr>
<td>Ser Ser Asp Tyr Val Lys Val Asn Lys Ala Thr Glu Ala Ala Met</td>
<td></td>
</tr>
<tr>
<td>Phe Leu Gly Trp Val Glu Gin Leu Val Tyr Asp Phe Thr Asp Glu Thr</td>
<td></td>
</tr>
<tr>
<td>Ser Glu Val Ser Thr Thr Asp Thr Ile Ala Asp Ile Thr Ile Ile Ile</td>
<td></td>
</tr>
<tr>
<td>Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gin Gin Met Leu Tyr Lys Asp</td>
<td></td>
</tr>
<tr>
<td>Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu</td>
<td></td>
</tr>
<tr>
<td>Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val</td>
<td></td>
</tr>
<tr>
<td>Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gin Thr Ile Asp Asn Ala</td>
<td></td>
</tr>
<tr>
<td>Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val</td>
<td></td>
</tr>
<tr>
<td>Thr Asn Trp Leu Ala Lys Val Asn Thr Gin Ile Asp Leu Ile Arg Lys</td>
<td></td>
</tr>
<tr>
<td>Lys Met Lys Glu Ala Leu Glu Asn Gin Ala Glu Ala Thr Lys Ala Ile</td>
<td></td>
</tr>
<tr>
<td>Ile Asn Tyr Gin Tyr Asn Gin Tyr Thr Glu Glu Lys Asn Asn Ile</td>
<td></td>
</tr>
<tr>
<td>Aam Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Aam</td>
<td></td>
</tr>
<tr>
<td>Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gin Cys Ser Val Ser</td>
<td></td>
</tr>
<tr>
<td>Tyr Leu Met Asn Ser Met Ile Pro Tyr Gin Val Lys Arg Leu Glu Asp</td>
<td></td>
</tr>
<tr>
<td>Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn</td>
<td></td>
</tr>
<tr>
<td>Arg Gly Thr Leu Ile Gly Gin Val Asp Arg Leu Lys Asp Lys Val Aam</td>
<td></td>
</tr>
<tr>
<td>Aam Thr Leu Ser Thr Asp Ile Pro Phe Gin Leu Ser Lys Tyr Val Asp</td>
<td></td>
</tr>
<tr>
<td>Aam Gin Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Aam Asn Ile Ile</td>
<td></td>
</tr>
<tr>
<td>Aam Thr Ser Ile Leu Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp</td>
<td></td>
</tr>
</tbody>
</table>
Leu Ser Arg Tyr Ala Ser Lys Ile Asn Ile Gly Ser Lys Val Asn Phe 580 585 590
Amp Pro Ile Amp Lys Asn Gln Ile Gln Leu Phe Asn Leu Glu Ser Ser 595 600 605
Lys Ile Glu Val Ile Leu Lys Asn Ala Ile Val Tyr Asn Ser Met Tyr 610 615 620
Glu Asn Phe Ser Thr Ser Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn 625 630 635 640
Ser Ile Ser Leu Asn Asn Glu Tyr Thr Ile Ile Asn Cys Met Glu Asn 645 650 655
Asn Ser Gly Trp Lys Val Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr 660 665 670
Leu Gln Asp Thr Gln Glu Ile Lys Gln Arg Val Val Phe Lys Tyr Ser 675 680 685
Gln Met Ile Ann Ile Ser Asp Tyr Ile Ann Arg Trp Ile Phe Val Thr 690 695 700
Ile Thr Asn Ann Arg Leu Asn Ann Ser Lys Ile Tyr Ile Ann Gly Arg 705 710 715 720
Leu Ile Asp Gln Lys Pro Ile Ser Ann Leu Gly Ann Ile His Ala Ser 725 730 735
Asn Ann Ile Met Phe Lys Leu Asp Gly Cys Arg Asp Thr His Arg Tyr 740 745 750
Ile Trp Ile Lys Tyr Phe Ann Leu Phe Asp Lys Glu Leu Ann Glu Lys 755 760 765
Glu Ile Lys Asp Leu Tyr Asp Ann Gln Ser Ann Ser Gly Ile Leu Lys 770 775 780
Asp Phe Trp Gly Asp Tyr Leu Gin Tyr Asp Lys Pro Tyr Tyr Met Leu 785 790 795 800
Asn Leu Tyr Asp Pro Ann Lys Tyr Val Asp Val Ann Val Gly Ile 805 810 815
Arg Gly Tyr Met Tyr Leu Lys Gly Pro Arg Gly Ser Val Met Thr Thr 820 825 830
Asn Ile Tyr Leu Ann Ser Ser Leu Tyr Arg Gly Thr Lys Phe Ile Ile 835 840 845
Lys Lys Tyr Ala Ser Gly Ann Asp Ann Ile Val Arg Ann Ann Asp 850 855 860
Arg Val Tyr Ile Ann Val Val Lys Ann Lys Glu Tyr Arg Leu Ala 865 870 875 880
Thr Ann Ala Ser Gln Ala Gly Val Glu Lys Ile Leu Ser Ala Leu Glu 885 890 895
Ile Pro Asp Val Gly Ann Leu Ser Gln Val Val Met Lys Ser Lys 900 905 910
Asn Asp Gln Gly Ile Thr Ann Lys Cys Lys Met Ann Leu Gln Asp Ann 915 920 925
Asn Gly Ann Asp Ile Gly Phe Ile Gly Phe His Gln Phe Ann Ann Ile 930 935 940
Ala Lys Leu Val Ala Ser Asn Trp Tyr Ann Arg Gln Ile Glu Arg Ser 945 950 955 960
Ser Arg Thr Leu Gly Cys Ser Trp Glu Phe Ile Pro Val Asp Gly 965 970 975
Trp Gly Glu Arg Pro Leu Gly Ala Gly Trp Ser His Pro Gin Phe Glu 980 985 990
Lys

`<210>` SEQ ID NO 6
`<211>` LENGTH: 1230
`<212>` TYPE: PRT
`<213>` ORGANISM: Artificial
`<220>` FEATURES:
`<222>` OTHER INFORMATION: Delta LC-GFP-BotNT/Atev Neurotoxin

`<400>` SEQUENCE: 6

```
1   Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
   5   10  15
20  Ser Tyr Ile Tyr Ala Ala His His His His His His His His His
   25  30
35  His Thr Arg Glu Asn Leu Tyr Phe Glu Gly Ala Gly Asp Ser Leu Ser
   40  45
50  Trp Leu Leu Arg Leu Asn Ala Arg Gly Ala Ser Val Ser Lys
   55  60
65  Gly Ala Glu Leu Phe Thr Gly Ile Val Pro Ile Leu Ile Glu Leu Asn
   70  75  80
85  Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly
   90  95
Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly
  100 105 110
115  Lys Leu Pro Val Pro Thr Pro Thr Leu Val Thr Leu Thr Ser Tyr Gly
120 125
130  Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Glu His Asp Phe
135 140
145  Phe Lys Ser Ala Met Pro Gln Gly Tyr Ile Glu Glu Arg Thr Ile Phe
150 155 160
165  Phe Glu Asp Asp Gly Asn Tyr Lys Ser Arg Ala Glu Val Lys Phe Glu
170 175
180  Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Thr Gly Thr Asp Phe Lys
185 190
195  Glu Asp Gly Asn Ile Leu Gly Asn Lys Met Glu Tyr Aen Tyr Aen Ala
200 205
210  His Asn Val Tyr Ile Met Thr Asp Asp Ala Lys Asn Gly Ile Lys Val
215 220
225  Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gin Leu Ala
230 235 240
245  Asp His Tyr Gin Gin Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu
250 255
260  Pro Asp Asn His Tyr Leu Ser Thr Gin Ser Ala Leu Ser Lys Asp Pro
265 270
275  Asn Glu Lys Arg Asp His Met Ile Tyr Phe Gly Phe Thr Val Thr Ala Ala
280 285
290  Ala Ile Thr His Gly Met Asp Glu Leu Tyr Lys Gly Thr Aen Gly Aen
295 300
305  Gly Asn Gly Asn Leu Arg Asn Thr Asn Ala Ala Aen Phe Aen
310 315 320
325  Gly Gin Aen Thr Glu Ile Aen Aen Met Aen Phe Thr Lys Leu Lys Aen
330 335
340  Phe Thr Gly Leu Phe Glu Phe Tyr Gly Glu Leu Leu Cys Val Arg Gly Ile
345 350
```
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>770</td>
<td>775</td>
<td>780</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ser</td>
<td>Thr</td>
<td>Phe</td>
</tr>
<tr>
<td>Thr</td>
<td>Thr</td>
<td>Glu</td>
<td>Tyr</td>
<td>Ile</td>
</tr>
<tr>
<td>Asn</td>
<td>Asn</td>
<td>Ile</td>
<td>Asn</td>
<td>Ile</td>
</tr>
<tr>
<td>Thr</td>
<td>Ser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>785</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>790</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>795</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Leu</td>
<td>Asn</td>
<td>Leu</td>
<td>Arg</td>
</tr>
<tr>
<td>Ser</td>
<td>Tyr</td>
<td>Glu</td>
<td>Ser</td>
<td>Asn</td>
</tr>
<tr>
<td>His</td>
<td>Leu</td>
<td>Ile</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>805</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>810</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>815</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Ala</td>
<td>Ser</td>
<td>Lys</td>
<td>Ile</td>
</tr>
<tr>
<td>Asn</td>
<td>Gly</td>
<td>Ser</td>
<td>Lys</td>
<td>Val</td>
</tr>
<tr>
<td>Val</td>
<td>Asn</td>
<td>Phe</td>
<td>Asp</td>
<td>Pro</td>
</tr>
<tr>
<td>Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>820</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>825</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>830</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Lys</td>
<td>Asn</td>
<td>Gln</td>
<td>Ile</td>
</tr>
<tr>
<td>Gln</td>
<td>Leu</td>
<td>Phe</td>
<td>Asn</td>
<td>Leu</td>
</tr>
<tr>
<td>Glu</td>
<td>Ser</td>
<td>Ser</td>
<td>Lys</td>
<td>Ile</td>
</tr>
<tr>
<td>Glu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>635</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>840</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>845</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Ile</td>
<td>Leu</td>
<td>Lys</td>
<td>Asn</td>
</tr>
<tr>
<td>Ala</td>
<td>Val</td>
<td>Tyr</td>
<td>Asn</td>
<td>Ser</td>
</tr>
<tr>
<td>Met</td>
<td>Tyr</td>
<td>Glu</td>
<td>Asn</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>950</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>855</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>860</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Thr</td>
<td>Ser</td>
<td>Phe</td>
<td>Trp</td>
</tr>
<tr>
<td>Ile</td>
<td>Arg</td>
<td>Ile</td>
<td>Pro</td>
<td>Lys</td>
</tr>
<tr>
<td>Tyr</td>
<td>Phe</td>
<td>Asn</td>
<td>Ser</td>
<td>Ile</td>
</tr>
<tr>
<td>Ser</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>865</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>870</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>875</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Asn</td>
<td>Glu</td>
<td>Tyr</td>
</tr>
<tr>
<td>Thr</td>
<td>Ile</td>
<td>Asn</td>
<td>Cys</td>
<td>Met</td>
</tr>
<tr>
<td>Glu</td>
<td>Asn</td>
<td>Asn</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>885</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>890</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>895</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>Lys</td>
<td>Val</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td>Asn</td>
<td>Tyr</td>
<td>Glu</td>
<td>Ile</td>
<td>Ile</td>
</tr>
<tr>
<td>Trp</td>
<td>Thr</td>
<td>Leu</td>
<td>Gln</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>905</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>910</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Gln</td>
<td>Glu</td>
<td>Ile</td>
<td>Lys</td>
</tr>
<tr>
<td>Gln</td>
<td>Arg</td>
<td>Val</td>
<td>Val</td>
<td>Phe</td>
</tr>
<tr>
<td>Lys</td>
<td>Tyr</td>
<td>Ser</td>
<td>Gln</td>
<td>Met</td>
</tr>
<tr>
<td>Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>915</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>920</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>925</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Ile</td>
<td>Ser</td>
<td>Asp</td>
<td>Tyr</td>
</tr>
<tr>
<td>Ile</td>
<td>Asn</td>
<td>Arg</td>
<td>Trp</td>
<td>Ile</td>
</tr>
<tr>
<td>Phe</td>
<td>Val</td>
<td>Thr</td>
<td>Ile</td>
<td>Thr</td>
</tr>
<tr>
<td>Asn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>930</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>935</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>940</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Arg</td>
<td>Leu</td>
<td>Asn</td>
<td>Ser</td>
</tr>
<tr>
<td>Lys</td>
<td>Ile</td>
<td>Tyr</td>
<td>Ile</td>
<td>Asn</td>
</tr>
<tr>
<td>Gly</td>
<td>Arg</td>
<td>Leu</td>
<td>Ile</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>945</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>950</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>955</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>960</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Lys</td>
<td>Pro</td>
<td>Ile</td>
<td>Ser</td>
</tr>
<tr>
<td>Asn</td>
<td>Leu</td>
<td>Gly</td>
<td>Asn</td>
<td>Ile</td>
</tr>
<tr>
<td>His</td>
<td>Ala</td>
<td>Ser</td>
<td>Asn</td>
<td>Asn</td>
</tr>
<tr>
<td>Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>965</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>970</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>975</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Phe</td>
<td>Lys</td>
<td>Leu</td>
<td>Asp</td>
</tr>
<tr>
<td>Gly</td>
<td>Cys</td>
<td>Arg</td>
<td>Asp</td>
<td>Thr</td>
</tr>
<tr>
<td>His</td>
<td>Arg</td>
<td>Tyr</td>
<td>Ile</td>
<td>Trp</td>
</tr>
<tr>
<td>Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>980</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>985</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>990</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Tyr</td>
<td>Phe</td>
<td>Asn</td>
<td>Leu</td>
</tr>
<tr>
<td>Leu</td>
<td>Phe</td>
<td>Asp</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Glu</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>Ile</td>
<td>Lys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>995</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Leu</td>
<td>Tyr</td>
<td>Asp</td>
<td>Asn</td>
</tr>
<tr>
<td>Gln</td>
<td>Ser</td>
<td>Asn</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>Ile</td>
<td>Leu</td>
<td>Lys</td>
<td>Asp</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1010</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>Gly</td>
<td>Asp</td>
<td>Tyr</td>
<td>Leu</td>
</tr>
<tr>
<td>Gln</td>
<td>Tyr</td>
<td>Asp</td>
<td>Lys</td>
<td>Pro</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyr</td>
<td>Met</td>
<td>Leu</td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1025</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1030</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1035</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Tyr</td>
<td>Asp</td>
<td>Pro</td>
<td>Asn</td>
</tr>
<tr>
<td>Lys</td>
<td>Tyr</td>
<td>Val</td>
<td>Asp</td>
<td>Val</td>
</tr>
<tr>
<td>Asn</td>
<td>Val</td>
<td>Gly</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1040</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1045</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Gly</td>
<td>Tyr</td>
<td>Met</td>
<td>Tyr</td>
</tr>
<tr>
<td>Leu</td>
<td>Lys</td>
<td>Gly</td>
<td>Pro</td>
<td>Arg</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Val</td>
<td>Met</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1055</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1060</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1065</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Ile</td>
<td>Tyr</td>
<td>Leu</td>
</tr>
<tr>
<td>Asn</td>
<td>Ser</td>
<td>Ser</td>
<td>Leu</td>
<td>Tyr</td>
</tr>
<tr>
<td>Arg</td>
<td>Gly</td>
<td>Thr</td>
<td>Lys</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1070</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1075</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1080</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Ile</td>
<td>Lys</td>
<td>Tyr</td>
<td>Ala</td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Asn</td>
<td>Lys</td>
<td>Asp</td>
</tr>
<tr>
<td>Asn</td>
<td>Ile</td>
<td>Val</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1085</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1090</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1095</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Asn</td>
<td>Asp</td>
<td>Arg</td>
<td>Val</td>
</tr>
<tr>
<td>Tyr</td>
<td>Ile</td>
<td>Asn</td>
<td>Val</td>
<td>Val</td>
</tr>
<tr>
<td>Lys</td>
<td>Asn</td>
<td>Lys</td>
<td>Glu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1105</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Thr</td>
</tr>
<tr>
<td>Asn</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Glu</td>
</tr>
<tr>
<td>Lys</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1115</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1120</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>Ala</td>
<td>Leu</td>
<td>Glu</td>
</tr>
<tr>
<td>Ile</td>
<td>Pro</td>
<td>Asp</td>
<td>Val</td>
<td>Gly</td>
</tr>
<tr>
<td>Asn</td>
<td>Leu</td>
<td>Ser</td>
<td>Gln</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1130</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1135</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1140</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Val</td>
<td>Met</td>
<td>Lys</td>
<td>Ser</td>
</tr>
<tr>
<td>Asn</td>
<td>Gln</td>
<td>Gly</td>
<td>Ile</td>
<td>Thr</td>
</tr>
<tr>
<td>Asn</td>
<td>Lys</td>
<td>Cys</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1145</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1150</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1155</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Met</td>
<td>Asn</td>
<td>Leu</td>
<td>Gln</td>
</tr>
<tr>
<td>Asp</td>
<td>Asn</td>
<td>Gly</td>
<td>Asp</td>
<td>Ile</td>
</tr>
<tr>
<td>Gly</td>
<td>Phe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1160</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1165</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Phe</td>
<td>His</td>
<td>Gln</td>
<td>Phe</td>
</tr>
<tr>
<td>Asn</td>
<td>Ile</td>
<td>Ala</td>
<td>Lys</td>
<td>Leu</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Ser</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1175</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1180</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1185</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Trp Tyr Asn Arg Gln Ile Glu Arg Ser Ser Arg Thr Leu Gly Cys
1190 1195 1200

Ser Trp Glu Phe Ile Pro Val Asp Asp Gly Trp Gly Glu Arg Pro
1205 1210 1215

Leu Gly Ala Gly Trp Ser His Pro Gln Phe Glu Lys
1220 1225 1230

<210> SEQ ID NO 7
<211> LENGTH: 1296
<212> TYPE: PRT
<213> ORGANISM: Clostridium Botulinum

<400> SEQUENCE: 7

Met Pro Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly
1 5 10 15

Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Glu Met Gln Pro
20 25 30

Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg
35 40 45

Asp Thr Phe Thr Asn Pro Glu Gly Asp Leu Asn Pro Pro Pro Glu
50 55 60

Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr
65 70 75 80

Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu
85 90 95

Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val
100 105 110

Arg Gly Ile Pro Phe Trp Gly Ser Thr Ile Asp Thr Glu Leu Lys
115 120 125

Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr
130 135 140

Arg Ser Glu Leu Asn Leu Val Ile Glu Gly Pro Ser Ala Asp Ile
145 150 155 160

Ile Gln Phe Glu Cys Lys Ser Phe Gly His Gln Val Leu Asn Leu Thr
165 170 175

Arg Asn Gly Tyr Gly Ser Thr Glu Tyr Ile Arg Phe Ser Pro Asp Phe
180 185 190

Thr Phe Gly Phe Glu Ser Leu Gln Thr Asp Thr Aasn Pro Leu Leu
195 200 205

Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Leu Ala His Glu
210 215 220

Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Aasn
225 230 235 240

Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu
245 250 255

Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys
260 265 270

Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Aasn
275 280 285

Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val
290 295 300

Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys
305 310 315 320

Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu
Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp 325 330 335
Aas Phe Val Lys Phe Phe Lys Val Leu Aas Arg Lys Thr Tyr Leu Aas 340 345 350
Phe Asp Lys Ala Val Phe Lys Ile Aas Ile Val Pro Lys Val Aas Tyr 355 360 365
Thr Ile Tyr Asp Gly Phe Aas Leu Arg Aas Thr Aas Leu Ala Ala Aas 370 375 380
Phe Aas Gly Glu Aas Thr Glu Aas Aas Met Aas Phe Thr Lys Leu 385 390 395 400
Lys Aas Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg 405 410 415
Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Aas Lys Gly Tyr Aas Lys 420 425 430 435
Aa Leu Aas Aas Leu Cys Ile Lys Val Aas Aas Trp Aas Leu Phe Phe 440 445 450 455
Ser Pro Ser Glu Aas Phe Thr Aas Aas Aas Leu Lys Gly Glu Glu 460 465 470 475 480
Ile Thr Ser Asp Thr Aas Ile Glu Ala Ala Glu Aas Ile Ser Leu 485 490 495
Aas Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Aas Phe Aas Aas Glu Pro 495 500 505 510
Glu Aas Ile Ser Ile Glu Aas Leu Ser Ser Aas Ile Ile Gln Gin Leu 520 525 530 535
Glu Leu Met Pro Aas Ile Glu Arg Phe Pro Aas Gin Lys Tyr Glu 540 545 550 555
Leu Aas Gln Tyr Thr Met Phe His Tyr Leu Arg Aas Gin Glu Phe Glu 560 565 570 575
His Gly Lys Ser Arg Ile Ala Leu Thr Aas Ser Val Aas Gin Ala Leu 580 585 590 595
Leu Aas Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Aas Tyr Val Lys 600 605 610 615
Lys Val Aas Al A Thr Glu Aa Aa Met Phe Leu Gln Tyr Trp Val Glu 620 625 630 635 640
Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr 640 645 650 655
Asp Lys Ile Ala Aas Ile Thr Ile Ile Pro Tyr Ile Ile Gln Gin Ala Leu 660 665 670
Leu Aas Ile Gly Gin Met Leu Tyr Lys Aas Aas Aas Phe Val Gin Ala Leu 680 685 690 695
Ile Phe Ser Gin Ala Val Ile Leu Leu Phe Gin Phe Gin Gin Ala Leu 700 705 710 715 720
Ile Pro Val Leu Gin Thr Phe Ala Leu Val Ser Tyr Ile Ala Aas Lys 720 725 730 735
Val Leu Thr Val Gin Thr Ile Aas Aas Ala Leu Ser Lys Arg Aas Glu 740 745 750 755
Lys Trp Aas Glu Val Tyr Lys Tyr Ile Thr Aas Trp Leu Ala Lys 760 765 770 775
Val Aas Gin Ile Gin Gin Leu Gin Gin Val Gin Gin Val Gin Gin Gin Ala Leu 780 785 790 795 800
Glu Aas Gin Ala Gin Ala Thr Lys Ala Ile Gin Tyr Gin Tyr Aas 810 815 820 825 830
Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp
755 760 765
Leu Ser Ser Lys Leu Asn Glu Glu Ile Asn Lys Ala Met Ile Asn Ile
770 775 780
Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met
785 790 795 800
Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys
805 810 815 820
Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly
820 825 830
Gln Val Asp Arg Leu Lys Asp Val Asn Thr Leu Ser Thr Asp
835 840 845 850
Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Gln Arg Leu Leu Ser
850 855 860
Thr Phe Thr Glu Tyr Ile Lys Asn Ile Asn Thr Ser Ile Leu Asn
865 870 875 880
Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu Ser Arg Tyr Ala Ser
885 890 895
Lys Ile Asn Ile Gly Ser Lys Val Asn Phe Asp Pro Ile Asp Lys Asn
900 905 910
Gln Ile Gln Leu Phe Asn Leu Glu Ser Ser Lys Ile Glu Val Ile Leu
910 915 920 925
Lys Asn Ala Ile Val Tyr Ser Met Tyr Glu Asp Phe Ser Thr Ser
930 935 940
Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn Ser Ile Ser Leu Asn
945 950 955 960
Glu Tyr Thr Ile Ile Asn Cys Met Glu Asn Ser Ser Gly Trp Lys Val
965 970 975
Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr Leu Gln Asp Thr Gln Glu
980 985 990
Ile Lys Gln Arg Val Val Phe Lys Tyr Ser Glu Met Ile Asn Ile Ser
995 1000 1005
Asp Tyr Ile Asn Arg Trp Ile Phe Val Thr Ile Thr Asn Asn Arg
1010 1015 1020
Leu Asn Asn Ser Lys Ile Tyr Ile Asn Gly Arg Leu Ile Asp Gln
1025 1030 1035
Lys Pro Ile Ser Asn Leu Gly Asn His Ala Ser Asn Asn Ile
1040 1045 1050
Met Phe Lys Leu Asp Gly Cys Arg Asp Thr His Arg Tyr Ile Trp
1055 1060 1065
Ile Lys Tyr Phe Asn Leu Phe Asp Lys Glu Leu Asn Glu Lys Glu
1070 1075 1080
Ile Lys Asp Leu Tyr Asp Asn Gln Ser Asn Ser Gly Ile Leu Lys
1085 1090 1095
Asp Phe Trp Gly Asp Tyr Leu Gln Tyr Asp Lys Pro Tyr Tyr Met
1100 1105 1110
Leu Asn Leu Tyr Asp Pro Asn Lys Tyr Val Asp Val Asn Asn Val
1110 1115 1120 1125
Gly Ile Arg Gly Tyr Met Tyr Leu Lys Gly Pro Arg Gly Ser Val
1130 1135 1140
Met Thr Thr Asn Ile Tyr Leu Asn Ser Ser Leu Tyr Arg Gly Thr
1145 1150 1155
<table>
<thead>
<tr>
<th>Lys</th>
<th>Phe</th>
<th>Ile</th>
<th>Ile</th>
<th>Lys</th>
<th>Tyr</th>
<th>Ala</th>
<th>Ser</th>
<th>Gly</th>
<th>Asn</th>
<th>Lys</th>
<th>Asp</th>
<th>Asn</th>
<th>Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1160</td>
<td></td>
<td>1165</td>
<td></td>
<td></td>
<td></td>
<td>1170</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Arg</td>
<td>Asn</td>
<td>Asp</td>
<td>Arg</td>
<td>Val</td>
<td>Tyr</td>
<td>Ile</td>
<td>Asn</td>
<td>Val</td>
<td>Val</td>
<td>Val</td>
<td>Lys</td>
<td>Asn</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1178</td>
<td>1180</td>
<td>1185</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Glu</td>
<td>Tyr</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Thr</td>
<td>Asn</td>
<td>Ala</td>
<td>Ser</td>
<td>Gln</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1200</td>
</tr>
<tr>
<td>Lys</td>
<td>Ile</td>
<td>Leu</td>
<td>Ser</td>
<td>Ala</td>
<td>Leu</td>
<td>Glu</td>
<td>Ile</td>
<td>Pro</td>
<td>Asp</td>
<td>Val</td>
<td>Gly</td>
<td>Asn</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1215</td>
</tr>
<tr>
<td>Gln</td>
<td>Val</td>
<td>Val</td>
<td>Val</td>
<td>Met</td>
<td>Lys</td>
<td>Ser</td>
<td>Lys</td>
<td>Asn</td>
<td>Asp</td>
<td>Gln</td>
<td>Gly</td>
<td>Ile</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1230</td>
</tr>
<tr>
<td>Lys</td>
<td>Cys</td>
<td>Lys</td>
<td>Met</td>
<td>Asn</td>
<td>Leu</td>
<td>Gln</td>
<td>Asp</td>
<td>Asn</td>
<td>Gly</td>
<td>Asn</td>
<td>Asp</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1238</td>
<td></td>
<td></td>
<td></td>
<td>1240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Ile</td>
<td>Gly</td>
<td>Phe</td>
<td>His</td>
<td>Gln</td>
<td>Phe</td>
<td>Asn</td>
<td>Asn</td>
<td>Ile</td>
<td>Ala</td>
<td>Lys</td>
<td>Val</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1260</td>
</tr>
<tr>
<td>Ser</td>
<td>Asn</td>
<td>Trp</td>
<td>Tyr</td>
<td>Asn</td>
<td>Arg</td>
<td>Gln</td>
<td>Ile</td>
<td>Glu</td>
<td>Arg</td>
<td>Ser</td>
<td>Ser</td>
<td>Arg</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Cys</td>
<td>Ser</td>
<td>Trp</td>
<td>Glu</td>
<td>Phe</td>
<td>Ile</td>
<td>Pro</td>
<td>Val</td>
<td>Asp</td>
<td>Gly</td>
<td>Trp</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Pro</td>
<td>Leu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 8
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Enterokinase cleavage site

<400> SEQUENCE: 8

Amp Asp Asp Amp Lys
1  5

<210> SEQ ID NO 9
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Spacer or linker sequence

<400> SEQUENCE: 9

Ala Arg Gly Gly Ala Ser Gly
1  5

<210> SEQ ID NO 10
<211> LENGTH: 8971
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aa4etv vector pFB1S3A TEV

<400> SEQUENCE: 10

gacggaagct gtacggaagc attaagcgcc gcggtgtgg tgtgtacgcc cagcgtgacc
60
getacacctg ccaagccgct acggcggccg cctttgagtt ttctcccccttttctctgcc
120
acgtcgccg gttttgagg tcaagcctta aatgcgggggc tcctctttagg gttccggttt
180
agttgttcac gcctggcagc ccctaaaaa attgattatgg ttggtttgtc ccttcgtttg
240
cctagcctgg gctaggctgt ttttcccgct tttgagttgtg agtcccaggg ctttaaatgt
300
gactttcctgg cccaaagctc aaccatctc ccgtccttctttgatatta
360
taagggagtct gcggattttg ttaaaaaatg agtggatttt aaaaaaaaatttt
420
aacggaatt ttacaaatatt ataaacgtttt acaatttcag gtgccactttt tgggggaat 480
gtgcgcggaat ccctatattt ttaatattttt taatactattt caaatatatgt ctggtctcag 540
agacaataac cctgataaa ggcctaaatt tattgaaata gaaagagtaat gactcatct 600
cattttcctgt ttcgcttact cccctttttt ttggtgaatt ggggcttatttt tttggctcact 660
cccagaagcc gggtggaacct aaagagttgt ggcgagcagc ctggtggctctgg aagggggtac 720
atcgaactg atcataacag ccgtaaagata ctgtagatgt tcggccgga aagagcgtttt 780
ccatgtcgtag gaaccttttg aagttcctgt ggtggcgcctg ctttcgctcag gaaaattgat 840
gggcaagac aacctcgcag cgcataacac tattctcaga atgacagctg cagatgtatca 900
ccgtaacag aaacgactct tcgagatgctc atgcagatgaa gagaacattc agctgctgcc 960
ataatttttct gtcagaacac tcggcgcacag ggacagatcg gggccctcgag 1020
agctaaagcc accttcctct ctgaacctcg cctgtgcaat ggtgataacct 1080
ccgagagtga atgaagcact acacacacgt caagcgtgca cacagctgcagc 1140
gccacatgct ctcggtatgtgttttgca ttcggtgtcag ttcggtgcttc ttcggtgctcag 1200
ctttaataagc ggtattgtgaacct ctcggtgtcag ttcggtgcttc ttcggtgctcag 1260
gctgggtgtg ttatagtgaact ccaagctgtc ttcggtgcttc ttcggtgctcag 1320
gcagagctgg gcggccggact cctggtgatgttgctgcgcatatcggggtggtcgggttt 1380
cggccactac ttcggtgacta cagcagcagact cccgctggcttctggcgc 1440
cattgtagcc ctgtaactgc acctttgtaaac ttaggtaagct ttcggtgcttc ttcggtgctcag 1500
ttttaaaaat ccaacctacca ggtggtgtcatttctcctg ttcggtgcttc ttcggtgctcag 1560
tactatgtgcc ctcggtgcttc gctggtgtcag ttcggtgcttc ttcggtgctcag 1620
ttgatgcttt ccaatggtgct ctcggtgcttc ttcggtgcttc ttcggtgctcag 1680
gcgggtttt ttttcggtgct cctttgtaa atatagctttt cccgctggcttctggcgc 1740
agccagcggc agatcagacaat cttggtgcttctg ctcggtgcttc ttcggtgctcag 1800
aagacaacttcttgtgacacgccacac ccctggctgct ctcggtgcttc ttcggtgctcag 1860
gctgggtgct cttggtgcttc gctggtgtcag ttcggtgcttc ttcggtgctcag 1920
gcggccggtg ggtggtgtcag ttcggtgcttc gctggtgtcag ttcggtgctcag 1980
taccgacgaa cggcagcagac cccgctggcttctggcgc 2040
agagagccgagg acatccgtgcc cccgctggcttctggcgc 2100
cttcgaaggg gcagcagcagac cccgctggcttctggcgc 2160
agcggctcggctc ctttggcttctg cggcagcagac cccgctggcttctggcgc 2220
ccggccttttct ctcggtgcttc gctggtgtcag ttcggtgcttc ttcggtgctcag 2280
cttccctgcttc cttggtgcttc gctggtgtcag ttcggtgcttc ttcggtgctcag 2340
ccggcagaccc cggcagcagac cccgctggcttctggcgc 2400
cggcagcagac cccgctggcttctggcgc 2460
cggcagcagac cccgctggcttctggcgc 2520
cataggctttt ctcggtgcttc ctcggtgcttc 2580
acgagcacttg aacagcaggc cccgctggcttctggcgc 2640
cggcagcagac cccgctggcttctggcgc 2700
cggcagcagac cccgctggcttctggcgc 2760
ttgaaacaag ccagctcaat cgggggtacg acagctccat tcgagtatat gaagaatgtt 5220
ttcaggggaa aatactctgt gcagagagat acctcgtgga acgtctctgt ggacaaactg 5280
aatgctgaca aacgcttcaac gactgtgacc gagaattata cgggaatataa cttctgtgaa 5340
tctctccaaag tcaacctagc gaaacactct cagaaacctgg acaagcctgt gcctcaagac 5400
aacatctcc acaaagttta ataacaatttc tatgagctgt tcaactctaga aaacaaccac 5460
tgggaccgca accctcaggg ccacaatacgc gaaattata acaagacttc caacaaactg 5520
aacatacctta cttgacgctgt caagctccat aagctctcct tgggtgcttgc cattcatcacc 5580
tcataattct acgtcttccaga cagaggttggc gagaacctgt aacttccagg tctctgcaac 5640
gatctgtgta tcaagcttgaata acetggctat cttcttcccta goccaaqcga gaatcactcc 5700
acgaaagctc tctcaacaaggg taacagagta acgcctcgata ccaaatagca aggctgtaaa 5760
gaaatattct cttggtccct catcaccgcaact tattacctga cttttaactt cggataacgg 5820
ccgggaaaac tctcctgtcag gaaactcaga tcagacatca ttgctgcagg ggagctgttg 5880
ccacacctct cagctcttcc aacgagyccg agaacaagtc aataaqgcag tcagacagtaa tayctgtt 5940
cattacatcc agcgcgagga gttttggcag cgacacgccc gctatgctt ctcactaactcc 6000
tgaataagaag cctcgcctcaac ttcgcctccag gttgacccat tctttagctct gcactatgtc 6060
aagaaatgcta acaaaagcagcg cagagctcag aagttcttgg gatggtttga acaacaggttc 6120
taagacattca cggagcgagcc tctctaggtg agcaacacgg cnnaggttgc aacaagatgg 6180
atacattctca cgttagttgctg aactcctctgta aataagggca aactgtccta ccaagacgt 6240
cccgttgggttctgcttctg cagccggtcg ctgctggctgt tggagctccat tctctgcaact 6300
gcatactcctg tgggtgggagat gttgctcttg tgcctacacta tgggagatgaa ggttggctc 6360
gtcgcagccaa gcataattgc cctcctgcaac ccgtaagagaa aatggagcag gctctctaca 6420	
tacatctgtg ccaactcgtgt ggcaagagtt acacccaccc ttgacgtctg ccgcaagaaga 6480
atgaagaggg ctctcggagaga cagggctctaa gcaataacaac cttttactcaa ctgccaggat 6540
aacatgattaa cagaagaggg aagaafatacc atcattctca acaatgcagta cttgcctctca 6600
aacgcacctg cagtccttcc acaagatctt agaacaactca acaaaattct gaacagtgctg 6660		tcgctgtctt acctgtgaaat cttctaggtc ctctacggtt tgaacgctct ggagacctctc 6720
pgctgcagccg tggagaagcgg aactgctccaa taacattcaac gtaatigor ctcctgttgat 6780
ggcaaggttg aacgcctgcag ggacacagtt ccaaatccct tgcacagcagc ttcctctcttt 6840
caacagcttc aatagcttga ttcacgcgcc ggttgtctgta ttttaatacgga atacatttaaac 6900
aacattatgta ctacatcagtttc cggagatgtc ccagtacatct ccaatacaatct ggtgtcatct 6960	
ttcgggttacg ccacaggaat cccactttagc agcaagttcag acctgtcagc caagcttcag 7020
aacccaatcgc aaggtcttgaac cctcaaggg ccaaaatctcg aaggtcctcc ggaagactgac 7080
agtctctcaac atcctctgta tgaataaacctc ttcagcttcat tgcgtgattc gctacaggaac 7140	
tctcttcactaat ctcctctctg gaaacagcagg taccaatacgt tcaacatgtgtag ggagaaac 7200
tctcgggtggaa aagtttctcc caacagttgga gaaatattct ggcagctctgc agaacaagcag 7260
ggagctcaacg agcgggtggttt cttttaatacgagcagatgta ttaacacttc tgcattacaac 7320
aaccagatgga ctcttggctgc cattaaaaaa aacctcctgta attacaacg ccatacatatc 7380
aacggctctgc gatgctaacaa gaaaccttt tcaaatccgcc gcacatccac tgcctcccaact 7440
aacatatgt ttaagttcggga tgggtgcttcgc gatcaccagcgt accatacttgc gtaaatggtat 7500
<table>
<thead>
<tr>
<th>149</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>ttcaatcgtg tcgcacaagc actcaatgac aacagagatca aagactgtga tgcatacag</td>
<td>7560</td>
</tr>
<tr>
<td>tcaaacctgc gcattctgga aagacttctgg ggggattccc accgatcaga taagcccatat</td>
<td>7620</td>
</tr>
<tr>
<td>tacattgta attccttatt cccatcacaatt tattgtaagg tgcacaatgt cggtatcggt</td>
<td>7680</td>
</tr>
<tr>
<td>ggctcaatgc tcaacacgac accagtgttg aagctatgta cacacacta ctacgtaaat</td>
<td>7740</td>
</tr>
<tr>
<td>agtcttctgt atcgcggtac gaaacctcatt atctaagaagt acgcttcaggg caacacagac</td>
<td>7800</td>
</tr>
<tr>
<td>aacacttgta ggacacatga tccggtatgc atcaacgtgc tgcgtggaaga taaggnatatc</td>
<td>7860</td>
</tr>
<tr>
<td>cgttggctgc ccaacagtctg tcaagctgca gttgacacag ttcgtgacag cttgacagatc</td>
<td>7920</td>
</tr>
<tr>
<td>ccacagctgt gcaacoccgac caagaagttgt gtagaaaaa gcacaagatta ccagggaaa</td>
<td>7980</td>
</tr>
<tr>
<td>accaacaact gcacaaagcaacctgcaagac aacacagggca atgacatcgg tttctatcgt</td>
<td>8040</td>
</tr>
<tr>
<td>ttccacocag ttcaccatat gcgcagacctg atcgggccaga actgtcataa cagggcgatt</td>
<td>8100</td>
</tr>
<tr>
<td>gagacgtcat cccgcctatt ggtatcgctt tggcgaattta tccgcgtagga cagttgtgga</td>
<td>8160</td>
</tr>
<tr>
<td>ggggggcga cccctggcgag ggttgcctcc caccctcagtt tgcagaaaat attgatataa</td>
<td>8220</td>
</tr>
<tr>
<td>gataataata tgtgagagca tgtgcgtcttc ctcagggaggg ctaaggtgca gag cacagtca</td>
<td>8280</td>
</tr>
<tr>
<td>gtggccgggc ctttccagat tagaagctgcg agtctgcagg ctagcgggtac caaggtggtc</td>
<td>8340</td>
</tr>
<tr>
<td>gagaagtaact aaggggacttt aatctccgtt accaccatttt taggggtttt actgtcttta</td>
<td>8400</td>
</tr>
<tr>
<td>aaaaactctc ccccttcctcc ctggagctctt aacataaaaa tgcagaatc tttggttggt</td>
<td>8460</td>
</tr>
<tr>
<td>aacgtgtttta tgtgctgattt taagtgggtac aaataagaagca atgacatcact aacattaoca</td>
<td>8520</td>
</tr>
<tr>
<td>aacaatagatt tttctcaact ggcattggttg ccaacacttc caagatatact</td>
<td>8580</td>
</tr>
<tr>
<td>tacatcgtct ggtatccgac actcggtcag cctaggggcat cggcaaccag gaattgaaat</td>
<td>8640</td>
</tr>
<tr>
<td>ctagctgata aacttttggt caatctttaaa tttggtttca aagtcggcag tcaacccag</td>
<td>8700</td>
</tr>
<tr>
<td>ttccacatca ttttctgttc cttccctaaa taactcttaa aacactctt ccaccacccctc</td>
<td>8760</td>
</tr>
<tr>
<td>ccagctttca aacttttggt cggccccagc cgggcttatt tttctctggt tatgttttta</td>
<td>8820</td>
</tr>
<tr>
<td>atccaagaact tggcacaactc atgtrgcaca aacgtaccct tgggtacttt tttctctgc</td>
<td>8880</td>
</tr>
<tr>
<td>acsgatgaag aatcttctct tctcttcttc gtattatttg tttgatattg aagtcgaatcc</td>
<td>8940</td>
</tr>
<tr>
<td>aacgtttact tgtgctgattt aaggggaagtg g</td>
<td>8971</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 11
<211> LENGTH: 7831
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Delta LC-Peptide-Bont/Aev Vector
pFD1SPepB2ATEV
<400> SEQUENCE: 11

gacgcgcgcc gtaggggccc atttaagcgc gcccgttgg tgggtacgg cagcgtgacc 60
gtcaacctgg gggggcgcct caatgcgtgc tgccttcctt ctctctgtgcc 120
acgtccgcg gttcctcggcc tcaaggtccta atcgcggtcgg gcctgccgtgctt 180
agtctggtg agccagtgcag cccaaagaa cccttgaggag gtattggttc aaggtgttgg 240
ccatctctgc gtagagcgtc tgggtccggcc agttcgcctgg ctcttattgt 300
ggcattcttg ctcaacatgag aacataactc aacccacttcc cgtcgctattt ctctgtatca 360
taagggattt tcgctgcttc ggcctttgtg ttaaataatg aagtgattta aacaaatatt 420
aagcgaaat ttacattctt ataatccctt cactttctac ggctgccctt tcggggaatt 480
gttgctggcc gcccctgtgg ttcatttttc tataatcct ctaatattgct gtcgctcttg 540
agacaataac cctgataaat gcttcaatatt gattgaaaaa gpaagagtat gagtatccaa 600
ccttccctgt tgcgccctatt tccttttttct gtcctatttt ttcctctagt tccttgccac 660
ccaacaaagc ccgtaaggtt aaagatgcaaat gaagatcagt tgggtgcagc agtgggttacgct 720
atcagatcgg atctccacgcc gcgtaagctc cctgagagttt tcggccggca gaaacgcttt 780
ccaatgatgaa gcacccctttt aagtctgtca cttggtgcagc tattacccct catttacgatc 840
gggcagagac aacctcagctcg cgcgataacac ttcttcctgca atgccttggc tgtgctctca 900
cctgctcagc aaagatcgctct taagagattac gtaacgatgaa gaaagattat tcgtgcgctg 960
atacccatctg tgtataacacg tcggcgcacac ttaaccttgta caacagatgcttgaggcaag 1020
agctactcgg gtttcttctctaa caacatggggtg gatgtatgtta cttgcttggaga 1080
cggatcgagc atgacgattta ccacaaacggc gacgctgacac ccacagatgc cttgactaat 1140
gcaacacaaaact attaacttggc aaactaccttc ctctagcttcg cccgaccaaa 1200
ttaatgtcg gcagtaaacgcg ttcacatcatt tcggcgccac cgcgcttcctgcccgggtggtct 1260
gctggtctgt gtatcgtggta ttaaatctgga gacgctgag tcggtctctcg ccgtgtaattt 1320
cgacaagctcg gcacacagagc taaaagctgcc gcctcgtctag ttactcaacgc acggggggagt 1380
agggacaactg gcgagctgctct aacattgcagc atgcgctgagc tttggcctgcc actggattaag 1440
catgtgatcg tcgatcaacag atgttattaacta tattacacac atgtacttcg cccgaccaaa 1500
ttttaaaattta aagatagtacg tgtgagacgct ctttcttgata atctcaacgc cccacaaact 1560	taacgagttc gccctgccca actgctctatgt cctgatcctt gataagatcactct 1620
tgagatcctt ttttcttctcg ccattactgc tggattgcacaa caatatttaaacc accgtaacc 1680
ggctggttttt gtttgggccc ttcaggccgta caacagcttct ttcacaggttc aatgtgctgtgat 1740
agccagagcg acgataacacag ttatgcttctt ctaagtgtagc ctgtagtagg ccaccccttc 1800
agaacactcttg ttcagcagctct tacaatacctgc atctctgtctgagcgctgcaga 1860
gccagtgcgctg atangaattgct ctctccagggt tggtaacacaa cggcataattt tggcccttatgctactctgat 1920
tacacgacgc gcaagatcag cctgagagaagc gccacacgctt ccggagacgcc 2040
agagaagcagc aaggtataccg gtaagcagcgg agcgtcgcggc ccacaggaggc 2100
ttcacagggg gaaacgctgctg ttaaatctttactctctatgctgacatcctgatctgctctgctct 2160
agggctagt gtttcttctttgctctgag gcggggggagg gttgggctgc ttcagatcattt 2220
ggggcctgct ttcgcaatgc atctctctttcgcttccagccagcctttagccctttttg 2280
ttaacccctctaccttggtgt acaccccttt tgcgtcagctt ggtgtgaggtg ctcggagatgctc 2340
cgcacggcc ccagcaggccgc ccagcagctcc gtagcggggag aaggggagcttgacggtgctct 2400
cggtagcatc gcgtctcgtt tcgtgactct atcccaaccgc gcagcagaccc gcgcagatactct 2460
ggcacaaatct gttagggttactgtaaatgc gttgagctgtc gttgggccgga 2520
cattatatgct ctaactgacct gcataatagct ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactgt cacaccagact ccataactactg
tgcctgagg gattgtagc cgcgggtgga cagctctgctgc tcccggctgc gcggggagc 3000
gcgagatcag gcgtatgcat ctcacactgc gcggctcag gcacctgcag gcgtgtgctgc 3060
gcgagacgc gcaccccagc tctccctgtcg aacgtggaac ccagcgtgac gcgtgtgctgc 3120
cggagagcaac tccccctgttg ccggtgtgctgc gccttggtgtag gcgtctgcatgc 3180
cggagagcaac gcaccccagc acgtggtctgc ccagcgcgctgc gcgtgtgctgc 3240
agccatagtgc tggagatagc ggcctacttc gccggtctgc gccttggtgtag gcgtctgcatgc 3300
cgccctgggtc tggagagcc gccgtgtagc gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3360
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3420
agccatagtgc tggagatagc ggcctacttc gccggtctgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3480
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3540
gggttgtagc acttgtagc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3600
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3660
gggttgtagc acttgtagc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3720
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3780
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3840
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3900
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 3960
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4020
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4080
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4140
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4200
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4260
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4320
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4380
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4440
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4500
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4560
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4620
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4680
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4740
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4800
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4860
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4920
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 4980
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 5040
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 5100
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 5160
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 5220
tcgcccaacgc ctggctggt tagctgcgtc gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc gccttggtgtag gcgtctgcatgc 5280
-continued

tacatcgtga cccactggtg gccaaggttt aaacaccaca aatgatctgt cegtaagaas 5340
atgaagaggg ctttgggaga ccagcctgaa gctactaaag ccattatccaa ctaccagatat 5400
aatccgata caagaagagg aagaataact atcaatttcca acatctgatga cttgttccctca 5460
aagtctgacg agttccatccaa ccaagctatgg atcaacatcc acaattttct gattcgtgct 5520
tcgctgttct acctgatgaa ctctatgact ccatacggtg tgagccgctg ggaggaacct 5580
gagtgcacgct tgaagaagccg acgtctcaaa tacatttacg ataatcgccg cacccttgat 5640
ggccacagtt gccgctgcttg gagaacagtt aacaatctct tgtcaacgga tattcctttt 5700
caccgcttcca aatactgtag ctcacacagc tgtctctcta ctttacacgca tataattaca 5760
acattaata caataataat ctcacacctg ogattagagt ccataaactct gatcgactctg 5820
tctggttaag ccggcagctt ccaatctgga acgaagaagt gaactgatccc gattcgoaag 5880
aaccacaaat ccgagcttca ctcgggaagc tccaaactcg aagtgactct ggaagaatggc 5940
atcgcttaca cccctacgta tgaanatttcc tcacattctat cttgtgtagat aatccggaat 6000
taacactaact caaatctctct gaatactgaa tcaacatca taaccttgtat ggagaataac 6060
tctgggagta aagttctttcg gaaacttgaga gaattactgg cagacttgca agataacgca 6120
gagatcaaac acgtctgtgg cttatatattac acgctctctag tcaacattct gcgtacactc 6180
aacagagttg ttggctccaa cttactaaaa acgctctctag tcaacatttcc taatacactc 6240
aagctgctgcc tgaagaacca gaacactatt tccaanctcgg gcacaaactct tcgctccaat 6300
aacacactgtg ttaagaaggg tgtgtctgcgc gataccacac gttacaattctg gatcaagat 6360
ttcactctgt tcggacacag aacctttaga aagagcactga acagctttga tgaactaca 6420
toaactcctg ctgttctgag aagacttccg ggcgttacct ctcagatcga taagcactat 6480
taactgtctga attcctctag cccacttaaa tattggaagc tggaacagtgt cgytatccgt 6540
ggcattacgtg acctcaggag aacccgggtg agcttgattga caacacacat ctaccctgat 6600
agctcttggt aacgctctcg aagaaagccg acaacagctg tcaacagtgaat 6660
aaccactagta ggacacactga ctggagttac acaccaactc tcaacactg 6720
cggctgggga ccagctttta cagggcggga gtgtgaagagc tccrgagacg ccggagactc 6780
ccagagcctg gcacacttgac accttctggc ttgagaaaaa caaagttgga ccaaggacacct 6840
accacaccaat gccaaggtga ctcggcagga caacagcggcg agtagccttgc ttctcctgtg 6900
ttcacacgat ttacacatgc tggcagcgtg tgtccgcaaca actgtctaca aatcgaggatg 6960
gagagctcct ctgactgtct acgggaatca tgggttaatt taacgggtgg cgaagtgggg 7020
ggcagagac ccggggcggc acgtgttgcct cccctcctct cggagaagtt ctgacatttta 7080
gataataa gctcgagagc tgtggagcct ctaaaggagc ctaaagcttg gatcgactact 7140
gtcggcggcc ctacctgact cagcctctg acgctctctag aagctgcttc gaactccgct 7200
gagaaagactt aacgaagctt acacacttgg tagagtctt tatttcttg tatgcttta 7260
aaacaaccc caccacctgg cctgaaacctg aaacataaaa tgaatgcaatt tgttggtttgt 7320
aacctgttta ttcgacctta taacttgacg aataaagccg atgacatcag ataatcttaca 7380
aatattatgt ttttttaact catttttact ctggtttgtttt cccagacactt caaatgactct 7440	
tatatctgtg ggctttgtag acgttttgag cttggaggat ccaacacggag ataagtgatat 7500
tatgcttaca actatttatac tattttatat tttgtatata gttcgcagct ctacaccagc 7560
ttcacctaa ttctttctac tcctctttaaa taatccacta aaacaccaat tcaaccctttc 7620
cagttccca acattttcatt cccggccagc cggggcattt tcgcttgctt tgaattttta 7680
atacgaatct cttgcaactc cagttgacaa aacgtcactt tcggctactt ttcttctgtc
7740
acagaatga aatattttctg cccaccccc tctatataag tgctttattc actgaatactc
7800
aacgtaatt tcgcagctga atggcaagt g
7831

<210> SEQ ID NO: 12
<211> LENGTH: 5542
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURES:
<223> OTHER INFORMATION: Delta LC-GFP-BoNT/Aenv Vector pFB1S GFP2ATEV

<400> SEQUENCE: 12

gacgoccct gcgtgcccgt ctttacgctc cctggtgtgg tgggtacgct caggtgaacc
60
getaactctc ccacgcocct acgcgctct tctttctctc tctcttctgcc tctctgctg
120
acgctgctg ctgctgctg ctgctgctg ctgctgctg ctgctgctg ctgctgctg
180
aggtgtctga ggcagtacta ccccaaaa ctgtttgagg gttggtcttc aegctgatgg
240
ccatgccct gcacgcagcg tcctctctct ctctctctct ctctctctct ctctctctct
300
ggctgtcttg tccacacgg aacacactc acacactc acacactc aacactcactc
360
taagggactt tgcgtcttgg tttttttttt tttttttttt tttttttttt tttttttttt
420
aacacgacat ttaacacatt aattctgcct cttttcctttt ctgtgtgtgtg tggggaatt
480
gtggcgcggaa cccctttttt tttttttttt tttttttttt tttttttttt tttttttttt
540
agcatcaat ctcgatatat gttcaattta tttttttttt tttttttttt tttttttttt
600
caatgcttgg tctcgttttt ctctctctct ctctctctct ctctctctct ctctctctct
660
ccatgcgcgg cgtttcttttt cttctctctct cttctctctct cttctctctct cttctctctct
720
atgcaatcgc acacactcag cctttgtttt gtttgcctgg ctttctctct cttctctctct
780
ccatgcgcat gccttttttt cttttttttt tttttttttt tttttttttt tttttttttt
840
ggcagcgagc aacactcttc cctttttttt tttttttttt tttttttttt tttttttttt
900
ccagacagcc aacactcttc cctttttttt tttttttttt tttttttttt tttttttttt
960
atagacaact ccacagctgg aatcagctgg aatcagctgg aatcagctgg aatcagctgg
1020
ggctgtcttu cttttttttt cttttttttt cttttttttt cttttttttt cttttttttt
1080
cggctgtctc gcacagctgg cagctgtctc gcacagctgg cagctgtctc gcacagctgg
1140
gcagctgtct gcacagctgg cagctgtctc gcacagctgg cagctgtctc gcacagctgg
1200
ttcctgcttg gcacagctgg cagctgtctc gcacagctgg cagctgtctc gcacagctgg
1260
gctggcgttg cattgttcag aacactcttc cctttttttt tttttttttt tttttttttt
1320
gcagcttcg gcagcttcg gcagcttcg gcagcttcg gcagcttcg gcagcttcg
1380
ccagcactgc tgggagagct aacactcttc aacactcttc aacactcttc aacactcttc
1440
catttgactgc gtttctctct gcacagctgg cagctgtctc gcacagctgg cagctgtctc
1500
tttaatagt cttttttttt cttttttttt cttttttttt cttttttttt cttttttttt
1560
taagccacttc ctcgataact gcacagctgg aacactcttc aacactcttc aacactcttc
1620
tgagatgtt cttttttttt cttttttttt cttttttttt cttttttttt cttttttttt
1680
ggctgtcttg gcacagctgg cagctgtctc gcacagctgg cagctgtctc gcacagctgg
1740
agcgacgtgc gcacagctgg cagctgtctc gcacagctgg cagctgtctc gcacagctgg
1800
aagacacttc ctcgacaact tcctatttc gcagcttgct cctatttc gcagcttgct
1860
-continued

gccagtggcg ataagtctgt tgttacccgg tgtgactcaa gacgatatgt aecgcataaag 1920
gggcagcgtg gggggtgaa gggggttctg tcgcacagac cagactttga ggaacaagcc 1980
tacaccaagtc tagatattc cagacagctg cattggaaa gagcacaagct cctccaaggg 2040
aggagggcgg aagaggtacgc gtataagcgg agggcatgag ccagagggag caggggagag 2100
ctcgcacagg gaaaaagctg gatttttatc agtctgtcag gtgtttgcgca ccttctgact 2160
gagcgtctgt tttttgttga ctctcagcgg gggcggagcc tattgaaaaa ccgacacaaa 2220
gggtcttttt aaagtcttct tgtccttttgc tcctaatcgttt cttcctgtcg 2280
attatocctt atccgtgtgga taacagcgatt aecgcctttg aagactgctg taacgcttcg 2340
cagcggcgaac gaggagcagcg cagagagact gtagagcgag aagcaggaaga gaggctgatg 2400
cgatattcttc tcctcaagct ttcgctcgcgt atttccaccc gggagccagc cggagtaacc 2460
ggcaattacg gttgaaggctg agttaaaaaa ggtagacccg ctaaagaggg tcttgggcca 2520
cataaatgc ttaatacgaa caaaataagc ttaataatcg acaataaagct ttaatactcg 2580
aagagtagaag ttgtaacgtc aacctataagt acaggagtca aagacaggtgct gaa 2640
ttgtagcc ttaagaaacac ctcctatatc tggagcaga tatttccagaa cttaaatcaatgt 2700
ccgctggtgag ccaagggcttg gtagagaccc cattgagcagc cttgctgacag 2760
aactagcgcgt cggagacgcc gtagctcgcc cttgaaggaatt tgaacgctgc ggaacttgctg 2820
tctgatataac gttgcacacg tcctctcgtg atgcoccaact tttatatagc agaacgtccg 2880
gagagcaca acgtatcgtg ttgcagcatc attacataaag cacaagacag cttggcctca 2940
tggaggtgag ccgggtgccc aagccgctgcc tccggtgtctc gcggagagct 3000
ccgagatata agataataagg ctcctacagc cggcgccaa cttggctgg 3000
gccgagagcg aaaacaacgcgc ttccttggtcg aaggcmgcaa gcggatgtaa tttttctgtc 3120
cggcgcaggt tccgctagcg atgaggactc ccttgcctgt gcccggagg 3180
cgaacatgcatca ggcggagaga ctaagagcagc ggcggctggg aatgtcatgtg 3240
agcctagctc tggagatgt gccaatctga tGGCTCAGA catttaaggg tgggcgatct 3300
ccgtgctggct agatactgtt gctctgctgct aagatcgttgct cttgtccata caaaaaaca 3360
tcgaccaacg gctgaacccg cttgtctctg ggtgacggca ggcatataag ctaacaaaaa 3420
acacccataac cagggcctag aagcggccaa agggcgcgtt cagcggctgc cttgctgca 3480
gttactgac gttgctgtcg aagcgtcagc ctcattctgct ctcagtttc aacgcaacc 3540
gggttactgc aatagggtgg ggtggtcctgt cttgactcgc gtcggagaca aecggcccaac 3600
cctgagacgc gactggcagc gggcttttct tcttgctgtg gcggacagac gcggaatttc 3660
gagaagcgcc agaggattgc cggcagttct cttgatgtct ccgggctagc aagggcatgtg 3720
caggatctgt ccctgcttctt aggacagctg aagcaacttg cccgggcgcgt gggcgcgggt 3780
gttggatgcc ggagataag ccgtatgctt ctggctcgag cttggggagt ctggatgtgt 3840
gtggtcctcgag acctactcact tagatcttag tgggtgctta cgtataactc ggaatattaa 3900
tagatcatg gttgactttaa aatggtctcc aatctgctaa aataaatttt ttatcagtgt 3960
ccggcagaatttt cttggataata aaatatctcc taaaacttcc ggatatacct aaccgtccca 4020
ccagtggcccc ggggttcggg gctcgtcatgg cagcattctg ggggataaa tctotaamaa 4080
acgaccaact tgaatgttcttg gcttgccttg tttatatttg atacatatgct 4140
tacatactag ggacggctta tcaaaacc cattgacacg accacacgag ggggaacacc 4200
gtttacttc aagggctcgc gtagcctcctt tttgggttgg gtnctttgct cagaagggccg 4260
cagttgttca acccgaaag ctcacaattc gaaagtgtac ccgaagaatgc cactgtctac 6660
aacocagtgt atgaaaaatct ttcacatcctctctccctcga taactcctcacc 6720
tcaaatctcctc gtagaaaaa ataacgtgatc attaactctgtac attggagaattc ctcctgggtg 6780
aagtttttcat ctagaacgt ctaaactgcg ccagatcac gctgatacaca 6840
cagctggtgct ttctaaaa cagcagatgtg ctaaacatatc cagcatacct ccaacagtgg 6900
atttttttga cattctcctt ccccccctcc ctaactccct aactactccat cccactgcccc 6960
cctgatgycc acgaacgcttt acaacacctt gcaaccacttt tggcttccaa tcctactctt 7020
ctcaaacag aactcaatgag aagaagatctt gtagtataactgctcttacctccgcttg 7080
gcagcccaac cgagacatctc accttagct cagtaacaagtt cttcactgctc 7140
aactctcatt acaccaacaa atatgtagcaag tggcaaccacgt tgggttcagct 7200
tactcctgaacc gacccagctg ctggtctctcg agaaccacca ctcattctgtg 7260
tactggtcct ttcttcacatg cagaagatg cttttatgttc gccaccaccct 7320
tatccggctg gagagagagtact tatcagttctccc gaccacacagt gccagctggct 7380
agggacacag tctccttac gtcgatgtg cagagaacgc cctttgttggt 7440
accacagctt cctagagagc ctttcggtttt cctcagatgct tggattgtt 7500
ggcctcatt gcagagagaggt gtagggagcc ccgagcagagc ccttttggagtt 7560
tgcaatactgacacagc agaagacgcgtttgctgctctctcctgcc 7620
ctccacactgc ctagctcgg gcccagcagc cccatccagc cctttggagtt 7680
tccctcctcct ccatcttccc aggtagtggtt cctgtgtccgc cctgagcc 7740
cctcttggc gcagctgtggc ctaacttctc gtagagatgc tggttcagtt 7800
agtcggcccg atggcctgatg ctctctatct ctagcttcgc cctcggtgctg 7860
gcttttagg tctctgagctc cctcctcgt ccagctgctg cctctctctc 7920
tactgacacat ttaacttcgct cccacaccttc gccaggtggtt aacactaaccc 7980
cctcctcctc ctctctcctc gacactacat ctagagatgct cttgtggttt 8040
attcagctct ttaatgggat ccaatcactc ccaactctctc cccactacacatg 8100
ctttttcctc ctagctcctg cattgtgtc tttcaccacct ctaatcaccat ccttttttctc 8160
tgagccgctg ccggaggtgtgtg gctgctgtg gctgctgtg gctgctgtg 8220
aactttttag ctatcctca tatcctcctc ccttctctctc aacttctctctc 8280
attctcttttag atatcactacat ccaatcactc cccactactc cccactactc 8340
aactttttag cattacagct ccggaggtgtgtg gctgctgtg gctgctgtg 8400
cctagcact ctaaggtttgtg ccgctgctgtg gctgctgtg gctgctgtg 8460
aatttttcct ttttctctctc cctttttagc ctatcctccct ccttttttctc 8520
tttcccgcttc ctgttttttt ttgcttttttt ttgcttttttt ttgcttttttt 8580
<210> SEQ ID NO 13
<211> LENGTH: 2890
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<222> OTHER INFORMATION: Vector pBluescriptII
<400> SEQUENCE: 13

cttccttactgc gttactgtggga cttttttggg gatgtccgc gsaacccctta tttggttttt 60
ttctcatacctcagcagc tgttccgctct cagacacactaccattcagctcactgctg 120
ataatatgga aaaggaaga gtagagatct ccacactttt ccgttggcc ttatcacttt 180
cttgggctg cccttgagtc ctaggttctct cggcctcctg gatctggtgta acgttgatga 240
tgtctcagt gcctgagcggg tattctgcc ctgctcctcc tggagccctgc tggagctgct 300
gatccctcag tgtttctcag gcggccagat gcggccagat cggccagat 360
gatgttgct gcgtcagcct gcggccagat gcggccagat cggccagat 420
acaactctct aagatgatagt gtgttggaga ctcacccgct cagacagatg atctacccgtgt 480
tggccatgca ctagatgagact atgcacttga tggccatatgcc atgatggatga acaactct 540
cacactttct ctgacaaagc ctggagacac gaaagagcta acggctcccc tggacacact 600
gggagatct ctagatgcttg gcggccagat gcggccagat cggccagat 660
cgcggagcct gccagacagat gcggccagat gcggccagat cggccagat 720
tggggagccct ctagatgcttg gcggccagat gcggccagat cggccagat 780
agttggagcc ctggccgctct cggccgctct cggccgctct cggccgctct 840
tggagccct ctagatgcttg gcggccagat gcggccagat cggccagat 900
ctgagccctct gtagatgact ctggccgctct cggccgctct cggccgctct 960
acagatgctgt gtagacatct ctggccgctct cggccgctct cggccgctct 1020
ctgatataaat ctttagatggt attacccctgt ctggatataat cagaaagacct cagaaagacct 1080
agaatttactt aagaccaaat ttaatttga taactgttata atttgtataattattatg 1140
aatatttatt tttacatcttg aatatttatt tttacatcttg aatatttatt 1200
aatcctagttt cggccgctct cggccgctct cggccgctct cggccgctct 1260
ctattaaaga ctggagccct cagacagatg atctacccgtgt cagacagatg atctacccgtgt 1320
ccacacgctc ctagacatct ctggccgctct cggccgctct cggccgctct 1380
aattcctaccag ctagacatctc ctggccgctct cggccgctct cggccgctct 1440
gaaagagacctc gcggccagat gcggccagat cggccagat 1500
cggtggcctg gccggccagat gcggccagat cggccagat 1560
atctggagtt gtagacatct ctggccgctct cggccgctct cggccgctct 1620
ccctccatttt ccctctgcgt tcggcctcatgcct cggccgctct cggccgctct 1680
cggtggctac ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 1740
ccagacattt ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 1800
ccacacgctc ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 1860
ccagacattt ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 1920
tctggcctac ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 1980
tgaaagagac ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2040
tacacccgtc ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2100
tacacccgtc ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2160
tgagctgcct ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2220
tgtggctggt ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2280
ttccctggcct ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2340
ccacacgctc ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2400
acacccgctc ctagacatct ccctacccctc ttttctctct ctttgctgtgtcctcttctctct 2460
| ctataggcca gatctgaaga cataagtcgg tccgttcgaa ccagaacctt ggagaacctg |
|-----------------|-----------------|-----------------|
| 2520            | 2580            | 2640            |
| cgccgcccgt atccatgcca caacgtgaac gccgagtttag gccctgctagc gcgcgtaccc |
| 2700            | 2760            | 2820            |
| cggtaacta agtagtatgat ccctcttcag ttacccgggc aagcctgctga tacaggatat |
| 2880            | 2880            |                 |
| ctcggaggct cggagctcgc ttcaggaaggcc tataaggtgt ctgtataacgc acctgaggcc |
|                 | 2820            | 2880            |
| gttttaacac gccgtagact gcgagcttcg ccgggacact gccagcacttg ccttgcttcg |
| 2910            | 2970            | 3030            |
| ctcggggccct cggccgcttc cggcggtgct gcgaagccgg cgggcgtctc cgggcggctg |
| 3090            | 3150            | 3210            |

SEQ ID NO: 14
LENGTH: 6925
TYPE: DNA
ORGANISM: Artificial
SEQUENCE: 14
ttcaactgag cgtcagaccc cgtagaaaaa atacaagagt atctctgaga tctctttttt 1680
c tgtcgcgtaa tgtctgtcct gcacaacaaa aacccacgcc tccacgcgtt ggttttgttg 1740
coggtaaag aggtaccaca ccttttttgg aagtttaagct tgggaacoag aagccagata 1800
ccaaatattct tctctttcatgt tgaagcttag ttagggcacc aacctcaagaa ctcctgagca 1860
cagctcact acctgcctct gcaaatcctg ttccacgctg ccgctgcaac aacggatatg 1920
tgcgttctta cccggtgtagg ctcgaaagca cttgattcgg ataaagcgca ggggtcggcg 1980
tgaacggggg gttcgctgac ccacccgacgg cttggagcga ccaacctcaac cgaactcgga 2040
tacactacgc gtggacgtatg agaaagcgcc accgtcctccg aagggagaaaa ggcggacag 2100
tacgctgtaaa ggcggaggg cccgcaacga gacccgcaag gggacgttcc aacggggaaac 2160
gcctgtaacct tttatagccg tgtctgggtt ccgtcgtattc gacttggcag tcgattttagt 2220
tgatgtctgt aggggctggcg gacgctactgg aaaaaacgcca gcacaagctgc ctttttaaccg 2280
ttcctggctc tttctgctgc tttctgtcact agttaagttg agttactgca ctctattagc 2340
acccacgagct tttactatcct tttcttcgcccg tctgtatgttg tgtgaaatgg ttagcggtta 2400
acaatctgcc acgggaacca gctatgcgac tgtaggctag aagctacgtta ataagctgca 2460
catagggcga gatctgataa cattaaagcgg tcgctgcaac ccgaactcttc ggaagccttaa 2520
cctttaaaaa acggccacca tgttaacttct gtcgacacgg gcccgttttg tttctgcgtt 2580
atacacatttc tacatctatgg cggctcgcttc tcaacaacat ctcacaactc acaacaccac 2640
gcgtggaaacc tctgatattc aagggggcgc tgtgctccctg tttctgctgc tctgtctgct 2700
cacggcgagcg ggtggagcag ggcggccttg cttcctcaaca cactttactt acaaggtacc 2760
tgtaagtattg tggagccattg ccttttttcc actctggaat ggcgcagcga tggacaccct 2820
gaaagccattc aagatacaca acaaffaacc tggctgtctcc gacggtgacca cttcaccacca 2880
ccctggagag gggcgaccca acctccccac aagggcacaac acctgctcgt ggtctttctc 2940
catcgagact acattgcggcc gccgttaaca gacggcggaggt ggtggaggtt acgtggccac 3000
gtggctggag aggtatatct ctctggtacct cccgagattg ctcgagcagta gcaatctggc 3060
cggtaaccc ctttggggcggt cggccttcttc acgaccacag aacaaatcgg tggcaccacca 3120
tgtacaact tgtagcctac cagattggcg aaggtctgccag ctgctgccag tggattgcat 3180
cattggtcga cccgaacgca gaaaatcctt gtcggtgtgcc gctctggtgg aagacaggtc 3240
cacacgtcag cggaacctct acgcccccag caaccctcaac cttccctcgg ggtctttatc 3300
attggcttc gcagagcccg tgtggaggttg cgccacaccc cttttgggct tgtgcaaggtt 3360
tgcacagct ccggcggctca ctttgctccc tgtcctgctgc acgccgctgg acacccctca 3420
tgctttgctg atacaaacag cccagctggg caaagctgat aaccaagcgct actgtgagat 3480
gagccgcttg aggctttctct tttgagattg gcaagcccttc gcggctgcaag atgcaagact 3540
cacgacacgc tgcagagaga atgaatcccg ctcttctact tacaccaagt ttaagaagct 3600
cggcttcaaca tgcacacgac aagagtctag cgagggctcag aagcttactg tacagttata 3660
gaaagatgtt ttcacaggga aatcatctct gcagacgagcc acggccccct gtctcctctgt 3720
ggcgacacag ctattgcatg aatcgtgctc cgctcttact tatccaacagat ttaagacatc 3780
cgggctaatc ccacaaaccgt ggcgttcgac ggcagggactaatttcag cccctcagata 3840
ttcctgaaa ttccttaaacg ccaacagcct gccagcttctg gcagagtcgtg acaagagcttg 3900
gtgaaggact gccattgccc tcaaggatct tatgtgtggc tcaactcttg 3960
<table>
<thead>
<tr>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>caaccaacttg cagagaacctta ctgagacttgtgctagttgctctc aagctgctct ggtgcctctgg</td>
</tr>
<tr>
<td>acctcaaoac agtcccagtagc ccaggcttat acagcagtac aagataagcg</td>
</tr>
<tr>
<td>tctgaaacgct ctgctgtcata agggtaataa ctgggcttcgt gcttttaggc caacgacgaga</td>
</tr>
<tr>
<td>taactctcaag aggctacatca aacaaagttgac agaatcagcg cctgtagaca atacgaaacgc</td>
</tr>
<tr>
<td>ggtgtagaag agatatttctt cggatctcat ccagcaaatat ttaactgac gtaacttga</td>
</tr>
<tr>
<td>taactctcaag aggctacatca aacaaagttgac agaatcagcg cctgtagaca atacgaaacgc</td>
</tr>
<tr>
<td>ggtgtagaag agatatttctt cggatctcat ccagcaaatat ttaactgac gtaacttga</td>
</tr>
<tr>
<td>gatggttcaag acctcaacttc ccggctgtcga aagagacggaca aacagagagc cggctgccttg</td>
</tr>
<tr>
<td>ttaactctcaag aggctacatca aacaaagttgac agaatcagcg cctgtagaca atacgaaacgc</td>
</tr>
<tr>
<td>ctctctcaacag aagactctgcac ccggctgtcga aagagacggaca aacagagagc cggctgccttg</td>
</tr>
</tbody>
</table>
|切割 - 结束
gtgcgcctag ggacactac gatagcttct tagtgggtcata gctataactcc ggaattttaa 3900
tagatcattgg agataataac aaatgataacc atctcgcacaa taataaagta ttttactgt 3960
tctgtggctg ttggttaaat aaaaagcaacct ttaataaatcac ggttactctaa ttaagctt 4020
ccatcgaggggc gcggattgcc tccccgagaa ccgaaactct ggaaagtttaa cttccaaaaa 4080
accggacaca tggaaatcttc tgcgtacgtt ctggctgtgtt tatgccattt atatcttctt 4140
tcatctctatgg cggcgctcct tcaacaccat ctcacccact accaaccaccc ggttgaacc 4200
ctgatttccc aggaggccctgg taactacaacctt ctgggctgctg tcctcgctgct ctggtg 4260
ggggcagcaagt gcggccggtt gtttaacaaa caaatttaacct acaaagatccc tgtcaatggt 4320
gtggacacttgc cttataatcc gctaccca aaggttcagaa ccaacgacccag cgggttcagc gtgtctacta cggagaaggg 4380
aagatctccaga acacaaaattgg ggtactctgact gaaagtgcaac ttaacacggt gttgaacattg 4440
ggcgacactaca acctctcacc cggattsagggact gtttactaca cggataagcaag 4500
tactttgtgca gcgtataaagc ggagaacaac actcctgaagg ggtgaccaac gttgcttggag 4560
aggactctact ttcggcagact cggacgcgat tctcgcgctca gccttcgctgt cggtctgact 4620
ctctgggggg gccgggatcag tcgtaacatc gttgactttaa ttgtaaatcccc gcggtattg 4680
gtttactcaac cggagacggag tttacgccct cggagatgta agcctctctacta ctgggtctca 4740
tcgccacagca ctcctctgctg tgaatgcgtta ttcctcgggt ctccgagctg ctgaaccgag 4800
cgcacaaggtga agcgacctcc ccaacgacccag cgggttcagc ctggtctact cggataagcaag 4860
gaggtggagcc gcggagcctgtc gcccggctcctg ggggtttctgt cggccgctc tggcactgtg 4920
ccctggggtaga cgcggctgtg cttgctcttc gaacggcagc gcgcgccact cggcgctctc 4980
atcagcttcc aggcggcttgctt caatgtgaaat acaggaagccct acataaggatt gaggggtctg 5040
ggggcttctc tctgggatcagata aaggggggcag actcctgaacc ggtgaccaac gttgcttggag 5100
tggccccgaatgtttctcg ccgcttactct taacactcagta ccaacgacccag cgggttcagc 5160
tggacacagcat ctggtgctctc tggagggctag cagagctcttt cggagaaggg 5220
ttcacgagga aatcgcgacg cccgcggctcctc gcgattttata gcgaagatgaa ctttgcttga 5280
aatactggac agacggcctg cgggtcagc gcggacagcgg cgaattctgg cttgctggga 5340
tctctcggag cggagcctgct tggactccttc gaacggcagc gcggaggtcct cggactgtgg 5400
aacaatgcttc ccaacgccactt catggtcctg cctgatcgag cggagctggag cggaggttgg 5460
tggccacagca ggtatatgctgcc cccagaccttt gcctatattac acataagcgtg cgggttcagc 5520
aagacaccctact caagctggtct cagagctcttt cggagaaggg 5580
tcccatctgct tccgctctgag cccgcggctcctc gcgattttata gcgaagatgaa ctttgcttga 5640
tcactatgccc gcggagctgc aagacgctgtc aagcataacgc gcgtcgtcagt cgtcttcagc 5700
accgcgtctcg ctggtgctctc tggagggctag cagagctcttt cggagaaggg 5760
aatctctgt tggatgatct cccgcggctcctc gcgattttata gcgaagatgaa ctttgcttga 5820
gaacacactc ctgcgcggag aaggtcggtgg cttgacacca gcgcgctctc gcgattttata gcgaagatgaa ctttgcttga 5880
aacaattgagc gttccaccctcc gcggaggtcct cgggttcagc ctggtctact cggataagcaag 5940	tcttctaggg cgcggaggtcct cgggttcagc ctggtctact cggataagcaag 6000
atatgaagctgg ccggctctctctcaagacgc gcggaggtcct cgggttcagc ctggtctact cggataagcaag 6060
aatctctggttg ccggctctctctcaagacgc gcggaggtcct cgggttcagc ctggtctact cggataagcaag 6120
gacacaccctcc gcggaggtcct cgggttcagc ctggtctact cggataagcaag 6180
-continued

attatcatct atttggacct gcggcttgatt attggcaaca tgctctccaa agagctatcctc 6240
gtggctgctgc gtagctctga catgctcttc atctctcttg agttctctcc tgaatctgcctc 6300
atccctcgtgc tggcgtgctgc gcgtctcttc tcatacatg gggatgagtt ctgctcttgctg 6360
ccgacattca tcattacatc tctccaacgct gaggccaaat gggagcggtg ccacaaatatc 6420
atctgacactc actgggtcggc aagggcattac accccacactg atctgacacg taagctaatgc 6480
aagggacctg tgcagacgac gccagcagatc accttgttgct actacatctg gccatctctctc 6540
cattcatacag aagagagaaact gataacatct accctcacaag tcgatgctgta ctgctctcag 6600
cgagcagcctg tgcgtctctgc tcgacttcgc aatctctgtgat cgaacatcctg cctctctctc 6660
tcgacgttac tcatgatgcc ctgctctgctg tggagccatc agagctcgttct ggcgtgtttgc 6720
gcggcgcgtgctg tgcgcaaggg aaggggctaa ccagctttgtc ctcagatgtaccactgcctg 6780
cagctggtgca ctgctaagag cactgtactg gttggttcac ccagcctctct ggcgtgtttgc 6840
cctctccactg tctctccata ctcgcttccta caccttgctt cccgctctct ggcgtgtttgc 6900
attatacata cttattcttc acatctctgctg ggtgcgtttgt cttctctccgt cggctgctttgt 6960
tgggcggcagc cactctgcttg ggtgcgtttgt ccagctttgtc cgctctttctt cttctctct 7020
cacactcttcg ttcaccttact gggagctgctg gggagctgctg gggagctgctg gggagctgctg 7080
gtctactttca ccagcagctgg ggtgcgtttgt ccagctttgtc ccagctttgtc ccagctttgtc 7140
tccgacttcgct ggcgtgtttgc aggagagcgctg ggtgcgtttgc aggagagcgctg ggtgcgtttgc 7200
tgcttggtgctgc gtcgcttgctgc cggctgtgctgc cggctgtgctgc cggctgtgctgc cggctgtgctgc 7260
acacactgtact ggtgcgtttgc aggagagcgctg ggtgcgtttgc aggagagcgctg ggtgcgtttgc 7320
agagctgttcag ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7380
gtctgtgtggg atgtgtgtgtgc tgggtagggagc ctgctctctct ctgctctctct ctgctctctct 7440
atactacctg ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7500
actctctctct ctgctctctct ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7560
actactctctct ctgctctctct ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7620
actactctctct ctgctctctct ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7680
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7740
ttccctctctct cggctgtgctgc tgggtagggagc ctgctctctct ctgctctctct ctgctctctct 7800
atgtgtgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7860
tgcttggtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7920
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 7980
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8040
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8100
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8160
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8220
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8280
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8340
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8400
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8460
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8520
agagctgtgtgctgc ttcacctctggggagc ctgctctctct ctgctctctct ctgctctctct ctgctctctct 8580
catgtctgga tcgtgatact gcttgacgct aggagatccg aaccagataaa gtgaatcta 8640
gtcccaaatc attttgttcat tttaatatat cgtatagct ttaggctca caccccgctt 8700
cacatcttta tgtcacttctt ccctaaataa toctaaaataa atccttttcg aaccctcccc 8760
gtcccaaatc attttgttcg cccacagcgg ggcatttttc tttctggtat gtttttaatc 8820
aacaatcctg ccacatcctaat gccaaacacc gcatactctgg gcatacttttt cctgctccaa 8880
gatgaaataa tttttgtgca tcctttctgt attaagtttt ataaagact gaaataacac 8940
gttattgtgc agctgtaaat gccaatgctg 8968

<210> SEQ ID NO 16
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide C1-1S
<400> SEQUENCE: 16

gatgtaagata ctaagttgcgg tccg ttgc gaacta cctg gga gctgg ccgt 60
atccatggca caggcttca gcgtgcttag ggcctatgca cgt 103

<210> SEQ ID NO 17
<211> LENGTH: 95
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide C1-1A
<400> SEQUENCE: 17

catagggcgc cagactgactc gaaagctgt atcatggata gcggccgctg caagttc 60

gagttctggt ta cagactgac cga cattg at gttca 95

<210> SEQ ID NO 18
<211> LENGTH: 108
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide C1-2S
<400> SEQUENCE: 18

ggcgtatgc gcgc aagcg ctggta ctgctgag ttac cggg 60
atgcactgca tgcctcgatg gcggc tcggggc tcggg 108

<210> SEQ ID NO 19
<211> LENGTH: 104
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide C1-2A
<400> SEQUENCE: 19

ccttcctggag gcggctcgat gcctcctagaa ttcgcatatg atggctcgctg cgggttaact 60
gcggagat atcctacgtg aggagccggc gttacg ccgg atag 104

<210> SEQ ID NO 20
<211> LENGTH: 98
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide CP1S
agcttacGct cttaaaacc gccacatga aatctttagt caacggtgcc cttgttttta 60
tggtgtata cattctttac atctatgc 88

<400> SEQUENCE: 21
gaggccgcatag tgtgaagaaa tgtatacgc cttaaacaca aggccaaagt tgactaagaa 60
tttcatgttg ggggttgggtt aggagta 88

<400> SEQUENCE: 22
gggcgtccat caccaccatc atcaccatca ccaccaca 38

cggcgtgtggt ggtgatggttg atgatgggtg tgtgagc 38

<400> SEQUENCE: 23
cgccgtgaaa cctgatattt caagggcccg gttactccct gtttgtggtg ctcggtgctc 60
tcaacgcgag ccgtgagcgc 78

<400> SEQUENCE: 24
cgccgtgaaa cctgatattt caagggcccg gttactccct gtttgtggtg ctcggtgctc 60
tcaacgcgag ccgtgagcgc 78

<400> SEQUENCE: 25
cgccgcccac gcgcgcgcgtt gcgcagacgg agcagccaag acagggaagt ccaccgccccc 60
tgaasaataa gggttctca 78

<400> SEQUENCE: 26
cgccgcccac gcgcgcgcgtt gcgcagacgg agcagccaag acagggaagt ccaccgccccc 60
tgaasaataa gggttctca 78
agttacat ggctcatcac caccatcata ccataccacca ccaca 45

gcgtgtgtgt gcgtgtgtgtgtagatgtgtggtgtgcgtgacca tgctga 45

agcttaagct gtctcaatcac caccatcata ccataccacca ccaca 45

gcgtgtgtgt gcgtgtgtgtgtagatgtgtggtgtgcgtgacca tgctga 45

taccctaggg 10

tcagcgcgta gcgtgaccaaggtgcgcgtcat g 31

tacgtgtcagc gtgaccaaggg gcgcgagct g 31

tataacagtc tctggatagc tcatcatttc g 32

tataagttacc ctgtgacagc tcacagctg ct 32
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Enhancer sequence

<400> SEQUENCE: 33
aacctctaa aaacgccac c

<210> SEQ ID NO: 34
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Strep Tag II Sequence

<400> SEQUENCE: 34
Trp Ser His Pro Gln Phe Glu Lys
1  5

<210> SEQ ID NO: 35
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: cis-DNA sequence

<400> SEQUENCE: 35
aacctctaa aaacgccac c

<210> SEQ ID NO: 36
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Honeybee mellitin signal peptide

<400> SEQUENCE: 36
Met Lys Phe Leu Val Asn Val Ala Leu Val Phe Met Val Val Tyr Ile
1  5  10  15
Ser Tyr Ile Tyr Ala Ala
20

<210> SEQ ID NO: 37
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aadek Light Chain N-Terminus

<400> SEQUENCE: 37
Gly Ala Gly Amp Ser Leu
1  5

<210> SEQ ID NO: 38
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aadek Heavy Chain N-Terminus

<400> SEQUENCE: 38
Ala Leu Asn Amp Leu Cys
1  5

<210> SEQ ID NO: 39
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aadek Light Chain Tryptic Peptide Sequence

<400> SEQUENCE: 39
Leu Leu Cys Val Arg
1  5

<210> SEQ ID NO: 40
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: BoNT/Aadek Heavy Chain Tryptic Peptide Sequence

<400> SEQUENCE: 40
Ala Leu Asp Leu Cys Ile Lys
1  5

<210> SEQ ID NO: 41
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Cleavage Product of BoNT/A Prepeptide

<400> SEQUENCE: 41
Gly Tyr Asn Lys
1

<210> SEQ ID NO: 42
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Excision Product of BoNT/A Precursor

<400> SEQUENCE: 42
Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys
1  5  10

<210> SEQ ID NO: 43
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Peptide Insert With Substrate Sequence for BACEL

<400> SEQUENCE: 43
Ser Glu Ile Ser Tyr Glu Val Glu Phe Arg Trp Lys Lys
1  5  10

<210> SEQ ID NO: 44
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Insert Between M447 and K448 of BoNT/Aadek Construct

<400> SEQUENCE: 44
Asp Asp Asp Asp
1
What is claimed:

1. An isolated *Clostridium botulinum* neurotoxin propeptide comprising:
   - a light chain region;
   - a heavy chain region, wherein the light and heavy chain regions are linked by a disulfide bond;
   - an intermediate region connecting the light and heavy chain regions and comprising a highly specific protease cleavage site, wherein said highly specific protease cleavage site has three or more specific adjacent amino acid residues that are recognized by the highly specific protease in order to enable cleavage; and
   - a cargo attachment peptide sequence to enable site-specific attachment of cargo, wherein the cargo attachment peptide sequence is positioned upstream of the light chain region and is separated from the N-terminus of the light chain region by an amino acid spacer sequence.

2. The propeptide according to claim 1, wherein the *Clostridium botulinum* is serotype A.

3. The propeptide according to claim 1, wherein the cargo attachment peptide sequence comprises an S6 peptide sequence of SEQ ID NO.2.
4. The propeptide according to claim 1 further comprising: a signal peptide coupled to the cargo attachment peptide sequence, wherein the signal peptide is suitable to permit secretion of the neurotoxin propeptide from a eukaryotic cell to a medium.

5. The propeptide according to claim 4 further comprising: a 10-His affinity tag positioned between and connecting the signal peptide to the cargo attachment peptide sequence.

6. The propeptide according to claim 5 further comprising: a TEV recognition sequence positioned between and connecting the 10-His affinity tag to the cargo attachment peptide.

7. The propeptide according to claim 5 further comprising: an 8 amino acid StrepTag II connected to the propeptide at the C-terminus.

8. The propeptide according to claim 1, wherein the highly specific protease cleavage site is selected from an enterokinase cleavage site and a TEV recognition sequence.

9. The propeptide according to claim 1, wherein the light and heavy chain regions are not truncated.

10. The propeptide according to claim 1, wherein the entire catalytic domain of the light chain region has been removed.

11. The propeptide according to claim 10 further comprising: a fluorophore connected to the N-terminus of the light chain region.

12. The propeptide according to claim 1, wherein the cargo is selected from the group consisting of lipid moieties, therapeutic agents, marker molecules, and targeting agents.

13. The propeptide according to claim 12, wherein the cargo is a lipid moiety selected from the group consisting of fatty acids, neutral lipids, phospholipids, and complex lipids.

14. The propeptide according to claim 12, wherein the cargo is a lipid moiety selected from the group consisting of palmitoyl-CoA, C-22 aliphatic CoA, and cholesterol CoA.

15. The propeptide according to claim 12, further comprising a neuron-specific protease cleavage site positioned between the N-terminus of the light chain region and the cargo.

16. The propeptide according to claim 1, wherein the amino acid spacer sequence comprises at least 7 amino acid residues.

17. An isolated, physiologically active Clostridium botulinum neurotoxin produced by cleaving the propeptide according to claim 1 at the highly specific protease cleavage site, wherein the light chain region and the heavy chain region are linked by a disulfide bond.

18. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the Clostridium botulinum is serotype A.

19. The isolated Clostridium botulinum neurotoxin according to claim 17 further comprising: a signal peptide coupled to the cargo attachment peptide sequence, wherein the signal peptide is suitable to permit secretion of the neurotoxin from a eukaryotic cell to a medium.

20. The isolated Clostridium botulinum neurotoxin according to claim 19 further comprising: a 10-His affinity tag positioned between and connecting the signal peptide to the cargo attachment peptide sequence.

21. The isolated Clostridium botulinum neurotoxin according to claim 20 further comprising: a TEV recognition sequence positioned between and connecting the 10-His affinity tag to the cargo attachment peptide sequence.

22. The isolated Clostridium botulinum neurotoxin according to claim 17 further comprising: an 8 amino acid StrepTag II connected to the neurotoxin at the C-terminus.

23. The isolated Clostridium botulinum neurotoxin according to claim 18 further comprising: E234>A and Y350>A mutations.

24. The isolated Clostridium botulinum neurotoxin according to claim 18 further comprising: one or more mutations in the light chain region selected from K43>S, K46>H, and K444>Q.

25. The isolated Clostridium botulinum neurotoxin according to claim 18 further comprising: a S571>N mutation in the heavy chain region.

26. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the light and heavy chain regions are not truncated.

27. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the entire catalytic domain of the light chain region has been removed.

28. The isolated Clostridium botulinum neurotoxin according to claim 27 further comprising: a fluorophore connected to the N-terminus of the light chain region.

29. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the cargo is selected from the group consisting of lipid moieties, therapeutic agents, marker molecules, and targeting agents.

30. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the neurotoxin is atoxic.

31. The isolated Clostridium botulinum neurotoxin according to claim 18, having a structure selected from BoNT/Ad<sup>5</sup> Αnt<sup>5</sup>, BoNT/Ad<sup>4</sup> Αnt<sup>4</sup>, ΔC-Peptide-BoNT/A<sup>4</sup> nt<sup>4</sup>, and ΔC-GFP-BoNT/A<sup>4</sup> nt<sup>4</sup>.

32. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the cargo attachment peptide sequence comprises at least 7 amino acid residues.

33. The isolated Clostridium botulinum neurotoxin according to claim 17, wherein the neurotoxin (a) has an LD<sub>50</sub> that is at least 75,000-fold higher than the LD<sub>50</sub> of wild-type Clostridium botulinum neurotoxin and/or (b) accumulates within neuronal cytosol in higher amounts than wild type Clostridium botulinum neurotoxin.

34. The propeptide according to claim 1, wherein the propeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6.

* * * * *