

# De novo subtype and strain identification of botulinum neurotoxin type B through toxin proteomics

Suzanne R. Kalb · Jakub Baudys · Jon C. Rees ·  
Theresa J. Smith · Leonard A. Smith ·  
Charles H. Helma · Karen Hill · Skadi Kull ·  
Sebastian Kirchner · Martin B. Dorner ·  
Brigitte G. Dorner · James L. Pirkle · John R. Barr

Received: 9 November 2011 / Revised: 5 January 2012 / Accepted: 18 January 2012 / Published online: 7 March 2012  
© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** Botulinum neurotoxins (BoNTs) cause the disease botulism, which can be lethal if untreated. There are seven known serotypes of BoNT, A–G, defined by their response to antisera. Many serotypes are distinguished into differing subtypes based on amino acid sequence, and many subtypes are further differentiated into toxin variants. Previous work in our laboratory described the use of a proteomics approach to distinguish subtype BoNT/A1 from BoNT/A2 where BoNT identities were confirmed after searching data against a database containing protein sequences of all known BoNT/A subtypes. We now describe here a similar approach to differentiate subtypes BoNT/B1, /B2, /B3, /B4, and /B5.

Additionally, to identify new subtypes or hitherto unpublished amino acid substitutions, we created an amino acid substitution database covering every possible amino acid change. We used this database to differentiate multiple toxin variants within subtypes of BoNT/B1 and B2. More importantly, with our amino acid substitution database, we were able to identify a novel BoNT/B subtype, designated here as BoNT/B7. These techniques allow for subtype and strain level identification of both known and unknown BoNT/B rapidly with no DNA required.

**Keywords** Botulinum neurotoxin · Botulism · Mass spectrometry · Proteomics

S. R. Kalb · J. Baudys · J. C. Rees · J. L. Pirkle · J. R. Barr (✉)  
Centers for Disease Control and Prevention, National Center  
for Environmental Health, Division of Laboratory Sciences,  
4770 Buford Hwy, N.E.,  
Atlanta, GA 30341, USA  
e-mail: jbarr@cdc.gov

T. J. Smith  
Integrated Toxicology, United States Army Medical  
Research Institute of Infectious Diseases (USAMRIID),  
Fort Detrick, MD 21702, USA

L. A. Smith  
Office of the Chief Scientist,  
Medical Research and Materiel Command (MRMC),  
Fort Detrick, MD 21702, USA

C. H. Helma · K. Hill  
Bioscience Division Los Alamos National Laboratory,  
Los Alamos, NM 87545, USA

S. Kull · S. Kirchner · M. B. Dorner · B. G. Dorner  
Robert Koch-Institut, Center for Biological Security,  
Microbial Toxins (ZBS3),  
13353 Berlin, Germany

## Introduction

Botulism is a potentially fatal disease caused by exposure to botulinum neurotoxins (BoNTs), 150 kDa protein neurotoxins. In vivo, BoNTs halt nerve impulses by cleaving proteins necessary for acetylcholine release into the neuromuscular junction. This leads to a flaccid paralysis, which can affect respiration and may necessitate ventilator support for the patient [1]. BoNT is the most lethal toxin known with an estimated oral LD<sub>50</sub> of approximately 70 µg for the average human [1]. This extreme toxicity has led, in part, to the BoNT's current CDC designation as a category A select agent for bioterrorism, making it one of the most likely agents for bioterrorism [1]. Botulism patients receive a serotype-specific immunoglobulin products that are most effective when administered within 24 h of exposure [2]. Therefore rapid determination of exposure and identification of the BoNT serotype is an important public health goal.

There are seven different serotypes of BoNT, A–G, which are defined by their ability to be neutralized by an antiserum that was produced using a specific BoNT. Nucleotide or amino acid sequence variation within BoNTs in strains of a serotype has led to the designation of subtypes. Historically, subtypes have been defined by distinct cultural/biochemical characteristics [3], functional differences [4], or differential binding of monoclonal antibodies (mAbs) [5–7]. Within the BoNT/B serotype, there are currently six known subtypes, B1–B6 [8–11]. The B1–B6 subtypes exhibit amino acid variation of 7% or less. In addition, there is amino acid variation within a subtype, such as 1.5% within B2 or 1.7% within B4 subtypes. This variation could be a new strain (*Clostridium* organism) or toxin variant (neurotoxin protein), with some the neurotoxin of some strains having as few as a single amino acid difference, or 0.08% difference.

Identification of the subtype of BoNT is important for several reasons. First, one definition of a subtype of BoNT indicates that different subtypes of toxin might have differential binding to monoclonal antibodies, and perhaps some polyclonal antibodies as well [4,11]. This becomes important as researchers search for an alternative treatment to the currently used equine immunoglobulin approach to treat botulism. Various mAb could be proposed as immunoglobulin treatments for botulism; however, if there is differential binding of these antibodies to different subtypes, care must be taken in choosing which antibodies to use as treatment, as the antibodies might not be effective at neutralizing all subtypes of BoNT within a serotype. Secondly, identification of the BoNT subtype could be important to epidemiology and forensic investigations attempting to trace the origin of the toxin and its spread in a botulism incident. Concurrent outbreaks of botulism could be identified as originating from the identical or diverse sources based upon the subtype of toxin present.

We previously described methods to identify the serotype A subtypes BoNT/A1 or /A2 and the serotype B subtypes BoNT/B1 or /B4 using mass spectrometry [12,13]. Our methods involve a tryptic digestion of the toxin and matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) or liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis of the tryptic fragments, followed by querying the data to a protein database that contains entries from the different BoNT subtypes. The subtypes BoNT/A1 and /A2 are approximately 10% different from each other, and the LC-MS/MS analysis was able to detect over half (58%) of the theoretical differences between BoNT/A1 and /A2 or 76 of the 131 residues [12]. MS/MS identification of the 76 amino acid differences provided a clear distinction of BoNT/A1 from /A2.

Currently, subtype identification is determined through DNA sequencing of the toxin's genes [10]. Other DNA analysis techniques such as pulsed field gel electrophoresis

[14], randomly amplified polymorphic DNA analysis [15], amplified fragment-length polymorphism analysis [16], *flaA* variable-region sequencing [17], multilocus sequence typing [18], multiple locus variable-number tandem repeat analysis [19], and a comparative genomic hybridization microarray [20] have also been used to differentiate strains. All of these methods rely on the presence of bacterial DNA in the sample material. However, BoNT can be present in a clarified or filtered sample in which the bacterium is absent. In such a situation, toxin subtype identification or strain characterization would be difficult, but perhaps possible, using traditional DNA-based methods [19,21] as clarified samples can contain small amounts of DNA depending on the degree of purity. If DNA is not present, then these DNA-based methods cannot be used for subtype identification.

Our method to distinguish the two BoNT/A subtypes from each other relied upon searching a database of existing proteins, which does not allow de novo identification of a new subtype or toxin variant of BoNT. De novo identification of a new subtype or variant within a BoNT subtype is just as important, if not more so, than classification of a sample as an existing subtype. In this work, we demonstrate our ability to first distinguish BoNT/B1, /B2, /B3, /B4, and /B5 subtypes from each other using mass spectrometry and the ability to distinguish multiple toxin variants within BoNT/B1 and /B2 subtypes using a amino acid substitution database that includes all possible amino acid substitutions of BoNT/B1. Finally, we describe the use of the amino acid substitution database approach to identify a novel subtype of BoNT/B, designated here as BoNT/B7. The study combines the use of mass spectrometry with the amino acid substitution database to provide a powerful tool that can be used to screen samples for new BoNT/B toxin variants or subtypes with no DNA needed.

## Experimental procedures

### Materials

BoNT is highly toxic and requires appropriate safety measures. All neurotoxins were handled in a class 2 biosafety cabinet equipped with HEPA filters. Commercially purified BoNT/B1 complex toxin was purchased (Metabio, Madison, WI, USA). Dynabeads® Protein G were purchased (Invitrogen, Carlsbad, CA, USA) at 1.3 g/cm<sup>3</sup> in phosphate-buffered saline (PBS), pH 7.4, containing 0.1% Tween® 20 and 0.02% sodium azide. Sequencing-grade modified trypsin at 0.5 mg/mL in 50-mM acetic acid and sequencing grade chymotrypsin at 1 µg/µL in 50-mM ammonium bicarbonate was purchased (Roche, Pleasanton, CA, USA). All chemicals were from Sigma-Aldrich (St. Louis, MO, USA) except where indicated.

## Production of BoNT/B culture supernatants

Crude culture supernatants representing specific BoNT/B strains were produced by incubating subcultures of each strain for 5 days at 30–35 °C. After centrifugation, supernatants were removed and filtered through 0.22- $\mu$ m filters. The filtered supernatants were tested for upper limits of toxicity using the mouse bioassay, which indicated that the toxins were all present at concentrations of  $\leq 10$   $\mu$ g/mL. Information on the strains used in these studies is listed in Table 1.

## Neurotoxin extraction and digestion

Monoclonal antibodies 1B18 and B12.1 [22] which bind the heavy chain of the toxin obtained from the laboratory of Dr. James Marks at the University of California, San Francisco, were immobilized and crosslinked to the Dynabeads® Protein G using 30  $\mu$ g of antibody diluted into 500  $\mu$ L of PBS for every 100  $\mu$ L of Dynabeads® Protein G. Cross-linked IgG-coated Dynabeads® were stored in PBS-Tween buffer (PBS with 0.05% Tween® 20) at 4 °C for up to 2 days. For the BoNT extraction assay, an aliquot of 20  $\mu$ L of antibody-coated beads was mixed for 1 h with a solution of 200  $\mu$ L of each culture supernatant and 300  $\mu$ L of PBS with 0.01% Tween (PBST) buffer. After mixing for 1 h with constant agitation at room temperature, the beads were washed twice in 1 mL each of PBST and once in 100  $\mu$ L of water. The beads were reconstituted in 15  $\mu$ L of 50-mM ammonium bicarbonate, pH=7.5 (tryptic buffer) and 2  $\mu$ L of stock trypsin, and digested for 5 min at 52 °C<sup>23</sup>. Following digestion, the supernatant was then removed from the beads and 1  $\mu$ L of 10% TFA was added to the supernatant. A second, independent digestion (5 min at 52 °C) was performed by reconstitution of the beads in 15  $\mu$ L of tryptic buffer and addition of 1  $\mu$ L of chymotrypsin diluted to 0.2 mg/mL in

the tryptic buffer. After digestion, each supernatant was then removed from the beads and 1  $\mu$ L of 10% TFA was added. Each digested sample was analyzed separately.

## Mass spectrometric analysis of BoNT

After digestion, 5  $\mu$ L of the sample was injected onto a Waters NanoAcquity C18 5  $\mu$ m UPLC trap column (180  $\mu$ m ID and 2 cm long) and separated on a Waters NanoAcquity C18 1.7  $\mu$ m UPLC column, 100  $\mu$ m ID, and 10 cm long (Waters, Milford, MA, USA). After a trap-loading flow rate of 5  $\mu$ L/min for 5 min, the peptides were eluted from the column at a flow rate of 500 nL/min using the following gradient conditions, where A is water with 0.1% formic acid and B is 100% acetonitrile with 0.1% formic acid: A=95% and B=5% at 0 min; A=85% and B=15% at 10 min; A=65% and B=35% at 60 min; A=30% and B=70% at 62 min; hold for 5 min, A=95% and B=5% at 70 min; hold for 10 min.

Peptides eluting from the column were introduced into a 7-Tesla linear trap (LTQ)-FT-ICR instrument (Thermo Electron, San Jose, CA) using an Advion TriVersa NanoMate nanoelectrospray ion source interface equipped with an LC coupler (Advion, Ithaca, NY, USA). The instrument was operated in data-dependent acquisition mode to automatically switch between MS and MS/MS analysis. The FT-ICR-MS was used for MS acquisition with a resolution of 25,000 from 400 to 1,600  $m/z$ . Simultaneously, the linear ion trap was used for MS/MS analysis of the five most abundant ions in each MS scan. Automatic gain control was used to accumulate ions for FT-ICR-MS analysis with a target value of 1,000,000 and 200,000 for MS and MS/MS analysis, respectively. Collision energy of 35% was utilized in the linear ion trap to fragment the tryptic digest fragments.

## Construction of an amino acid substitution database

FASTA databases were generated using the proteolytic cleavage products of the sequence of BoNT/B1 as database entries, and single substitutions in every proteolytic cleavage product were included as an entry in the FASTA database (19 other possible amino acids per site). First, in silico digestions of BoNT/B1 based upon the known cleavage rules of trypsin and chymotrypsin, respectively, were completed and collated into cleavage product lists. Subsequently, each peptide was looped through one amino acid at a time, with each particular amino acid changed one at a time to one of the 19 other amino acids, with each changed peptide written to file as an entry in the NCBI FASTA format. Thus, using trypsin as the enzyme, the BoNT/B1 sequence generated 157 possible cleaved peptides, and with the change of one amino acid at a time, the trypsin FASTA database contained 23,899 theoretical sequences. Using chymotrypsin as the proteolytic

**Table 1** Strain information on culture supernatants used for this study

Sample	Strain	NCBI accession no.
B1	Okra	AB232927
B1	CDC 1656	EF028396
B1	CDC 1758	EF033127
B2	CDC 1828	EF051571
B2	Prevot 59	EF033128
B2	Prevot 25	EF033129
B3	CDC 795	EF028400
B4 (nonproteolytic)	Eklund 17B	EF051570
B5 (bivalent)	An436	EF028397
B7	Bac-04-07755	JQ354985
B7	NCTC3807	JN120760

enzyme, the algorithm generated 86 possible cleaved peptides, with 24,033 theoretical sequences. All machine code to generate the two FASTA databases was written in-house using Microsoft Visual C# 2008.

#### Searching of data against databases

Peak lists for all MS/MS spectra were extracted using Mascot Distiller (Matrix Science Inc., Boston, MA, USA, version 2.2.0) and merged together using Mascot Daemon (Matrix Science Inc, Boston, MA, USA, version 2.2.2). Database searches were conducted using Mascot (Matrix Science Inc, Boston, MA, USA, version 2.2.0) against a database generated by extracting entries from the NCBI non-redundant database using “Clostridia” as the extraction parameter. Mascot parameters were applied as follows for standard protein database searching: nine maximum missed cleavages, non-specific enzyme, peptides with up to +4 charge, MS1 tolerance of 20 ppm and MS/MS tolerance of 0.8 Da, and carbamidomethyl, deamidation, and oxidation modifications allowed. Mascot parameters were applied as follows for amino acid substitution database searching: nine maximum missed cleavages, semi-tryptic (K and R) and semi-chymotryptic enzyme selectivity, peptides with up to +4 charge, MS1 tolerance of 20 ppm and MS/MS tolerance of 0.5 Da, semi-specific searches included one non-specific cleavage site located on either the N or C terminus of the peptide. MS/MS-based peptide and protein identifications were validated using Scaffold (version Scaffold\_2\_00\_06, Proteome Software Inc., Portland, OR, USA). Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability as specified by the Protein Prophet algorithm, and if they contained at least three peptides with unique amino acid sequences.

#### DNA sequencing and phylogenetic analysis

Strains were grown anaerobically in tryptone–peptone–glucose–yeast broth and DNA was isolated as reported [23]. The *bont/b* gene was amplified and subjected to Sanger sequencing. DNA alignment of the obtained full-length sequences of the *bont/b* genes with already published sequences followed by phylogenetic analysis and construction of the dendrogram was carried out as described previously [10].

## Results

Peptides derived from the tryptic and chymotryptic digestion of BoNT/B1 Okra, B2 Prevot 25, BoNT/B3 CDC 795, BoNT/B4 Eklund 17B, and BoNT/B5 An436 were analyzed

by MS/MS with an LTQ-FT-ICR mass spectrometer and the data were compared against a standard protein database. All five BoNT proteins were correctly identified as BoNT/B1, B2, B3, B4, or B5 despite the high level of sequence similarity (as high as 98.4%). Table 2 is a list of peptides from the digests of these neurotoxins which were identified by MS/MS and are unique for each subtype. The overall percent coverage for each of the proteins was 76%, 76%, 66%, 75%, and 74%, respectively. Because these listed peptides are unique for each of the subtypes, they serve as biomarkers that identify each of the different BoNT/B subtypes.

Although these five BoNT proteins show as much as 98.4% identity, there are still many differences in the amino acid sequence among these toxins, and these differences can be exploited to distinguish toxins from each other by mass spectrometry. For example, there are 56 amino acids that differ between BoNT/B1 Okra and B2 Prevot 25. Upon MS examination of BoNT/B1 Okra and B2 Prevot 25, we obtained evidence for 34 of these amino acids or approximately 60% of the differing amino acid residues.

The 34 amino acid differences were identified by querying the MS/MS data against a protein database that included a protein sequence (BoNT/B2 Prevot 25) containing all 56 differing amino acids. However, protein sequences of any novel BoNT/B subtypes would not be present in the protein database and therefore, a unique new subtype of BoNT/B would simply be identified as the closest match that exists within the database. To optimize identification of novel subtypes, we developed an amino acid substitution database in which every amino acid within the BoNT/B1 Okra sequence was mutated in silico to 19 other possibilities. It is important to note that this amino acid substitution database was not intended to include all amino acid differences, as it does not account for more than one mutation within a tryptic or chymotryptic fragment, but rather suggests the presence of a new subtype or variant, and assists in the differentiation of toxins from botulism outbreaks. Although the database would certainly be more inclusive through the inclusion or more than one mutation within a tryptic or chymotryptic fragment, it would also require substantially more time to search against, thus lengthening the time frame for identification.

We first tested the use of the amino acid substitution database by digesting the BoNT/B2 Prevot 25 and treating it as an unknown while searching the data in the /B1 Okra amino acid substitution database. BoNT/B1 Okra and BoNT/B2 Prevot 25 are fairly similar (95.6%) with 56 amino acids that differ. Upon querying the data from the digests of BoNT/B2 Prevot 25 against the /B1 Okra amino acid substitution database, MS/MS data were found to support the identity of ten amino acid differences; the MS/MS spectra which demonstrate the presence of some of those

**Table 2** Peptides from the digests of BoNT/B1–/B5 which are unique for each subtype and were identified by MS/MS

Subtype	Peptide sequence
BoNT/B1 (Okra)	384NLLDNEIYTIEEGFNISDK <b>D</b> MEK <sup>406</sup>
	595QIV <b>N</b> DFVIEANK <sup>606</sup>
	692W <b>S</b> DMYGLIVAQWLSTVNTQFYTIK <sup>715</sup>
	721ALNYQAQALEEIIK <b>Y</b> R <sup>736</sup>
	759LNEGINQAIDNINNFING <b>G</b> CSVSYLMK <sup>784</sup>
	833T <b>I</b> MPFDLS <b>I</b> Y <sup>842</sup>
	843T <b>N</b> D <b>T</b> ILIE <b>M</b> FNK <sup>854</sup>
	869 <b>Y</b> KDNNLIDLSGY <sup>880</sup>
	1003FVTITNNL <b>N</b> NAK <sup>1014</sup>
	1021LES <b>N</b> T <b>D</b> IKDIR <sup>1031</sup>
	1057YFSIFNTELSQS <b>N</b> IEER <sup>1073</sup>
	1242FYESGIV <b>F</b> E <b>E</b> YK <sup>1253</sup>
	1029 <b>N</b> IGEVIANGEI <b>H</b> FK <sup>1042</sup>
BoNT/B2 (Prevot 25)	1162 <b>K</b> EDYIYLDFFNS <b>R</b> <sup>1175</sup>
	1179VYAYKDFKE <b>E</b> E <b>K</b> <sup>1190</sup>
	1242FYESGIVLKD <b>Y</b> K <b>N</b> YF <sup>1256</sup>
BoNT/B3 (CDC 795)	255FFMQSTADIQAEELYTFGGQ <b>D</b> PR <sup>277</sup>
	884VEV <b>Y</b> NGVELNDKNQ <b>F</b> K <sup>899</sup>
	910VTQ <b>N</b> Q <b>D</b> I <b>H</b> F <sup>918</sup>
BoNT/B4 (Eklund 17B)	254KFFMQSTD <b>T</b> IQAELYTFGGQDPSI <b>S</b> PSTDK <sup>285</sup>
	329FVEDSEGKYSIDVES <b>F</b> N <b>K</b> <sup>346</sup>
	350SLMFGFTE <b>I</b> NAEN <b>Y</b> K <sup>365</sup>
	384NLLDNEIYTIEEGFNISDK <b>N</b> MG <b>K</b> <sup>406</sup>
	441 <b>V</b> PGICIDVD <b>N</b> EN <b>L</b> FFIAD <b>K</b> <sup>459</sup>
	472 <b>V</b> EYNTQNNY <b>I</b> GNDFPINELILD <b>T</b> DLISK <sup>499</sup>
	529 <b>V</b> FTDENTIFQY <b>L</b> Y <sup>541</sup>
	542 <b>S</b> QTFPL <b>N</b> IR <sup>550</sup>
	623IGLALNV <b>G</b> D <b>E</b> TAK <sup>635</sup>
	737YNIYSE <b>E</b> EKSNNIN <b>F</b> NDINS <b>K</b> <sup>758</sup>
	759LNDGINQAM <b>D</b> NINDFINECSVSYLMK <sup>784</sup>
	814LYLIG <b>S</b> VE <b>D</b> E <b>K</b> <sup>824</sup>
	833T <b>H</b> IPFDLSTYTN <b>N</b> E <b>I</b> L <b>I</b> K <sup>850</sup>
	936 <b>Y</b> R <b>N</b> DD <b>I</b> QNYIHNEY <b>T</b> I <b>N</b> CMK <sup>956</sup>
	1015 <b>I</b> YING <b>T</b> LES <b>N</b> M <b>D</b> IK <sup>1028</sup>
	1029D <b>I</b> GEVIV <b>N</b> GE <b>I</b> T <b>F</b> K <sup>1042</sup>
	1042 <b>K</b> LDGD <b>V</b> DR <b>T</b> Q <b>F</b> <sup>1052</sup>
	1057YFSIFNT <b>Q</b> L <b>N</b> Q <b>S</b> NIK <sup>1071</sup>
	1111 <b>L</b> V <b>K</b> DSSVGE <b>I</b> L <b>R</b> <sup>1123</sup>
	1163 <b>E</b> DY <b>I</b> H <b>L</b> DFVNS <b>N</b> E <b>E</b> WR <sup>1178</sup>
1184 <b>N</b> FKE <b>Q</b> E <b>Q</b> K <sup>1191</sup>	
1192 <b>L</b> FL <b>S</b> I <b>I</b> YDSNEFYK <sup>1205</sup>	
1226 <b>K</b> DEESTD <b>D</b> IGLIGI <b>H</b> R <sup>1241</sup>	
1242FYESG <b>V</b> L <b>R</b> <sup>1249</sup>	
1272 <b>K</b> SNLGCN <b>W</b> Q <b>F</b> IPKDEGW <b>T</b> E <sup>1290</sup>	
BoNT/B5 (An436)	472 <b>I</b> AYNTQNNY <b>I</b> D <b>N</b> DFINELILD <b>T</b> DLISK <sup>499</sup>
	500 <b>I</b> ELPSENTESLTD <b>F</b> N <b>V</b> Y <b>V</b> PE <b>Y</b> K <sup>521</sup>
	677 <b>I</b> E <b>T</b> INSAL <b>T</b> K <sup>687</sup>
	746 <b>S</b> NINIDFND <b>V</b> NS <b>K</b> <sup>758</sup>

**Table 2** (continued)

Subtype	Peptide sequence
	794LLDFD <b>N</b> TL <b>R</b> <sup>802</sup>
	833 <b>T</b> S <b>I</b> PFDLST <b>Y</b> <sup>842</sup>
	855YNS <b>D</b> ILNNIILN <b>L</b> R <sup>868</sup>
	983 <b>S</b> VFFE <b>S</b> SI <b>K</b> <sup>991</sup>
	1032 <b>E</b> VIAN <b>D</b> E <b>I</b> IF <b>K</b> <sup>1042</sup>
	1043 <b>L</b> DG <b>N</b> IDRTQ <b>F</b> I <b>W</b> M <b>K</b> <sup>1056</sup>

Residues which make each peptide unique for the given subtype are bolded

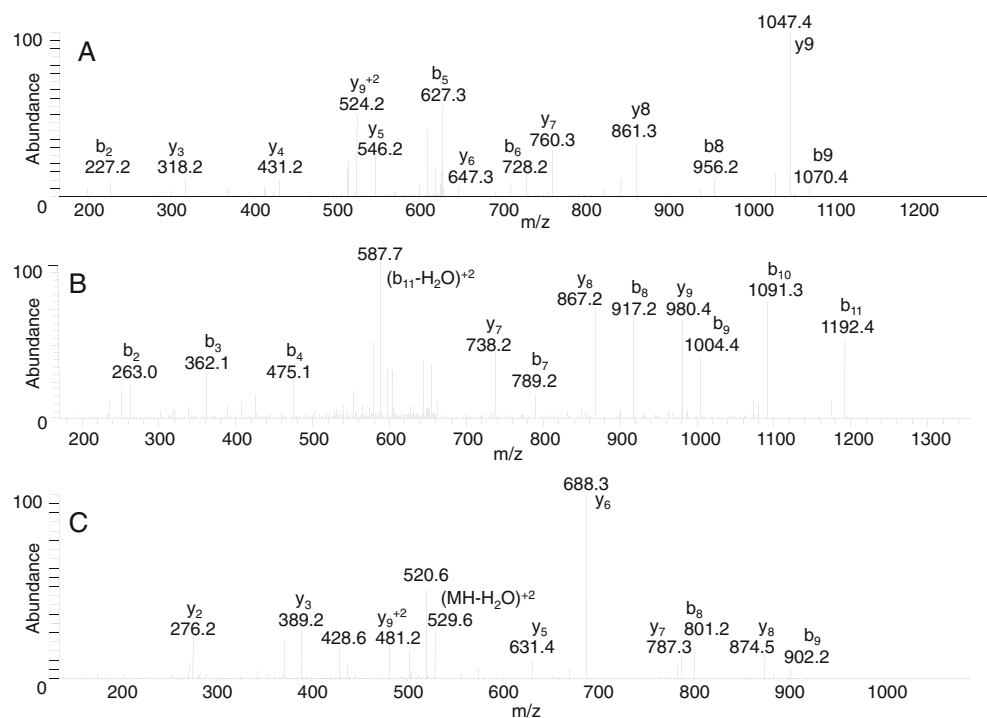
differences are shown in Fig. 1. The ten amino acids represent 18% (10 of 56) of the total known differences

Table 3 lists the peptides with amino acid differences that were discovered through the use of the amino acid substitution database. Figure 1a is the MS/MS of the sequence <sup>970</sup>IHWTLTDINGK<sup>980</sup> which includes the difference from I to T in position 975. In addition to the altered intact peptide ion at  $m/z$  637.36, the mass difference between  $y_5$  at  $m/z$  546.2 and  $y_6$  at  $m/z$  647.3 is 101 Da, which corresponds to a threonine rather than an isoleucine. Additionally, the mass difference between  $b_5$  at  $m/z$  627.3 and  $b_6$  at  $m/z$  728.2 is 101 Da, which corresponds to a threonine rather than an isoleucine. This MS/MS spectrum indicates the existence of the amino acid substitution from I to T in position 975.

Figure 1b is the MS/MS of the sequence <sup>599</sup>DFVIEANKSSTM<sup>610</sup> with the mutation from N to S in position 608. The altered intact peptide doubly charged ion at  $m/z$  671.32, and the mass difference of 87 Da between  $b_9$  at  $m/z$  1,004.4 and  $b_{10}$  at  $m/z$  1,091.3, demonstrate the existence of a serine rather than an asparagine. Figure 1c is the MS/MS of the sequence <sup>1114</sup>DSSVGEILTR<sup>1123</sup> which includes the mutation from P to S in position 1116. In addition to the altered intact peptide ion at  $m/z$  538.79, the mass difference between  $y_7$  at  $m/z$  787.3 and  $y_8$  at  $m/z$  874.5 is 87 Da, which corresponds to a serine rather than a proline; this confirms that this peptide has an amino acid substitution in position 1116. The discovery of these ten amino acids indicates that the amino acid substitution database can be used for de novo discovery of many of the differences present in the B2 Prevot 25 sequence. It also demonstrates that a previously unknown subtype of BoNT/B could be identified as “novel” through the use of the amino acid substitution database.

We then used this approach to distinguish BoNT/B1 toxin variants, whose amino acid sequences are more similar to BoNT/B1 Okra than that of a different BoNT/B subtype, like BoNT/B2 Prevot 25. Two strains identified as producing BoNT/B1, but whose protein sequences were not

**Fig. 1** MS/MS spectra of the precursor ions at  $m/z$  637.36 (IA), 671.32 (IB), and 538.79 (IC) corresponding to doubly-charged peptides IIWTLTDINGK (IA), DFVIEANKSSTM (IB), and DSSVGEILTR (IC) from the digests of BoNT/B2. The spectra were obtained by LC-MS/MS on an FT-ICR mass spectrometer, with MS analysis in the FT-ICR cell and MS/MS analyses within the LTQ portion. Residues which are set in *bold* indicate amino acid substitutions



entered into the protein database, were tested. The BoNT/B1 strain CDC 1656 has nine amino acid differences compared to BoNT/B1 strain Okra, and two of the nine differences were discovered through the use of the amino acid substitution database, as seen in Table 4. Strain CDC 1758 has only three amino acid differences for a variation of less than 0.3%, and one of those differences was discovered through the use of the amino acid substitution database as seen in Table 4.

Two additional strains expressing BoNT/B2, with the sequence of their *bont* genes confirmed by full-length sequencing, but whose protein sequences were not entered into the protein database, were also analyzed using this approach. BoNT/B2 strain CDC 1828 has 58 amino acid differences compared to BoNT/B1 strain Okra, and, as indicated in Table 5, seven of those differences were discovered through the use of the amino acid substitution

database. This corresponds to 23% of the total number of peptides which have a single amino acid difference. BoNT/B2 strain Prevot 59 has 56 amino acid differences, and as indicated in Table 5, seven of those differences were discovered by querying the amino acid substitution database, corresponding to 26% of the total number of peptides which have a single difference.

As a final demonstration of our ability to identify a BoNT/B sample as “novel”, we analyzed several samples which were determined to contain BoNT/B after testing by mouse bioassay and Endopep-MS [22]. However, these samples had not yet been subjected to DNA analysis, so there was no subtype or strain identification on the samples. After digestion, MS/MS analysis, and searching of the standard protein database, one of the samples was tentatively identified as BoNT/B2 L-590 with 48% sequence coverage, with the chromatogram shown in Fig. 2. After searching the

**Table 3** Peptides from the digest of both BoNT/B1 and /B2 with amino acid differences that were discovered through the use of the amino acid substitution database

Amino acid position	Difference (B2 Prevot 25)	Normal (B1 Okra)
472-484	IEYNTQSNYIEN <b>R</b>	IEYNTQSNYIEN <b>D</b>
599-610	DFVIEANKS <b>S</b> TM	DFVIEANKS <b>N</b> TM
850-868	E <b>I</b> FNKYNSEILNNIILNLR	E <b>M</b> FNKYNSEILNNIILNLR
876-887	DLSGYGAN <b>V</b> EVY	DLSGYGA <b>K</b> EVY
910-923	VTQNQNIIFNS <b>M</b> FL	VTQNQNIIFNS <b>V</b> FL
970-980	IIWTL <b>T</b> DINGK	IIWTL <b>I</b> DINGK
976-992	DINGKTKSVFFEYS <b>I</b> RE	DINGKTKSVFFEY <b>N</b> IRE
1057-1071	YFSIFNTELSQS <b>N</b> I <b>K</b>	YFSIFNTELSQS <b>N</b> I <b>E</b>
1072-1084	E <b>I</b> YKIQSYSEYLK	E <b>R</b> YKIQSYSEYLK
1114-1123	D <b>S</b> S <b>V</b> GEILTR	D <b>S</b> <b>P</b> VGEILTR

**Table 4** Locations and type of mass-spectrometric identified amino acid differences in two additional strains of BoNT/B2 as compared to BoNT/B1 Okra

B1 Okra residue	B1 CDC 1656 residue	B1 CDC 1758 residue	B1 Okra peptide identified by MS/MS	B1 CDC 1656 peptide identified by MS/MS	B1 CDC 1758 peptide identified by MS/MS
70C	70W	70C	DV <u>C</u> EYYDPDYLN <u>T</u> NDKK	DV <u>W</u> EYYDPDYLN <u>T</u> NDK	DV <u>C</u> EYYDPDYLN <u>T</u> NDK
1250E	1250K	1250K	FY <u>E</u> SGIVF <u>E</u> EYK	FY <u>E</u> SGIVF <u>K</u> EYK	FY <u>E</u> SGIVF <u>K</u>

MS/MS data from this sample against the amino acid substitution database, five differences compared to BoNT/B2 L-590 were found, indicating that this protein is not BoNT/B2 L-590. Specifically, these five differences were found to be V597I/L, S746I/L, L877S, S1010I/L, and A1035V, and MS/MS evidence for some of these differences are shown in Fig. 3. Before this work, amino acid positions 597, 746, and 877 were not known to have differences in any strain or subtype of BoNT/B, so these differences are completely new for BoNT/B.

Figure 3a is the MS/MS of the sequence <sup>595</sup>QIIDDVIEANK<sup>606</sup> which includes the mutation from V to I in position 597. In addition to the altered intact peptide ion at *m/z* 702.9, the mass difference between *y*<sub>9</sub> at *m/z* 1,050.6 and *y*<sub>10</sub> at *m/z* 1,163.6 is 113 Da, which corresponds to a leucine or isoleucine rather than a valine. Residue 1010, an I or L in this sample, is an L in some BoNT/B1 and B4, but this difference is new for any BoNT/B2 where this residue is typically a serine. Figure 3b is the MS/MS of the sequence <sup>1001</sup>WFFVTITNLDNAK<sup>1014</sup> which includes the mutation from S to L in position 1010. The intact peptide ion at *m/z* 841.93 is different from *m/z* 828.90, which appears in all other BoNT/B2. The mass difference between *y*<sub>4</sub> at *m/z* 447.4 and *y*<sub>5</sub> at *m/z* 560.4 is 113 Da, which corresponds to a leucine or isoleucine. Furthermore, the mass difference between *b*<sub>9</sub> at *m/z* 1,123.6 and *b*<sub>10</sub> at *m/z* 1,236.5 is 113 Da, which corresponds to a leucine or isoleucine. Residue 746 is an S in all other BoNT subtypes and was not known until now to be mutated. Similarly, residue 877 is an L in all other

BoNT/B subtypes. Residue 1035 is a V only in BoNT/B4 subtypes, and this difference is new for any BoNT/B2.

After discovering these novel point mutations in the BoNT/B sample, the *bont/b* gene was sequenced confirming the above point mutations as well as 45 other amino acid differences, defining this sample as a new subtype of BoNT/B, BoNT/B7. After adding this new BoNT/B sequence to our standard database and searching the MS/MS data against the updated database, the top hit was the new BoNT/B7 Bac-04-07755 strain. The percent coverage increased from 48% coverage with BoNT/B2 L-590 to 68% coverage with BoNT/B7 Bac-04-07755 as seen in Fig. 4, further verifying that this is a more accurate identification of the BoNT/B in this sample.

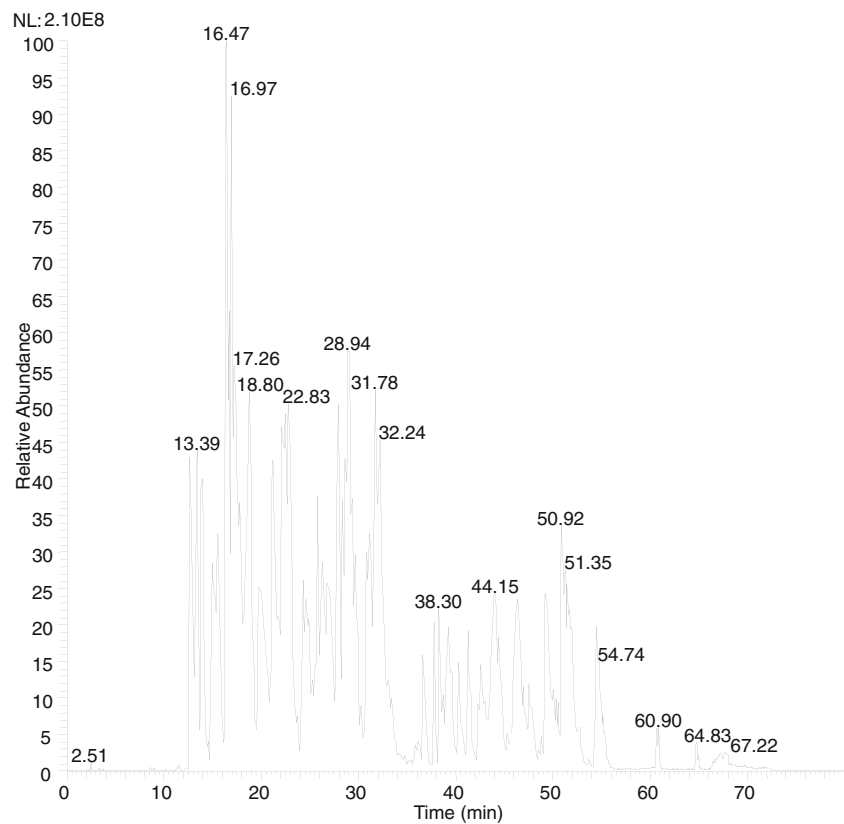
## Discussion

Our experiments demonstrate the ability to identify the subtype (B1–B5, B7) of a BoNT/B using both the MS/MS with a standard protein database or with our amino acid substitution database. The combination of the two methods also can identify the subtype, a toxin variant within a subtype, or identify a unique subtype, not previously discovered. In previous work, we demonstrated that BoNT subtypes could be distinguished from each other by MS [12,13]. While Kull et al. used MALDI-TOF MS to differentiate between BoNT/B1 and B4, we previously analyzed the tryptic peptides of BoNT/A1 and A2 by LC-MS/MS.

**Table 5** Locations and type of mass spectrometric identified amino acid differences in two additional strains of BoNT/B2 as compared to BoNT/B1 Okra

B1 Okra residue	B2 CDC 1828 residue	B2 Prevot 59 residue	B2 CDC 1828 peptide identified by MS/MS	B2 Prevot 59 peptide identified by MS/MS
608N	608S	608S	DFVIEANKS <u>S</u> TM	DFVIEANKS <u>S</u> TM
851M	851I	851I		E <u>I</u> FNKYNSEILNNIILNLR
870K	870R	870R	YNSEILNNIILNLR <u>Y</u> R	
883K	883N	883N	DLSGYGAN <u>N</u> VEVY	
921V	921M	921M	VTQNQNIIFNS <u>M</u> FL	VTQNQNIIFNS <u>M</u> FL
975I	975T	975T	IIWTL <u>T</u> DINGK	
1031R	1031G	1031G		D <u>I</u> G <sup>+</sup> EVIANG <sup>+</sup> EIIFK
1071E	1071K	1071K	SIFNTELSQS <u>N</u> IK	YFSIFNTELSQS <u>N</u> IK
1073R	1073I	1073I		E <u>I</u> YKIQSYSEYLK
1116P	1116S	1116S	D <u>S</u> SVGEILTR	D <u>S</u> SVGEILTR

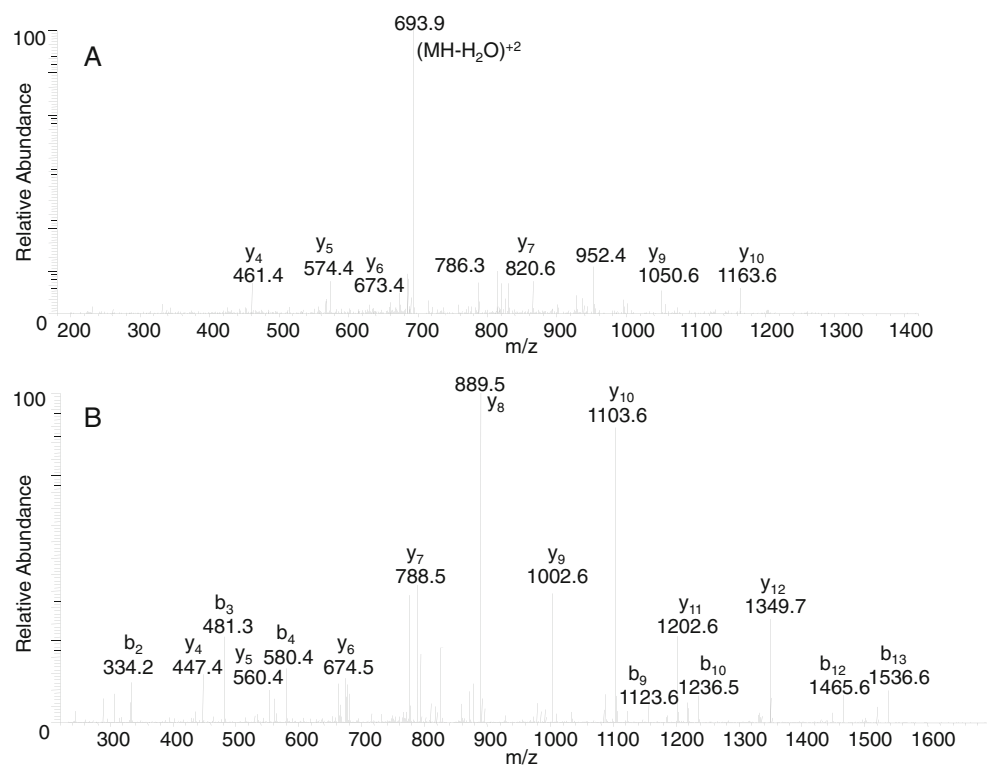
**Fig. 2** Chromatogram of the tryptic digest of a BoNT/B obtained by LC-MS on an FT-ICR mass spectrometer



Although these two A subtypes are approximately 90% homologous, we demonstrated that we could exploit the 10% of the protein that was unique, to identify the toxin as

BoNT/A1 or /A2. The amino acid sequences of BoNT/B1–B5 are more conserved than A1 and A2. They differ by as little as 1.6% (in the case of BoNT/B2 and /B3), while the

**Fig. 3** MS/MS spectra of precursor ions at  $m/z$  702.89 (3A) and 841.93 (3B) corresponding to doubly charged peptides QIIDDFVIEANK (3A) and WFFVTITNLDNAK (3B) from the digests of BoNT/B7 Bac-04-07755. The spectra were obtained by LC-MS/MS on an FT-ICR mass spectrometer, with MS analysis in the FT-ICR cell and MS/MS analyses within the LTQ portion. Residues which are set in *bold* indicate amino acid substitutions





PVTINNFNYN DPIDNNIIM MEPPFARGTG RYYKAFKITD RIWIIPERYT  
 FGYPEDFNK SSGIFNRDVC EYDPPDYLNT NDKKNIFLQT MIKLFNRIKS  
 KPLGEKLEEM IINGIPYIGD RRVPLEEFNT NIASVTVNKL ISNPGEVERK  
 KGIFANLIIF GPGPVLNENE TIDIGIQNHF ASREGFGGIM QMKFCPEYVS  
 VFNVQENKG ASIFNRRGYF SDPALILMHE LIHVLHGLYG IKVDDLPIVP  
 NEKKFFMQST DAIQAEELYT FGGQDPSIIT PSTDKSIYDK VLQNFRRGIVD  
 RLNKVLVLCIS DPNININIYK NKFKDKYKVF EDSEGKYSID VESFDKLYKS  
 LMFGFETNI AENYKIKTRA SYFSDSLPPV KIKNLLDNEI YTIEEGFNIS  
 DKDMEKEYRG QNKAINQAY EEISKEHLAV YKIQMCKSVK APGICIDVDN  
 EDLFFIADKN SFSDDLKNE RIEYNTKNIY IENYFSINEL ILDTDLISGI  
 ELPSENTESE TDFNVDVVPV EKQPAIKKIF TDENTIFQYL YSQTFLDIR  
 DISLTSSFFD ALLFSNKVYS FFSMDYIKTA NKVVEAGLFA GWVQIIDDV  
 VIEANKSSTM DKIAISLIV PYIGLALNVG NETAKGNFEN AFEIAGASIL  
 LEFIPPELLIP VVGAFLLSEY IDNKNKI IKT IDNALTKRVE KWIDMYGLIV  
 AQWLSTVNTQ FYTIKEGMYK ALNYQQALE EIIKYKYNII SEKEKLNINI  
 DFNIDNSKLN EGINQAIDNI NNFINECSVS YLMKKMIPLA IEKLLDFDNA  
 LKKNLLNYID ENKLYLIGSV EEEKSKVDKF FKTIIPFDLS MYTNTTILIE  
 MVNKYNSLIL NNIILNLRYS DNNLIDSSGY GAKVEVYNGV ELNDKNQFKL  
 TSSANSKIKV TQNQNITFNS MFLDFSVSFW IRI PKYKNDG IQNYIHNEYT  
 IINCMMNSG WKISIRGNRI IWTLDINGK TKSVFVEYSI REDISDYINR  
 WFFVTITNNL DNAKIYINGK LESNIDIRDI REVIVNGEII FKLDEGIDRT  
 QFIWMKYPSI FNTELSQSNV KEIYKIQSYS KYLKDFWGNP LMYNKEYYMA  
 NAGNNSYIK LVKDSSVGEI LTRSKYQNS NYINYRNLYI GEKPIIRKKS  
 SSQSISDDIV RKEDYIYLDV FNSNREWRVY AYKNFKGQEE KLFLANIYDS  
 NEFYKTIQIK EYDEQPTYSC QLLFKKDEES TDEIGLIGIH NPYESGILFK  
 DYKDYFCISK WYLKEVKKKP YSSNLGCNWQ FIPKDEGWTE

**Fig. 4** Amino acid sequence of BoNT/B7 Bac-04-07755. Residues in red comprise the 68% sequence coverage for which MS/MS evidence was obtained

difference between /A1 and /A2 is 10.1%. Although these five BoNT/B subtypes are very similar, they still differ by as few as 21 amino acids, which was sufficient for subtype identification through enzymatic digestion and subsequent mass spectrometric analysis.

The use of two different enzymes (trypsin and chymotrypsin) to digest the samples contributes to the high percent coverage of these proteins with a mass above 150 kDa. A high percentage of coverage is critical for novel BoNT subtype and toxin variant identification as with increasing coverage (amino acid sequence for which there is mass spectrometric evidence), there is a higher likelihood of identifying more amino acid differences. Through the use of the combined enzymatic digestions and LC-MS/MS analysis of the resultant peptides, we were able to obtain 77%, 76%, 66%, 75%, and 74% amino acid coverage of BoNT/B1, /B2, /B3, /B4, and /B5, respectively. This high sequence coverage allowed us to use mass spectrometry to distinguish each of these highly similar proteins from each other, and in the case of /B1 and /B2, to identify 34 of the 56 amino acids which differed between these two proteins. Although less than 40% of the differing residues could not be detected by mass spectrometry; nonetheless, the identification of 34 amino acid differences provides an effective method allowing for a definitive identification of BoNT/B as either the /B1 or /B2 subtype. Furthermore, these studies represent the first protein sequencing method applied to BoNT/B2, B3, B4, and /B5.

BoNT/B5 is a bivalent toxin producer, producing both BoNT/B5 and a lesser amount of BoNT/F or A. Due to the

specificity of the capture antibodies, we can examine bivalent toxin strains without interference from the second toxin. The detection of two toxins in a preparation can be difficult for some other identification methods. Identification of the toxin as a /B5 through proteomic analysis might indicate that there is a second toxin present which could have an effect on botulism treatment. Additionally, the /B4 subtype is associated with nonproteolytic BoNT/B strains. These strains are not always reliably detected using other methods, but the proteomic technique discussed here can easily identify BoNT/B4.

If the amino acid sequence of BoNT/B2 Prevot 25 had not been in the protein database, as would be expected for an unknown subtype or toxin variant of BoNT/B, we still would have identified the protein as discrete from that of BoNT/B1 Okra through the use of the amino acid substitution database. The creation of the database allowed the “de novo” identification of ten amino acid differences. Although not all of the theoretical differences were discovered using this method, there were enough differences to assign the identity of BoNT/B2 as similar to BoNT/B1, but not as BoNT/B1—a more precise identification than the de facto assignment of BoNT/B1. Since our goal is not to identify all possible amino acid differences, but identify a BoNT protein as an existing subtype or a novel toxin variant and to identify differences in toxins from ongoing botulism outbreaks; these results demonstrate that this goal can be achieved through the use of this technique.

To test the ability of the method to detect differences within subtypes (<1.4% amino acid variation which is ~21 amino acids), we obtained two BoNT/B1 samples whose amino acid sequences were not in the protein database, digested those proteins, and analyzed the resultant fragments by LC-MS/MS. Searching the data against an amino acid substitution database of BoNT/B1 Okra allowed for identification of two of nine possible differences in one strain and one of three possible differences in another strain which is sufficient to indicate differences in outbreak samples.

The identities of residues 70 and 1250 are critical for identification of these strains as three separate proteins. According to data listed in Table 4, BoNT/B1 Okra contains 70C and 1250E, BoNT/B1 CDC1656 contains 70W and 1250K, and BoNT/B1 CDC 1758 contains 70C and 1250K. Because each of these samples has a unique combination of identities of residues 70 and 1250, these three proteins can be identified as three separate protein identities. Again, although not all differences were discovered through this technique, the goal of strain differentiation of BoNT/B1 by mass spectrometry was achieved, as these three samples were identified as having very similar yet distinctly unique protein sequences. Additionally, this work reports the first protein comparison of multiple strains of the same BoNT subtype by MS techniques.

Strain variance exists in most of the BoNT subtypes, and BoNT/B2 is not an exception. We obtained two BoNT/B2 samples whose sequences were not in the protein database, digested those samples, and the resultant peptides were analyzed by LC-MS/MS. The data were then queried against the amino acid substitution database generated from the BoNT/B1 Okra sequence. Seven to ten of the possible 56–58 amino acid differences were discovered using this technique. Note that searching the data through the amino acid substitution database yielded one to two amino acid differences for toxin variants (BoNT/B1) and seven to ten amino acid differences for subtype (BoNT/B2) differences. Therefore, this technique could allow for a tentative identification of a novel protein or a new BoNT/B subtype or toxin variant, depending on the number of amino acid differences discovered through this technique.

In fact, the number of amino acid differences was critical toward our tentative identification of the BoNT/B7 sample as a new subtype rather than a new BoNT/B2 strain/toxin variant, as this sample contained five novel amino acid differences, or greater than the number of one to two that we saw for strain differentiation. BoNT/B7 strain Bac-04-07755 originates from a recent infant botulism case in New York. Interestingly, the identical novel subtype BoNT/B7 was independently discovered by BGD and co-workers by “classical” DNA sequencing of strain NCTC 3807 (National Collection of Type Cultures, UK). Strain NCTC 3807 was originally isolated in 1921 from soil from the Blue Ridge Mountains of Virginia and later deposited into the NCTC strain collection [24], meaning that this “old” strain is still circulating within the USA. Noteworthy, this new subtype BoNT/B7 has been reported to differ from other BoNT/B molecules, as the toxin was not recognized by a mAb in a specific ELISA [5], again justifying that this toxin should be named a novel BoNT/B subtype.

Indeed, DNA sequencing and subsequent phylogenetic analysis [10] showed that strains Bac-04-07755 and NCTC 3807 form a separate cluster, here designated as BoNT/B7, adjacent to the BoNT/B4 cluster as seen in the dendrogram in Fig. 5.

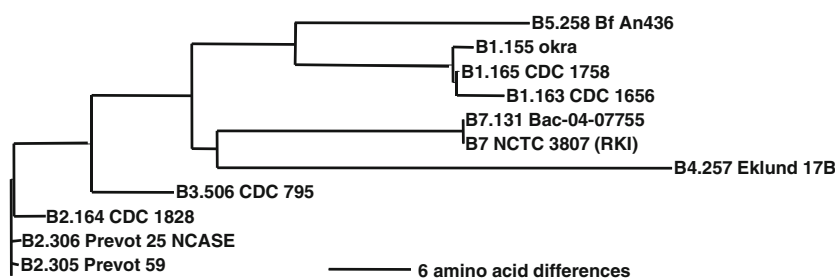
Although the dendrogram shows that its closest relative is BoNT/B4, searching the MS/MS data of BoNT/B7 against the known protein database yielded BoNT/B2 as the most

closely related protein. This is simply because unlike DNA sequencing, the mass spectrometer did not identify 100% of the protein and as a result could not be used solely to determine evolutionary relatedness. The portion of the protein identified by mass spectrometry was most similar to BoNT/B2. This is not problematic as the goal of this work is not to identify the closest neighbor of the toxin, but rather to identify the toxin as a new subtype or toxin variant.

Because our technique identified single-point mutations within a peptide, conservative data interpretation is a key component of correct identification. Mass spectrometry data do not always provide evidence for every amino acid within a peptide, so a mass spectrum could be misinterpreted and provide misleading results. Specifically, it is not enough for the intact peptide mass and the fragmentation pattern to match a new peptide; it is critical that product ions be obtained to prove the identity of the single mutated residue. As an example, it was initially suspected that the identity of a peptide at  $m/z$  888.94 was the doubly charged peptide YFSIFNTELSQSVIE, with a single amino acid substitution from the sequence of YFSIFNTELSQSNIE. Product ions from  $y_4$  to  $y_{11}$  appeared to confirm this finding. However, with the absence of the  $y_2$  and  $y_3$  or  $b_{12}$  and  $b_{13}$  ions, this finding was not confirmed. After PCR sequencing, this peptide was found to have the identity of YFSIFNTELSQNVK, or two differences within the peptide. This peptide also produced an  $m/z$  of 888.94 and the same  $y_4$  to  $y_{11}$  product ions as produced with the peptide sequence YFSIFNTELSQSVIE. Because our amino acid substitution database only supplied single-point mutations within a peptide, the correct possibility was not considered, and a less conservative data analysis would have produced a misleading result.

By examining the tryptic and chymotryptic digests of BoNT/B by mass spectrometry, a toxin can be first identified as BoNT/B and then further identified as subtype /B1, /B2, /B3, /B4, or /B5 similar to previous experiments which identified BoNT/A as A1 or A2 [12]. Additionally, we have demonstrated that novel BoNT/B subtypes or toxin variants could be reported using the point mutation search approach, rather than incorrectly identified as the most similar protein match in the database of known protein sequences. Although the use of an amino acid substitution database does

**Fig. 5** Dendrogram of proteins used in this study. Because the horizontal line represents the distance for six amino acids, the closer the subtypes in distance the fewer differences in amino acid between them



not identify all of the amino acid differences present in a novel toxin variant or subtype of BoNT/B, it identifies enough of the amino acid differences to enable classification of a toxin as a new BoNT/B subtype or even a new toxin variant. Such information is especially important for forensics and epidemiological purposes in order to identify the possible source of an outbreak or the spread of BoNT-producing *Clostridium* species.

Because this information can be obtained without DNA, it is possible to identify BoNT/B to or below the subtype level using protein-based analysis. This technique is, to our knowledge, the only protein-based technique for BoNT subtyping and toxin variant identification, providing a relatively rapid and accurate protein subtyping tool for analysis of botulism samples. Additionally, such information can be obtained in a few days rather than a few weeks, making it a valuable tool for epidemiologists who are tracking botulism outbreaks on a real-time basis. Currently, the amino acid substitution database is searched manually, and one future goal is to automate this process. We are optimistic that automation of data searching would allow for novel BoNT subtype and toxin variant identification within 1 day as sample preparation and data acquisition occurs in only a few hours.

**Acknowledgments** This research was partially supported through Interagency Agreement B18-120 from the National Institute of Allergy And Infectious Diseases. The opinions, interpretations, and recommendations are those of the authors and are not necessarily those of the Centers for Disease Control and Prevention, the US Army, the Robert Koch-Institut, the National Institute of Allergy and Infectious Diseases, or the National Institutes of Health.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Arnon SS, Schechter R, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Hauer J, Layton M, Lillibridge S, Osterholm MT, O'Toole T, Parker G, Perl TM, Russell PK, Swerdlow DL, Tonat K (2001) Botulinum toxin as a biological weapon: medical and public health management. *JAMA* 285(8):1059–1070
2. Centers for Disease Control and Prevention (1998) Botulism in the United States, 1899–1996. Handbook for epidemiologists, clinicians, and laboratory workers. Atlanta, GA
3. Wagner E, Meyer KF, Dozier CC (1925) Studies on the metabolism of *B. Botulinus* in various media. *Xxvi. J Bacteriol* 10(4):321–412
4. Kozaki S, Kamata Y, Nishiki T, Kakinuma H, Maruyama H, Takahashi H, Karasawa T, Yamakawa K, Nakamura S (1998) Characterization of *Clostridium botulinum* type B neurotoxin associated with infant botulism in Japan. *Infect Immun* 66(10):4811–4816
5. Gibson AM, Modi NK, Roberts TA, Hambleton P, Melling J (1988) Evaluation of a monoclonal antibody-based immunoassay for detecting type B *Clostridium botulinum* toxin produced in pure culture and an inoculated model cured meat system. *J Appl Bacteriol* 64(4):285–291
6. Kozaki S, Nakaue S, Kamata Y (1995) Immunological characterization of the neurotoxin produced by *Clostridium botulinum* type A associated with infant botulism in Japan. *Microbiol Immunol* 39(10):767–774
7. Smith TJ, Lou J, Geren IN, Forsyth CM, Tsai R, Laporte SL, Tepp WH, Bradshaw M, Johnson EA, Smith LA, Marks JD (2005) Sequence variation within botulinum neurotoxin serotypes impacts antibody binding and neutralization. *Infect Immun* 73(9):5450–5457
8. Carter AT, Paul CJ, Mason DR, Twine SM, Alston MJ, Logan SM, Austin JW, Peck MW (2009) Independent evolution of neurotoxin and flagellar genetic loci in proteolytic *Clostridium botulinum*. *BMC Genomics* 10:115
9. Chen S, Kim JJ, Barbieri JT (2007) Mechanism of substrate recognition by botulinum neurotoxin serotype A. *J Biol Chem* 282(13):9621–9627
10. Hill KK, Smith TJ, Helma CH, Ticknor LO, Foley BT, Svensson RT, Brown JL, Johnson EA, Smith LA, Okinaka RT, Jackson PJ, Marks JD (2007) Genetic diversity among botulinum neurotoxin-producing clostridial strains. *J Bacteriol* 189(3):818–832
11. Umeda K, Seto Y, Kohda T, Mukamoto M, Kozaki S (2009) Genetic characterization of *Clostridium botulinum* associated with type B infant botulism in Japan. *J Clin Microbiol* 47(9):2720–2728
12. Kalb SR, Goodnough MC, Malizio CJ, Pirkle JL, Barr JR (2005) Detection of botulinum neurotoxin A in a spiked milk sample with subtype identification through toxin proteomics. *Anal Chem* 77(19):6140–6146
13. Kull S, Pauly D, Stormann B, Kirchner S, Stammer M, Dörner MB, Lasch P, Naumann D, Dörner BG (2010) Multiplex detection of microbial and plant toxins by immunoaffinity enrichment and matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* 82(7):2916–2924
14. Lin WJ, Johnson EA (1995) Genome analysis of *Clostridium botulinum* type A by pulsed-field gel electrophoresis. *Appl Environ Microbiol* 61(12):4441–4447
15. Hielm S, Björkroth J, Hyttia E, Korkeala H (1999) Ribotyping as an identification tool for *Clostridium botulinum* strains causing human botulism. *Int J Food Microbiol* 47(1–2):121–131
16. Keto-Timonen R, Nevas M, Korkeala H (2005) Efficient DNA fingerprinting of *Clostridium botulinum* types A, B, E, and F by amplified fragment length polymorphism analysis. *Appl Environ Microbiol* 71(3):1148–1154
17. Paul CJ, Twine SM, Tam KJ, Mullen JA, Kelly JF, Austin JW, Logan SM (2007) Flagellin diversity in *Clostridium botulinum* groups I and II: a new strategy for strain identification. *Appl Environ Microbiol* 73(9):2963–2975
18. Jacobson MJ, Lin G, Whittam TS, Johnson EA (2008) Phylogenetic analysis of *Clostridium botulinum* type A by multi-locus sequence typing. *Microbiology* 154(Pt 8):2408–2415
19. Macdonald TE, Helma CH, Ticknor LO, Jackson PJ, Okinaka RT, Smith LA, Smith TJ, Hill KK (2008) Differentiation of *Clostridium botulinum* serotype A strains by multiple-locus variable-number tandem-repeat analysis. *Appl Environ Microbiol* 74(3):875–882
20. Raphael BH, Joseph LA, McCroskey LM, Luquez C, Maslanka SE (2010) Detection and differentiation of *Clostridium botulinum* type A strains using a focused DNA microarray. *Mol Cell Probes* 24(3):146–153
21. De Medici D, Anniballi F, Wyatt GM, Lindstrom M, Messelhauser U, Aldus CF, Delibato E, Korkeala H, Peck MW, Fenicia L (2009) Multiplex PCR for detection of botulinum neurotoxin-producing clostridia in clinical, food, and environmental samples. *Appl Environ Microbiol* 75(20):6457–6461

22. Kalb SR, Smith TJ, Moura H, Hill K, Lou JL, Geren IN, Garcia-Rodriguez C, Marks JD, Smith LA, Pirkle JL, Barr JR (2008) The use of Endopep-MS to detect multiple subtypes of botulinum neurotoxins A, B, E, and F. *Int J Mass Spectrom* 278(2–3):101–108
23. Kirchner S, Kramer KM, Schulze M, Pauly D, Jacob D, Gessler F, Nitsche A, Dörner BG, Dörner MB (2010) Pentaplexed quantitative real-time PCR assay for the simultaneous detection and quantification of botulinum neurotoxin-producing clostridia in food and clinical samples. *Appl Environ Microbiol* 76(13):4387–4395
24. Meyer KF, Dubovsky BJ (1922) The distribution of the spores of *B. botulinus* in the United States. IV. *J Infect Dis* 31:559–594