APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Aug. 1992, p. 2345–2354 Vol. 58, No. 8
0099-2240/92/082345-10$02.00/0

Molecular Cloning of the *Clostridium botulinum* Structural Gene Encoding the Type B Neurotoxin and Determination of Its Entire Nucleotide Sequence

SARAH M. WHELAN, MICHAEL J. ELMORE, NICOLA J. BODSWORTH, JOHN K. BREHM, TONY ATKINSON, AND NIGEL P. MINTON*

Molecular Genetics Group, Division of Biotechnology, PHLS Centre for Applied Microbiology and Research, Porton Down, Salisbury, Wiltshire SP4 0JG, United Kingdom

Received 11 February 1992/Accepted 5 May 1992

DNA fragments derived from the *Clostridium botulinum* type A neurotoxin (BoNT/A) gene (*botA*) were used in DNA-DNA hybridization reactions to derive a restriction map of the region of the *C. botulinum* type B strain Danish chromosome encoding *botB*. As the one probe encoded part of the BoNT/A heavy (H) chain and the other encoded part of the light (L) chain, the position and orientation of *botB* relative to this map were established. The temperature at which hybridization occurred indicated that a higher degree of DNA homology occurred between the two genes in the H-chain-encoding region. By using the derived restriction map data, a 2.1-kb *Bgl*II-*Xba*I fragment encoding the entire BoNT/B L chain and 108 amino acids of the H chain was cloned and characterized by nucleotide sequencing. A contiguous 1.8-kb *Xba*I fragment encoding a further 623 amino acids of the H chain was also cloned. The 3′ end of the gene was obtained by cloning a 1.6-kb fragment amplified from genomic DNA by inverse polymerase chain reaction. Translation of the nucleotide sequence derived from all three clones demonstrated that BoNT/B was composed of 1,291 amino acids. Comparative alignment of its sequence with all currently characterized BoNTs (A, C, D, and E) and tetanus toxin (TeTx) showed that a wide variation in percent homology occurred dependent on which component of the dichain was compared. Thus, the L chain of BoNT/B exhibits the greatest degree of homology (50% identity) with the TeTx L chain, whereas its H chain is most homologous (48% identity) with the BoNT/A H chain. Overall, the six neurotoxins were shown to be composed of highly conserved amino acid domains interceded with amino acid tracts exhibiting little overall similarity. In total, 68 amino acids of an average of 442 are absolutely conserved between L chains and 110 of 845 amino acids are conserved between H chains. Conservation of Trp residues (one in the L chain and nine in the H chain) was particularly striking. The most divergent region corresponds to the extreme carboxy terminus of each toxin, which may reflect differences in specificity of binding to neurone acceptor sites.

Botulinum neurotoxin (BoNT) and tetanus toxin (TeTx) are high-molecular-weight proteins (approximately 150,000 Da) which exert potent neuroparalytic effects on vertebrates (14, 42). They are elaborated by anaerobic gram-positive bacteria belonging to the genus *Clostridium*. TeTx is synthesized by *Clostridium tetani*, whereas the majority of clostridia which produce BoNT are classified as *C. botulinum*. In recent years, however, isolates which resemble *C. barati* and *C. butyricum* have been shown to produce BoNT (15, 24). On the basis of antigenicity, BoNT has been subdivided into seven distinct types, designated A to G. All eight neurotoxins (BoNT/A to BoNT/G and TeTx) are synthesized as a single-chain, 150,000-Da molecule which subsequently becomes nicked to the more potent dichain form, composed of a heavy (H) (approximately 100,000 Da)- and a light (L) (50,000 Da)-chain polypeptide linked by at least one disulfide bridge (40).

It has been proposed (38, 39) that action of both BoNT and TeTx involves three distinct phases. In the first phase the toxins become bound to acceptors on the external surface of the targeted neural cells. This is followed by an energy-dependent internalization step in which the toxin, or part of it, enters the cell. Thereafter, an unidentified active moiety of the toxin causes nerve cell dysfunction by blocking the intracellular release of neurotransmitters. The two classes of toxins differ, however, in that BoNT preferentially inhibits acetylcholine release at the nerve periphery, whereas TeTx blockades the release of inhibitory amino acids principally in the central nervous system. On the basis of a number of pieces of experimental evidence, and by analogy to the characterized binary toxins (e.g., diphtheria and ricin), it is generally assumed that the L chain possesses the catalytic activity responsible for cell poisoning (1, 5, 41) and that the H chain delivers this moiety to the cell cytoplasm by mediating binding of the toxin to the cell and subsequent internalization. The dual role of the H chain in toxicity has been rationalized by the suggestion that the amino-terminal portion mediates internalization (29, 32, 33) and the carboxy terminus plays a crucial role in binding to nerve acceptors (20, 21, 30, 37).

To clarify structural and functional relationships, clostridial neurotoxin gene cloning programs have been initiated in a number of laboratories. As a result of nucleotide sequence analysis of cloned genes, the complete primary sequences of TeTx (11), BoNT/A (4, 44), BoNT/C (17), BoNT/D (3), and BoNT/E (45) are known. In the present report, we describe the cloning of the gene encoding BoNT/B (*botB*) and the derivation of the entire amino acid sequence of the neurotoxin by nucleotide sequencing.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. The source of chromosomal DNA was *C. botulinum* Danish, and the recombinant host used for cloning experiments was

* Corresponding author.

Escherichia coli TG1 [Δ(*lac-pro*) *supE thi hsdD5*/F'-*traD36 proA*$^+$ *B*$^+$ *lacI*^{q} *lacZ*Δ*M15*]. Cloning vectors employed were plasmids pMTL32 (this study), pMTL23 (7), and pCR1000 (26) and the M13 bacteriophages mp18 and mp19 (46). *C. botulinum* was cultivated in USA II broth (2% peptone, 1% yeast extract, 1% N-Z amine, 0.05% sodium mercaptoacetate, 1% glucose [pH 7.4]), and *E. coli* was cultivated in L broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl). Solidified medium (L agar) consisted of L broth with the addition of 2% (wt/vol) Bacto Agar (Difco Laboratories). Antibiotic concentrations used for the maintenance and the selection of transformants were 50 μg of ampicillin (pMTL32) and 50 μg of kanamycin (pCR1000) per ml. Restriction endonucleases and DNA modifying enzymes were purchased from Northumbria Biochemicals Ltd., Taq polymerase was from United States Biochemical Corporation, and radiolabel was from Amersham International.

Purification and manipulation of DNA. Transformation of *E. coli* and large-scale plasmid isolation procedures were as described previously (27). Small-scale plasmid isolation was by the method of Holmes and Quigley (19), while chromosomal DNA from *C. botulinum* was prepared essentially as described by Marmur (23). Restriction endonucleases and DNA modifying enzymes were used under the conditions recommended by the suppliers. Digests were electrophoresed in 1% agarose slab gels on a standard horizontal system (Bethesda Research Laboratories model H4), employing Tris borate-EDTA (0.09 M Tris borate, 0.002 M EDTA) buffer. Fragments were isolated from gels by electroelution (25). All primary cloning procedures were undertaken under United Kingdom ACGM C2 containment conditions, and total cell lysates of all recombinants carrying cloned material were tested in mice for the absence of toxic polypeptides.

DNA-DNA hybridization experiments. DNA restriction fragments were transferred from agarose gels to Zeta Probe nylon membrane by the procedure of Reed and Mann (34). After partial depurination with 0.25 M HCl (15 min), DNA was transferred in 0.4 M NaOH by capillary elution for 4 to 16 h. Bacterial colonies were screened for desired recombinant plasmids by in situ colony hybridization (13), using nitrocellulose filter disks (0.22 μm; Schleicher and Schuell). The gel-purified *botA* DNA fragments were labelled with [α-^{32}P]dATP, using a multiprime kit supplied by Amersham International. Hybridizations were carried out as described previously (44) at temperatures ranging from 45 to 60°C.

Nucleotide sequence of pCBB plasmid inserts. The insert of plasmid pCBB1 was excised by cleavage with *Bam*HI and *Bgl*II and circularized by treatment with T4 ligase, and size-fractionated 500- to 1,000-bp fragments generated by sonication were cloned into the *Sma*I site of M13mp18 (for experimental conditions, see reference 28). Approximately 50 templates were then sequenced by the dideoxynucleotide method of Sanger et al. (35), using a modified version of bacteriophage T7 DNA polymerase, Sequenase (43). Experimental conditions used were those stated by the supplier (United States Biochemical Corp.). The inserts of plasmids pCBB2 and pCBB3 were sequenced by using templates derived by subcloning the entire region between the appropriate sites of M13mp18 and M13mp19. Sequence data obtained by employing universal primer were then sequentially extended by the use of custom-synthesized oligonucleotide primers. In certain instances, templates were generated by the insertion of *Dra*I restriction subfragments into the *Sma*I site of M13mp18. In all cases the sequence was determined on both DNA strands. The chromosomal DNA region amplified with primers X1 and X2 (Fig. 1) was cloned directly into ddT-tailed, *Sma*I-cut M13mp8 (prepared by incubating *Sma*I-cut DNA with terminal transferase in the presence of dideoxy TTP), and the resultant template was sequenced with universal primer. DNA sequence data were analyzed by using the computer software of DNASTAR Inc.

Amplification of DNA by PCR. Amplification of *C. botulinum* DNA was undertaken by polymerase chain reaction (PCR), using an M J Research Inc. thermal cycler. Reaction mixtures contained 10 mM Tris-HCl, 50 mM KCl, 3 mM $MgCl_2$, 0.1 mM deoxynucleoside triphosphate, 30 nmol of each primer, 2.5 U of *Taq* polymerase, and 10 ng of strain Danish genomic DNA, in a final volume of 0.1 ml. Amplification was for 30 cycles, as follows: 1.5 min at 93°C, 3 min at 37°C, and 3 min at 72°C. For inverse PCR, 140 ng of chromosomal DNA, cleaved with an appropriate restriction endonuclease, was ligated overnight at 14°C in a 50-μl volume and a 10-μl portion of the resultant concatenated DNA was used in PCR.

Nucleotide sequence accession number. The nucleotide sequence has been submitted to the GenBank/EMBL data banks, with the accession number M81186.

RESULTS AND DISCUSSION

Southern blot analysis of the *botB* gene. Previous studies have shown that BoNT appears to conform to the classical A-B binary toxin model (12). Thus, both L and H chains are required for toxicity (14, 39). The risk of generating an *E. coli* clone with the capacity to produce a neuroparalytic polypeptide may therefore be alleviated by cloning genomic restriction fragments which encode principally only one component of the dichain molecule. To identify such fragments, we exploited DNA homology between *botB* and the previously cloned *botA* (44).

A 389-bp *Hpa*I-*Xho*II *botA* fragment, encoding amino acids 216 through 346 of the BoNT/A L chain, and a 628-bp *Hae*II-*Hin*dIII fragment, coding for amino acids 526 through 736 of the H chain (44), were radiolabelled and used in DNA-DNA hybridizations with type B chromosomal DNA cleaved with various restriction enzymes. Reactions were performed in aqueous solution over a range of temperatures. "Weak" hybridization between the two genes was found to occur at 53 and 56°C with the L- and H-chain probes, respectively (data not shown). The strength of the signal observed and the relatively low stringency required were indicative of a fairly low level of DNA homology between *botA* and *botB*. Furthermore, these results suggest that the L-chain-encoding regions of the two genes are less homologous than the H-chain-encoding region, at least in the areas probed. The conditions under which hybridization occurred having been established, the type B genomic DNA was cleaved with various combinations of restriction endonucleases and the nylon membranes carrying the resultant fragments were sequentially hybridized with the two probes. The data obtained allowed the derivation of a restriction map of the region of the type B genome encoding *botB*. Furthermore, the use of the two probes enabled the assignment of both the position of *botB* and its relative orientation with respect to the derived map (Fig. 1).

Cloning and sequencing of the *botB* L chain. The restriction map derived by the Southern blot experiments (Fig. 1)

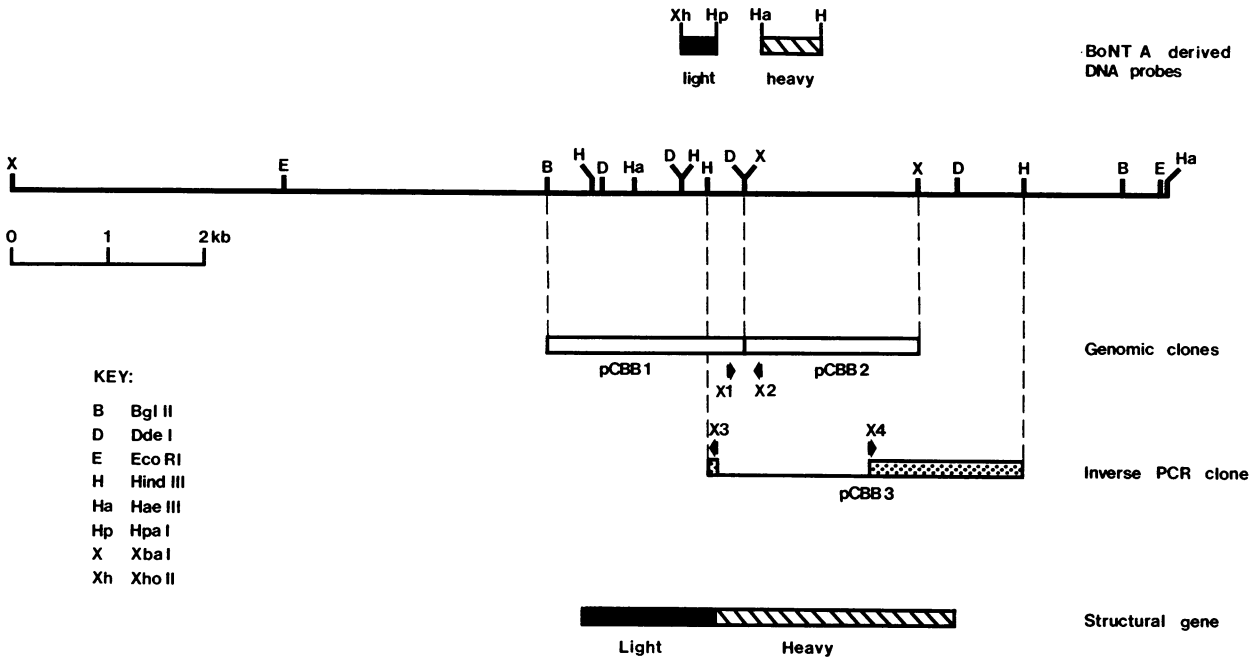


FIG. 1. Strategy employed in cloning the *botB* gene. The illustrated restriction map of the *C. botulinum* genome was generated by using the indicated *botA* DNA fragments as probes in Southern blots. Regions of the strain B/Danish chromosome that were cloned in recombinant plasmids pCBB1 and pCBB2 are represented by open boxes below the restriction map. The cloned inserts of these plasmids were shown to be contiguous on the genome by PCR amplification of the region of the chromosome spanning their common *Xba*I site, using primers X1 (5′-CCAAGTGAAAATACAGAATCAC-3′) and X2 (3′-CCCACTTTGTCTATCATTTA-5′) and sequencing across this junction. The insert of pCBB3 was derived by PCR amplification of *Hin*dIII-cut, concatenated chromosomal DNA, using primers X4 (5′-ATAGAGATTTATATAT TGGAG-3′) and X3 (5′-TTATATACAGCCAAATGCTCCTTGC-3′).

indicated that a 2.1-kb *Bgl*II-*Xba*I fragment principally encoded the L chain of BoNT/B. To clone this DNA, and to minimize the risk of cloning contiguous BoNT/B-encoding regions, the targeted fragment was purified by a two-stage gel isolation procedure. *C. botulinum* type B chromosomal DNA was cleaved with *Xba*I, and fragments of approximately 7.5 kb were purified from agarose gels by electroelution. The isolated DNA was then subjected to digestion with *Bgl*II, DNA fragments of around 2.1 kb were gel purified and ligated to pMTL32 vector DNA (Fig. 2) cut with *Xba*I and *Bam*HI, and the resultant TG1 transformants were screened for the presence of recombinant clones, using the *botA* L-chain probe. Vector pMTL32 was specifically constructed for the purposes of cloning the *botB* DNA (Fig. 2). Based on the pMTL1003 backbone (6), it carries multiple cloning sites flanked on either side by tandem copies of transcriptional terminators. Heterologous genes inserted into the multiple cloning sites will therefore only be expressed if they carry indigenous transcriptional elements recognized by the RNA polymerases of *E. coli*.

The recombinant plasmid obtained, designated pCBB1, was shown by digestion with appropriate endonucleases to contain restriction enzyme recognition sites consistent with the map illustrated in Fig. 1. Its entire insert was excised by digestion with *Bam*HI and *Bgl*II, and M13 recombinant templates containing random inserts were derived by using a sonication procedure (28). By using these templates and custom synthesized oligonucleotides, the entire nucleotide sequence of the insert was determined on both strands. Translation of the resultant sequence indicated the presence of an open reading frame encoding a polypeptide of 549 amino acids in size. The amino terminus of this polypeptide exhibited perfect conformity to that experimentally determined for purified BoNT/B L chain (36). Amino acids 442 through 459 were identical to those determined for purified BoNT/B H chain (36). Thus, the insert carried by pCBB1 was deemed to encode the entire L chain of BoNT/B and 108 amino acids from the H chain.

Cloning and sequencing of the *botB* H chain. After it was determined that the 2.1-kb *Bgl*II-*Xba*I fragment encoded the entire BoNT/B L chain and the amino terminus of the H chain, it was apparent that the adjacent 1.8-kb *Xba*I fragment (Fig. 1) should encode the majority of the remaining H chain. Type B chromosomal DNA was cleaved with *Hin*dIII, fragments of approximately 3.5 kb were isolated and digested with *Xba*I, and fragments of around 1.8 kb were gel purified. The isolated DNA was ligated with *Xba*I-cleaved pMTL32 and transformed into *E. coli* TG1, and recombinant plasmids were identified by probing with the radiolabelled *botA* H-chain probe. One such plasmid was designated pCBB2, and the nucleotide sequence of its insert was determined, following its insertion into M13mp18, by employing custom-synthesized oligonucleotide primers.

Translation of the nucleotide sequence obtained revealed the presence of a continuous open reading frame of 623 codons, in the same reading frame relative to the *Xba*I site of that of the insert of plasmid pCBB1. To confirm that the two sequences were indeed contiguous, a 289-bp region of DNA encompassing the *Xba*I site was amplified from type B genomic DNA by using primers X1 and X2 (Fig. 1) in a PCR and cloned directly into ddT-tailed *Sma*I-cut M13mp8. Nu-

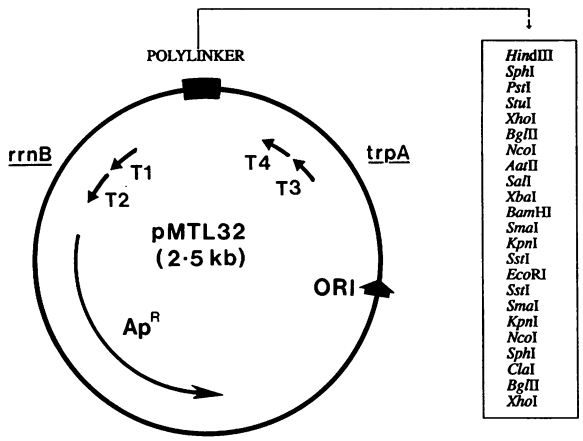


FIG. 2. Cloning vector pMTL32. This plasmid was derived as follows. A synthetic DNA fragment (5'-AGCCCGCCTAATGAGCG GGCTTTTTTT-3'), corresponding to the *E. coli trpA* transcriptional terminator, was ligated to *Stu*I-cleaved pMTL23 (7), and a recombinant plasmid (pTRP23) was selected in which two tandem copies of *trpA* had been inserted. The resultant double terminator, together with part of the pMTL23 polylinker region, was excised as a 107-bp *Nru*I-*Eco*RI fragment and inserted between the *Eco*RI and *Eco*RV sites of plasmid pMTL1003 (6). As the ca. 350-bp *Eco*RI-*Eco*RV fragment of pMTL1003 is deleted during this manipulation, the resultant plasmid, pMTL32, does not carry a copy of the *trp* promoter.

cleotide sequencing of a derivative template, using universal primer, demonstrated that the inserts of plasmids pCBB1 and pCBB2 were contiguous in the *C. botulinum* type B chromosome.

Completion of the *botB* sequence. By combining the two sequences of pCBB1 and pCBB2, the derived contiguous open reading frame encoded 1,170 amino acids, indicating that some 120 or so codons of the *botB* gene were missing. A DNA region encompassing the remaining 3' end of the gene was cloned by inverse PCR. Type B chromosomal DNA was cleaved with *Hin*dIII and incubated with T4 ligase, and the resultant concatenated DNA was used as a template in PCR with oligonucleotides X3 and X4 (Fig. 1). The 1.6-kb fragment generated was cloned directly into the specialized vector pCR1000, and the recombinant plasmid obtained was designated pCBB3. A plasmid sequence reaction, undertaken with a primer previously employed in the determination of the nucleotide sequence of the insert of plasmid pCBB2, confirmed the presence of the *botB* gene. Thereafter, the nucleotide sequence of the region of pCBB3 encompassing the 3' end of *botB* was determined by subcloning selected overlapping fragments into M13. To rule out the possibility that the insert of pCBB3 may have contained PCR-induced errors, a second version of this plasmid recombinant was derived by cloning the amplified DNA product from a further independent inverse PCR. Nucleotide sequencing of the appropriate regions of this second plasmid gave a sequence identical to that already derived from the primary isolate of pCBB3.

The entire nucleotide sequence of the *botB* gene (Fig. 3) was obtained by splicing the individual sequence information derived from the inserts of pCBB1, pCBB2, and pCBB3 into a contiguous sequence. The gene is composed of 1,291 codons, initiating with an AUG codon at position 55 and terminating with a UAA stop codon at position 3928 (Fig. 3). The choice of these particular translational codons is typical of clostridial genes (47). As with all other *bot* genes characterized to date, the high A+T content of the DNA (74.6%) results in an extreme bias towards the use of codons ending in A or T and the frequent use of codons recognized as modulators in *E. coli*. The translational start codon is preceded by a sequence typical of clostridial ribosome binding sites (47).

Alignment of the nucleotide sequences of the two *botA*-derived DNA probes used in Southern blot mapping with the equivalent regions of *botB* confirmed that the greater degree of homology existed in the respective H-chain-encoding regions over those encoding L chain. Specifically, the 628-bp *Hae*III-*Hin*dIII *botA* fragment demonstrated 65% homology with *botB*, whereas the 389-bp *Hpa*I-*Xho*II *botA* fragment had 54.8% homology with *botB*. Comparative alignment demonstrated that, in general, the overall DNA homology (Table 1) between the H- and L-chain-encoding regions of all sequenced neurotoxin genes reflected the level of amino acid sequence homology (Table 2) and averaged between 50 and 60% identity. One consequence of this relative dissimilarity between genes is that DNA probes specific to each toxin gene may be easily designed. However, although there is sufficient homology in certain regions to derive a generalized probe for the generic detection of neurotoxin genes, it has not proven possible to design a probe which hybridizes to all *bot* genes and not to the TeTx gene (unpublished data).

Predicted amino acid sequence of BoNT/B. The deduced primary sequence of BoNT/B demonstrates that the toxin is composed of 1,291 amino acid residues. By comparison to partial amino acid sequences derived from purified polypeptides from other *C. botulinum* type B strains, it is apparent that variations in toxin structure occur. Thus, although amino acid residues 2 through 17 exhibit perfect conformity to the sequence derived by Edman degradation of purified BoNT/B L chain of strain B/Okra (36), the amino acid at position 23 of the H chain was determined (10) to be Arg rather than the Ser residue seen here (position 464, Fig. 4). Similarly, the BoNT/B of strain B/657 possesses a Met amino acid at position 30 of the L chain (9) compared with Thr in the case of BoNT/B from both strain Danish and strain B/Okra. Variations in the primary amino acid sequence of other types of BoNT have been noted, e.g., between BoNT/A of strains 62A (4) and NCTC 2916 (44) and between BoNT/E of strains Beluga, Mashike, Iwanai, Otaru, and NCTC 11219 (see reference 45). In the case of BoNT/B, such variations help to explain observed dissimilarity in the immunological properties of BoNT/B isolated from different strains (16, 31).

Pairwise comparisons of the respective L- and H-chain components of all six toxins were undertaken, and the results are summarized in Table 2. From this it can be seen that, with notable exceptions, the overall level of identity between L chains varies from around 30 to 35%. The three exceptions are the degrees of homology seen between BoNT/E and TeTx (40%), BoNT/C and BoNT/D (47%), and BoNT/B and TeTx (50%). The last homology is particularly striking and serves to illustrate the close relationships between the pharmacological action of BoNT and TeTx. In contrast to the situation with the L-chain subunit, the H chains of BoNT/B and TeTx represent one of the most divergent pairings. The greatest level of homology (48% identity) to BoNT/B in this region is with BoNT/A. A similar relationship exists between the dichain components of BoNT/E and TeTx and BoNT/A. These observations sug-

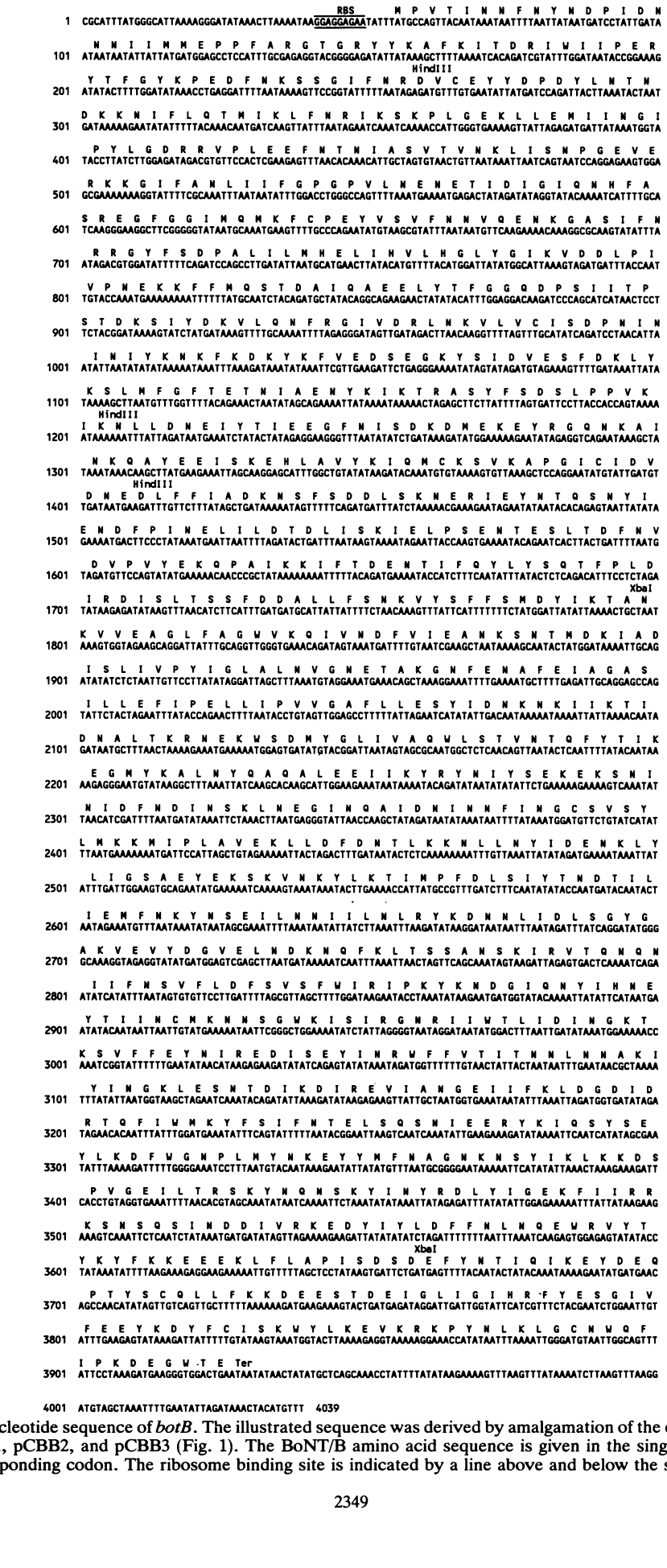
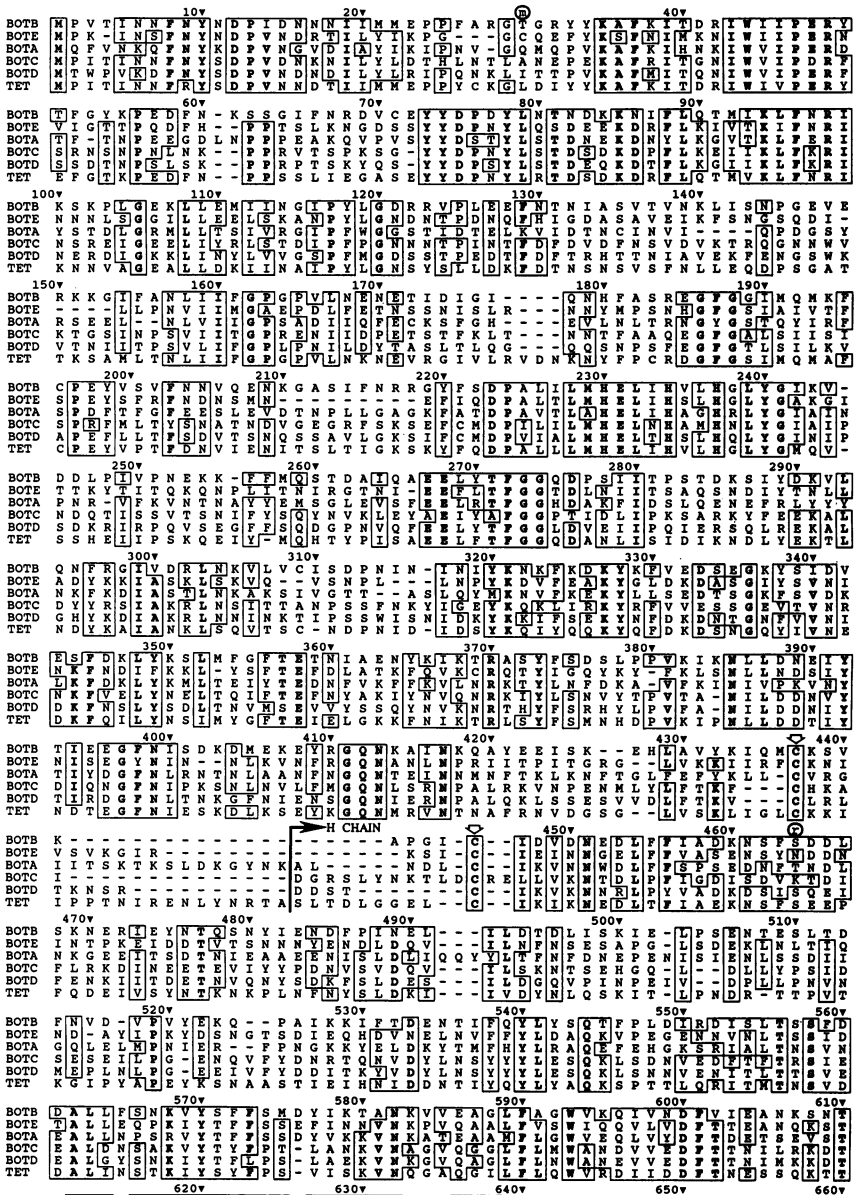


FIG. 3. Complete nucleotide sequence of *botB*. The illustrated sequence was derived by amalgamation of the derived nucleotide sequences of the inserts of pCBB1, pCBB2, and pCBB3 (Fig. 1). The BoNT/B amino acid sequence is given in the single-letter code above the first nucleotide of the corresponding codon. The ribosome binding site is indicated by a line above and below the sequence.



gest that either L- and H-chain domains of an individual neurotoxin have evolved at disproportionate rates or at various stages toxins have arisen by fusion of distinct H and L chains.

A full alignment of all presently characterized clostridial neurotoxins is illustrated in Fig. 4. This analysis demonstrates that they are composed of highly conserved domains interceded with tracts of amino acids exhibiting little overall relatedness, although considerable identity between the components of a specific pair is apparent in certain of these regions. The close relationship that exists between these neurotoxins is further exemplified by their arrangement of polar and nonpolar amino acids. Thus, analysis by the method of Kyte and Doolittle (22) demonstrates that all six neurotoxins possess a highly characteristic hydrophilicity profile (Fig. 5). From these profiles it is apparent that a substantial region of hydrophobicity is centrally conserved in all toxins (Fig. 5). This region has previously been suggested, in the case of TeTx and BoNT/A (11, 44), to play some role in the channel-forming properties of toxin H chain.

Within the L-chain region (average size, 442 amino acids), 68 amino acids are totally conserved, including a Cys amino acid in the carboxy termini. The Cys residues at this position represent one of only two positions where this particular amino acid is absolutely conserved, the second occurring in the amino termini of the H chains. These two Cys residues are therefore undoubtedly involved in the formation of the disulfide bridge linking the two dichain components of all six neurotoxins. Eleven of the conserved amino acids of the

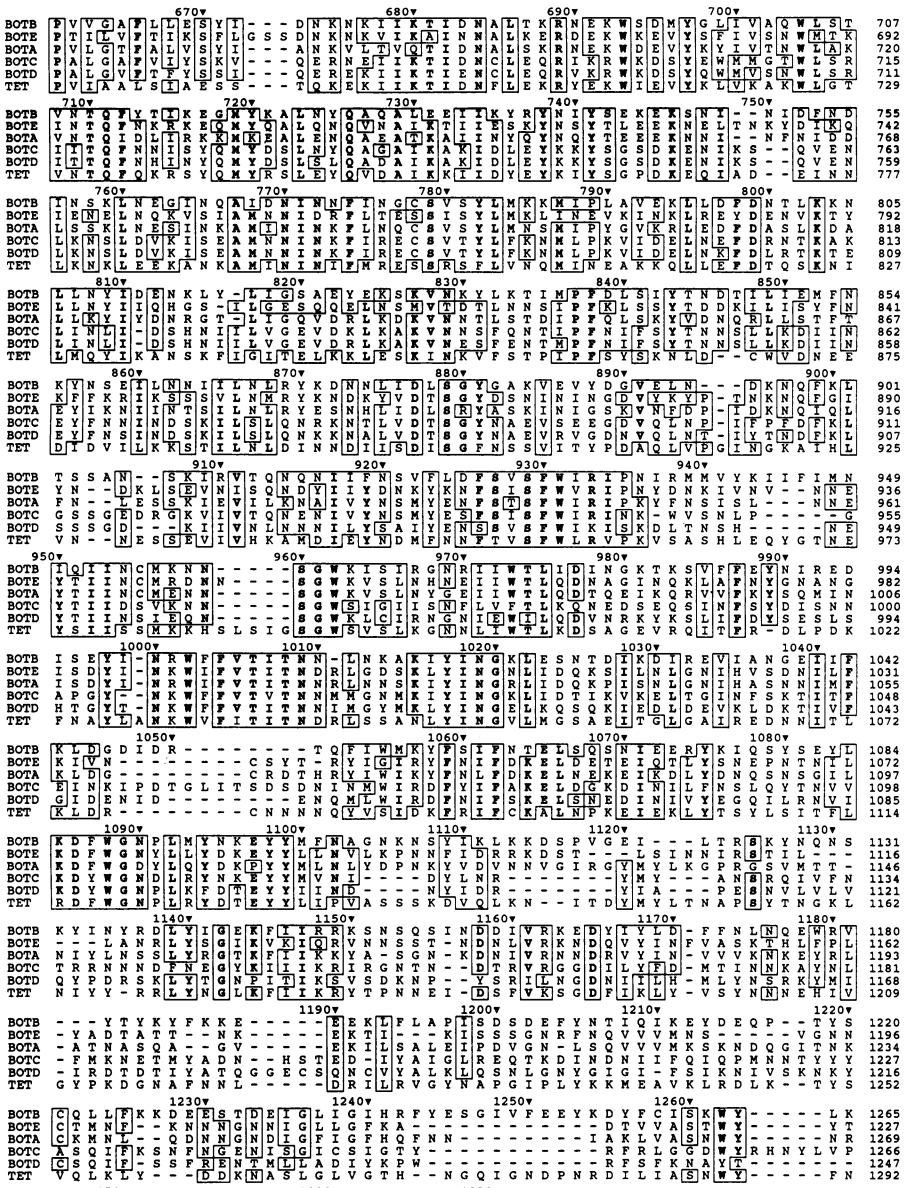


FIG. 4. Comparative alignment of clostridial neurotoxins. The illustrated alignment was essentially derived by using computer program CLUSTAL (18) and has been gapped to maximize similarity. Regions highly conserved among all six neurotoxins have been boxed and include areas in which conservative replacements have occurred, in addition to sequence identity. Amino acids absolutely conserved in five of six toxins are in boldface type. Numbering above the alignment corresponds to that of BoNT/B. Differences from the partial amino acid sequence of BoNT/B of strain B/657 (Met-30 and Arg-464) are circled and indicated above the strain B/Danish BoNT/B sequence. The Cys amino acids presumed to be involved in the formation of the disulfide bridge between the neurotoxin L and H chains are marked by downward facing open arrows.

neurotoxin L chains reside in a region (positions 223 to 241 of BoNT/B) which encompasses a histidine-rich motif. The three conserved His residues of this region, on the basis of their conservation in BoNT/A, BoNT/E, and TeTx, have previously been suggested to play some role in the presumed catalytic activity of the L chain (4). Their conservation in all six neurotoxins does not detract from this hypothesis. Preliminary work, however, in which site-directed mutagenesis has been used to effect amino acid substitutions at all three His positions did not affect the toxicity of a BoNT/A subunit in an *Aplysia californica* buccal ganglian model system (2).

A total of 110 amino acids are absolutely conserved within the H-chain region (average size, 845 amino acids). Most notable is the high degree of conservation of Trp amino acids. Of the 13 Trp residues which occur in the BoNT/B H

TABLE 1. Nucleotide sequence homology between characterized *bot* structural genes[a]

Gene	% Identity among H-chain- and among L-chain-encoding regions[b]					
	botA	*botB*	*botC*	*botD*	*botE*	TET
botA		58.6	52.8	55.2	62.7	54.0
botB	50.1		54.1	55.3	58.4	56.3
botC	49.8	52.5		70.0	53.0	48.5
botD	50.8	52.1	61.5		52.5	53.8
botE	51.6	51.2	51.2	52.3		53.3
TET	49.3	65.2	51.1	53.5	49.8	

[a] A, B, C, D, and E refer to the respective gene; TET represents the TeTx gene.
[b] Identities between H-chain-encoding and between L-chain-encoding regions are given above and below the diagonal, respectively.

TABLE 2. Degree of homology between the respective L- and H-chain components of characterized clostridial neurotoxins[a]

Neurotoxin	% Identity among H chains and among L chains[b]					
	A	B	C	D	E	TET
A		48	34	35	46	35
B	31		39	40	44	36
C	32	32		56	36	32
D	35	35	47		37	36
E	33	33	32	33		35
TET	30	50	34	34.5	40	

[a] A, B, C, D, and E refer to the respective BoNT; TET represents TeTx.
[b] Identities between H chains and between L chains are given above and below the diagonal, respectively.

chain, 9 are absolutely conserved in all toxins. In the majority of the four other positions, where a difference does exist in a particular toxin in six of nine cases, the substitution of a chemically similar amino acid has occurred. The only Trp that occurs in the BoNT/B L chain is conserved in all neurotoxins. The functional significance of the apparent evolutionary pressure for maintaining this relatively rare amino acid, or chemically similar residues, at these positions in BoNT and TeTx remains unknown. However, previous studies in which BoNT Trp residues have been selectively modified by chemical means has established a potential role in both toxicity and immunogenicity (8). Indeed, in one study it was reported that the modification of a single Trp resulted in nearly complete detoxification (Shibaeva et al., 1981, cited in reference 8). The selective disruption of conserved Trp amino acids in BoNT by site-directed mutagenesis should clarify which residue(s), if any, is important in toxicity and antigenicity.

The most striking area of sequence divergence between toxins occurs in carboxy-terminal areas of their H chains from, in the case of BoNT/B, around residue 1100 onwards. Given that this part of the toxin plays a major role in cell binding and that different toxins bind to distinct cell acceptor molecules, the finding that none of the toxins are alike in this region is perhaps not surprising. In view of the preceding region of divergence, the conservation of a sequence motif conforming to the consensus W-X-F-I/V-P/S-X-D/E-X-G-W-X-E/N (BoNT/B positions 1280 through 1291) at the extreme carboxy terminus is particularly intriguing.

ACKNOWLEDGMENTS

We thank Helen King and Howard Tranter for providing cell paste, Cliff Shone and Jack Melling for assisting with toxicity testing

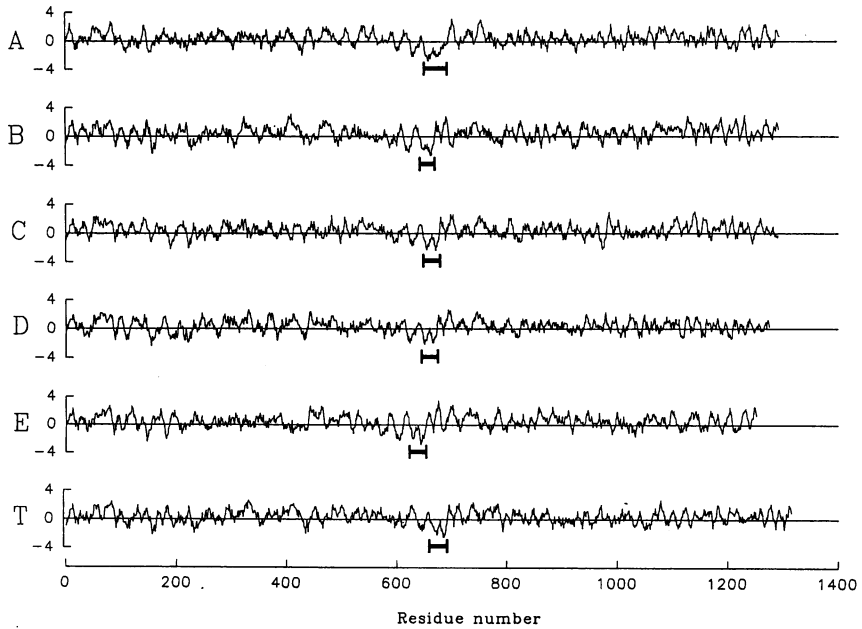


FIG. 5. Hydrophobicity plots of all currently characterized clostridial neurotoxins. Hydrophobicity was calculated by using the computer program of Kyte and Doolittle (22) with a window size of nine amino acids. The average value for each toxin was as follows: BoNT/A, −0.37; BoNT/B, −0.42; BoNT/C, −0.41; BoNT/D, −0.36; BoNT/E, −0.45; TeTx., −0.37. The conserved hydrophobic region is indicated below each profile by a barred line. The respective residues involved are 652 through 687 (BoNT/A), 642 through 671 (BoNT/B), 648 through 678 (BoNT/C), 646 through 674 (BoNT/D), 624 through 654 (BoNT/E), and 660 through 691 (TeTx).

of recombinant clones, and Nicola Minion for typing the manuscript.

This work was supported by the Public Health Laboratory Service, the U.K. Ministry of Defence (grant 2238/002), and the U.S. Army Medical Research and Development Command (grant DAMD17-90-Z-0033).

REFERENCES

1. **Ahnert-Higler, G., U. Weller, M. E. Dauzenroth, E. Habermann, and M. Gratzl.** 1989. The tetanus toxin light chain inhibits exocytosis. FEBS Microbiol. Lett. **242:**245–248.
2. **Binz, T., O. Grebenstein, H. Kurazono, J. Thierer, U. Eisel, K. Wernars, M. Popoff, S. Mochida, B. Poulain, L. Tauc, S. Kozaki, and H. Niemann.** 1992. Molecular biology of the L chains of clostridial neurotoxins, p. 56–65. *In* B. Witholt, J. E. Alouf, G. J. Boulnois, P. Cossart, B. W. Dijkstra, P. Falmagne, F. J. Fehrenbach, J. Freer, H. Niemann, R. Rappuoli, and T. Wadström (ed.), Bacterial protein toxins. Gustav Fischer Verlag, New York.
3. **Binz, T., H. Kurazono, M. R. Popoff, M. W. Eklund, G. Sakaguchi, S. Kozaki, K. Krieglstein, A. Henschen, D. M. Gill, and H. Niemann.** 1990. Nucleotide sequence of the gene encoding *Clostridium botulinum* neurotoxin type D. Nucleic Acids Res. **18:**5556.
4. **Binz, T., H. Kurazono, M. Wille, J. Frevert, K. Wernars, and H. Niemann.** 1990. The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. J. Biol. Chem. **265:**9153–9158.
5. **Bittner, M. A., B. DasGupta, and R. W. Holz.** 1989. Isolated light chains of botulinum neurotoxins inhibit exocytosis. J. Biol. Chem. **264:**10354–10360.
6. **Brehm, J. K., S. P. Chambers, K. J. Bown, T. Atkinson, and N. P. Minton.** 1991. Molecular cloning and nucleotide sequence determination of the *Bacillus stearothermophilus* NCA 1503 superoxide dismutase genes and its overexpression in *Escherichia coli*. Appl. Microbiol. Biotechnol. **36:**358–363.
7. **Chambers, S. P., S. E. Prior, D. A. Barstow, and N. P. Minton.** 1988. The pMTL *nic*$^-$ cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene **68:**139–149.
8. **DasGupta, B. R.** 1990. Structure and biological activity of botulinum neurotoxin. J. Physiol. Paris. **84:**220–228.
9. **DasGupta, B. R., and A. Datta.** 1987. Botulinum neurotoxin type B (strain 657); partial sequence and similarity with tetanus toxin. Fed. Proc. Fed. Am. Soc. Exp. Biol. **45:**2289.
10. **DasGupta, B. R., and A. Datta.** 1988. Botulinum neurotoxin type B (strain 657), partial sequence and similarity to tetanus neurotoxin. Biochimie **70:**811–817.
11. **Eisel, U., W. Jarausch, K. Goretzki, A. Henschen, J. Engels, U. Weller, M. Hudel, E. Habermann, and H. Niemann.** 1986. Tetanus toxin: primary structure, expression in *E. coli* and homology with botulinum toxins. EMBO J. **5:**2495–2502.
12. **Gill, D. M.** 1978. Seven toxic peptides that cross cell membranes, p. 291–332. *In* J. Jeljaszewicz and T. Wadstrom (ed.), Bacterial toxins and cell membranes. Academic Press, Inc., New York.
13. **Grunstein, M., and D. Hogness.** 1975. Colony hybridisation: a simple method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72:**3961–3965.
14. **Habermann, E., and F. Dreyer.** 1986. Clostridial neurotoxins: handling and action at the cellular and molecular level. Curr. Top. Microbiol. Immunol. **129:**93–179.
15. **Hall, J. D., L. McCroskey, B. J. Pincomb, and C. L. Hatheway.** 1985. Isolation of an organism resembling *Clostridium barati* which produces type F botulinum toxin from an infant with botulism. J. Clin. Microbiol. **21:**654–655.
16. **Hatheway, C. L., L. M. McCroskey, G. L. Lombard, and V. R. Dowell, Jr.** 1981. Atypical toxin variant of *Clostridium botulinum* type B associated with infant botulism. J. Clin. Microbiol. **14:**607–611.
17. **Hauser, D., M. W. Eklund, H. Kurazono, T. Binz, H. Niemann, D. M. Gill, P. Boquet, and M. R. Popoff.** 1990. Nucleotide sequence of the *Clostridium botulinum* C1 neurotoxin Nucleic Acids Res. **18:**4924.
18. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene **73:**237–244.
19. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids Anal. Biochem. **114:**193–197.
20. **Kozaki, S., A. Miki, Y. Kamata, J. Ogasawara, and G. Sakaguchi.** 1989. Immunological characterization of papain-induced fragments of *Clostridium botulinum* type A neurotoxin and interaction of the fragments with brain synaptosomes. Infect. Immun. **57:**2634–2639.
21. **Kozaki, S., J. Ogasawara, Y. Shimote, Y. Kamata, and G. Sakaguchi.** 1987. Antigenic structure of *Clostridium botulinum* type B neurotoxin and its interaction with gangliosides, cereboside, and free fatty acids. Infect. Immun. **55:**3051–3056.
22. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
23. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. **3:**208–218.
24. **McCroskey, L. M., C. L. Hatheway, L. Fenicia, B. Pasolini, and P. Aureli.** 1985. Characterization of an organism that produces type E botulinum toxin but which resembles *Clostridium butyricum* from the feces of an infant with type E botulism. J. Clin. Microbiol. **23:**210–202.
25. **McDonnell, M. W., M. N. Simon, and F. W. Studier.** 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. **110:**119–146.
26. **Mead, D. A., N. C. Pey, C. Herrnstadt, R. A. Marcil, and L. M. Smith.** 1991. A universal method for the direct cloning of PCR amplified nucleic acid. Bio/Technology **9:**657–663.
27. **Minton, N. P., T. Atkinson, and R. F. Sherwood.** 1983. Molecular cloning of the *Pseudomonas* carboxypeptidase G_2 gene and its expression in *Escherichia coli* and *Pseudomonas putida*. J. Bacteriol. **156:**1222–1227.
28. **Minton, N. P., H. M. S. Bullman, M. D. Scawen, T. Atkinson, and H. J. Gilbert.** 1986. Nucleotide sequence of the *Erwinia chrysanthemi* 1066 L-asparaginase gene. Gene **46:**25–35.
29. **Mochida, S., B. Poulain, U. Weller, E. Habermann, and L. Tauc.** 1989. Light chain of tetanus toxin intracellularly inhibits acetylcholine release at neuro-neuronal synapses, and its internalisation is mediated by heavy chain. FEBS Microbiol. Lett. **253:**47–51.
30. **Morris, N. P., E. Consiglio, L. D. Kohn, W. H. Habig, M. C. Hardegree, and T. B. Helting.** 1981. Interaction of fragments B and C of tetanus toxin with neural and thyroid membranes and with gangliosides. J. Biol. Chem. **255:**6071–6076.
31. **Notermans, S., S. Kozaki, Y. Kamata, and G. Sakaguchi.** 1984. Use of monoclonal antibodies in enzyme linked immunosorbent assay (ELISA) for detection of botulinum type B toxins. Jpn. J. Med. Sci. Biol. **37:**138–140.
32. **Poulain, B., S. Mochida, J. D. F. Wadsworth, U. Weller, B. Hogy, E. Habermann, J. O. Dolly, and L. Tauc.** 1990. Inhibition of neurotransmitter release by botulinum neurotoxins and tetanus toxin at *Aplysia* synapses: role of the constituent chains. J. Physiol. (Paris) **84:**247–261.
33. **Poulain, B., S. Mochida, U. Weller, B. Hogy, E. Habermann, J. D. F. Wadsworth, C. C. Shone, J. O. Dolly, and L. Tauc.** 1991. Heterologous combinations of heavy and light chains from botulinum neurotoxin A and tetanus toxin inhibit neurotransmitter release in *Aplysia*. J. Biol. Chem. **266:**9580–9585.
34. **Reed, K. C., and D. A. Mann.** 1985. Rapid transfer of DNA from agarose to nylon membranes. Nucleic Acids Res. **13:**7207–7221.
35. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143:**161–178.
36. **Sathymoorthy, V., and B. R. DasGupta.** 1985. Separation, purification, partial characterisation and comparison of the heavy and light chains of botulinum neurotoxin types A, B and E. J. Biol. Chem. **260:**10461–10466.
37. **Shone, C. C., P. Hambleton, and J. Melling.** 1985. Inactivation

of *Clostridium botulinum* type A neurotoxin by trypsin and purification of two tryptic fragments. Eur. J. Biochem. **151:**75–82.

38. **Simpson, L. L.** 1981. The origin, structure and pharmacological activities of botulinum toxin. Pharmacol. Rev. **33:**155–188.
39. **Simpson, L. L.** 1986. Molecular pharmacology of botulinum and tetanus toxin. Annu. Rev. Pharmacol. Toxicol. **26:**427–453.
40. **Simpson, L. L.** 1989. Botulnum neurotoxin and tetanus toxin. Academic Press, Inc., New York.
41. **Stechner, B., U. Weller, E. Habermann, M. Gratzl, and G. Ahnert-Higler.** 1989. The light chain but not the heavy chain of botulinum A toxin inhibits exocytosis from permeablised adrenal chromaffin cells. FEBS Microbiol. Lett. **255:**391–394.
42. **Sugiyama, H.** 1980. *Clostridium botulinum* neurotoxin. Microbiol. Rev. **44:**419–448.
43. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with modified T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:**4767–4771.
44. **Thompson, D. E., J. K. Brehm, J. D. Oultram, T.-J. Swinfield, C. C. Shone, T. Atkinson, J. Melling, and N. P. Minton.** 1990. The complete amino acid sequence of the *Clostridium botulinum* type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene. Eur. J. Biochem. **189:**73–81.
45. **Whelan, S. M., M. J. Elmore, N. J. Bodsworth, T. Atkinson, and N. P. Minton.** 1992. The complete amino acid sequence of the *Clostridium botulinum* type E neurotoxin, deduced by nucleotide sequence analysis of the encoding gene. Eur. J. Biochem. **204:**657–667.
46. **Yanisch-Perron, C., and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene **19:**103–119.
47. **Young, M., N. P. Minton, and W. L. Staudenbauer.** 1989. Recent advances in the genetics of clostridia. FEMS Microbiol. Rev. **63:**301–326.