

Molecular characterization of two forms of nontoxic-nonhemagglutinin components of *Clostridium botulinum* type A progenitor toxins

Ryuji Fujita^{a,b}, Yukako Fujinaga^a, Kaoru Inoue^a, Hiroshi Nakajima^a, Hiromi Kumon^b,
Keiji Oguma^{a,*}

^aDepartment of Bacteriology, Okayama University Medical School, 2-5-1 Shikata-cho, Okayama 700, Japan

^bDepartment of Urology, Okayama University Medical School, 2-5-1 Shikata-cho, Okayama 700, Japan

Received 12 October 1995

Abstract The entire sequences of the type A nontoxic-nonhemagglutinin gene and an adjacent open reading frame designated as *orf 22-a*, which are located between the neurotoxin and the HA-35 genes were determined. SDS-PAGE and N-terminal amino acid sequence analyses of the purified type A progenitor toxins (12S, 16S and 19S) indicate that the nontoxic-nonhemagglutinins of 16S and 19S are single peptides of approximately 120k, but that of 12S has a cleavage at the site between Pro-144 and Phe-145 of this protein.

Key words: *Clostridium botulinum*; Botulinum toxin; Neurotoxin; Progenitor toxin; Nontoxic-nonhemagglutinin; PCR

1. Introduction

Toxigenic strains of *Clostridium botulinum* produce one or more of the seven immunologically distinct neurotoxins (type A-to-G). The neurotoxins, approximately 150 kDa proteins, exist in the bacterial cultures as stable, large complexes designated as progenitor toxins, which are found in three forms: 12S toxin ($M_r \sim 300k$), 16S toxin ($M_r \sim 500k$), and 19S toxin ($M_r \sim 900k$). Type A progenitor toxin involves three forms, 19S, 16S, and 12S. Type B, C, and D involve two forms, 16S and 12S. Type E, F, and G involve a single form, 12S, and 16S, respectively. 19S and 16S toxins have hemagglutinin (HA) activity, but 12S toxin does not. It was postulated that the 12S toxin is formed by association of a neurotoxin with a nontoxic component ($M_r \sim 120k$) having no HA activity which is designated here as nontoxic-nonHA. Whereas, the 16S and 19S toxins are formed by conjugation of the 12S toxins with HAs. In this study we show the SDS-PAGE profiles and N-terminal amino acid sequences of purified type A HA-positive (16S and 19S) and HA-negative (12S) progenitor toxins in addition to the complete gene sequences of type A nontoxic-nonHA gene and adjacent open reading frame (*orf 22-a*), and the results of comparative analyses of these amino acid sequences with those of other types of *C. botulinum* progenitor toxins.

2. Materials and methods

2.1. Purification of progenitor toxins

The HA-positive (19S and 16S) and HA-negative (12S) progenitor toxins were purified from the culture fluid of *C. botulinum* type A strain,

A-NIH. The organisms were cultured by a cellophane tube procedure [1], and the culture medium was centrifuged ($6000 \times g$, 30 min). The supernatant was precipitated with 60% saturated ammonium sulfate. After removing RNA in the precipitate by protamine treatment [2], this preparation was subjected to a SP-Toyopearl 650M (Tosoh, Tokyo, Japan) column equilibrated with 50 mM sodium acetate buffer (pH 4.2). The column was washed with 100 ml of this solution, and the concentration of NaCl was then increased linearly to 0.5 M. The fractions which showed both toxic and hemagglutinating activity, and the fraction which showed only toxic activity were collected separately, concentrated by 70% saturation of ammonium sulfate, and then subjected to a Sephacryl S-300 (Pharmacia Biotechnology AB, Uppsala, Sweden) column equilibrated with 50 mM sodium acetate buffer (pH 4.2) containing 0.5 M NaCl, to obtain purified HA-positive (16S and 19S) and HA-negative (12S) progenitor toxins.

2.2. SDS-PAGE and N-terminal amino acid sequencing

SDS-PAGE and N-terminal amino acid sequencing analyses were performed as described previously [3].

2.3. Preparation of DNA and PCR amplification

Total DNA was isolated from A-NIH strain as described previously [4]. PCR amplification was performed in a 50 μ l reaction mixture containing 200 ng of template DNA, 100 nM of each primer, $1 \times$ Ex Taq buffer (Takara Shuzo Co. Ltd., Kyoto, Japan), 200 μ M of each dNTP, and 1.25 unit of Takara Ex Taq DNA polymerase (Takara Shuzo Co. Ltd.), by using a GeneAmp PCR System (Model 9600-R; Perkin-Elmer, Foster City, CA, USA). The PCR profile was 27 cycles of 94°C for 30 s, 53°C for 30 s, and 72°C for 180 s.

2.4. Cloning and nucleotide sequencing

PCR products were cloned directly into pT7 Blue T-vector (Novagen, Madison, WI, USA) according to the manufacturer's instructions and nested deletion of the cloned inserts was constructed by partial unidirectional hydrolysis by using a Deletion kit (Takara Shuzo Co. Ltd.). The sequence was determined with an automated model 373A DNA sequencer and the PRISM sequencing kit (Perkin-Elmer) as outlined in the manufacturer's instructions. The sequence was determined on two cloned fragments derived from different PCR experiments. Where the two clones differed in sequence, a third, independently amplified fragment was cloned and sequenced. Sequence data were analysed with the GENETYX-MAC sequence analysis software (version 7.3; Software Development Co. Ltd., Tokyo, Japan).

2.5. Nucleotide sequence accession number

The GSDB/DBJ/EMBL/NCBI accession number for the nucleotide sequence reported in this paper is D67030.

3. Results and discussion

3.1. SDS-PAGE and N-terminal amino acid sequence analyses of progenitor toxins

HA-positive and HA-negative progenitor toxins were purified as described in section 2. As reported previously [2], the HA-positive progenitor toxin preparation is a mixture of 16S and 19S toxins and HA-negative progenitor toxin preparation is 12S toxin. These preparations were subjected to SDS-PAGE

*Corresponding author. Fax: (81) (86) 222-1815.

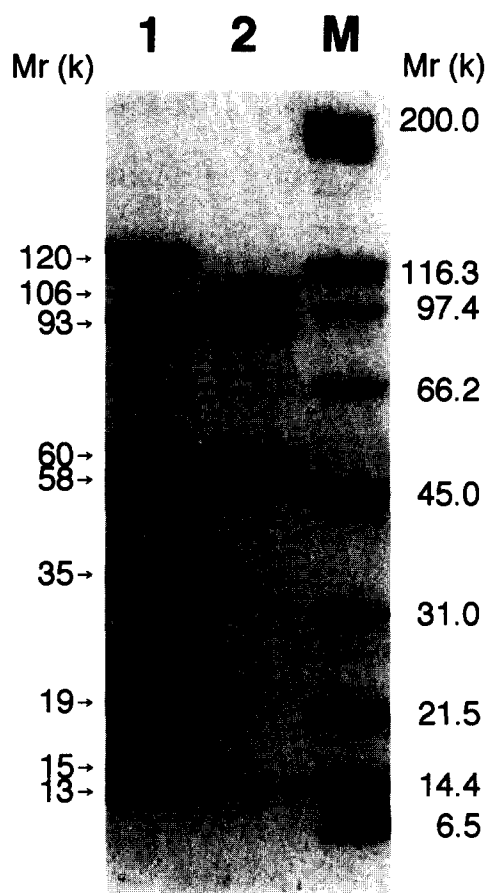


Fig. 1. SDS-PAGE of HA positive (lane 1) and HA-negative (lane 2) progenitor toxin fractions separated on Sephacryl S-300. Each toxin fraction (10 μ g) was heated at 100°C for 7 min in the presence of 2-ME and subjected to SDS-PAGE (5–15% acrylamide gradient gel). The gel was stained with CBB R-250. The positions of nontoxic-nonHA (120k), 106k component of 12S toxin, heavy chain (93k) and light chain (55k) of neurotoxin, HA-52 (52k), HA-35 (35k), HA-19 (19k), HA-15 (15k), and 13k component of 12S toxin are indicated by arrows. M = molecular weight standards, broad range (Bio-Rad).

with a reducing agent (Fig. 1). HA-positive progenitor toxins demonstrated seven major bands with M_r s of approximately 120k, 93k, 55k, 52k, 35k, 19k, and 15k. Based on previous reports [3,5–7,9], it can be concluded that the band of 120k is nontoxic-nonHA, and those of 93k and 55k are the heavy and light chains of the neurotoxin, respectively, and the remaining bands are the subcomponents of HA. On the contrary, the HA-negative progenitor toxin showed two bands with M_r of 106k and 13k in addition to the heavy and light chains of neurotoxin, indicating that the nontoxic-nonHA of 12S toxin consists of two fragments. In an attempt to confirm this hypothesis, the N-terminal amino acid sequences of these nontoxic components were determined by a direct protein microsequencing procedure (Table 1). The N-terminal amino acid sequences of nontoxic-nonHA, HA-52, and HA-15 were highly homologous to those of type C (64–82% identity), whereas the sequences of HA-35 and HA-19 showed only 30–40% identity to those of type C [3,6,8]. These N-terminal sequences are consistent with those of type A HA subcomponents reported by Somer and DasGupta [9].

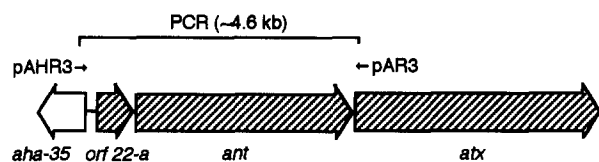


Fig. 2. Diagrammatic representation of nontoxic-nonHA gene and its surrounding region showing position of primers used in PCR for cloning. The respective genes coding for neurotoxin (*atx*), nontoxic-nonHA (*ant*), HA-35 (*aha-35*), and *orf 22-a* are shown by open boxes driving their directions of transcription. The sequences of primers were as follows: *pAHR3* 5'-ATCATTTAATGAATTTGTATTAC-3'; *pAR3* 5'-ATATCAACACCATTTACAGG-3'.

3.2. Sequence determination of nontoxic-nonHA gene and *orf 22-a*

From the reported construction of the type C and D 16S progenitor toxin gene clusters [3,6–8,10], we deduced that also in type A the nontoxic-nonHA gene is located between the HA-35 gene and the neurotoxin gene. Therefore, to amplify the nontoxic-nonHA gene (*ant*) we designated PCR primers *pAHR3* and *pAR3*, corresponding to the coding regions for N-terminals of the HA-35 gene (*aha-35*) and the neurotoxin gene (*atx*), respectively (Fig. 2). The *pAHR3* was designed with the N-terminal protein sequence data of HA-35. Nucleotides in the positions of codon degeneracy were chosen on the basis of those most commonly found in clostridial genes [11]. *pAR3* was designed with the nucleotide sequence of the 5'-end of the type A neurotoxin gene [12,13]. The PCR products (about 4.6 kb) were cloned and sequenced as described in section 2. The complete nucleotide sequences of the *ant* and the adjacent region are shown in Fig. 3. The gene encodes a protein of 1193 amino acids with a calculated M_r of 138.1 kDa. The *ant* ends 46 nucleotides before the start of *atx*. Upstream of the *ant*, there exists an open reading frame (*orf 22-a*) coding for 178 amino acids with a calculated M_r of 21.7 kDa. The amino acid sequence encoded here shows 52.2% identity to that of the *orf-22* of type C [8]. The N-terminal amino acids of the nontoxic-nonHAs of the HA-positive progenitor toxins determined by protein sequencing were located within the open reading frame starting at Met-1, and the those of the 13k and 106k components of the HA-negative progenitor toxin were located within the open reading frame starting at Met-1 and Phe-145, respectively. These results indicate that the nontoxic-nonHA of 12S toxin have a nick at the site between Pro-144 and Phe-145. Deduced 13k and 106k components consists of 144 and 1049 amino acids, and is calculated to be 15.9 kDa and 122.2 kDa,

Table 1
N-Terminal amino acid sequences of nontoxic components

Protein	N-terminal amino acid sequence
<i>HA-positive progenitor toxins</i>	
nontoxic-nonHA	MNINDNLSINSPV
HA-52	VSSTQRVLPYXNGLYVIN
HA-35	VIQNSLNDKIVTI
HA-19	SDTIDLADGNY
HA-15	SVERTFLPNGNYNIKSIFSG
<i>HA-negative progenitor toxin</i>	
106 k component	FPYAGYRETNV
13 k component	MNINDNLSINXPV

X = not determined.

ataacataacataatcaaatatttttggtaaacctaaaatttaatatatacaatt	60	tctcttaagataaatcttttagtgattcttcttctcttaagttaatttttttaga	2520
ttattagtagttttacataattgattatgtgattatctgtaaaatggcttataaaaat	120	S S K D K S L V Y S F L S N V M F Y L D	544
taaaagcaattagttttattttatgataataaaaaataatgatatattatggaagg	180	ttccataaaggataatagtcctattgatacagataaaaaatattattatggttaagaga	2580
agtggttaaatgaataaattgtttttcaaaataaaattgtaaaaaatgacaaagg	240	S I K D N S P I D T D K K Y Y L W L R E	564
SD M N K L F L Q I K M L K N D N R	16	gatttttagaaattattcttttgatattactgcaactcaagaaatataactaattgtgg	2640
orf 22-a-		I F R N Y S F D I T A T Q E I N T N C G	584
gttccaagaatttttaagcattttgaaaaaactataaattatttactagaaaatata	300	tattataaagttagtaactgtttgttggaaagcattaaaatttttaataacatctgattc	2700
F Q E I F K H F E K T I N I F T R K Y	36	I N K V V T W F G K A L N I L N T S D S	604
tatatatgataattacaatgatattttgaccattttatgtatcaccttaaaaaagttg	360	tttttagaagaatttcaaaatttaggggcaatttctacttataaaaaaagaaatttt	2760
I Y D N Y N D I L Y H L W Y T L K K V	56	F V E E F Q N L G A I S L I N K K E N L	624
tttgagcaatttcaatacacaaaatgatttagagagataattatgtaggactttaaaaa	420	aagtatgcccaattgagagattatgaaatccctaacgataattgtaggattaccactaaa	2820
L S N F N T Q N D L E R Y I S R T L K	76	S M P I I E S Y E I P N D M L G L P L N	644
atattgcttagatatttgcataaaaagaagattgataagaaaataatataaattcag	480	tgatttaaaatgaaaattatttaacatatattcttaaaacacagctatttttaaaaaaat	2880
Y C L D I C N K R K I D K K I I Y N S	96	D L N E K L F N I Y S K N T A Y F K K I	664
aatgttagataagaaattagccttaatagcaaatgttttcaagttatttagaatttg	540	ctactataatttcttagatcagtggtggacacaaattattatgctcaaatattttgttaat	2940
I V D K K L S L I A N S Y L E F	116	Y Y N F L D Q W W T Q Y Y S Q Y F D L I	684
atttaattgatttaataatccattactgtatgatacaaaaatattatataataaatt	600	ttgtatggcctaagatcagtggttagctcaagaacttttaaaaaagataatacaaaa	3000
F N D L I S I L P D D Q K K I I Y M K	136	C M A K R S V L A Q E T L I K R I Q K	704
tggtgaagataataaggagatagatatagctaaaaacttaataaagtgcgcaattctg	660	aaaattgagttatttaataaggaaattctaatatcatctgtaaaccttagcattgatgaa	3060
V E D I K E I D I A K K L N I S R Q S	156	K L S Y L I G N S N I S D N L A L M N	724
atataaaaataaataatggcttttagagagattagaacccatattgaaaaataatta	720	tcttaacaacaataacattagagataatttcaaacgaatcacaaatagcaatgaataa	3120
Y K N K I M A L E R L E P I L K K L I	176	L T T T N T L R D I S N E S Q I A M N N	744
tatgtatgttatttttttaaaaaatttaggtttacaaaaaattagtggtctatgttat	780	tgtagatagtttttttaataatggcgttatattgttttgaagatgaatataatcttaa	3180
I M *		V D S F L N N A A I C V F E S N I Y P K	764
tataaatgataagaataactgaaataatgtatccaaatttaagggggcgtgtatagta	840	atttttttctttttaggaacaattgataataataataataataaagcaaaaagatttat	3240
ataataaagtagtggtgtgaaataaatttaggggtgttagatgaatataaa	900	F I S F M E Q C I N N I N I K T K E F I	784
SD M N I N	4	acaaaagtgactataatgaatgaagataaataacataataacacgaatgtttt	3300
ant-		Q K C T N I N E D E K L Q L I N Q N V F	804
gacaacttaagtataaattccccggtagataataaaaattgtgttagttagagctag	960	taatagcttagattttgaattcttaaatattcaaaatagaaagttatttttagtcaga	3360
D N L S I N S P V D N K N V V V V R A	24	N S L D F E F L N I Q N M K S L F S S E	824
aaaactgatacaggttttttaaggttttagtcaccaatatttggtggcgccaga	1020	gacagcattactataaaggagaacttgcccttagtaactagttgttatgtcttttaa	3420
K T D T V F K A F K V A P N I W V A P E	44	T A L L I K E E T W P Y E L V L Y A F K	844
agataattatggcgaattctttgagtagatagaagaataaagaattgtatgggggaata	1080	ggacaccggaataatgttatcggagatgcattcgtgtaaaatcatcaatagaatattc	3480
R Y Y G E S L S I D E E Y K V D G G I Y	64	E P G N N V I G D A S G K N T S I E Y S	864
gattcctaattttcttcacaagtagtgaagaaagataagttcttcaagccattattac	1140	taaggacataggttttagtttaggaataaattgtagatcattatataatggatctaa	3540
D S N F L S Q D S E K D K F L Q A I I T	84	K D I G L V Y G I N S D A T A L Y L N G S N	884
ttgttaaaaagaattatagtaactgacgtgggaaagttattattctttgattttctac	1200	tcaaagataagtttttctaagattcttcttgagaatggatttaactaacagtttttcaat	3600
L L K R I N S T N A G E K L L S L I S T	104	Q S I S F S N D F F E N G L T N S F S I	904
a ctattccatttctcttagtatatataggtggagggtattatgcacctaattatgattac	1260	tatttttgggttgagaatttgggcaagataactataaacttaagtttaattagtagtaa	3660
A I P F P Y G Y I G G G Y Y A P N M I T	124	Y F W L R N L G K D T I K S K L I G S K	924
t ttggatcagcaccataatctataaaaaattgaattcttcttaatttcaagttactattcc	1320	ggaaagataatgttggtgggaaatttatttcaagatgggttggttttcaatgat	3720
F G S A P K S N K K L N S L I S T I P	144	E D N C G W E I Y F Q D T G L V F N M I	944
a ttcttatgtaggataatagagaacaaatattcttctcctgaagataataaaagttt	1380		
F P Y A G Y R E T N Y L S S E D N K S F	164	agattctaagtggaatgagaagaatataatctctgatttcttaataatagttggca	3780
a tgcattctaatatgatttatttttggccaggagcaaacatagatagaacaaactatgt	1440	D S N G N E K N I Y L S D V S N N S W H	964
Y A S N I V I F G P G A N I V E N N T V	184	ctataataactatactgtagatcggttttaagaacaaattataatattttagtgataa	3840
t ttataaaaaggagatgcagaaaattggaatgggaacaaatgactgaataatggtttcca	1500	Y I T I S V D R L K E Q L L I F I D D N	984
F Y K K E D A E N G M G T M T E I W F Q	204	tttagtggtcattgaagatttaaggaaatttttaattcttctcaagtaataataattc	3900
a catttctaactataaataatagacgaattttatattgctgcgaatagaataataaa	1560	L V A N E S I K E I L N I Y S S N I I S	1004
P F L T Y K Y D E F Y I D P A I E L I K	224	tttatgaagcagataaatacgaagttattattgagggttaactattttaaataacccac	3960
a gtttaataaaattcttcttattcttatatggtataaaaactagtgatgatttagttat	1620	L S E N N P S Y I E G L T I L N K P T	1024
C L I K S L Y F L Y G I K P S D D L V I	244	tacaagtcaggaagtttgagtaatttttgaagttcctaataattcatatataagaga	4020
t catatagattgaagtggaattagagaataatagataactcaaatgaattatagttga	1680	T S Q E V L S N Y F E V L N N S Y I R D	1044
P Y R L R S E L E N I E Y S Q L N I V D	264	cagtaaggaagcattagataacataagacataatataatataattatgatttttc	4080
t tactagtagtctggagcattgacttaatttataaataacagatccattatgttttac	1740	S N E E R L E Y N K T Y Q L Y N Y V F S	1064
L L V S G G I D P K F I N T D P Y W F T	284	agataagcctatagtgaagtttaacaaaataataatataatttatacaataataatad	4140
a ataattattttctcaaatgcaaaaaaggtgtttgaagatcataggaattatttatgaac	1800	D K P I C E V K O N N N I Y L T I N N T	1084
D N Y F S N A K K V F E D H R N I Y E T	304	aaacaaatttaaatctacaagcttctaaattttaaatttagtatataaatacaataaca	4200
a aaattgaaggaataatgcattgttgatataaaattgagatttaaaacaaaagtt	1860	N N L N L Q A S K F K L S I N P N K Q	1104
E I E G N N A I G N D I K L R L K Q K F	324	atatgttcaaaaacttgatgaggttaataatttctgttagtagataataggaataatata	4260
t gaatcaatcatgaatgatataatgggaattaaatttaaattatttctctaaagagtttag	1920	Y V Q K L D E V I I S V L D N M E K Y I	1124
R I N I N D I W E L N L N Y F S K E F S	344	agatatatcgaagataatagattgcaacttaagatacaacaaaataacgcgaagaagat	4320
c ttatgatgcagatagatttaataatgcacttaaacattttttatagaaaacatacta	1980	D I S E D N R L Q L I D N K N N A K K M	1144
I M M P D R F N N A L K H F Y R K Q Y Y	364	gataattagtaatgatatatttttccaaattgttttaacccattcttataacggttaata	4380
a aaatagattatccagaaaattatagataaattggtttttgatttggtcgaataatgc	2040	I I S N D I F I S N C L T L S Y N G K Y	1164
K I D Y P E N Y S I A G F V N G Q I N A	384	tatatgttttatctagaagatgaaacataattggtgatgataatgaatgdtgc	4440
t aattatctttatcagatagaatacaagattataaaaactgaagaaataattaa	2100	I C L S M K D E N H N W I C N N D M S	1184
Q L S L S D R N Q D I I N K P E E I I N	404	aaagattttgtattttatggtcatttaataataataatttataatttttaaatattat	4500
t tattaatggaataatgtttcattatgagaagtaattattatggtgatgattaaa	2160	K Y L Y L W S F K *	
L L N G N N V S L M R S N I Y G D G L K	424	aaaaggttataaataatgcatttggtaataaacaatttaataaagatcctgtaaatg	4560
a gcactgtagatgatttttacagtaattataaaatcccatataatagagcctatgaata	2220	SD M P F V N K Q F N Y K D P V N	15
S T V D D F Y S N Y K I P Y N R A Y E Y	444	atx-	-pAR3
t atttttaaaattccaattgattcttctttagataaattgtaaacattggagtaatagacaa	2280	gtagtgaat	4570
H F N N S N D S S L D N V N I G V I D N	464	G V D	
t ttccagagattatagatgtaaatccttataagaaaattgtgataagttttcaccggt	2340		
I P E I I D V N P Y K E N C D K F S P V	484		
a agaaaattacaagttactagagaataatacaaatataccatggcctataaaattattt	2400		
Q K I T S T R E I N T N I P W P I N Y L	504		
a aagctcaaaataactacaataaataaatttagtttactcagattttgtagaagtagt	2460		
Q A Q N T N N E K F S L S S D F V E V V	524		

Fig. 3. Nucleotide sequence of nontoxic-nonHA gene (*ant*), *orf* 22-a and surrounding region. The translated amino acid sequence is given under the second nucleotide of each codon. The position of the pAR3 primer is indicated by dotted underline. The putative Shine-Dalgarno (SD) regions complementary to the 3'-end of *C. perfringens* 16S rRNA [14] are underlined. The amino acid sequence corresponding to that determined by protein N-terminal analysis is underlined. Stop codons are indicated by asterisks.

Type A	1:MNINDNLNSINSPVDNKNVVVRARKTDTVFKAFAVAPNIWVAPERYYYGSLSIDEYKVD	60
Type C	1:MDINDDLNINSFVDNKNVVIVRARKTNTFFKAFKVAVAPNIWVAPERYYYGPLDIAEYKLD	60
Type D	1:MDINDDLNINSFVDNKNVVIVRARKTNTFFKAFKVAVAPNIWVAPERYYYGPLDIAEYKLD	60
Type E	1:MKINGNLNIDSPVDNKNVAIVRSRN-QMFFKAFQVAPNIWVAPERYYYGSLKINEDQKFD	59
Type F	1:MKINDDLNINSFVDNKNVVIVRARKTNIFFKAFQVAPNIWVAPERYYYGPLNISDQEKSD	60
Type A	61:GGIYDSNFLSQDSEKDKFLQAIITLLKRINSTNAGEKLLSLISTAIPFPYGYIGGGYYAP	120
Type C	61:GGIYDSNFLSQDSERENFLQAIITLLKRINNTISGKQLLSLISTAIPFPYGYIGGGYSSP	120
Type D	61:GGIYDSNFLSQDSERENFLQAIITLLKRINNTISGKQLLSLISTAIPFPYGYIGGGYSSP	120
Type E	60:GGIYDSNFLSTNNEKDDFLQATIKLLQRINNNVVGAKLLSLISTAIPFPYENNT-----	113
Type F	61:GGIYDENFLKENSEKEEFLQAIITLLKRINNNIIGQKLLSIMCTSIPLHEYKQ-----	114
Type A	121:NMITFGSAPKSNKKLNSLISSTIPFPYAGYRETNYSSEDNKSFYASNIVIFGPGANIVE	180
Type C	121:NIFTFGKTPKSNKKLNSLVTSTIPFPFGGYRETNIESONNKNFYASNIIIFGPGSNIVE	180
Type D	121:NIFTFGKTPKSNKKLNSLVTSTIPFPFGGYRETNIESONNKNFYASNIVIFGPGSNIVE	180
Type E	114:-----EDYRQTNYSKNEHYTANLVIFGPGSNIK	146
Type F	115:-----GDYRQSNYLGSKNSEYLYSANIVIFGPGSNIVK	147

Fig. 4. Amino acid alignment of the N-terminal regions of nontoxic-nonHA from type A, type C [7], type D [10], type E [15], and type F [16] strains. Identical amino acids in all, or all except one, of the sequences are shown in bold. The repeated motifs are underlined. The arrowheads indicate N-terminal amino acids of ~100k components of 12S toxins determined experimentally by protein sequencing.

respectively, which are in good agreement with the 13 k and 106 k estimated by SDS-PAGE.

3.3. Comparison of type A nontoxic-nonHA with other nontoxic-nonHAs

The amino acid identity and similarity values of type A nontoxic-nonHA with those of other types of *C. botulinum* progenitor toxins published previously are shown in Table 2. The nontoxic-nonHA of type A shows 65.4%, 65.4%, 65.8%, and 70.8% identity with nontoxic-nonHA of types C, D, E, and F. A multiple alignment of N-terminal regions of the nontoxic-nonHAs of types A, C, D, E, and F is shown in Fig. 4. As described above, the type A nontoxic-nonHAs of the 16S and 19S toxins are single peptides, but that of 12S toxin has a cleavage. The same phenomena were observed in type D [10], and in type C (unpublished data). N-Terminal sequencing of the nontoxic-nonHA fragments of purified 12S and 16S of type A, C, and D toxins revealed that these processing occurs after the Pro-144, Thr-140, and Thr-140, respectively. These residues are contained in a region that includes a short repeat of the sequence, L-L/N-S-L-I/M/V-S/T-T/S-A/T-I-P-F-P/L-Y/F-G/A, and the cleavage occur the second repeat. The nontoxic-nonHAs from type E and F strains, which produce only HA-negative progenitor toxins (12S), have a common deletion of the second repeat with respect to those from type A, C, and D strains, which produce both HA-positive and HA-negative progenitor toxins. These results indicate that the region in and around the second repeat motif may have a critical role in forming HA-positive progenitor toxins, probably by the binding of the HA to the nontoxic-nonHA component.

Table 2
Percentage amino acid identities (lower left-hand triangle) and similarities (upper right-hand triangle) of nontoxic-nonHA

	A	C	D	E	F
A	—	94.1	94.0	89.6	90.7
C	65.4	—	99.7	88.4	88.6
D	65.4	99.6	—	88.7	88.9
E	65.8	55.8	55.8	—	93.2
F	70.8	58.2	58.0	71.1	—

Acknowledgements: This work was supported in part by the Grant 07670311 from the Ministry of Education of Japan. We thank Naomi Uchiage and Kiyomi Hamada for technical assistance. We also thank Dr. Hiroyuki Ohmori, Department of Urology, Okayama University Medical School, for his critical reading of this manuscript.

References

- [1] Stern, M. and Wentzel, L.M. (1950) *J. Immunol.* 65, 175–183.
- [2] Sugii, S. and Sakaguchi, G. (1975) *Infect. Immunol.* 12, 1262–1270.
- [3] Fujinaga, Y., Inoue, K., Shimazaki, S., Tomochika, K., Tsuzuki, K., Fujii, N., Watanabe, T., Ohyama, T., Takeshi, K., Inoue, K. and Oguma, K. (1994) *Biochem. Biophys. Res. Commun.* 205, 1291–1298.
- [4] Fujinaga, Y., Takeshi, K., Inoue, K., Fujita, R., Ohyama, T., Moriishi, K. and Oguma, K. (1995) *Biochem. Biophys. Res. Commun.* 213, 737–745.
- [5] DasGupta, B.R. and Sugiyama, H. (1972) *Biochem. Biophys. Res. Commun.* 48, 108–112.
- [6] Tsuzuki, K., Kimura, K., Fujii, N., Yokosawa, N., Indoh, T., Murakami, T. and Oguma, K. (1990) *Infect. Immunol.* 58, 3173–3177.
- [7] Tsuzuki, K., Kimura, K., Fujii, N., Yokosawa, N. and Oguma, K. (1992) *Biochem. Biophys. Res. Commun.* 183, 1273–1279.
- [8] Hauser, D., Eklund, M.W., Boquet, P. and Popoff, M.R. (1994) *Mol. Gen. Genet.* 243, 631–640.
- [9] Somers, E. and DasGupta, B.R. (1991) *J. Protein Chem.* 10, 415–425.
- [10] Ohyama, T., Watanabe, T., Fujinaga, Y., Inoue, K., Sunagawa, H., Fujii, N., Inoue, K. and Oguma, K. (1995) *Microbiol. Immunol.* 39, 457–465.
- [11] Young, M., Minton, N.P. and Staudenbauer, W. L. (1989) *FEMS Microbiol. Rev.* 63, 301–326.
- [12] Binz, T., Kurazono, H., Popoff, M.R., Eklund, M.W., Sakaguchi, G., Kozaki, S., Kriegstein, K., Henschen, A., Gill, D.M. and Niemann, H. (1990) *J. Biol. Chem.* 265, 9153–9158.
- [13] Thompson, D.E., Brehm, J.K., Oultram, J.D., Swinfield, T.J., Shone, C.C., Atkinson, T., Melling, J. and Minton, N.P. (1990) *Eur. J. Biochem.* 189, 73–81.
- [14] Garnier, T., Canat, B. and Cole, S.T. (1991) *J. Bacteriol.* 173, 5431–5438.
- [15] Fujii, N., Kimura, K., Yokosawa, N., Yashiki, T., Tsuzuki, K. and Oguma, K. (1993) *J. Gen. Microbiol.* 139, 79–86.
- [16] East, A.K. and Collins, M.D. (1994) *Curr. Microbiol.* 29, 69–77.