THE PRIMARY STRUCTURE OF TOXIN B FROM THE VENOM OF THE INDIAN COBRA *Naja naja*

Mitsuhiro OHTA, Toyosaku SASAKI* and Kyozo HAYASHI**
Department of Biological Chemistry, Faculty of Pharmaceutical Sciences, Kyoto University, Kyoto, Japan

Received 27 September 1976

1. Introduction

Venoms of many species of snakes belonging to the family *Elapidae* (cobras, mambas, tiger snakes, black snakes, taipan, etc.) and *Hydrophiidae* (sea snakes) are highly toxic and produce flaccid paralysis and respiratory failure [1]. These effects have been attributed to the so-called 'neurotoxins' contained in the venoms. The toxins contained in the venom of *Naja naja* have been demonstrated to show a great affinity to the acetylcholine receptors on the motor endplate. Unlike the crude cobra venom, cobra neurotoxin produces neuromuscular block without causing muscle contracture or inhibition of the muscle response to direct stimulation even which high concentrations. The blockade produced by the neurotoxin can be reversed by neostigmine. Pretreatment with *d*-tubocurarine can protect the chick biventer cervicis muscle from the neuromuscular blocking action of cobra neurotoxins.

In 1971, Karlsson et al. [2] reported that the venom of *Naja naja naja*, or spectacled Indian cobra contained two principal neurotoxins, which differ only by a serine/isoleucine substitution. In the course of the study of biologically active principles in the venom of *Naja naja*, several neurotoxic proteins including Types I and II were isolated by gel filtration on Sephadex G-50 followed by chromatography on CM-cellulose [3]. Among these neurotoxins, toxins A and B were the major neurotoxins and the content of these toxins varied considerably with the preparation of the crude venom.

Studies on the chemical properties of toxin B have established its amino acid composition and revealed that toxins B and A had the same amino acid composition except that the former had one more serine residue and one less isoleucine residue than the latter. Accordingly, toxin A [4] corresponded to *Naja naja naja* 3 neurotoxin which was isolated by Karlsson et al. [2] and toxin B to *Naja naja naja* 4 neurotoxin, although the subspecies of their origins were different. Knowledge of the sequence of toxin B is required in order to determine the location of the essential amino acid residues required for the maintenance of active conformation of this type. The purpose of this communication is to report the complete covalent structure of toxin B.

2. Materials and methods

The crude lyophilized *Naja naja* venom was the product of the Haffkine Institute, Bombay, India, and of Sigma Chemical Co., Ltd., USA. Trypsin and α-chymotrypsin were the products of Worthington Biochemical Corp., USA. Acid cardoxypeptidase produced by *Penicillium janthinellum* was given by Dr E. Ichishima to Tokyo Noko University.

Toxin B used in this study was prepared from the venom of *Naja naja* by gel filtration on Sephadex G-50 followed by CM-cellulose. The venom dissolved in 1% acetic acid was applied to a column of Sephadex G-50 equilibrated with the same elution medium. Due to the molecular weight of other known snake venom neurotoxins and its strongly lethal toxicity, the protein fraction with molecular weight of about 6000–7500 was freeze-dried and applied to a CM-cellulose column

* Himeji Institute of Technology, Himeji, Japan
**To whom inquiries about this article should be directed

equilibrated with a 0.005 M sodium acetate buffer, pH 5.8. A gradient (from 0.005 M sodium acetate buffer, pH 5.8, to 0.5 M sodium acetate buffer, pH 6.5) was then applied at the top of the column. Each fraction was pooled, lyophilized and then gel-filtered to remove sodium acetate. Homogeneity was examined by polyacrylamide disc-gel electrophoresis using 15% gels at pH 2.3 according to the method of Davis [5], and by sodium dodecyl sulfate disc gel electrophoresis according to the method of Weber and Osborn [6]. The isoelectric point was determined by the method of Vesterberg and Svensson [7] using 0.5% ampholytes, pH 3–10, and a sucrose gradient at 4°C for 20 h. Its LD_{50} in mice (NH strain) by intraperitoneal injection was determined according to the method of Litchfield and Wilcoxon [8]. Rat phrenic nerve-diaphragm preparation [9] was used for testing its action on the neuromuscular junction. The reductin and *S*-carboxymethylation (RCM) of toxin B was performed as described by Crestfield et al. [11]. Amino acid analysis of the toxin and the tryptic and chymotryptic peptides were carried out by standard method using a Hitachi KLA 3B automatic amino acid analyzer. Sequential degradations were conducted by the Edman procedure modified by Iwanaga et al. [10], and the phenylthiohydantoin amino acids were identified by thin layer chromatography [12,13]. The carboxy-terminal sequence of RCM-toxin B was determined by the use of acid carboxypeptidase [14]. The carboxy-termini of chymotryptic peptide (C-I) and RCM-toxin B were determined by hydrazinolysis [15]. Trypsin and chymotrypsin digestions were carried out at 37°C for 3 h in 0.1 M ammonium bicarbonate buffer, pH 7.9, using a weight ratio of enzyme to substrate of 1:50. These peptides derived by proteolytic enzymes were separated by a combination of gel filtration on a column of Sephadex G-15 using 0.2% acetic acid as the eluent and high voltage paper electrophoresis using a buffer (pyridine/acetic acid/water=1:10:289, by vol) pH 3.6, and descending paper chromatography using a solvent system (1-butanol/acetic acid/pyridine/water=15:3:10:12, by vol.)

3. Results and discussion

Toxin B was obtained in a yield of about 7% from the crude venom of *Naja naja*. When injected intraperitoneally in mice, its LD_{50} was 0.09 μg/g body weight. The purified toxin B was homogeneous by the criteria of polyacrylamide disc-gel electrophoresis and the isoelectric point was pH 9.3. The toxin decreased the single twitches of the diaphragm elicited by phrenic nerve stimulation. The molecular weight of toxin B was estimated by gel filtration to be about 7000. Based on this value and on the amino acid composition, one molecule of the toxin was composed of 71 amino acid residues: Asp 8.61, Thr 8.30, Ser 3.60, Glu 1.13, Pro 5.38, Gly 4.60, Ala 1.85, Half-Cys 8.40, Val 3.40, Ile 3.60, Leu 1.00, Tyr 1.01, Phe 2.96, Trp 1, Lys 3.64, His 0.82, Arg 5.60. Tryptophan was determined spectrophotometrically.

The thirty three stepwise Edman degradations of RCM-toxin B revealed the amino terminal sequence to be

H–Ile–Arg–Cys–Phe–Ile–Thr–Pro–Asp–Ile–
Thr–Ser–Lys–Asp–Cys–Pro–Asn–Gly–His–Val–
Cys–Tyr–Thr–Lys–Thr–Trp–Cys–Asp–Gly–
Phe–Cys–Ser–Ser–Arg– – –

The carboxy-terminal sequence was examined by the use of acid carboxypeptidase isolated from the culture fluid of *Penicillium janthinellum* and by hydrazinolysis. From the results, the carboxy-terminal sequence was determined to be

–Arg–Lys–Arg–Pro–OH

Then, RCM-toxin B was digested by chymotrypsin and the resulting peptides were separated by gel filtration on a Sephadex G-15 column and the fractions were assayed for absorbances at 226 and 280 nm. Further purification of heterogeneous fraction (C-IV) was made by high voltage paper electrophoresis. In the chymotrypsin digestion of RCM-toxin B, it was observed that the peptide bond $-\overset{67}{\text{Thr}}-\overset{68}{\text{Arg}}-$ was hyrolyzed considerably but no other threonyl peptide bonds were. The amino acid compositions and the yields of the chymotryptic peptides are given in table 1, and the amino acid sequences of these peptides elucidated by the Edman degradation are shown in table 2. The peptide numbers do not refer to the order in which their location in the sequence starting from the amino-terminal end but to the order of elu-

Table 1
Amino acid compositions of chymotryptic peptides derived from RCM-toxin B

Amino acid	C-I	C-II	C-III	C-IV-1	C-IV-2	C-V
CM-Cysteine	5.90 (6)	2.06 (2)		1.02 (1)	1.02 (1)	
Aspartic acid	4.90 (5)	2.90 (3)			1.08 (1)	
Threonine	4.92 (5)	2.01 (2)				1.93 (2)
Serine	2.54 (3)	1.00 (1)				
Glutamic acid	1.12 (1)					
Proline	2.80 (3)	1.85 (2)	1.14 (1)			
Glycine	2.68 (3)	1.00 (1)			0.90 (1)	
Alanine	1.96 (2)					
Valine	2.66 (3)	1.07 (1)				
Isoleucine	1.13 (1)	1.64 (2)		0.94 (1)		
Leucine	0.96 (1)					
Tyrosine		0.77 (1)				
Phenylalanine	1.00 (1)			1.00 (1)	1.00 (1)	
Lysine	0.85 (1)	0.90 (1)	1.00 (1)			1.00 (1)
Histidine		0.74 (1)				
Arginine	2.69 (3)		2.04 (2)	0.99 (1)		
Tryptophan						+ (1)
Yield (%)	69	67	75	88	89	85
Position	30–67	5–21	68–71	1–4	26–29	22–25

The values of the amino acids underlined were taken as 1.0. The numbers in parentheses represent the nearest integers. Tryptophan was determined spectrophotometrically.

Table 2
Chymotryptic peptides from RCM-toxin B

Peptide No.	Sequence
C-I	Cys–Ser–Ser–Arg–Gly–Lys–Arg–Val–Asp–Leu–Gly–Cys–Ala–Ala– Thr–Cys–Pro–Thr–Val–Arg–Thr–Gly–Val–Asp–Ile–Gln–Cys–Cys– Ser–Thr, (Asp, Asp, Cys, Asp, Pro, Phe, Pro)–Thr
C-II	Ile–Thr–Pro–Asp–Ile–Thr–Ser–Lys–Asp–Cys–Pro–Asn–Gly–His Val–Cys–Tyr
C-III	Arg–Lys–Arg–Pro
C-IV-1	Ile–Arg–Cys–Phe
C-IV-2	Cys–Asp–Gly–Phe
C-V	Thr–Lys–Thr–Trp

Right- and left-handed arrows showed that the sequence was elucidated, respectively, by Edman degradation or by the hydrazinolysis.

tion from a Sephadex column. The various peptides were designated by the symbol C–.

Chymotryptic peptides C-I and C-II with large molecular weight were subjected to the digestion by the use of trypsin. The resulting peptides were isolated by high voltage paper electrophoresis and their purities were examined by paper chromatography. The amino acid compositions of the tryptic peptides are given in table 3, and the amino acid sequences elucidated by the Edman degradation are shown in table 4. The peptides derived from the trypsin digestion were designated by the symbol T–.

Table 3
Amino acid compositions of tryptic peptides derived from chymotryptic peptides C-II and C-I

Amino acid	C-I-T-1	C-I-T-2	C-I-T-3	C-I-T-4	C-II-T-1	C-II-T-2
CM-Cysteine	1.01 (1)		2.15 (2)	3.04 (3)		1.57 (2)
Aspartic acid			1.31 (1)	4.41 (4)	1.12 (1)	2.08 (2)
Threonine			2.00 (2)	3.22 (3)	2.10 (2)	
Serine	1.80 (2)			1.10 (1)		
Glutamic acid				1.32 (1)		
Proline			0.95 (1)	2.22 (2)	1.18 (1)	0.77 (1)
Glycine		0.96 (1)	1.00 (1)	1.30 (1)		1.00 (1)
Alanine			1.54 (2)			
Valine			1.90 (2)	1.00 (1)		0.86 (1)
Isoleucine				0.90 (1)	1.71 (2)	
Leucine			0.75 (1)			
Tyrosine						0.62 (1)
Phenylalanine				0.80 (1)		
Lysine		1.00 (1)			1.00 (1)	
Histidine						0.89 (1)
Arginine	1.00 (1)	1.03 (1)	1.00 (1)			
Tryptophan						
Yield (%)	25	47	18	30	32	28
Position	30–33	34–36	37–49	50–67	5–12	13–21

The values of the amino acids underlined were taken as 1.0. The numbers in parentheses represent the nearest integers.

Table 4
Tryptic peptides from chymotryptic peptides C-I and C-II

Peptide No.	Sequence
C-I-T-1	Cys–Ser–Ser–Arg
C-I-T-2	Gly–Lys–Arg
C-I-T-3	Val–Asp–Leu–Gly–Cys–Ala (Ala–Thr–Cys–Pro–Thr–Val–Arg)
C-I-T-4	Thr–Gly–Val–Asp–Ile–Gln–Cys–Cys–Ser–Thr–Asp–Asp–Cys–Asp–Pro–Phe–Pro–Thr
C-II-T-1	(Ile–Thr–Pro–Asp–Ile–Thr–Ser–Lys)
C-II-T-2	(Asp–Cys–Pro–Asn–Gly–His–Val–Cys–Tyr)

Right-handed arrows showed that the sequence was elucidated by Edman degradation. The sequences in parentheses were determined from the amino acid compositions of the peptides and from data of the sequence analysis of RCM-toxin B and chymotryptic peptide C-I.

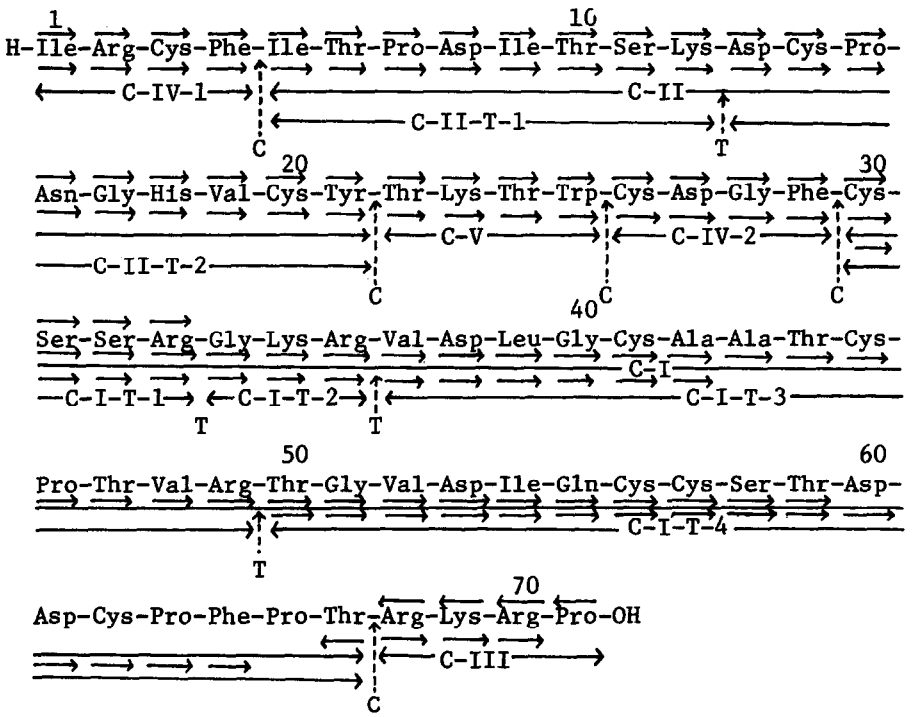

Fig.1. Amino Acid Sequence of Toxin B from Indian Cobra Venom (*Naja naja*). Horizontal arrows below amino acid residues denote the sequences of chymotryptic and tryptic peptides. Right- and left-handed arrows show that the sequence was elucidated, respectively, by Edman degradation by the action of acid carboxypeptidase or by the hydrazinolysis. C and T represent the peptide bonds which were hydrolyzed by the action of chymotrypsin, respectively.

There are overlaps between the tryptic and chymotryptic peptides which allowed the unequivocal placement of the remainder of the peptides. Referring to the amino- and the carboxy-terminal sequences of RCM-toxin B, the complete covalent structure of toxin B was elucidated as shown in fig.1. The sequence of toxin B indicated that the molecular weight of the protein consisting of 71 amino acid residues, cross-linked by five disulfide bridges, was 7812.

There is a remarkable similarity in the primary structure between toxins B and toxin A [4]. Toxin B differs from toxin A only in the presence of serine residue in place of isoleucine residue at position 32. The replacement can be explained by a single base replacement in the triplet codes (AGU to AUU or AGC to AUC).

Acknowledgements

The authors are indebted to Professor Ikuo Yamashina of our laboratory for his interest and valuable advice in this work. We are also grateful to Professor E. Ichishima of Toyo Noko University for the gift of acid carboxypeptidase. This work was aided in part by a grant from the Ministry of Education.

References

[1] Lee, C. Y. (1970) Clinical Toxicology 3, 457–472.
[2] Karlsson, E., Arnberg, H. and Eaker, D. (1971) Eur. J. Biochem. 21, 1–16.
[3] Ohta, M. and Hayashi, K. (1975) Tampakushitsu Kakusan Koso (in Japanese) 20, 53–69.
[4] Nakai, K., Sasaki, T. and Hayashi, K. (1971) Biochem. Biophys. Res. Commun. 44, 893–897.
[5] Davis, B. J. (1964) Ann. N.Y. Acad. Sci. 121, 404–427.
[6] Weber, K. and Osborn, M. (1969) J. Biol. Chem. 244, 4406–4412.
[7] Vesterberg, O. and Svensson, H. (1966) Acta Chem. Scand. 20, 820–834.

[8] Litchfield, J. T. and Wilcoxon, F. (1948) J. Pharmacol. Exp. Ther. 96, 99–113.
[9] Bülbring, E. (1946) Brit. J. Pharmacol. 1, 38–61.
[10] Iwanaga, S., Wallen, P., Gröndahl, N. J., Henschen, A. and Blömback, B. (1969) Eur. J. Biochem. 8, 189–199.
[11] Crestfield, A. M., Moore, S. and Stein, W. H. (1963) J. Biol. Chem. 238, 622–627.
[12] Jeppsson, J. O. and Sjöquist (1969) Anal. Biochem. 18, 264–269.
[13] Summers, M. R., Smythers, G. W. and Oroszlan, S. (1973) Anal. Biochem. 53, 624–628.
[14] Yokoyama, S. and Ichishima, E. (1972) Arg. Biol. Chem. 36, 1259–1261.
[15] Goetzl, E. J. and Metzger, H. (1970) Biochemistry 9, 3862–3871.