AMINO ACID SEQUENCE OF A SNAKE NEUROTOXIN FROM THE VENOM OF *LAPEMIS HARDWICKII* AND THE DETECTION OF A SULFHYDRYL GROUP BY LASER RAMAN SPECTROSCOPY

Jay W. FOX*, Marshall ELZINGA**, and Anthony T. TU*+

**Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523, USA, **Department of Biology, Brookhaven National Laboratory, Upton, Long Island, New York 11973, USA*

Received 9 June 1977

1. Introduction

Venoms of Elapidae and Hydrophiidae snakes contain potent postsynaptic neurotoxins [1]. The neurotoxins present in Hydrophiidae venoms are the short, Type I toxins containing 60 to 62 amino acids and four disulfide bonds [2,3].

The isolation, chemical properties and conformation of the major neurotoxin from *Lapemis hardwickii* have been extensively investigated [4,5]. Because the amino acid sequence of this toxin has not been reported, it is difficult to fully interpret these data in correlation with the structure–function relationship. In this paper the complete sequence of the major *Lapemis hardwickii* toxin is presented. Laser Raman spectroscopy was utilized to ascertain the free cysteine residue not participating in the disulfide bonds.

2. Materials and methods

The neurotoxin was isolated by a modification of the procedure of Tu and Hong [4] using gel filtration and ion-exchange chromatography. The purified major toxin was reduced and alkylated following the method of Crestfield as described by Elzinga [6]. Prior to tryptic digestion the carboxymethylated toxin (CM-toxin) was reacted with citraconic anhydride so as to 'block' the lysine residues [7]. The citraconylated CM-toxin (CM-CA toxin) was then subjected to tryptic digestion using an enzyme to substrate ratio of 1 : 30 (w/w). After 16 h at 23°C the digestion was terminated by the addition of soybean trypsin inhibitor (w/w trypsin) and glacial acetic acid to 25%. The mixture was then passed through a Sephadex G-50 column equilibrated with 25% acetic acid which served to separate the products and also 'deblock' the lysines (fig.1). Amino acid analyses were then performed on the tryptic peptides (table 1).

Carboxypeptidase A digestion of the CM-toxin was carried as previously described [6] and the digest then subjected to amino acid analysis.

All sequencing was performed on a Beckman 890C Sequencer. Using the 1 M Quadrol program, the partial sequence of the whole toxin was completed. To decrease extractive losses, the toxin was first reacted in the spinning cup with a 100-fold excess (total amino groups) of 3-isothiocyano-1,5-napthalene disulfonic acid (Braunitzer Reagent III) [8]. The T-1 tryptic peptide was partially sequenced using the 1 M Quadrol program without prior reaction with the Braunitzer reagent so that the N-terminal could be recovered. The tryptic peptide mixture T-2 + T-3, and the T-4 peptide were sequenced using the DMAA (dimethylallyamine) peptide program. The phenylthiohydantoins (PTHs) were identified by thin-layer chromatography and amino acid analysis after acid and base hydrolysis [9,10].

Laser Raman spectra were obtained using a Spex 1401 double monochromator, Coherent Radiation Model 52G argon ion beam and photon counting electronics. The spectra were obtained with 514.5 nm line excitation and 4 cm^{-1} spectral slit width resolution.

+To whom correspondence should be addressed

3. Results and discussion

The separation of the tryptic peptides is illustrated in fig.1. Peak I was determined to be the trypsin–trypsin inhibitor complex. Peaks II and III were the incomplete digestion products of the CM-CA toxin. The remaining ninhydrin and/or 280 nm absorbance peaks (T-1, T-2 + T-3, T-4) were identified by amino acid analyses as the resultant tryptic peptides from the toxin (table 1). The peptides T-2 and T-3 co-eluted as a mixture.

Amino acid analysis of the carboxypeptidase digestion mixture identified the amino acid residues released as Asp, Thr, Glu, and CMC in the ratio 3 : 1 : 1 : 1 : 1.

Edman degradation on the CM-toxin was carried out to 33 residues (fig.2). The N-terminal was lost due to the reaction with Braunitzer reagent. This portion of the sequence provided the overlap into the T-2 region. Four cycles of Edman degradation on the T-1 peptide identified it as the N-terminal tryptic peptide from the toxin. It also identified the N-terminal amino acid as Met. The peptide mixture T-2 + T-3 was sequenced as such. Each cycle of Edman degradation simultaneously yielded two amino acids; one from

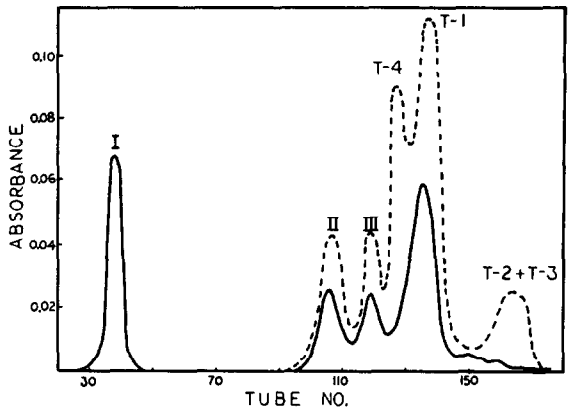

Fig.1. Fractionation of tryptic peptides from the major toxin. The solid line is the absorbance at 280 nm. The discontinuous line is the absorbance at 570 nm. The volume of each fraction is 6.5 ml/tube.

Table 1
Amino acid composition of *Lapemis hardwickii* major toxin and tryptic peptides

Amino acid	*Lapemis hardwickii* CM-Toxin		Peptide T-1	Peptide mixture T-2, T-3	Peptide T-4
Lys	5	5	3.38 (3)		3.23 (2)
His	2	2	1.20 (1)		1.03 (1)
Arg	3	3	0.92 (1)	2.23 (2)	
CMC	8	9	3.74 (4)		4.65 (5)
Asp	6	6	3.11 (3)		3.40 (3)
Thr	8	7	4.60 (5)	1.08 (1)	1.03 (1)
Ser	6	5	4.76 (5)		
Glu	8	8	3.51 (3)	1.02 (1)	3.31 (3)
Pro	3	3	1.50 (1)		1.66 (2)
Gly	4	4		1.28 (1)	2.46 (3)
Ala	1	1	1.34 (1)		
Val	1	1			0.99 (1)
Met	1	1			
Ile	2	2	0.71 (1)	1.01 (1)	0.85 (1)
Leu	1	1	0.86 (1)		1.10 (1)
Tyr	1	1	1.05 (1)		
Trp*	1	1	(1)		
Phe	0	0			
Total Residues	61[a]	60[b]	31	6	23

*Tryptophan determined spectrophotometrically. Brackets, ( ), indicated nearest integer

[a]Tu and Hong [3]
[b]From present sequence analysis.

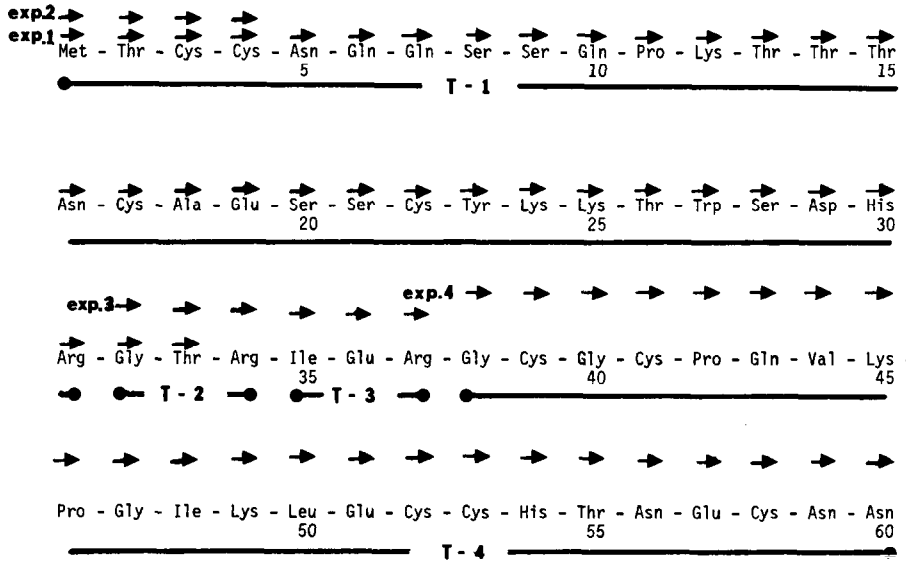

Fig.2. Complete sequence of the major toxin from *Lapemis hardwickii*. Arrows above the amino acids indicate the direction and length of each Edman degradation experiment. Sequencer programs used in experiments 1 through 4 are: 1 M Quadrol on whole CM-toxin, DMAA on T-1 peptide, DMAA on T-2 + T-3 peptide mixture, and DMAA on T-4 peptide, respectively. (•————•) represents tryptic peptides from toxin.

each tripeptide. The PTH-amino acids from the first degradation cycle of the mixture were identified as Gly and Ile. The second and third cycles yielded Thr and Gly and two Arg residues, respectively. The amino acid residues from 23 cycles of Edman degradation on peptide T-4 are shown in fig.2. This data, when considered in conjunction with the overlap provided by the partial CM-toxin sequence, positioned the two tripeptides as following T-1 in the order (T-1)–Gly–Thr–Arg–Ile–Glu–Arg. T-4 was positioned as following T-3 by comparing it with the amino acid compositions of the other tryptic peptides and in conjunction with the carboxypeptidase A digestion of the whole toxin.

The amino acid composition from the sequence yielded a total of 60 rather than 61 amino acids as previously reported [4,5]. The number of Thr and Ser residues was found to be 7 and 5, respectively. The total number of half-cystines from the sequence was 9 compared to 8 from the original amino acid composition. This is in contrast to most of the Type I neurotoxins which usually contain 8 half-cystines involved in disulfide bonds.

In the Raman spectra of alkyl thiols the S–H stretching vibrations appear in the 2570–2580 cm^{-1} region [11,12]. The Raman spectra of proteins normally originate from: C–C, C=C, C–N, C=N, C–H, C–S, C–O, C=O, N–H, O–H and S–H vibrations. Among all these vibrational modes, only the S–H vibration will appear in the 2500–2600 cm^{-1} region; therefore, very little interference and ambiguity will be involved in the detection of a sulfhydryl group. The spectrum of the native *L. hardwickii* toxin did show the typical S–H stretching vibration at 2585 cm^{-1} (fig.3). This indicated the presence of a free sulfhydryl group in addition to the 4 disulfide bonds. The previous report of 8 half-cystines [4,5] determined by the methods of reduction, aklylation and amino acid analysis was in error due to the lesser sensitivity and accuracy of these methods compared to Raman spectroscopy and sequence determination. The detection of a free sulfhydryl group in the toxin demonstrates the usefulness of Raman spectroscopy as an analytical tool in protein studies.

It is of great interest that snakes of different species and geographical origins contain similar and identical toxins. Some common toxins are apparently present in the same subfamily Hydrophiinae. For instance,

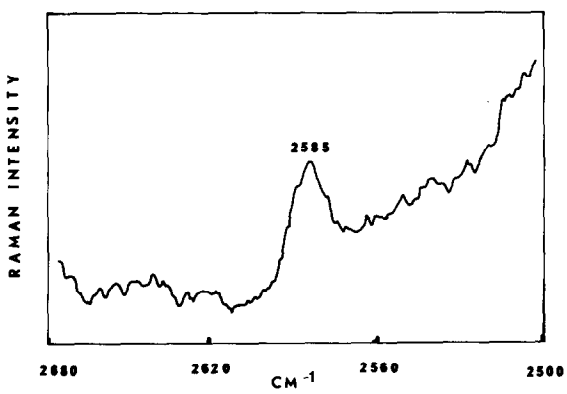

Fig.3. Laser Raman spectra of *Lapemis hardwickii* major toxin (lyophilized powder) in the region of the S–H stretching vibration.

schistosa 4 toxin of *Enhydrina schistosa* (Strait of Malacca) and the major toxin of *L. hardwickii* (Gulf of Thailand) are identical [13]. Schistosa 5 toxin of *E. schistosa* is also identical to pelamitoxin a from *Pelamis platurus* (Strait of Formosa) [13,14]. It is significant that taxonomically different snakes have identical or very similar toxins. This suggests that different species do not necessarily have totally different toxins. In view of the molecular evolution of venom proteins, this is logical because venoms of snakes from related families originate from common ancestral protein molecules (15). Some of the proteins are probably conserved in the venom very near to their ancestral sequence; however, some have naturally evolved slightly different sequences.

Acknowledgements

This investigation was supported by N.I.H. grant 5 R01 GM191172. The expert technical assistance of Dr N. T. Yu and Mr Nicholas Alonzo is most gratefully acknowledged.

Research was carried out at Brookhaven National Laboratory under contract with the US. Energy Research and Development Center.

References

[1] Lee, C. Y. (1972) Ann. Rev. Pharmacology 12, 265–286.
[2] Tu, A. T. (1973) Ann. Rev. Biochemistry 42, 235–258.
[3] Tu, A. T. (1977) Venoms: Chemistry and Molecular Biology, pp. 152–155, John Wiley, New York.
[4] Tu, A. T. and Hong, B.-S. (1971) J. Biol. Chem. 246, 2772–2779.
[5] Yu, N.-T., Lin, T.-S. and Tu, A. T. (1975) J. Biol. Chem. 250, 1782–1785.
[6] Elzinga, M. (1970) Biochemistry 9, 1365–1374.
[7] Atassi, M. Z. and Habeeb, A. F. S. A. (1972) Met. Enzymol. 25, 546–553.
[8] Braunitzer, G., Schrank, B. and Ruhfus, A. (1971) Hoppe-Seyler's Z. Chem. Bd. 352, 1730–1733.
[9] Mendez, E. and Lai, C. Y. (1975) Anal. Biochem. 68, 47–53.
[10] Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. and Ballantyne, D. L. (1971) Biochemistry 10, 4912–4921.
[11] Silverstein, R. and Bassler, G. (1966) Spectrometric Identification of Organic Compounds, pp. 56–57, John Wiley, New York.
[12] Dollish, F., Fateley, W. and Bently, F. (1973) Characteristic Raman Frequencies of Organic Compounds, pp. 46–47, John Wiley, New York.
[13] Fryklund, L., Eaker, D. and Karlsson, E. (1972) Biochemistry 11, 4633–4640.
[14] Wang, C., Liu, C., Hung, Y. and Blackwell, R. (1976) Toxicon 14, 459–466.