Hartshorne, D. J., & Mueller, H. (1969) *Biochim. Biophys. Acta 175*, 301.
Hartshorne, D. J., & Gorecka, A. (1979) *Handbook of Physiology* (in press).
Hartshorne, D. J., Gorecka, A., & Aksoy, M. O. (1977) in *Excitation-Contraction Coupling in Smooth Muscle* (Casteels, R., Godfraind, T., & Ruegg, J. C., Eds.) p 377, Elsevier, Amsterdam.
Hirs, C. H. W., Moore, S., & Stern, W. H. (1956) *J. Biol. Chem. 219*, 623.
Itzhaki, R. F., & Gill, D. M. (1964) *Anal. Biochem. 9*, 401.
Jarrett, H. W., & Penniston, J. T. (1977) *Biochem. Biophys. Res. Commun. 77*, 1210.
Kakiuchi, S., Yamazaki, R., & Nakajima, H. (1970) *Proc. Jpn. Acad. 46*, 587.
Kuo, I. C. Y., & Coffee, C. J. (1976) *J. Biol. Chem. 251*, 1603.
Lin, Y. J., & Cheung, W. Y. (1975) *Int. J. Biochem. 6*, 271.
Lin, Y. M., Liu, Y. P., & Cheung, W. Y. (1974a) *Methods Enzymol. 38*, 262.
Lin, Y. M., Liu, Y. P., & Cheung, W. Y. (1974b) *J. Biol. Chem. 249*, 4943.
Liu, Y. P., & Cheung, W. Y. (1976) *J. Biol. Chem. 251*, 4193.
Perrie, W. T., & Perry, S. V. (1970) *Biochem. J. 119*, 31.
Righetti, P. G., & Drysdale, J. W. (1974) *J. Chromatogr. 98*, 271.
Schulman, H., & Greengard, P. (1978) *Nature (London) 271*, 478.
Stevens, F. C., Walsh, M., Ho, H. C., Teo, T. S., & Wang, J. H. (1976) *J. Biol. Chem. 251*, 4495.
Teo, T. S., & Wang, J. H. (1973) *J. Biol. Chem. 248*, 5950.
Vandermeers, A., Vandermeers-Piret, M. C., Rathe, J., Kutzner, R., Delforge, A., & Christophe, J. (1977) *Eur. J. Biochem. 81*, 379.
Wang, J. H. (1977) in *Cyclic Nucleotides: Mechanisms of Action* (Cramer, H., & Schultz, J., Eds.) Vol. 5, p 179, Raven Press, New York.
Wang, J. H., Teo, T. S., Ho, H. C., & Stevens, F. C. (1975) in *Advances in Cyclic Nucleotide Research* (Drummond, G. I., Greengard, P., & Robison, G. A., Eds.) p 179, Raven Press, New York.
Watterson, D. M., Harrelson, W. G., Jr., Keller, P. M., Sharief, F., & Vanaman, T. C. (1976a) *J. Biol. Chem. 251*, 4501.
Watterson, D. M., Van Eldik, L. J., Smith, R. E., & Vanaman, T. C. (1976b) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2711.
Weber, A., & Murray, J. M. (1973) *Physiol. Rev. 53*, 612.
Yerna, M.-J., Aksoy, M., Hartshorne, D. J., & Goldman, R. D. (1978) *J. Cell Sci. 31*, 411.
Yerna, M.-J., Dabrowska, R., Hartshorne, D. J., & Goldman, R. D. (1979) *Proc. Natl. Acad. Sci. U.S.A.* (in press).

Amino Acid Sequence and Disulfide Bond Assignment of Myotoxin *a* Isolated from the Venom of Prairie Rattlesnake (*Crotalus viridis viridis*)†

Jay W. Fox, Marshall Elzinga, and Anthony T. Tu*

ABSTRACT: The primary structure of myotoxin *a*, a myotoxin protein from the venom of the North American rattlesnake *Crotalus viridis viridis*, was determined and the position of the disulfide bonds assigned. The toxin was isolated, carboxymethylated, and cleaved by cyanogen bromide, and the resultant peptides were isolated. The cyanogen bromide peptides were subjected to amino acid sequence analysis. In order to assign the positions of the three disulfide bonds, the native toxin was cleaved sequentially with cyanogen bromide and trypsin. A two-peptide unit connected by one disulfide bond was isolated and characterized, and a three-peptide unit connected by two disulfide bonds was isolated. One peptide in the three-peptide unit was identified as Cys-Cys-Lys. In order to establish the linkages between the peptides and Cys-Cys-Lys, one cycle of Edman degradation was carried out such that the Cys-Cys bond was cleaved. Upon isolation and analysis of the cleavage products, the disulfide bonds connecting the three peptides were determined. The positions of the disulfide bridges of myotoxin *a* were determined to be totally different from those of neurotoxins isolated from snake venoms. The sequence of myotoxin *a* was compared with the sequences of other snake venom toxins using the computer program RELATE to determine whether myotoxin *a* is similar to any other types of toxins. From the computer analysis, myotoxin *a* did not show any close relationship to other toxins except crotamine from the South American rattlesnake *Crotalus durissus terrificus*.

Muscle degeneration is commonly observed upon rattlesnake envenomation. Rattlesnake venoms exhibit high levels of proteolytic activity and therefore it has been suggested that the myonecrotic effects of the venom were the results of the action of venom proteases (Porges, 1953). However, a heat-stable component of *Timersurus flavoviridis* venom and a basic polypeptide from black widow spider venom (Lee et al., 1974) (neither of which has proteolytic activity) have been shown to elicit myonecrotic effects. Thus, venom-induced myonecrosis appears to be the result of more complex action than that attributable solely to venom proteases. It is likely that myonecrosis results from both independent and synergistic effects of venom components.

† From the Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523 (J.W.F. and A.T.T.), and the Department of Biology, Brookhaven National Laboratory, Upton, New York 11973 (M.E.). *Received July 21, 1978.* This work is based on the dissertation submitted by J. W. Fox in partial fulfillment of the degree of Ph.D. at Colorado State University. This work was supported by grants from the National Institutes of Health, 2 R01 GM15991 (A.T.T.) and HL-21471 (M.E.). The research carried out at Brookhaven National Laboratory was under contract with the U.S. Energy Research and Development Administration.

0006-2960/79/0418-0678$01.00/0

Recently the presence of two myonecrotic fractions has been demonstrated in the venom of the North American prairie rattlesnake (*Crotalus viridis viridis*); one of these, myotoxin *a*, has been isolated and characterized (Cameron & Tu, 1977). The molecular weight and isolectric point were estimated to be 4100 and 9.6, respectively. The mode of muscle degeneration by myotoxin *a* was examined by electron microscopy and it was found to damage the endoplasmic reticulum as the primary target and it eventually effected a disorganization of muscle filaments (Ownby et al., 1976).

Amino acid sequence analyses provide a variety of information including the possible reconstruction of genetic events involved in the evolutionary history of the species and the protein structure inherent in its sequence (Wu et al., 1974). In order to expand our understanding of the function of myotoxin *a* and compare it with other myonecrotic and neurotoxic peptides from snake venoms, we deemed that determination of its covalent structure was essential. From the sequence analysis and the position of the disulfide bonds, consideration of the genetic and evolutionary relationships among myotoxin *a* and similar peptides could be undertaken.

Methods and Materials

Preparation of Myotoxin a. Lyophilized crude venom was purchased from Miami Serpentarium Laboratories. Myotoxin *a* was purified by a modification of the procedure of Cameron & Tu (1977). The homogeneity of the protein was confirmed by polyacrylamide gel electrophoresis on β-alanine gels. The modification was such that all isolation steps were performed at pHs of 6 or less. Aliquots of the purified myotoxin *a* were subjected to performic acid oxidation (Moore, 1963). The oxidized toxin was hydrolyzed with 6 N hydrochloric acid at 110 °C, followed by amino acid analysis. Alkylation of myotoxin *a* was performed as described by Bamburg et al. (1973). The alkylated myotoxin *a* (CM myotoxin *a*) was digested with cynogen bromide (CNBr), and the resultant peptides were isolated by gel filtration on Sephadex G-50.

Sequence Determination. Sequential Edman degradations were performed on a Beckman 890C sequencer operated with either the 1.0 M Quadrol or dimethylallylamine (DMAA)[1] program. The sequencer run on CM myotoxin *a* was carried out following prior reaction with Braunitzer reagent III (3-isothiocyano-1,5-naphthalenedisulfonic acid), as described previously, in order to reduce extractive losses (Fox et al., 1977). The anilinothiazolinones were converted to phenylthiohydantoins (Pth) and identified by two independent methods: thin-layer chromatography (Laursen, 1971) and amino acid analysis following regeneration of free amino acids by hydrolysis in 6 N HCl + 0.1% $SnCl_2$ at 150 °C for 4 h or hydrolysis in 0.01 M NaOH at 150 °C for 4 h (Mendez & Lai, 1975).

Disulfide Bond Assignment. Inspection of the sequence indicated that tryptic digestion of myotoxin *a* would yield a variety of products including a two-peptide unit connected by one disulfide bond, and a three-peptide unit connected by two disulfides. Initial studies indicated the best way to digest native myotoxin *a* was to first cleave the toxin at residue 28 (Met) with CNBr, which apparently opens the molecule to tryptic action. Upon isolation and characterization of the two-peptide unit, the first disulfide assignment could be made. In order to establish the disulfide bonds within the three-peptide unit, one cycle of Edman degradation would be performed. Separation and identification of the peptides that resulted would permit the assignment of the disulfide bonds.

Trypsin (TRTPCK) and soybean trypsin inhibitor were purchased from Worthington, Freehold, NJ. Tryptic digestions were carried out at an enzyme to substrate ratio of 1:15 in 0.2 M Pipes buffer at pH 6.5 for 24 h at 37 °C. The digestion was terminated by the addition of soybean trypsin inhibitor (trypsin to inhibitor ratio 1:2). Glacial acetic acid was added to make the reaction mixture 25% in acetic acid.

The tryptic digestion mixture was loaded onto a Sephadex G-50 column equilibrated with 25% acetic acid and gel filtration carried out as described earlier (Fox et al., 1977). The elution of the peptides was monitored by their absorbance at 280 nm and by estimating the quantity of ninhydrin positive material in the tubes using a cadmium–ninhydrin paper technique as described by Elzinga & Collins (1975). Absorbance at 570 nm was also followed after prior reaction of an alkaline-hydrolyzed aliquot with ninhydrin (Hirs, 1967). The peaks were hydrolyzed with 6 N HCl for 22 h at 110 °C for amino acid analysis.

Aliquots of both the performic acid oxidized and unoxidized two-peptide unit were dried onto a sheet of Whatman no. 3 MM paper in preparation for high voltage electrophoresis in 7% (v/v) formic acid. A thin strip of the sheet was developed with the cadmium–ninhydrin stain to locate the peptide spots. The peptides on the remainder of the sheet were eluted with 1% acetic acid, hydrolyzed in 6 N HCl, and subjected to amino acid analysis.

One cycle of Edman degradation (DMAA program) was then performed on the three-peptide unit. The cycle was terminated prior to the butyl chloride extraction step; the reaction mixture was removed from the sequencer cup and gel filtered to separate the peptides from the reagents, and the peptides were isolated by chromatography on SP-Sephadex C-25. The pure peptides were oxidized with performic acid, hydrolyzed for 22 h at 110 °C, and subjected to amino acid analysis.

Computer Comparison of Myotoxin a Sequence with Other Toxin Sequences. The computer program RELATE (Schwartz et al., 1975) was used in an attempt to detect relationships between myotoxin *a* and members of the superfamily snake venom toxins as grouped together by Hunt (Hunt & Dayhoff, 1976). The program compared 15 residue segments of the myotoxin with all fragments of 15 residues from one of the other toxin sequences. From these comparisons a distance score is derived which gives a measure of relatedness between the two sequences compared. The toxins with which myotoxin *a* was compared are long neurotoxins (type II), Formosan banded krait (*Bungarus multicinctus*), Indian cobra (*Naja naja*); short neurotoxins (type I), Cape cobra (*Naja nivea*), common sea snake (*Enhydrina schistosa*), broad-banded blue sea snake (*Laticauda semifasciata*), banded Egyptian cobra (*Naje haje*). Myotoxin *a* was also compared with the sequence of crotamine (Laure, 1975) which is not presently listed in a family. Cytotoxins are Cambodian cobra (*Naja naja*) and Indian cobra (*Naja naja*). (For detailed explanation of the RELATE program or the sequences of the other toxins, see Hunt & Dayhoff, 1976.)

Results

Characterization of Myotoxin a and Cyanogen Bromide Peptides. The amino acid composition of performic acid oxidized myotoxin *a* is shown in Table I. Sequence analysis of the Braunitzer reagent III coupled intact CM-myotoxin *a* was carried out for 34 cycles with the 1.0 M Quadrol program.

[1] Abbreviations used: DMAA, dimethylallylamine; Pth, phenylthiohydantoins; Pipes, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); SP, sulfopropyl; BuCl, butyl chloride.

Table I: Apparent Amino Acid Composition of Oxidized Myotoxin *a* and CNBr Peptides[a]

	myotoxin *a*			CB-I		CB-II	
Lys	9.68	(10)	(10)	5.55	(6)	3.69	(4)
His	1.76	(2)	(2)	1.77	(2)	0	
Arg	0.96	(1)	(1)	0.13		0.87	(1)
CMC			(4)	2.48	(3)	2.61	(3)
Cys acid	5.56	(6)					
Asp	2.10	(2)	(2)	1.02	(1)	0.96	(1)
Met			(1)				
Met-SO_2	0.89	(1)					
Thr	0						
Ser	2.64	(3)	(3)	1.62	(2)	0.66	(1)
Glu	1.89	(2)	(2)	2.12	(2)	0	
Pro	2.94	(3)	(3)	2.92	(3)	0	
Gly	5.06	(5)	(5)	2.98	(3)	0	
Ala	0			0		0	
Val	0			0		0	
Ile	1.76	(2)	(2)	1.73	(2)	0	
Leu	0.97	(1)	(1)	1.01	(1)	0	
Tyr	0.61	(1)	(1)	0.86	(1)	0	
Trp		(2)[d]	(1)	ND[e]		ND	(2)[f]
Phe[g]	0.82	(1)	(1)	0.89	(1)	0	
Hse				0.96	(1)	0	
total residues		42[b]	39[c]		28		14

[a] Samples were hydrolyzed for 22 h in 6 N HCl at 110 °C. Numbers in parentheses indicate integer value. [b] From present amino acid analysis. [c] Cameron & Tu, 1977. [d] Tryptophan determined spectrophotometrically. [e] ND, not determined. [f] Estimated value. [g] Homoserine.

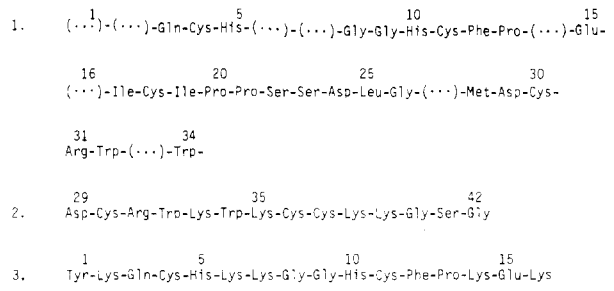

FIGURE 1: Sequence analysis of myotoxin *a*. (1) One molar Quadrol program degradation of CM-myotoxin *a*; (2) DMAA program degradation CB-III peptide; (3) DMAA program degradation of CM-myotoxin *a*. (···) denotes Pth not recovered.

The identified residues are shown in Figure 1 and Table II. Because of the initial reaction of the toxin with Braunitzer reagent III, no sequence information was obtained for the NH_2 terminus and residues 2, 6, 7, 14, 16, 27, and 33.

Sixteen cycles of degradation of unmodified CM-myotoxin *a* with the DMAA program permitted identification of residues 1, 2, 6, 7, 14, and 16 (see Figure 1). Residues 27 and 33 were identified as Lys from the amino acid composition of subsequently purified tryptic peptides.

Gel filtration of the cyanogen bromide digest is illustrated in Figure 2. From the amino acid compositions, peaks CB-I and CB-II were determined to be the NH_2-terminal and COOH-terminal cyanogen bromide peptides, respectively (Table I).

A sequencer run on the peptide in peak CB-II (DMAA program) yielded the sequence of the COOH-terminal peptide. This in conjunction with the other sequence information allowed for the assignment of the complete sequence of the toxin.

Disulfide Bond Assignments. Gel filtration of the tryptic digest of native myotoxin *a* is illustrated in Figure 3. From the amino acid analyses, the five peaks were identified:

T-I: Trypsin-soybean trypsin inhibitor.

T-II: Undigested myotoxin *a*.

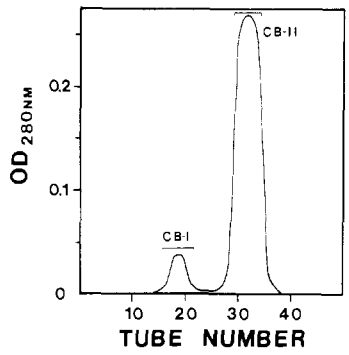

FIGURE 2: Gel filtration of a CNBr digest of myotoxin *a* on Sephadex G-50-80. Column size was 200 × 1.9 cm, 23 °C, and the solvent was 25% (v/v) acetic acid. Fraction size was 6.5 mL and a 40-mL forefraction was collected.

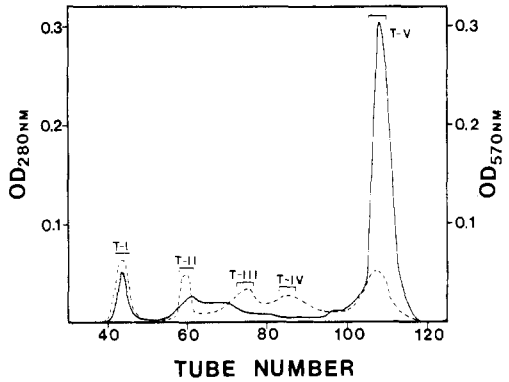

FIGURE 3: Gel filtration of a tryptic digest of CNBr myotoxin *a* on Sephadex G-50-80. Conditions were identical with those described in the legend to Figure 2. (Solid line) OD_{280nm}; (dashed line) OD_{570nm}.

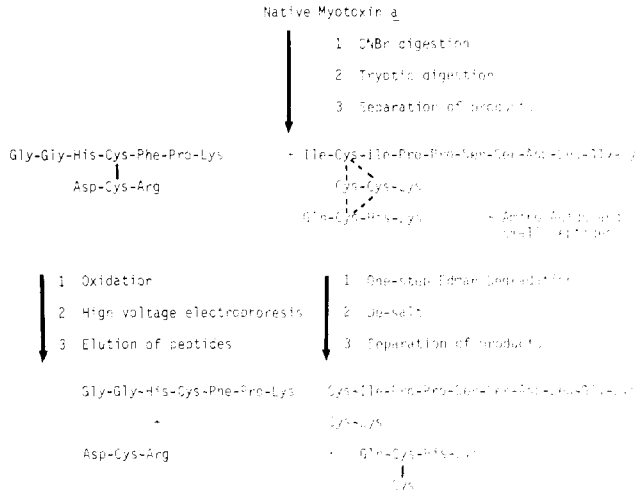

FIGURE 4: Scheme of disulfide bond assignment. The peptides were identified by amino acid analysis and comparison of the compositions with the sequence of the toxin. Dashed lines indicate disulfide bond assignments unknown at that stage of analysis.

T-III: Comparison of the amino acid composition of the peak (Table III) with the sequence of the toxin indicated that the peak contained three peptides connected by two disulfide bonds (Figure 4).

T-IV: The amino acid composition of this peak indicated the presence of the two-peptide unit with one disulfide bridge (Figure 4).

T-V: The amino acid composition of this peak (Table IV) was consistent with the presence of the free lysine and the small peptides that were expected in the tryptic digest, based upon

Table II: Automated Edman Degradation[a]

CM myotoxin a[b]						CB-II[c]			CM myotoxin a[d]		
step	amino acid	% yield	step	amino acid	% yield	step	amino acid	% yield	step	amino acid	% yield
1	(Tyr)[e]		18	Cys[g]	30	1	Asp	98	1	Tyr	99
2	(Lys)		19	Ile	22	2	Cys[g]	69	2	Lys	76
3	Gln[f]	68	20	Pro	30	3	Arg	50	3	Gln[f]	68
4	Cys[g]	60	21	Pro	27	4	Trp[j]	52	4	Cys[g]	49
5	His	34	22	Ser[j]	17	5	Lys	41	5	His	42
6	(Lys)		23	Ser[i]	13	6	Trp[j]	32	6	Lys	39
7	(Lys)		24	Asp	25	7	Lys	28	7	Lys	38
8	Gly	42	25	Leu	20	8	Cys[g]	28	8	Gly	29
9	Gly	40	26	Gly	19	9	Cys[g]	24	9	Gly	27
10	His	27	27	(Lys)		10	Lys	18	10	His	20
11	Cys[g]	37	28	Met	15	11	Lys	15	11	Cys[g]	24
12	Phe	35	29	Asp	ND[k]	12	Gly	16	12	Phe	21
13	Pro	32	30	Cys[g]	ND	13	Ser[j]	12	13	Pro	20
14	(Lys)		31	Arg	ND	14	Gly	14	14	Lys	19
15	Glu	33	32	Trp[j]	ND				15	Glu	17
16	(Lys)		33	(Lys)					16	Lys	16
17	Ile[h]	27	34	Trp[j]	ND						

[a] Amino acids recovered after $HCl \cdot SnCl_2$ hydrolysis of Pth's obtained from sequencer. [b] 1 M Quadrol program degradation on 500 nmol of CM myotoxin a reacted with Braunitzer reagent. [c] DMAA program degradation on 200 nmol of CB-II peptide. [d] DMAA program degradation on 250 nmol of CM myotoxin a. [e] Parentheses denote Pth-amino acid not recovered. [f] Gln residues recovered as Glu. [g] Recovered as alanine. [h] Recovered as isoleucine and alloisoleucine. [i] Recovered as alanine. [j] Recovered as glycine and alanine. [k] ND, not determined.

Table III: Amino Acid Composition of Tryptic Peptides from CNBr Digested Myotoxin a[a]

	T-III	T-IV	T-V
Lys	2.95 (3)	0.67 (1)	5.88 (6)
His	0.90 (1)	0.98 (1)	0
Arg	0	1.02 (1)	0
Cys acid	3.67 (4)	1.97 (2)	0
Asp	1.01 (1)	0.99 (1)	0
Thr	0	0	0
Ser	1.44 (2)	0	0.46 (1)
Glu	1.05 (1)	0	0.72 (1)
Pro	1.74 (2)	0.77 (1)	0
Gly	0.96 (1)	2.10 (2)	2.11 (2)
Ala	0	0	0
Val	0	0	0
Ile	1.99 (2)	0	0
Leu	1.05 (1)	0	0
Tyr	0	0	0.57 (1)
Trp	ND[c]	ND	ND (2)[b]
Phe	0	0.98 (1)	0
Hse	0	0	0.68 (1)
total residue	18	10	14

[a] Footnote same as Table II. [b] Estimated value. [c] ND, not determined.

the overall sequence. This result indicates that the tryptic digestion proceeded satisfactorily.

High voltage electrophoresis of T-IV after performic acid oxidation yielded two ninhydrin-positive spots. The amino acid compositions of the two eluted spots (HV-I, HV-II) are seen in Table IV. They coincide with those of the two tryptic peptides, Gly-Gly-His-Cys-Phe-Pro-Lys and Asp-Cys-Arg, which therefore must be connected by a disulfide bond in the native toxin.

Gel filtration of the tryptic digest G-50 peak T-III (three-peptide unit) following the single sequence step in which the Cys-Cys bond in one of the three peptides was cleaved is shown in Figure 5. This served to separate the peptides, peak G-I from the reagents, and by-products, peak G-II (Figure 5 and Table IV). The absorbance at 280 nm was due to anilinothiazolinones.

The SP-Sephadex chromotograph of G-1 is shown in Figure 6. Amino acid analyses of these peaks are presented in Table

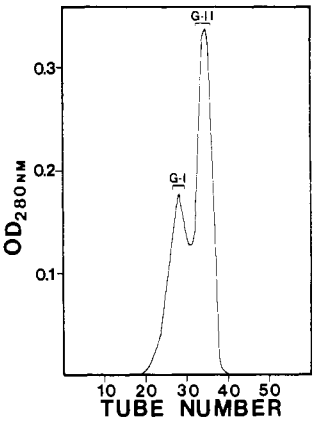

FIGURE 5: Gel filtration of three-peptide unit following one step of Edman degradation, on Sephadex G-50-80. Column size was 200 × 0.8 cm, 23 °C, and the solvent was 25% (v/v) acetic acid. Fractions (4 mL) were collected.

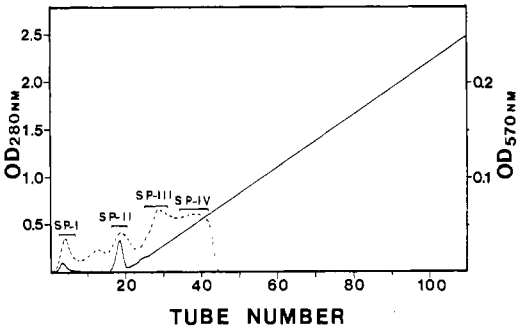

FIGURE 6: Chromatography of Sephadex G-50 peak G-1 on a 0.9 × 13 cm column of SP-Sephadex C-25. Fractions (3.5 mL) were collected. A gradient in pyridine concentration was established using reservoirs containing 200 mL of 25% (v/v) acetic acid and 200 mL of 10:25:65 (v/v/v) of pyridine, acetic acid, and water. Solid line is OD_{280nm}; dashed line is OD_{570nm}.

IV. The composition of SP-III in conjunction with the sequence data for the toxin identified the peptide as

Cys
|
Gln-Cys-His-Lys

Gln was not lost during the Edman degradation because it

Table IV: Amino Acid Compositions of Peptides from Various Steps in Disulfide Bond Analysis[a]

	HV-I[b]	HV-II	G-I[d]	SP-II	SP-III	SP-IV
Lys	0.94 (1)	0.28	2.66 (3)	1.22 (1)	0.99 (1)	1.80 (2)
His	0.95 (1)	0.22	0.75 (1)		0.87 (1)	0
Arg	0	0.78 (1)	0		0	0
Cys acid	1.10 (1)	1.11 (1)	3.72 (4)	1.2 (1)	1.71 (2)	1.56 (2)
Asp	0	0.98 (1)	1.21 (1)	0.72 (1)	0	1.24 (1)
Thr	0	0	0	0	0	0
Ser	0	0	1.57 (2)	0	0	1.67 (2)
Glu	0	0	0.84 (1)	0.45	1.41 (1)	0.18
Pro	0.94 (1)	0	1.72 (2)	1.82 (2)	0.21	2.13 (2)
Gly	2.11 (2)	0	1.08 (1)	3.21 (3)	0.13	1.48 (1)
Ala	0	0	0	0	0	0
Val	0	0	0	0	0	0
Ile	0	0	0.99 (1)	1.11 (1)	0	1.07 (1)
Leu	0	0	0.85 (1)	0.97 (1)	0.10	0.97 (1)
Tyr	0	0	0	0	0	0
Trp	ND[c]	ND	ND	ND	ND	ND
Phe	0.90 (1)	0	0	0	0	0
total residues	7	3	17	10	5	12

[a] Footnote same as Table II. [b] HV denotes peptides eluted after high voltage paper electrophoresis. [c] ND, not determined. [d] G-1 denotes G-50 desalting column peak 1 following one step of Edman degradation.

Table V: Computer Analysis Sequence Comparison Scores[a] between Myotoxin *a* and Various Snake Venom Toxins

family[b]	snake	real mean	random mean	SD	distance
long neurotoxins (type II)	Formosan banded krait	20.429	17.807	3.967	0.661
	Indian cobra	17.280	16.670	3.800	0.153
short neurotoxins (type I)	Cape cobra	15.321	17.328	3.565	−0.563
	common sea snake	12.357	17.420	3.312	−1.528
	broad-banded blue sea snake	17.107	16.345	3.657	0.208
	banded Egyptian cobra	18.429	16.074	3.350	0.703
cytotoxins	Cambodian cobra	16.036	14.218	3.521	0.516
	Indian cobra	16.679	14.326	3.758	0.626
	crotamine[c]	93.464	15.517	3.893	20.023

[a] Scores determined using computer program **RELATE** (Schwartz et al., 1975). [b] Families grouped as per Dayhoff, 1976. [c] Crotamine from the South American rattlesnake has not been assigned to a sequence family.

had cyclized to pyrrolidonecarboxylic acid and its amino group was therefore unavailable for reaction. Similar consideration of SP-IV identified the peptide in this peak as

Cys-Ile-Pro-Pro-Ser-Ser-Asp-Leu-Gly-Lys
|
Cys-Lys

The isoleucine which was the original NH_2 terminus of the 11-residue peptide was removed as expected by the Edman degradation.

These results indicated the disulfide arrangement in the three-peptide group to be

Ile-Cys-Ile-Pro-Pro-Ser-Ser-Asp-Leu-Gly-Lys
|
Cys-Cys-Lys
|
Gln-Cys-His-Lys

The complete disulfide assignments are illustrated in Figure 7.

Sequence Comparisons by Computer Analysis. Table V shows the values for the real mean, random mean, standard deviation, and sequence comparison score (distance) between myotoxin *a* and members of various snake venom toxin families. The real mean is the mean high fragment score. The random mean is the mean of comparisons in which myotoxin *a* is compared with the sequence which has been scrambled. The standard deviation of mean real score from the random

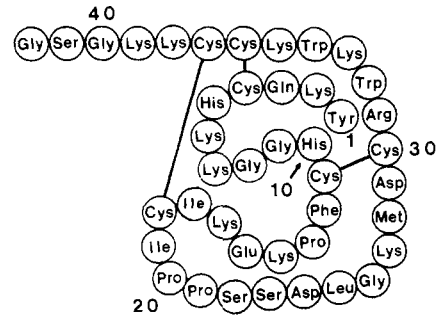

FIGURE 7: Scheme of disulfide bond arrangement in myotoxin *a*.

comparisons is calculated. From this a statistical equal to the number of standard deviations of the real mean score above the mean random score is calculated. This is termed distance in Table V and comes from

$$\frac{\text{real mean} - \text{random mean}}{\text{standard deviation}} = \text{distance}$$

As seen from Table V there is no close relationship between myotoxin *a* and any of the members of the various families of superfamily snake venom toxins. Myotoxin *a* is therefore unrelated to the toxins found in Elapidae and Hydrophiidae venoms. The only sequence with which a high distance score

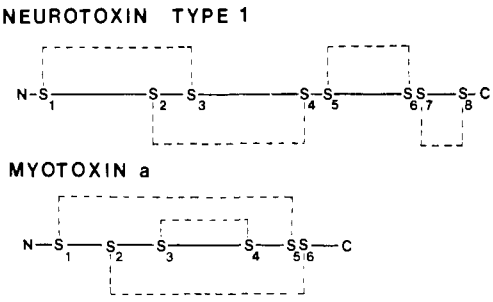

FIGURE 8: Comparison of the relative positions of disulfide bonds in myotoxin *a* and type I neurotoxins from Elapidae and Hydrophiidae venoms.

was found was crotamine, a toxin from the South American rattlesnake *Crotalus durissus terrificus*, which is identical with the toxin studied here at all but three positions.

Discussion

Upon comparison of the complete sequence and disulfide bond arrangement of myotoxin *a* with type I (short) snake neurotoxins, it is apparent that they are unrelated. Myotoxin *a* and the neurotoxins both have high contents of basic amino acids in their sequences and this is reflected by their alkaline isoelectric points (>9.0); however, their molecular weights are quite different—4621 vs. ~6800 for type I neurotoxins. Furthermore, the arrangement of the three disulfide bonds of myotoxin *a* determined in this work (Figure 7) is not homologous to those of the neurotoxins (see Figure 8).

The sequence of myotoxin *a* is very similar to that of crotamine. Crotamine is a relatively weak, basic nonneurotoxic protein comprising 42 amino acids; it is isolated from the venom of the South American rattlesnake *Crotalus durissus terrificus*. Both myotoxin *a* and crotamine show hemolytic and myotoxic activity. The myotoxin *a* is observed at an ultrastructural level to cause vacuolation of skeletal muscle fiber and dilation of the sarcoplasmic reticulum (Ownby et al., 1976). The main physiological response of muscle to crotamine is a depolarization of muscle fiber. Crotamine acts in the muscle fiber membrane effecting a change in Ca^{2+} or Na^+ ion permeability causing a K^+ ion efflux (Cheymol et al., 1971a,b). Similar physiological experiments have not been performed with myotoxin *a*. From the amino acid composition, circular dichroic spectra, and the biological activities mentioned above, Cameron & Tu (1978) concluded that myotoxin *a* and crotamine are different toxins; however, they are chemically and functionally homologous. The present study indicates that myotoxin *a* differs from crotamine (Laure, 1975) at three positions: myotoxin → crotamine, Ile (19) → Leu, Leu (25) → Phe, and Lys (33) → Arg. These replacements are chemically conservative substitutions.

The determination of the covalent structure of myotoxin *a* from this investigation will facilitate the assignment of secondary and tertiary structure of myotoxin *a* serving to further the understanding of its mechanism of action.

From the comparison of myotoxin *a* and crotamine it is of interest that two genotypically different snakes from different geographical habitats contain strikingly similar toxins in their venoms. This suggests a close evolutionary relationship between the two snakes.

Comparison of the sequence of myotoxin *a* with various members of the snake toxin sequence families permits an evaluation of the structural and evolutionary relationships among these proteins. An evaluation of these possible relationships was carried out by computer analysis. The sequence comparison score (distance) between myotoxin *a* and crotamine served as a basis for comparison of myotoxin *a* with members of the snake venom sequence families, since myotoxin *a* and crotamine have nearly identical sequences. Therefore, a relatively high distance score would be expected. As seen in Table V the distance score between the two is 20.023. The distance scores between myotoxin *a* and representative toxins of the long neurotoxin, short neurotoxin, and cytotoxin sequence families are very low, indicating that no close relationships between myotoxin *a* and any of those sequence families exist (Barker, 1978).

The computer comparison of the sequence of myotoxin *a* and other snake venom toxin sequence families did not indicate any relationships among them. Several other basic proteins have been isolated from the venoms of other rattlesnake species (Sulkowski et al., 1975; Bonilla & Fiero, 1971). Also, other basic proteins of low toxicity have been isolated from snake venoms of which there is some question about the exact nature of their mode of action (Tu, 1977).

Further computer sequence evaluations of these basic nonneurotoxic proteins, once their sequences are determined, may indicate that these proteins are related to one another. Then they could possibly be considered as comprising their own sequence family under the sequence superfamily of snake venom toxins proposed by Hunt & Dayhoff (1976). These types of evaluations in conjunction with biochemical investigations should further the understanding of the structural, functional, and evolutionary relationships among snake toxins.

Acknowledgments

We wish to acknowledge the excellent technical assistance of Nicholas Alonzo in carrying out the amino acid analyses. We also express thanks to Dr. Winona Barker for her assistance with the computer analyses.

References

Bamburg, J. R., Shooter, E. M., & Wilson, L. (1973) *Neurobiology 3*, 162–168.

Barker, W. (1978) personal communication.

Bonilla, C. A., & Fiero, M. K. (1971) *J. Chromatogr. 56*, 253–263.

Cameron, D. L., & Tu, A. T. (1977) *Biochemistry 16*, 2546–2553.

Cameron, D. L., & Tu, A. T. (1978) *Biochim. Biophys. Acta 532*, 147–154.

Cheymol, J., Gonçalves, J. M., Bourillet, F., & Roch-Arveiller, M. (1971a) *Toxicon 9*, 279–286.

Cheymol, J., Gonçalves, J. M., Bourillet, F., & Roch-Arveiller, M. (1971b) *Toxicon 9*, 287–289.

Dayhoff, M. O. (1976) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, pp 9–15, National Biomedical and Research Foundation, Washington, D.C.

Elzinga, M., & Collins, J. H. (1975) *J. Biol. Chem. 250*, 5897–5905.

Fox, J. W., Elzinga, M., & Tu, A. T. (1977) *FEBS Lett. 80*, 217–220.

Hirs, C. H. W. (1967) *Methods Enzymol. 11*, 325–330.

Hunt, L. T., & Dayhoff, M. O. (1976) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 2, pp 147–164, National Biomedical Research Foundation, Washington, D.C.

Laure, C. J. (1975) *Hoppe-Seylers Z. Physiol. Chem. 356*, 213, 215.

Laursen, R. A. (1971) *Eur. J. Biochem. 20*, 89–102.

Lee, C. K., Chan, T. K., Ward, B. C., Howell, D. E., & O'Dell, G. V. (1974) *Arch. Biochem. Biophys. 164*, 341–350.

Mendez, E., & Lai, C. Y. (1975) *Anal. Biochem. 68*, 47–53.
Moore, S. (1963) *J. Biol. Chem. 238*, 235–329.
Ownby, C. L., Cameron, D. L., & Tu, A. T. (1976) *Am. J. Pathol. 85*, 149–158.
Porges, N. (1953) *Science 117*, 47–51.
Schwartz, R. M., Barker, W. C., & Dayhoff, M. O. (1975) *NBR Report No. 08710-750215*, National Biomedical Research Foundation, Washington, D.C.
Sulkowski, E., Kress, L. F., & Kaskowski, M., Sr. (1975) *Toxicon 13*, 149–157.
Tu, A. T. (1977) *Venoms: Chemistry and Molecular Biology*, pp 218–219, Wiley, New York.
Wu, T. T., Fitch, W. M., & Margoliash, E. (1974) *Annu. Rev. Biochem. 43*, 539–566.

Mercury Binding to Hemerythrin. Coordination of Mercury and Its Effects on Subunit Interactions†

S. E. Clarke, L. C. Sieker, R. E. Stenkamp, and J. S. Loehr*

ABSTRACT: The reaction of hemerythrin, an oxygen-carrying protein in marine invertebrates, with K_2HgI_4 was investigated to determine the mode of mercurial binding to the protein in the heavy atom derivative used for solving the crystal structure of *Themiste dyscritum* hemerythrin. Spectrophotometric titrations established that the *T. dyscritum* protein reacts with approximately 1.5 molecules of K_2HgI_4 per protein subunit and that the reaction involves mercury–sulfur bond formation. Crystallographic studies using difference density maps showed that one mercury is coordinated to the sulfur of cysteine-50 as HgI and that a second mercury is shared between two cysteine-9 sulfurs of adjacent subunits, resulting in a linear S–Hg–S coordination. Gel filtration chromatography of K_2HgI_4-treated hemerythrin demonstrated that the protein retains its octameric structure in solution after reacting with K_2HgI_4, as had been observed in the crystalline state. The roles of the cysteine-9 and -50 residues in the subunit interactions of *T. dyscritum* hemerythrin were determined from the effects of the bulkier sulfhydryl reagents, *p*-hydroxymercuribenzoate and *N*-ethylmaleimide. Binding of either reagent to cysteine-50 resulted in the disruption of all subunit interactions and the conversion of the native octamers to monomers. Binding of either reagent to cysteine-9 caused the isolated monomers to reassociate as dimers. The critical effect of cysteine-50 modification on subunit interactions was found to be due to its proximity to the major points of subunit contact, as seen in the crystal structure of the protein.

The nature of the interaction of mercury compounds with proteins is of interest from a number of points of view. Mercury binding to proteins and consequent poisoning of active sites or disruption of protein structure play an important role in the toxic effects of mercurials (Bremner, 1974). As protein sulfhydryl groups are the primary targets of mercury binding, reactions with mercurials are useful for elucidating the involvement of cysteine sulfhydryl groups in enzymatic catalysis and in the maintenance of tertiary and quaternary structure (Webb, 1966). On the other hand, protein crystallography depends on mercurials which do not disrupt protein structure in order to obtain heavy atom derivatives by the isomorphous replacement technique (Blundell & Johnson, 1976). Thus, the effects of mercury compounds vary considerably depending on the nature of the mercurial and the protein to which it binds.

One protein which has been shown to undergo dissociation upon treatment with sulfhydryl reagents is hemerythrin, a respiratory protein found in marine invertebrates. Hemerythrin from the sipunculid, *Phascolopsis gouldii*, has been studied in considerable detail (Klotz, 1971). The native, octameric protein can be dissociated by reaction with sulfhydryl reagents such as *p*-hydroxymercuribenzoate (PHMB),[1] *N*-ethylmaleimide (NEM), or salyrganic acid (Kereztes-Nagy & Klotz, 1963). As the protein contains only a single cysteine at residue 50, this must be the residue whose modification results in subunit dissociation.

Hemerythrin from *Themiste dyscritum* appears to be very similar to *P. gouldii* hemerythrin in its molecular weight and subunit composition (Loehr et al., 1975), amino acid sequence (Loehr et al., 1978), spectroscopic properties (Dunn et al., 1977), and protein conformation as judged by X-ray crystallography at approximately 5-Å resolution (Stenkamp et al., 1976; Ward et al., 1975). However, in addition to the cysteine at position 50, there is a second cysteine at position 9 in the polypeptide chain (Loehr et al., 1978). The crystal structure of K_2HgI_4-treated hemerythrin at 2.8-Å resolution indicates that the isomorphous derivative has one mercury binding site close to cysteine-50 on each subunit and another mercury site close to cysteine-9 residues on adjacent subunits related by noncrystallographic twofold axes (Stenkamp et al., 1978b).

The above information leads to the conclusion that, if K_2HgI_4 has reacted with cysteine sulfurs in *T. dyscritum* hemerythrin, it has not caused any disturbance of the protein structure, in contrast to the results observed with mercurials like PHMB (Kereztes-Nagy & Klotz, 1963). However, mercurials used in protein crystallography are quite often bound to groups other than cysteine (Blundell & Johnson, 1976). The present study was undertaken to determine

† From the Department of Chemistry, Portland State University, Portland, Oregon 97207 (S.E.C. and J.S.L.), and the Departments of Biological Structure and Biochemistry, University of Washington, Seattle, Washington 98195 (R.E.S. and L.C.S.). *Received August 3, 1978.* This work was supported by the U.S. Public Health Service, National Institutes of Health, Grants GM 18865 and AM 3288.

[1] Abbreviations used: PHMB, *p*-hydroxymercuribenzoate; NEM, *N*-ethylmaleimide; Tris, 2-amino-2-hydroxymethyl-1,3-propanediol.

0006-2960/79/0418-0684$01.00/0