SNAKE VENOM TOXINS

THE PRIMARY STRUCTURES OF TWO NOVEL CYTOTOXIN HOMOLOGUES FROM THE VENOM OF FOREST COBRA (*NAJA MELANOLEUCA*)

Fritz H.H. Carlsson

National Chemical Research Laboratory, Council for Scientific and Industrial Research, Pretoria 0001, Republic of South Africa

Received May 9,1974

SUMMARY

The primary structures of $V^{II}2$ and $V^{II}3$, two closely related basic proteins of low toxicity from *Naja melanoleuca* venom are described. Both consist of single 61-residue chains cross-linked intramolecularly by four disulphide bridges. They differ by a lysine-asparagine substitution at position 36, $V^{II}3$ being the more basic. A high degree of homology with the cytotoxin (cardiotoxin) group is evident. The unusual features of the sequences when compared to the cytotoxins are, the radical substitutions at positions 19, 25, 34, 36 and 41 and particularly the sequence of residues 38-44 *i.e.* Gly-Cys-Ala-Ala-Thr-Cys-Pro which occurs at the corresponding position in all the 71-74 (long) neurotoxins. The possible implication of certain amino acid substitutions is discussed.

The cytotoxins (cardiotoxins) are 60/61 residue polypeptides which have been found in the venoms of the elapid genera *Naja* and *Hemachatus*. They are highly basic proteins containing a high proportion of hydrophobic amino acids arranged in prominent blocks. The eight half-cystine residues form four intramolecular disulphide bridges. They make up a considerable proportion of the weight of dry venom and are sometimes the chief protein constituents of a venom. Their toxicities (intravenous LD_{50}) vary, but are usually numerically an order of magnitude greater than those of the neurotoxins when assayed in mice. Their biochemistry and pharmacology have been extensively reviewed (1,2).

In our studies on the low toxicity components of *Naja melanoleuca* venom, three basic proteins designated $V^{II}1$, $V^{II}2$ and $V^{II}3$ have been

isolated and the primary structure of the major protein, $V^{II}1$, described (3). The present study reports the amino acid sequences of the two minor cytotoxin type proteins, $V^{II}2$ and $V^{II}3$.

MATERIALS AND METHODS

The experimental techniques for sequence determination were all as previously detailed (3) except that a modified Edman-Begg degradation was used for manual sequencing of peptides (4). Further modifications to the procedure were the use of 3-dimethylamino-1-propyne as coupling buffer (5) and three extractions with 5:1 v/v n-hexane-benzene replacing the single benzene extraction after the coupling stage. This latter modification minimized losses of peptides at this stage and greatly facilitated the sequencing of hydrophobic peptides such as T-9 ($V^{II}2$) and T-8 ($V^{II}3$) which was impossible using the normal procedure. The identification procedures for phenylthiohydantoyl amino acids were the same as previously employed.

RESULTS AND DISCUSSION

Digestion of the reduced, S-carboxymethylated proteins $V^{II}2$ and $V^{II}3$ with trypsin and subsequent fractionation of digests on DEAE-cellulose, yielded the chromatograms shown in Figs. 1 and 2, respectively. The peptides were purified from the appropriate fractions by paper chromatography and high voltage paper electrophoresis (3) and the amino acid analyses were as given in Tables 1 and 2. Sequence studies, on the peptides using manual degradation methods, and on the intact reduced, S-carboxymethylated proteins using the Beckman Sequencer, are summarized in Fig. 3. Alignment of the peptides was facilitated using the sequence data depicted in Fig. 3.

Protein $V^{II}2$: Peptides T-1 to T-8 could be aligned, in that order, from the NH_2-terminal sequence obtained using the Beckman Sequencer. Peptide T-10b was the COOH-terminal peptide since it did not contain lysine or arginine (The asparginyl COOH-terminal residue was identified following

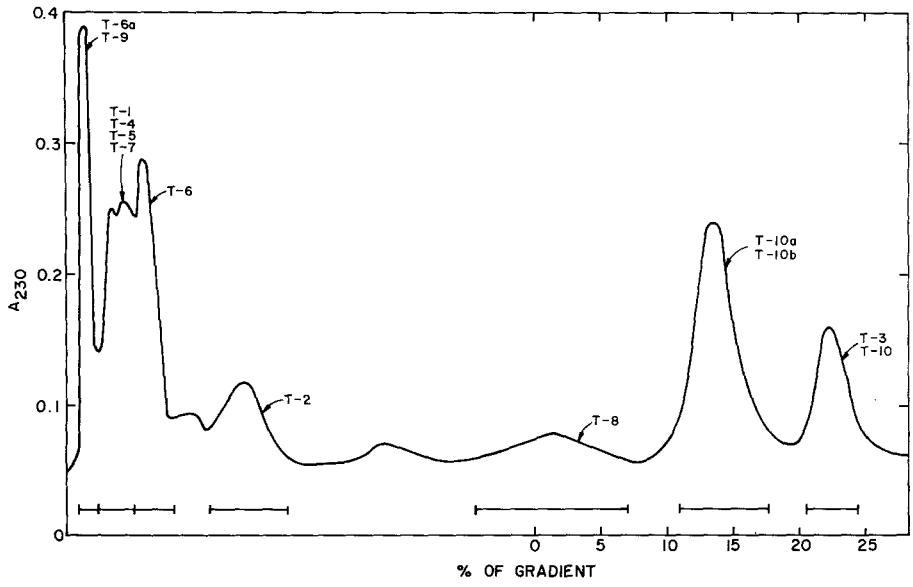

Fig. 1 Chromatography of 2 h tryptic digest of reduced, S-carboxymethylated $V^{II}2$ on DEAE-cellulose (1.9 x 50 cm) using 2 litre linear gradient of NH_4HCO_3 solution from 0.05 to 0.60 M at a flow rate of 100 ml h^{-1} and monitoring at 230 nm.

aminopeptidase M digestion). Peptide T-10a was placed NH_2-terminal to T-10b by the amino acid composition of T-10, and T-9 placed NH_2-terminal to T-10 by difference.

Protein $V^{II}3$: The peptide alignment was achieved by analogous reasoning.

Proteins $V^{II}2$ and $V^{II}3$ adhere to the general amino acid composition and sequence pattern established for the cytotoxins (2). Some striking differences are evident nevertheless. The intravenous toxicities are very much lower than has been found for the other members of the cytotoxin group (2,3). The structural changes introduced by particular amino acid substitutions could be responsible directly or indirectly for this. Rydén *et al.* (6) point out that ½Cys 3, 17, 24, 45, 49, 60, 61, 66; Tyr 25; Gly 44; Pro 50 and Asn 67 (numbering according to alignment chart in ref. 6) are common to all known snake venom toxin sequences and are probably necessary for determining the gross tertiary structure of the chain. In Fig. 3 these positions correspond to the eight half-cystines,

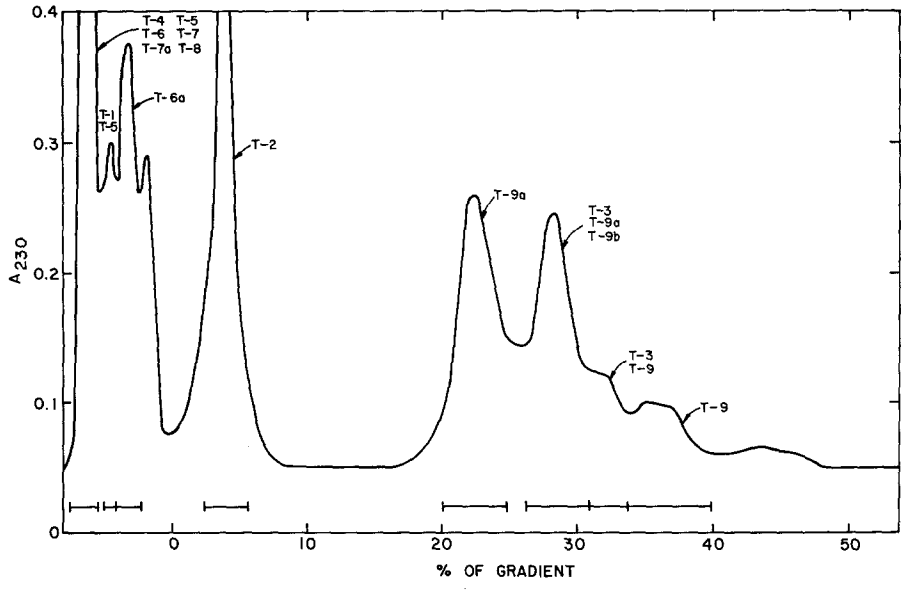

Fig. 2 Chromatography of 2 h tryptic digest of reduced, S-carboxymethylated $V^{II}3$ using identical conditions to Fig. 1.

Phe 23, Gly 38, Pro 44 and Asn 61. All these residues have been conserved except for Phe 23 which probably represents a conservative replacement of tyrosine in spite of the loss of a phenolic hydroxyl group. Rydén *et al.* (6) further refer to the invariant grouping found exclusively in the neurotoxins. Such a grouping is not as well-defined for the cytotoxins, but comparing all known cytotoxin sequences (3,6,7,8,9) with the neurotoxins (6), a group of residues, exclusively invariant in the cytotoxins, emerges *i.e.* (see alignment chart ref. 6) Leu 9, Lys 21, Leu 23, Met 27, Pro 40, Lys 42, Asp 47, Ser 53. In Fig. 3 these correspond to positions 7, 19, 21, 25, 34, 36, 41 and 47 respectively. This group, like the invariant group in the neurotoxins, could be important for the function of the cytotoxins as they have been conserved during evolution. It is evident that positions 19, 25, 34, 36 and 41 have undergone radical substitution. Position 19 shows the neutral glutamine for the basic lysine. Position 25 shows glycine in place of the hydrophobic methionine

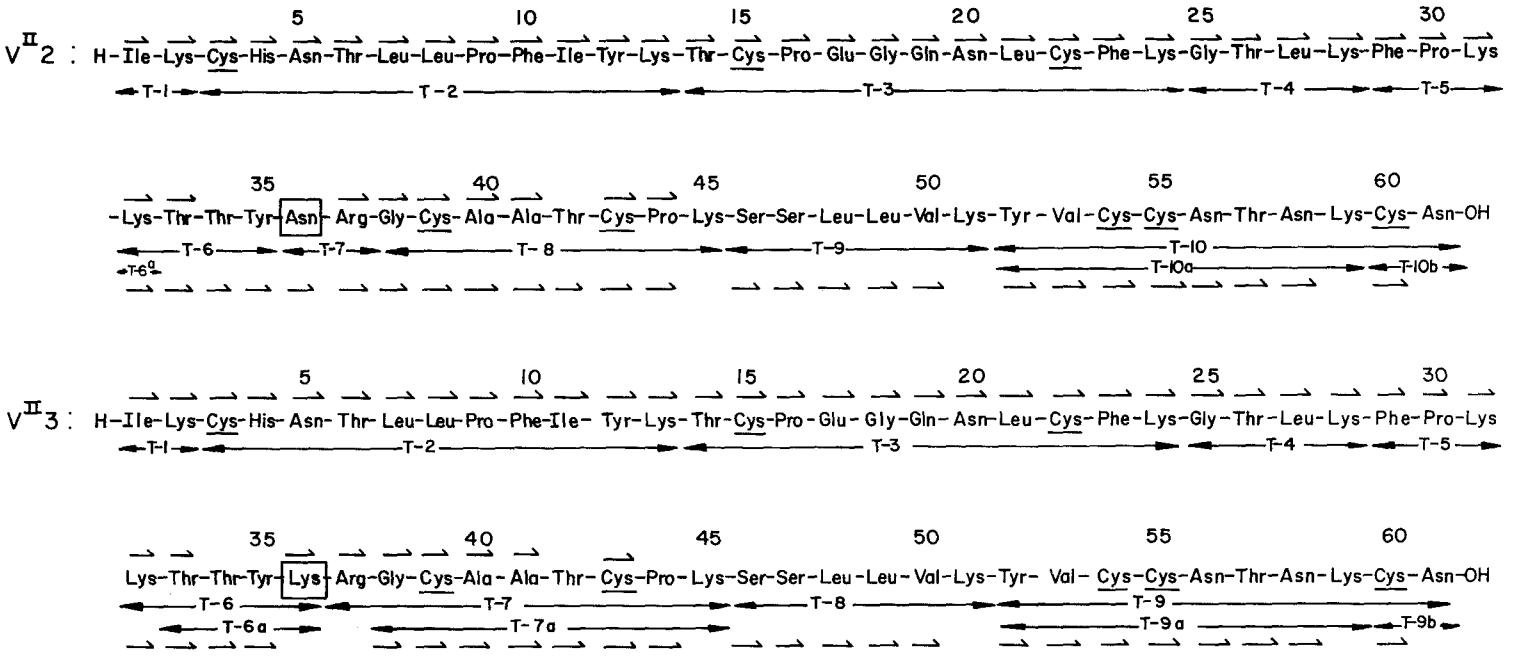

Fig. 3 The amino acid sequences of proteins $V^{II}2$ and $V^{II}3$. The blocked residues (36) indicate the difference between the two sequences. The double-headed arrows indicate the tryptic peptides. The upper half-arrows show the residues identified by degradation of the reduced, S-carboxymethylated proteins using the Beckman Sequencer. The lower half-arrows indicate the extent of manual degradation on individual peptides by the modified Edman-Begg technique.

TABLE 1

The amino acid compositions of the tryptic peptides of reduced, S-carboxymethylated $V^{II}2$.

Numbers in parentheses are the assumed integral values.

Amino acid	$V^{II}2$	T-1	T-2	T-3	T-4	T-5	T-6	T-6a	T-7	T-8	T-9	T-10	T-10a	T-10b
Lys	8.79(9)	1.02(1)	1.29(1)	0.97(1)	1.08(1)	0.93(1)	1.08(1)	1.00(1)	0.14	1.02(1)	1.07(1)	1.05(1)	1.04(1)	
His	0.77(1)		0.61(1)											
Arg	1.04(1)								1.01(1)					
Cys (Cm) *	7.68(8)		0.69(1)	1.96(2)						2.00(2)		2.85(3)	1.94(2)	1.00(1)
Asp	5.90(6)		0.94(1)	1.04(1)					0.99(1)			2.91(3)	2.03(2)	1.00(1)
Thr	7.06(7)		0.90(1)	0.95(1)	0.98(1)		1.91(2)		0.15	0.99(1)		0.97(1)	0.99(1)	
Ser	1.85(2)										1.77(2)			
Glu	2.05(2)			1.99(2)										
Pro	3.80(4)		0.87(1)	0.94(1)		1.00(1)				0.94(1)				
Gly	2.93(3)			1.05(1)	0.91(1)					1.04(1)				
Ala	1.98(2)									1.92(2)				
Val	2.10(2)										0.99(1)	0.97(1)	0.98(1)	
Met	0													
Ile	1.87(2)	0.98(1)	0.83(1)											
Leu	5.97(6)		1.98(2)	1.05(1)	1.11(1)						2.01(2)			
Tyr	2.65(3)		0.93(1)				1.12(1)					0.81(1)	0.74(1)	
Phe	3.02(3)		0.89(1)	0.97(1)		0.90(1)								
Total	61	2	11	11	4	3	4	1	2	8	6	10	8	2

* S-carboxymethylcysteine

TABLE 2

The amino acid compositions of the tryptic peptides of reduced, S-carboxymethylated $V^{II}3$.

Numbers in parentheses are the assumed integral values

Amino Acid	$V^{II}3$	T-1	T-2	T-3	T-4	T-5	T-6	T-6a	T-7	T-7a	T-8	T-9	T-9a	T-9b
Lys	9.81(10)	1.11(1)	1.09(1)	1.07(1)	1.12(1)	1.13(1)	1.95(2)	1.02(1)	1.37(1)	1.40(1)	0.97(1)	1.15(1)	1.06(1)	
His	1.01(1)		0.92(1)											
Arg	0.96(1)								1.00(1)					
* Cys (Cm)	8.11(8)		1.08(1)	1.88(2)					1.97(2)	1.90(2)		2.81(3)	1.89(2)	0.98(1)
Asp	5.24(5)		1.02(1)	1.02(1)								3.16(3)	2.03(2)	1.02(1)
Thr	6.23(7)		0.96(1)	1.00(1)	0.96(1)		2.04(2)	1.97(2)	1.09(1)	1.12(1)		1.03(1)	1.08(1)	
Ser	1.79(2)										1.88(2)			
Glu	1.92(2)			2.07(2)										
Pro	4.04(4)		0.90(1)	0.94(1)		0.94(1)			0.98(1)	1.14(1)				
Gly	3.03(3)			1.03(1)	1.00(1)				1.18(1)	0.96(1)				
Ala	2.04(2)								2.10(2)	1.90(2)				
Val	2.10(2)										1.01(1)	0.91(1)	0.94(1)	
Met	0													
Ile	1.95(2)	0.89(1)	1.20(1)											
Leu	6.13(6)		1.87(2)	0.98(1)	1.00(1)						2.12(2)			
Tyr	2.61(3)		0.93(1)				1.02(1)	1.01(1)				0.93(1)	0.85(1)	
Phe	2.65(3)		0.96(1)	1.01(1)		0.93(1)								
Total	61	2	11	11	4	3	5	4	9	8	6	10	8	2

* S-carboxymethylcysteine

with its potentially important sulphur containing side-chain. Besides changing the hydrophobicity, glycine, lacking a side-chain, probably introduces a change in geometry at position 25. The substitution of threonine 34 for a proline would also result in a conformational change. At position 36 the two proteins differ (Fig. 3) by a lysine-asparagine substitution representing a polarity difference as well as the loss of a potentially functional lysine in the case of $V^{II}2$. This, however, does not result in a toxicity difference (3). The substitution of an alanyl for an aspartyl residue at position 41 results in the loss of a site for electrostatic interactions. This region (38-44) is remarkable because it has the same sequence as the corresponding region in the 71-74 residue (long) neurotoxins *i.e.* Gly-Cys-Ala-Ala-Thr-Cys-Pro. Other interesting features are the NH_2-terminal isoleucines, the absence of methionine at position 27 and the asparagine at position 58.

The general conclusion is that the loop between half-cystines 22 and 43 has probably suffered conformational changes relative to the corresponding region in the other cytotoxins. It has also lost the potentially functional thio-ether side-chain of methionine 25. If this loop is important, which it is in the neurotoxins, where it carries all the functionally invariant residues except one (6), then these changes could be responsible for the low toxicity of proteins $V^{II}2$ and $V^{II}3$ relative to the other cytotoxins.

ACKNOWLEDGEMENTS

The author thanks Drs. A.I. Louw and D.J. Strydom for valuable discussions.

REFERENCES

1. Jimenez-Porras, J.M. (1970) Clin. Toxicol. 3, 389-431.
2. Lee, C.Y. (1972) Ann. Rev. Pharmacol. 12, 265-286.
3. Carlsson, F.H.H. and Joubert, F.J. (1974) Biochim. Biophys. Acta 336, 453-469.
4. Petersen, J.D., Nehrlich, S., Oyer, P.E. and Steiner, D.F. (1972) J. Biol. Chem. 247, 4866-4871.
5. Braunitzer, G. and Schrank, B. (1970) Z. Physiol. Chem. 351, 417.
6. Rydén, L., Gabel, D. and Eaker, D. (1973) Int. J. Peptide Protein Res. 5, 261-273.
7. Weise, K.H.K., Carlsson, F.H.H., Joubert, F.J. and Strydom, D.J. (1973) Z. Physiol. Chem. 354, 1317-1326.
8. Louw, A.I. (1974) Biochim. Biophys. Acta 336, 481-495.
9. Louw, A.I. (1974), unpublished.