AMINO ACID SEQUENCE OF NEUROTOXIN V FROM THE SCORPION *LEIURUS QUINQUESTRIATUS QUINQUESTRIATUS*

Charles KOPEYAN, Gérard MARTINEZ and Hervé ROCHAT

Groupe U 172 INSERM, Laboratoire de Biochimie, Faculté de Médecine, Secteur Nord, Boulevard Pierre Dramard, 13326 Marseille Cédex 3, France

Received 2 March 1978

1. Introduction

From the 5 neurotoxins active on mammals found in the venom of *Leiurus quinquestriatus quinquestriatus*, toxin V is quantitatively the most important and qualitatively the most potent [1]. It consists of a single polypeptide chain of 64 amino acid residues cross-linked by 4 disulfide bridges. Taking into account preliminary results obtained with the reduced and *S*-methylated protein (26 residues from the N-terminus) it was classified in the same group of scorpion neurotoxins as neurotoxin II of *Androctonus australis* Hector [2]. We report here the complete amino acid sequence of neurotoxin V of *Leiurus quinquestriatus quinquestriatus* which was worked out using exclusively a liquid phase protein sequencer. The N-terminal sequence (49 residues) was determined by degradation of the reduced and *S*-carboxymethylated protein. Other positions, i.e., residues 50–64, were identified by sequencing some tryptic fragments.

2. Materials and methods

The venom of *Leiurus quinquestriatus quinquestriatus* obtained from scorpions collected in the area of Khartoum (Sudan) was provided by F. G. Celo (Zweibrücken, FRG). Toxin V (Lqq V) was purified as in [1]. Trypsin (EC 3.4.21.4) treated with L-tosylamido-2-phenylethylchloromethyl ketone and carboxypeptidase A (EC 3.4.12.2) treated with diisopropylfluorophosphate were obtained from Worthington (Freehold, NJ). Sephadex G-25 was from Pharmacia (Uppsala), Biogel P4 and P6 (200–400 mesh) from Bio-Rad Laboratories (Richmond, CA). Iodoacetic acid, dithioerythritol, Quadrol and dimethylbenzylamine sequanal grade reagents were purchased from Pierce Chemical Company (Rockford, IL), succinic anhydride and *N*-ethylmorpholine (purum grade) from Fluka (Buchs, SG Switzerland). L-Aspartic diamide was provided by Bachem (Liestal, Switzerland). Hake parvalbumin was prepared as in [3].

2.1. *Chemical modification of the toxin*

The reduction, *S*-carboxymethylation and succinylation have been detailed [4]. In each case, after modification, the protein was recovered by gel filtration on Sephadex G-25 equilibrated with ammonium acetate, pH 8.6.

2.2. *Digestion by trypsin*

The reduced and *S*-carboxymethylated protein (0.9 μmol) was digested by trypsin after the lysine residues have been modified by succinylation. The digestions were carried out in 3 ml 0.2 M ammonium bicarbonate buffer pH 8.0, for 16 h at 37°C. In the first step the modified protein was digested by 4% enzyme (w/w) and the digest was fractionated on Biogel P4 (two 2 × 200 cm columns, connected in series, equilibrated with 0.1 M ammonium acetate, pH 8.6). As the recovery of peptides was low, the head fraction obtained by this way was submitted, after lyophilisation, to a second digestion by trypsin with an enzyme to substrate ratio of 7% (w/w). The digest was filtrated on Biogel P6 (2 × 200 cm column) again equilibrated with a 0.1 M ammonium acetate buffer, pH 8.6.

Elsevier/North-Holland Biomedical Press

2.3. *Amino acid analyses*

Amino acid analyses (corresponding to 20 h hydrolyses) of the modified protein as well as of tryptic peptides were obtained using a Beckman 120 C automatic amino acid analyzer using the single column procedure (M 82 resin, sodium citrate buffers, 4 h run) according to manufacturer's instructions.

2.4. *Automatic sequential degradations*

The N-terminal sequence of the reduced and *S*-carboxymethylated toxin (4.5 mg) was determined with a Socosi PS 100 protein sequencer as in [5]. The degradations of the peptides were performed using the same apparatus with a dimethylbenzylamine buffer [6] and the hake parvalbumin [3] as peptide carrier. Concerning peptide T_4, two runs were necessary. In the first run, two cycles were performed. The second run consisted of one cycle only followed by an extraction from the cup of the last residue (extraction by 2 X 1 ml 0.01 M acetic acid which does not dissolve parvalbumin). After freeze drying the extract was identified on the Beckman 120 C. It was also submitted to high voltage electrophoresis (40 V/cm, pH 6.5, 2 h) on Whatman 3 MM paper and its mobility compared to aspartic diamide.

Identifications of phenylthiohydantoin amino acids were carried out by gas chromatography [7] with a Beckman CG-65 apparatus, by thin layer chromatography on silica gel plates [5,8] and on polyamide double-faced sheets [9]. Acidic conversions of phenylthiohydantoins derivatives into free amino acids were used, when necessary, as in [10] and the resulting amino acids identified on the Beckman 120 C. Arginine residues were determined by Sakagushi's reaction [8].

2.5. *Digestion with carboxypeptidase A*

The reduced and *S*-carboxymethylated protein was submitted to the action of carboxypeptidase A.

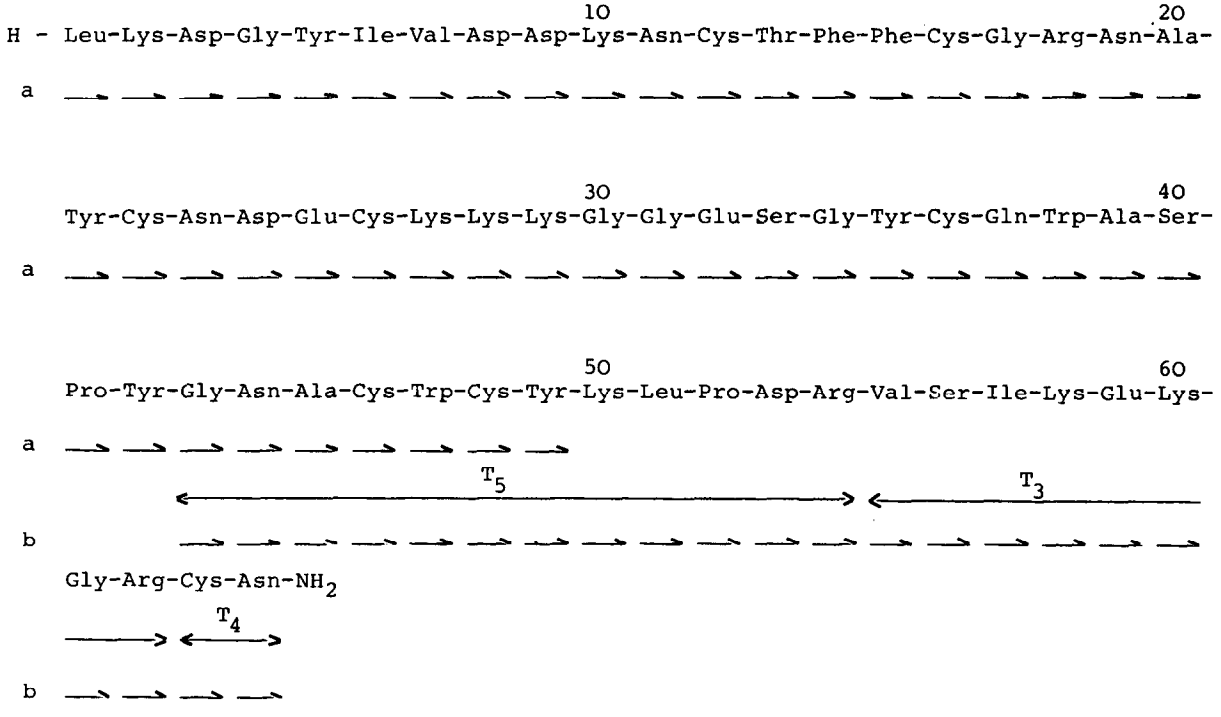

Fig. 1. Primary structure of *Leiurus quinquestriatus quinquestriatus* neurotoxin V. Underlined amino acids (⟶) were determined by automated sequencing: a = reduced and *S*-carboxymethylated protein submitted to the Quadrol program; b = tryptic peptides submitted to the dimethylbenzylamine-parvalbumin program. Half-cystine residues were identified as *S*-carboxymethylated derivatives.

The protein (120 nmol) was dissolved in a *N*-ethylmorpholine buffer (0.2 M, pH 8.0) and digested at 37°C with an enzyme to substrate ratio of 5% (w/w). Aliquots were removed after 15 min, 60 min, 120 min, 180 min and 360 min, acidified with 1.0 N HCl (0.1 ml) and submitted to amino acid analysis.

3. Results and discussion

Figure 1 gives the amino acid sequence of neurotoxin V of *Leiurus quinquestriatus quinquestriatus* as it has been determined by automatic phenylisothiocyanate degradation. From the degradation of 4.5 mg reduced and *S*-carboxymethylated toxin, using the protein program and the Quadrol buffer, the 49 first residues, from the N-terminal end, were identified unambiguously. When the tryptic digest of the reduced, *S*-carboxymethylated and succinylated toxin was fractionated on Biogel P4, two peptides, T_3 and T_4, were obtained in a pure form. Their amino acid compositions are given in table 1. They both were degradated using the peptide program and hake parvalbumin. T_3 was sequenced throughout using 100 nmoles : this peptide must be placed as the last but one peptide in the sequence (fig.1): it contains one of the three arginine residues. T_4 is the C-terminal peptide since it does not contain any arginine residue. This peptide was sequenced twice using 100 nmol each time. The first run (two steps of degradations) gave successively the carboxymethylcysteine and the asparagine derivatives. However, when one step of degradation was made, in the course of the second run, and then the cup was extracted with 0.01 N acetic acid, carboxymethylcysteine derivative was again obtained but aspartic diamide and not asparagine was found in the cup extract as identified by chromatography on the Beckman 120 C (aspartic diamide elutes between histidine and lysine in the conditions used) and by electrophoretic mobility. From these results it was concluded that the C-terminal end of the protein was amidated. This is in good agreement

Table 1
Amino acid compositions of the reduced and *S*-carboxymethylated protein and of some tryptic peptides

	Reduced and *S*-carboxymethylated protein		T_3		T_4		T_5	
Aspartic acid	9.6	(10)	–		1.1	(1)	1.6	(2)
Threonine	0.9	(1)	–		–		–	
Serine	2.8	(3)	1.1	(1)	–		–	
Glutamic acid	4.0	(4)	1.2	(1)	–		–	
Proline	2.3	(2)	–		–		0.9	(1)
Glycine	6.7	(7)	1.2	(1)	–		1.2	(1)
Alanine	3.0	(3)	–		–		1.0	(1)
CM-cysteine	7.3	(8)	–		0.9	(1)	1.3	(2)
Valine	2.0	(2)	1.0	(1)	–		–	
Methionine	–		–		–		–	
Isoleucine	1.7	(2)	0.9	(1)	–		–	
Leucine	2.0	(2)	–		–		0.9	(1)
Tyrosine	4.9	(5)	–		–		0.9	(1)
Phenylalanine	1.7	(2)	–		–		–	
Histidine	–		–		–		–	
Lysine	8.1	(8)	1.7	(2)	–		1.1	(1)
Arginine	3.0	(3)	0.9	(1)	–		1.0	(1)
Tryptophan	1.8[a]	(2)	–		–		1.1[a]	(1)
Total		64		8		2		12
Sequence positions			55–62		63–64		43–54	

[a] Spectrophotometric determination

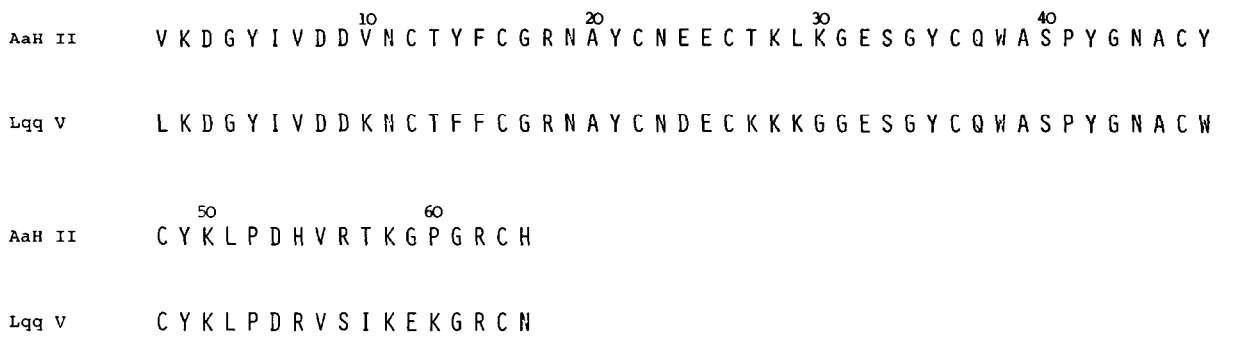

Fig.2. Amino acid sequences of *Androctonus australis* Hector neurotoxin II (AaH II) and *Leiurus quinquestriatus quinquestriatus* neurotoxin V (Lqq V). The IUPAC one-letter notation for amino acids sequences is used. see Eur. J. Biochem. (1978) 5, 151–153.

with the completely negative result obtained when the reduced and *S*-carboxymethylated protein was tentatively submitted to carboxypeptidase A digestion. The gap between residue 49 and peptide T_3 was filled with the help of an unexpected tryptic peptide formed during the digestion of the head fractions obtained by filtration on Biogel P4 of the first tryptic digest. This peptide results from the abnormal cleavage of the peptidyl bond Tyr_{42}–Gly_{43}. It was found, in a pure form, in the Biogel P6 fractions. Its amino acid composition is given in table 1. It was sequenced throughout using 100 nmol and the peptide program. With this last result the complete amino acid sequence of neurotoxin V of *Leiurus quinquestriatus quinquestriatus* was established (fig.1). This protein is made of 64 amino acid residues which is in good agreement with the amino acid composition of the reduced and *S*-carboxymethylated protein (table 1).

Neurotoxin V of *Leiurus quinquestriatus quinquestriatus* is homologous to neurotoxin II of *Androctonus australis* Hector, a North African scorpion: 51 residues out of 64, are identical. Some differences are conservative: positions 1, 14, 24, 47, 54. These two neurotoxins belong to the second group of scorpion neurotoxins as defined by N-terminal amino acid sequence determinations [2]. They both have a blocked C-terminal end: it has been established that the last residue of *Androctonus australis* Hector neurotoxin II is not free and probably amidated [11]. This feature seems to be extended to neurotoxins of the third group (J. Gregoire, personal communication). On the contrary neurotoxins of the first group have free C-terminal end: that is the case for neurotoxins I, I′ [12] and III (C.K. et al., to be published) of *Androctonus australis* Hector.

Acknowledgements

This work was supported in part by the Centre National de la Recherche Scientifique (ERA 070617), the Direction des Recherches, Etudes et Techniques (contrat 77/247), the Direction Générale de la Recherche Scientifique et Technique and the Recherche Médicale Française. The authors wish to thank Miss B. Ceard for expert technical assistance in determination of amino acid compositions and preparation of solvents used for sequencing. They are obliged to Mrs J. Gomez for establishment and typewriting of this paper. They are greatly indebted to Professor F. Miranda for his constant support and encouragement.

References

[1] Miranda, F., Kopeyan, C., Rochat, H., Rochat, C. and Lissitzky, S. (1970) Eur. J. Biochem. 16, 514–523.
[2] Rochat, H., Rochat, C., Kopeyan, C., Miranda, F., Lissitzky, S. and Edman, P. (1970) FEBS Lett. 10, 349–351.
[3] Rochat, H., Bechis, G., Kopeyan, C., Grégoire, J. and Van Rietschoten, J. (1976) FEBS Lett. 64, 404–408.
[4] Kopeyan, C., Miranda, F. and Rochat, H. (1975) Eur. J. Biochem. 58, 117–122.
[5] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.

[6] Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H. and Walsh, K. A. (1972) Biochemistry 11, 4493–4502.
[7] Pisano, J. J. and Bronzert, J. J. (1969) J. Biol. Chem. 244, 5597–5607.
[8] Edman, P. and Henschen, A. (1975) in: Protein Sequence Determination (Needleman, S. B. ed) pp. 232–279.
[9] Kulbe, K. D. (1974) Anal. Biochem. 59, 564–573.
[10] Mendez, E. and Lai, C. Y. (1975) Anal. Biochem. 68, 47–53.
[11] Rochat, H., Rochat, C., Sampieri, F., Miranda, F. and Lissitzky, S. (1972) Eur. J. Biochem. 28, 381–388.
[12] Rochat, H., Rochat, C., Miranda, F., Lissitzky, S. and Edman, P. (1970) Eur. J. Biochem. 17, 262–266.