Primary structure of toxin IV of *Leiurus quinquestriatus quinquestriatus*

Characterization of a new group of scorpion toxins

Charles Kopeyan, Gérard Martinez and Hervé Rochat

INSERM U. 172 et CNRS UA 553, Laboratoire de Biochimie, Faculté de Médecine, Secteur Nord, Boulevard Pierre-Dramard, 13326 Marseille Cédex 15, France

Received 15 November 1984

We report the sequence of toxin IV of *Leiurus quinquestriatus quinquestriatus*, one of the five toxins lethal for the mouse purified from the venom of this scorpion. Automatic sequencing of the *S*-carboxymethylated protein and derived peptides obtained by action of the *Staphylococcus aureus* protease allowed us to work out the complete sequence of 65 residues. The toxin which was found to be blocked at the C-terminal end has an extra residue at the N-terminus as compared with other toxins from Asian and African scorpions. Moreover, toxin IV must be classified as an α-scorpion toxin since it was shown to displace competitively ^{125}I-toxin II of *Androctonus australis Hector* (^{125}I-AaH II) from its binding site on rat brain synaptosomes. Nevertheless, it cannot be classified in one of the groups of scorpion toxins previously defined on the level of their antigenic properties. Again the classifications of scorpion toxins based on sequence homologies or antigenic properties support one another.

Scorpion toxin *Amino acid sequence* *Automatic sequencing* *Scorpion toxin classification*

1. INTRODUCTION

Scorpion toxins isolated from venoms of animals collected in North Africa were first classified into three groups of toxins according to structural homologies [1,2]. More recently, due to their binding properties to rat brain synaptosomes [3] and their action on sodium channels of excitable membranes, two types of scorpion toxins were characterized: α-toxins and β-toxins [4]. α-Toxins and β-toxins bind to rat brain synaptosomes on two different sites both related to the sodium channel, the former binds in a voltage-dependent manner, while the latter binds independently of membrane potential. Up to now all the toxins purified from the Buthinae collected in North Africa and Asia are α-toxins, while those purified from the Centrurinae *Centruroides suffusus* (Central America) are β-toxins [5] and finally those purified from Tityinae are either α- or β-toxins ([6], unpublished). In the case of *Centruroides sculpturatus* (North America) contradictory results have been published [7–9].

From the venom of *Leiurus quinquestriatus quinquestriatus* originating from Sudan, 5 toxins were isolated [10]. They present the general features of scorpion toxins; i.e., they are low-molecular-mass basic proteins cross-linked by 4

Abbreviations: AaH, *Androctonus australis Hector*; Bot, *Buthus occitanus tunetanus*; Amm, *Androctonus mauretanicus mauretanicus*; Lqq, *Leiurus quinquestriatus quinquestriatus*; Be, *Buthus epeus*; CsE, *Centruroides sculpturatus Ewing*; Css, *Centruroides suffusus suffusus*; Ts, *Tityus serrulatus*; AaH I, AaH I′ (28) - AaH I′′ (29) - AaH II (30) - AaH III (19) - Bot I and Bot II (25) - Bot III (31) - Lqq V (11) - Amm V (32) - Be M10 (33) - CsE I (34) - Css II (5) - Ts VII (22)

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

disulfide bridges. The structure of toxin V, the most represented in the venom and the most potent (LD_{50} on mice, 0.5 μg), was already determined [11]. It belongs to the same group as toxin II of *Androctonus australis Hector*, the most potent α-scorpion toxin known (LD_{50} on mice, 0.18 μg). We report here the primary structure of the least potent of the *L. quinquestriatus quinquestriatus* toxins, i.e. toxin IV (LD_{50} = 1.40 μg). These data are part of a structure-activity relationship study on scorpion toxins.

2. MATERIALS AND METHODS

Toxin IV was purified according to Miranda et al. [10] from a batch of venom provided by F.G. Celo (Zweibrücken, FRG). This venom was obtained from scorpions collected in the area of Kartoum (Sudan).

Staphylococcus aureus V8 protease (EC 3.4.21) was from Miles Laboratories; diisopropylphosphofluoridate carboxypeptidase A (EC 3.4.12.3) and carboxypeptidase B (EC 3.4.12.3) from Worthington (Freehold, NJ). Biogel P6 was purchased from Bio-Rad Laboratories (Richmond, CA) and Sephadex G-25 from Pharmacia (Uppsala, Sweden). Iodoacetic acid, dithioerythritol, Quadrol and dimethylbenzylamine sequanal grade from Pierce Chemical Company (Rockford, IL). *N*-Ethylmorpholine (purum grade) was provided by Fluka (Switzerland) and hake parvalbumin was prepared according to Rochat et al. [12].

2.1. *Activity tests*

LD_{50} values were established on C57BL/6 mice from CSEAL (Orléans-La Source, France) by subcutaneous or intracerebro-ventricular injections [13] according to Behrens and Karber [14]. Competitive experiments were conducted on rat brain synaptosomes using a ^{125}I derivative of toxin II of *A. australis Hector* (^{125}IAaH II) and native toxins according to Jover et al. [15]. The radioactive toxin was prepared according to [16]. Its specific radioactivity was 700 Ci/mmol.

2.2. *Chemical modifications of toxin IV*

After reduction of the protein (20 h at 40°C under nitrogen), *S*-methylation by methyl iodide or *S*-carboxymethylation by iodoacetic acid were performed according to Rochat et al. [17] and Crestfield et al. [18], respectively. In both cases 5 M guanidine–HCl and dithioerythritol were used as denaturing and reducing agents [19]. The modified protein was dialyzed against water in Spectra Por 3 membrane.

2.3. *Automatic sequential degradations*

The N-terminal sequence of the reduced and *S*-carboxymethylated toxin and the peptides derived from the action of *S. aureus* protease were established using a SOCOSI PS 100 protein sequencer as in [11,19]. The *S*-carboxymethylated protein (400 nM) was sequenced in the presence of 1 M Quadrol, as recommended by Edman and Begg [20].

The degradation of the peptides was performed with dimethylbenzylamine buffer prescribed by Hermodson et al. [21] in the presence of hake parvalbumin as protein carrier [12] to avoid losses of material. Phenylthiohydantoins amino acids (PTHaa) were identified on an HPLC Waters System, as described by Bechis et al. [22].

2.4. *Amino acid analysis*

Protein or peptides were hydrolyzed by 5.7 N hydrochloric acid in sealed evacuated tubes at 110°C for 20 h. The analyses were then achieved on a Beckman 120C amino acid analyzer.

2.5. *Digestion by Staphylococcus aureus protease*

The conditions recommended by Houmard and Drapeau [23] for the cleavage of glutamoyl bonds were closely followed, i.e., a 0.05 M ammonium bicarbonate buffer (pH 7.8) was used, the enzyme:substrate ratio being 10% (w/w). The digest was submitted to gel filtration either on Biogel P6 in an ammonium bicarbonate buffer (0.1 M, pH 8.5) or on Sephadex G-25 equilibrated in 0.5 M acetic acid.

2.6. *Digestions by carboxypeptidases A and B*

S-methylated as well as *S*-carboxymethylated toxin samples (100 nmol) were digested at 37°C in an *N*-ethylmorpholine acetate buffer (0.2 M, pH 8.0). In the case of carboxypeptidase A, the enzyme:substrate ratio was 5% (w/w), whereas this ratio was 10% for digestion by carboxypeptidase B. Aliquots were removed at various intervals 5 15, 60, 90, 180 min for amino acid analyses.

3. RESULTS AND DISCUSSION

When the digest by the *S. aureus* protease of 1 μmol *S*-carboxymethylated toxin IV was filtered on Biogel P6, four peptides P_{1-4}, P_{5-26}, P_{27-33} and P_{34-54} were obtained in a pure form (fig.1, table 1). The last fraction, F5, in which no UV absorbance was detectable, due to a very large dilution (400 ml), was concentrated by lyophilisation and further fractionated on a column of Sephadex G-25 equilibrated in 0.5 M acetic acid. Thus, peptide P_{55-65} was obtained in a pure form (fig.2, table 1). This peptide contains two arginine residues and several hydrophobic residues which may well explain its strong adsorption on Biogel P6: it elutes in a volume greater than the total volume of the column. P_{55-65} is, most probably for the above reason, the peptide obtained with the lowest yield (table 1).

The sequence of toxin IV presented on fig.3 is deduced from the sum of the results obtained by automatic Edman degradation runs made on the protein and the peptides derived from it. The 'Quadrol program' was used with the reduced and *S*-carboxymethylated toxin (400 nmol). Forty residues were so determined without any ambiguity. The five peptides obtained by *S. aureus* protease digestion were degraded (40–100 nmol each) using the peptide program and dimethylbenzylamine as coupling buffer. For the C-terminal peptide, the automatic sequencing showed an arginine residue in the last position. This sequence of 65 amino acids is in complete agreement with the amino acid composition of the protein. Peptides P_{1-4} and P_{34-54} were unexpected as they were obtained by cleavage of peptide bonds Asp_4–Ala_5 and Asp_{54}–Lys_{55} although conditions that were described to specifically split glutamoyl bonds [23] were used. After action of carboxypeptidase A on the *S*-carboxymethylated toxin (100 nmol) no free amino acid could be detected which is due to the enzyme specificity. However, the same result was obtained when the digestion by carboxypeptidase B was done either on the *S*-carboxymethylated or on the *S*-methylated derivatives of toxin IV. Thus, the C-terminal end of this protein is not free: it is blocked most probably by amidation of the α-carboxylate of the C-terminal arginine residue. Indeed, many scorpion toxins have been shown to have their C-terminal end blocked by amidation: it is the case for *A. australis Hector* toxin II [24], *L. quinquestriatus quinquestriatus* toxin V [11], *Buthus occitanus tunetanus* toxins I and II [25] and *Tityus serrulatus* toxin VII [22].

Considering the classification based on structural homologies (fig.4) obtained when half-

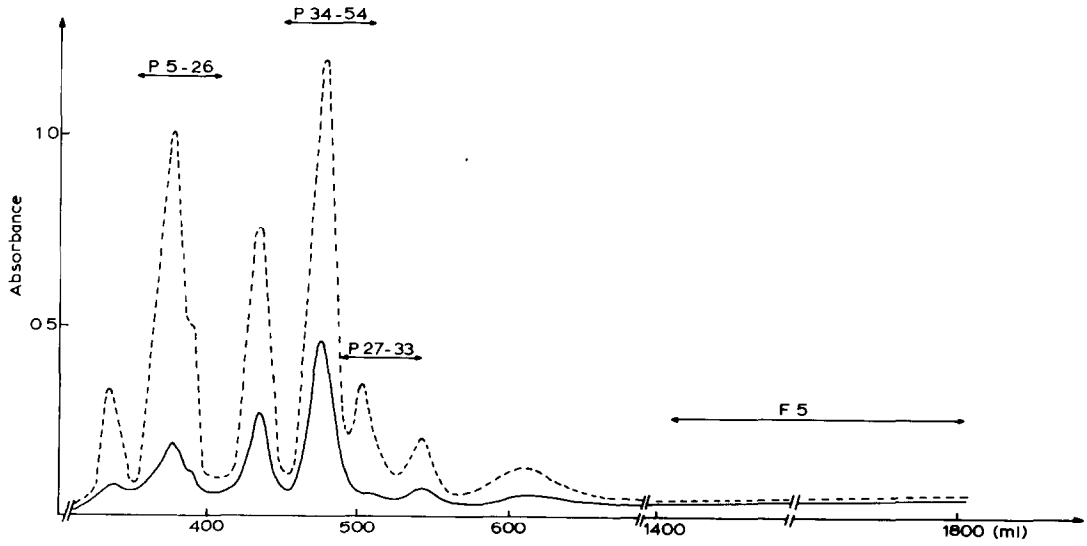

Fig.1. Gel filtration on Biogel P6 of peptides obtained by action of *Staphylococcus aureus* V8 protease. Two columns in series (2 × 200 cm) equilibrated in 0.1 M ammonium bicarbonate (pH 8.5). Full line absorbance at 280 nM; dotted line, absorbance at 230 nm; flow rate, 10 ml/h.

Table 1

Amino acid composition of peptides obtained by action of *Staphylococcus protease* on *S*-carboxymethylated toxin IV

Amino acids	*S*-Carboxymethylated protein	Peptide 1–4	Peptide 5–26	Peptide 27–33	Peptide 34–54	Peptide 55–65
Carboxymethylcysteine	7.5 (8)		2.6 (3)	1.2 (1)	2.6 (3)	0.4 (1)
Aspartic acid	8.9 (9)	1.0 (1)	4.9 (5)	0.8 (1)	2.2 (2)	
Threonine	2.9 (3)		1.9 (2)	1.0 (1)		
Serine	2.7 (3)		2.0 (2)		1.0 (1)	
Glutamic acid	3.0 (3)		1.2 (1)	1.1 (1)	1.1 (1)	
Proline	3.0 (3)				0.9 (1)	2.1 (2)
Glycine	7.0 (7)	1.1 (1)	1.5 (1)	1.3 (1)	3.0 (3)	1.0 (1)
Alanine	4.0 (4)		2.2 (2)	1.1 (1)	1.3 (1)	
Valine	2.8 (3)	0.9 (1)	1.1 (1)			1.0 (1)
Methionine	0					
Isoleucine	3.7 (4)		0.9 (1)		0.8 (1)	1.8 (2)
Leucine	2.0 (2)				1.5 (2)	
Tyrosine	4.5 (5)		2.9 (3)		1.3 (2)	
Phenylalanine	0					
Histidine	0					
Lysine	6.0 (6)		1.0 (1)	0.9 (1)	1.9 (2)	1.6 (2)
Arginine	3.3 (3)	1.0 (1)				2.1 (2)
Tryptophan	1.8[a] (2)				2.0[a] (2)	
Number of residues	65	4	22	7	21	11
Yield of purification		43%	42%	50%	28%	20%

[a] Spectrophotometric determination

cystine residues are aligned [2], it clearly appears that toxin IV has an extra residue at the N-terminus, which makes it different from other scorpion toxins. However, one can observe some homologies with toxins of group II: except the glycine 1 and alanine 5, there is a great homology with Be M10 for the ten first residues. The sequence 12–15 (NCVY) is found in groups I and

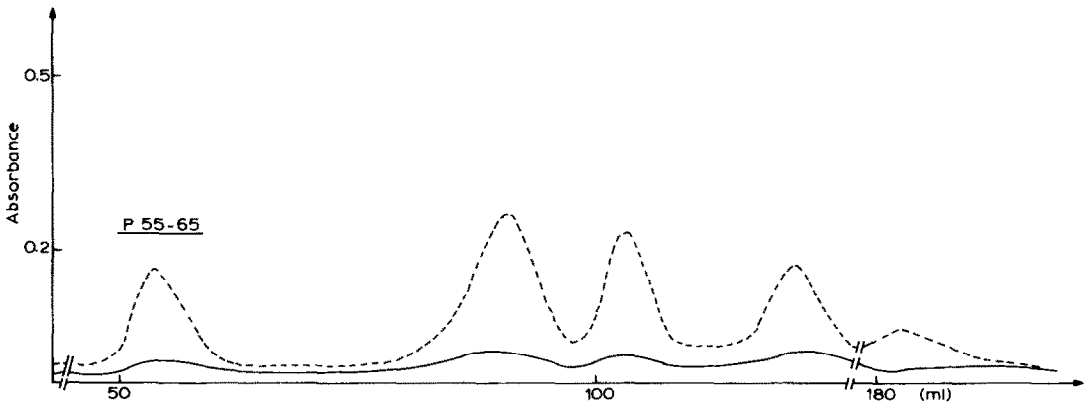

Fig.2. Gel filtration on Sephadex G-25 of fraction P_{1-4} obtained by Biogel P6 filtration. The column (1 × 150 cm) is equilibrated in 0.5 M acetic acid. Flow rate, 5 ml/h.

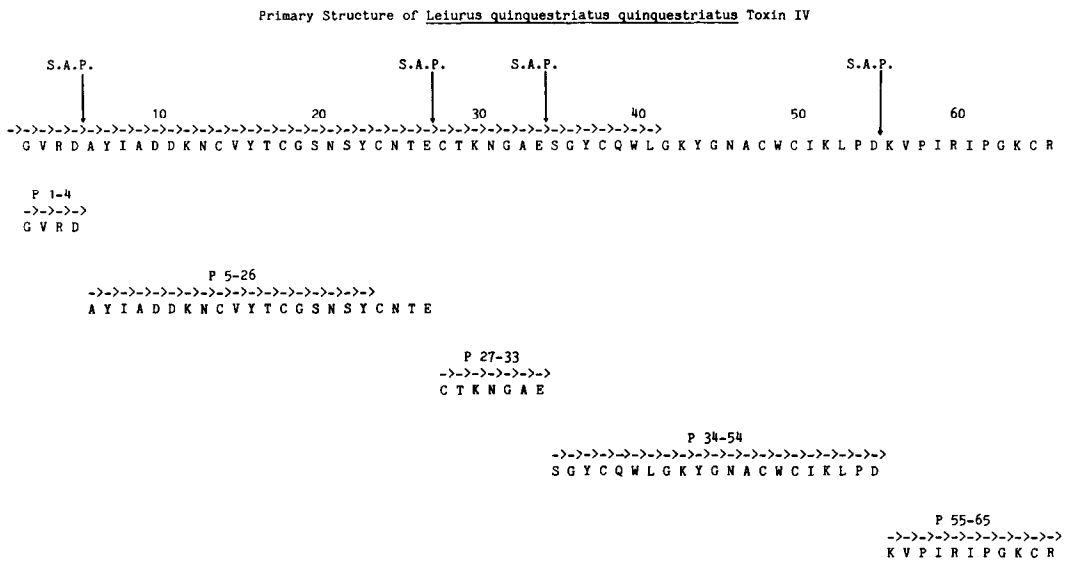

Fig.3. Amino acid sequence of the toxin IV of *Leiurus quinquestriatus quinquestriatus*. Positions determined by automatic sequencing. Vertical arrows indicate splitting by *Staphylococcus aureus* protease. The IUPAC one-letter notation for amino acid sequences is used [(1968) Eur. J. Biochem. 5, 151–153].

```
                    1         10        20        30        40        50        60        70

    AaH I           KRDGYIVYPN-NCVYHCVP--P---CDGLCKKN-GGSSGS-CSFLVPSGLACWC-KDLP-DNVPIKDT--SHKCT-
    AaH I'          KRDGYIVYPN-NCVYHCIP--P---CDGLCKKN-GGSSGS-CSFLVPSGLACWC-KDLP-DNVPIKDT--SRKCT-
I   AaH I''          KRDGYIVYPN-NCVYHCVP--P---CDGLCKKN-GGSSGS-CSFLVPSGLACWC-KDLP-DNVPIKDT--SRKCTR
    AaH III         VRDGYIVNSK-NCVYHCVP--P---CDGLCKKN-GAKSGS-CGFLIPSGLACWCV-ALP-DNVPIKDP--SYKCHS

    AaH II          VKDGYIVDDV-NCTYFCGR---NAYCNEECTKL-KGESG-YCQWASPYGNACYCYK-LP-DHVRTKGP---GRCH-
    BoT III         VKDGYIVDDR-NCTYFCGR---NAYCNEECNKL-KGESG-YCQWASPYGNACYCYK-VP-DHVRTKGP---GRCN-
II  Lqq V           LKDGYIVDDK-NCTFFCGR---NAYCNDECKKK-GGESG-YCQWASPYGNACWCYK-LP-DRVSIKEK---GRCN-
    Amm V           LKDGYIIDDL-NCTFFCGR---NAYCDDECKKK-GGESG-YCQWASPYGNACWCYK-LP-DRVSIKEK---GRCN-
    Be M10          VRDGYIADDK-DCAYFCGR---NAYCDEECKKG-GAESG-KCWYAGQYGNACWCYK-LP-DWVPIKQKV-SGKCN-

IV  Lqq IV        GVRDAYIADDK-NCVYTCGS---NSYCNTECTKN-GAESG-YCQWLGKYGNACWCIK-LP-DKVPIRIP---GKC-R

    BoT I           GRDAYIAQPE-NCVYECAQ---NSYCNDLCTKN-GATSG-YCQWLGKYGNACWC-KDLP-DNVPIRIP---GKCHF
III BoT II          GRDAYIAQPE-NCVYECAK---NSYCNDLCTKN-GAKSG-YCQWLGRWGNACYCI-DLP-DKVPIRIE---GKCHF

    Css II          -KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAF-----ACWC-THLY-EQAVVWPLPNKT-CN-
V   CsE I           -KDGYLVEK-TGCKKTCYKLGENDFCNRECKWKHIGGSYGYCYGF-----GCYC-EGLP-DSTQTWPLPNK--CT-
    Ts  VII         -KEGYLMDHE-GCKLSCF-IRPSGYCGRECGIKKG-SSG-YC-AW-P---ACYCY-GLP-NWVKVWDR-ATNKC--
```

Fig.4. Amino acid sequences of scorpion toxins classified according to their structural homologies.

III, while phenylalanine or glutamic acid in position 15 is replaced by a threonine residue in Lqq IV. The presence of glutamic acid in positions 30 and 37 is a characteristic feature of group II, whereas the sequence 33–37 (KNGA) is in common in the groups I and III. The region 42–62 of Bot I (group III) is very close to the corresponding region of Lqq IV, except that it does not show the deletion in position 58 as observed in toxins of group II and for Lqq IV. Lastly, the C-terminal region of Lqq IV more closely resembles the C-terminal region of group III toxins. These observations are summarized by the systematic classification proposed by Dufton and Rochat [26] on the basis of amino acid compositions as well as sequence comparisons, which suggests the evolutionary history of the family of scorpion toxins. This work confirms the above-mentioned homologies since Lqq IV differs from Bot I and Bot II (group III) only by 16 and 10 amino acids,

respectively, and from BeM 10 (group II) by 18 amino acids, the resemblance being even more evident when comparing the sequences. The alternate homologies with toxins of one or another group may explain that Lqq IV toxin cannot be classified in the already known groups of scorpion toxins. This is confirmed by immunological tests of immunodiffusion since scorpion toxins have been classified according to their antigenic properties [27]. When Lqq IV was tested by immunodiffusion with antisera prepared against toxin belonging to group I, II and III, no cross reactions could be observed ([27], P. Dolori, personal communication). In the same way, no cross reaction was observed when Lqq IV antisera were used against toxins of other groups.

As for other scorpion toxins, the toxicity of Lqq IV for the mouse was dependent on the mode of injection. The ratio between the values of LD_{50} obtained by subcutaneous and intracerebroventricular injections depends on the toxin affinities for its receptor present in the periphery and in the central nervous systems. In the case of Lqq IV the ratio is 52 (1400 ng SC/27 ng ICV) a value close to values found for Bot I and Bot II from group III (Martin, M.F., personal communication). Lqq V from the same venom shows a ratio of 200 (500 ng SC/2.5 ng ICV) a higher value closer to the ratio of AaH II (250 ng SC/0.5 ng ICV), both toxins belonging to the same group II.

Finally, in fig.5, are represented the results obtained when ^{125}I-AaH II was tentatively displaced by AaH II and Lqq IV: it is obvious that both toxins are able to fully displace the labelled α-toxin AaH II from its binding site on rat brain synaptosomes. The calculated $K_{0.5}$ values (i.e., the concentrations of native toxins which displace 50% of the labelled toxin) are, respectively, 3.3×10^{-10} M for AaH II and 3×10^{-8} M for Lqq IV. The value found for AaH II is in good accordance with the values previously found for this toxin [15]. One can calculate the ratio between the $K_{0.5}$ values (= 90) and the ratio between the LD_{50} (= 57) of the two toxins: two close numbers that show a good correlation between the lethal activity of this molecule and its ability to interact with the sodium channel of rat brain synaptosomes. This is in favour of the same mode of action for AaH II and Lqq IV. Thus, Lqq IV must be considered as an α-scorpion toxin. However, it cannot be classified on

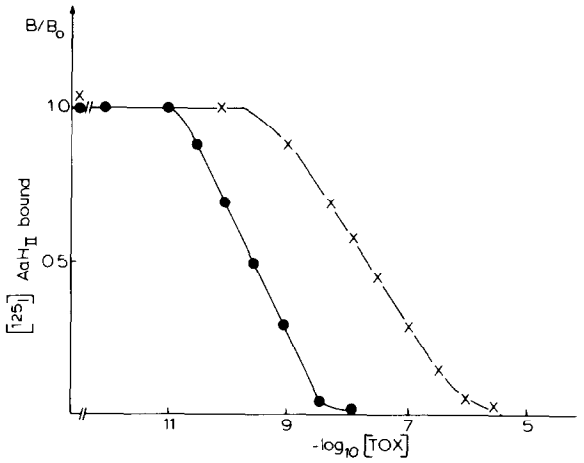

Fig.5. Displacement of ^{125}I-AaH II (700 Ci/mmol) by native AaH II (●) and Lqq IV (×). Synaptosomal P_2 fraction (0.3 mg protein per ml, i.e., a binding sites concentration of 10^{-10} M) was incubated 30 min at 37°C in the presence of 0.2 nM ^{125}I-AaH II and various amounts of AaH II or Lqq IV. B_0 is the binding of ^{125}I-AaH II in the absence of native toxin and B the binding in the presence of the indicated concentrations of native toxin. The non-specific binding (22% of the total binding) has been subtracted.

the level of its sequence and its immunological properties within the α-toxin groups previously described: it represents the first α-scorpion toxin of group IV [35].

ACKNOWLEDGEMENTS

The authors wish to thank Mrs M.F. Martin for the purification of Lqq IV and Dr H. Darbon for his help in the achievement of pharmacological tests. They are obliged to Mrs C. Roussarie for fine secretarial work and Mrs Th. Brando for skillful assistance in amino acid analysis and phenylthiohydantoin residues identification on a high-performance liquid chromatography system.

REFERENCES

[1] Rochat, H., Rochat, C., Kopcyan, C., Miranda, F., Lissitsky, S. and Edman, P. (1970) FEBS Lett. 10, 349–351.

[2] Rochat, H., Bernard, P. and Couraud, F. (1979) Adv. Cytopharmacol. 3, 325–334.

[3] Couraud, F., Jover, E., Dubois, J.M. and Rochat, H. (1982) Toxicon 4, 145–152.
[4] Jover, E., Couraud, E. and Rochat, H. (1980) Biochem. Biophys. Res. Commun. 95, 1607–1674.
[5] Martin, M.F., Garcia y Perez, L.G., El Ayeb, M., Kopeyan, C., Bechis, G., Jover, E. and Rochat, H. (1984) J. Biol. Chem., submitted.
[6] Barhanin, J., Giglio, J.R., Leopold, P., Schmid, A., Sampaio, S.V. and Lazdunski, M. (1982) J. Biol. Chem. 257, 12553–12558.
[7] Meves, H., Rubly, N. and Watt, D.D. (1982) Pflügers Arch. 393, 56–62.
[8] Barhanin, J., Ildefonse, M., Rougier, O., Sampaio, S.V., Giglio, J.R. and Lazdunski, M. (1984) Pflügers Arch. 400, 22–27.
[9] Meves, H., Rubly, N. and Watt, D.D. (1984) Pflügers Arch. 402, 24–33.
[10] Miranda, F., Kopeyan, C., Rochat, H., Rochat, C. and Lissitsky, S. (1970) Eur. J. Biochem. 16, 514–523.
[11] Kopeyan, C., Martinez, G. and Rochat, H. (1978) FEBS Lett. 89, 54–58.
[12] Rochat, H., Bechis, G., Kopeyan, C., Gregoire, J. and Van Rietschoten, J. (1976) FEBS Lett. 64, 404–408.
[13] Haley, T.J. and McCormick, W.G. (1957) Biochem. Biophys. Res. Commun. 95, 1607–1614.
[14] Behrens, B. and Karber, C. (1935) Arch. Exptl. Pathol. Pharmakol. 177, 379–388.
[15] Jover, E., Martin-Moutot, N., Couraud, F. and Rochat, H. (1978) Biochem. Biophys. Res. Commun. 85, 377–382.
[16] Rochat, H., Tessier, M., Miranda, F. and Lissitsky, S. (1977) Anal. Biochem. 82, 332–548.
[17] Rochat, C., Rochat, H. and Edman, P. (1970) Anal. Biochem. 37, 259–267.
[18] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[19] Kopeyan, C., Martinez, G. and Rochat, H. (1979) Eur. J. Biochem. 94, 609–615.
[20] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[21] Hermodson, M.A., Ericsson, L.H., Titani, K., Neurath, H. and Walsh, K.A. (1972) Biochemistry 11, 4493–4502.
[22] Bechis, G., Sampieri, F., Yuan, P.M., Brando, T., Martin, M.F., Dinitz, C. and Rochat, H. (1984) Biochem. Biophys. Res. Commun. 122, 1146–1153.
[23] Houmard, G. and Drapeau, G. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[24] Rochat, H., Rochat, C., Sampieri, F., Miranda, F. and Lissitsky, S. (1972) Eur. J. Biochem. 28, 381–388.
[25] Gregoire, I. and Rochat, H. (1983) Toxicon 21, 153–162.
[26] Dufton, M.J. and Rochat, H. (1984) J. Mol. Evol. 20, 120–127.
[27] Delori, P., Van Rietschoten, J. and Rochat, H. (1981) Toxicon 19, 393–407.
[28] Rochat, H., Rochat, C., Miranda, F., Lissitsky, S. and Edman, P. (1970) Eur. J. Biochem. 17, 262–266.
[29] Martin, M.F. and Rochat, H. (1984) Toxicon, in press.
[30] Rochat, H., Rochat, C., Sampieri, F., Miranda, F. and Lissitsky, S. (1972) Eur. J. Biochem. 28, 381–388.
[31] El Ayeb, M., Martin, M.F., Delori, P., Bechis, G. and Rochat, H. (1983) Mol. Immun. 20, 223–232.
[32] Rosso, J.P. and Rochat, H. (1984) Toxicon, in press.
[33] Grishin, E.V., Soldatova, L.N., Shakjparonov, M.I. and Kazakov, V.K. (1980) Bioorgan. Khim. 6, 714–723.
[34] Babin, D.R., Watt, D.D., Goos, S.M. and Mlesnek, R.V. (1974) Arch. Biochem. Biophys. 164, 694–706.
[35] Kopeyan, C., Martinez, G. and Rochat, H. (1982) Toxicon 20, 71.