Scorpion toxins from *Centruroides noxius* and *Tityus serrulatus*

Primary structures and sequence comparison by metric analysis

Lourival D. POSSANI,*|| Brian M. MARTIN,† Ib SVENDSEN,‡§ Gustavo S. RODE* and Bruce W. ERICKSON§

*Departamento de Biología Molecular, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Apartado Postal 70-247, México D.F., 04510 México, †National Institutes of Health, National Institute of Neurological and Communicative Disorders and Stroke (Developmental and Metabolic Neurology Branch), Bethesda, MD 20205, U.S.A., ‡Department of Chemistry, Carlsberg Laboratory, DK-2500 Copenhagen, Denmark, and §The Rockefeller University, New York, NY 10021, U.S.A.

(Received 21 December 1984/22 March 1985; accepted 15 April 1985)

The complete primary structures of toxin II-14 from the Mexican scorpion *Centruroides noxius* Hoffmann and toxin γ from the Brazilian scorpion *Tityus serrulatus* Lutz and Mello have been determined. Cleavage of toxin γ after Met-6 with CNBr produced the 55-residue peptide 7–61, which maintained the four disulphide bonds but was not toxic to mice at a dose 3 times the lethal dose of native toxin γ . Pairwise comparison by metric analysis of segment 1–50 of toxin γ and the corresponding segments from two other South American scorpion toxins, five North American scorpion toxins, nine North African scorpion toxins and one Central Asian scorpion toxin showed that the three Brazilian toxins are intermediate between the North American and North African toxins. This result is consistent with the hypothesis that the South American and African continents were joined by a land connection in the distant past.

Scorpion toxins are basic polypeptide chains of 36–70 amino acid residues bridged by four disulphide bonds (Kopeyan *et al.*, 1974; Zhdanova *et al.*, 1977; Rochat *et al.*, 1979; Fontecilla-Camps *et al.*, 1980). These toxins have attracted scientific attention because they interfere with several normal events that take place in biological membranes (Narahashi *et al.*, 1972; Cahalan, 1975; Rochat *et al.*, 1979; Catterall, 1980; Carbone *et al.*, 1982; Ovchinnikov & Grishin, 1982). The best-studied toxins are those from scorpions collected in North Africa and North America (Babin *et al.*, 1975; Rochat *et al.*, 1979; Possani *et al.*, 1982). We have purified to homogeneity several toxins from the venom of the Mexican scorpion *Centruroides*

noxius Hoffmann (Possani *et al.*, 1981b). Toxin Cn II-10 opens the Na⁺ channels of the squid giant axon, and toxin Cn II-11 blocks its K⁺ channels (Carbone *et al.*, 1982; Sitges *et al.*, 1984). The venom from the Brazilian scorpion *Tityus serrulatus* Lutz and Mello (range Southeastern Brazil) affects the cholinergic systems of mammals and causes lung oedema and pancreatitis in experimental animals (Gomez *et al.*, 1975; Lima & Freire Maya, 1977; Machado & Silveira-Filho, 1976; Fletcher & Fletcher, 1984). We have also purified and partially characterized several toxins from this venom (Possani *et al.*, 1977, 1981a, 1982). This work has been confirmed by others (Sampaio *et al.*, 1983; Bechis *et al.*, 1984). Toxin Ts γ acts upon the Na⁺ channels of excitable membranes (Norman *et al.*, 1983).

The present paper describes the complete amino acid sequences of toxins Cn II-14 and Ts γ and additional partial sequence data for toxins Ts III-8 and Ts IV-5. The rigorous mathematical methods of metric analysis (Erickson & Sellers, 1983) were used to compare the first two of these sequences with each other and with that of toxin variant 3 from *Centruroides sculpturatus* Ewing (CsE), for

Abbreviations used: AaH, *Androctonus australis* Hector; Amm, *Androctonus mauretanicus mauretanicus*; Be, *Buthus eupeus*; Bot, *Buthus occitanus tunetanus*; Cn, *Centruroides noxius*; CsE, *Centruroides sculpturatus* Ewing; Css, *Centruroides suffusus suffusus*; Lqq, *Leiurus quinquestriatus quinquestriatus*; Ts, *Tityus serrulatus*; mbc, minimum base changes; PE, pyridylethylated; RCM, reduced and carboxymethylated.

|| To whom correspondence should be sent.

which an X-ray-crystallographic structure has been determined (Fontecilla-Camps *et al.*, 1980). The structurally conserved *N*-terminal segments of 18 scorpion toxins, three from Brazil, five from North America, nine from North Africa and one from Central Asia, were also compared by metric analysis. The results show that the Brazilian scorpion toxins form a cluster that is intermediate between the clusters of North American and North African scorpion toxins.

Materials and methods

Source of materials

C. noxius venom was obtained by electrical stimulation of the telsons of anaesthetized animals, which were maintained in the laboratory as previously described (Possani *et al.*, 1981*b*). The amount of venom obtained from a single scorpion by this procedure is in the order of 100 µg. *T. serrulatus* venom was a gift from the Instituto Butantan, São Paulo, Brazil. All reagents, solvents and chromatographic resins were analytical-grade materials obtained as previously described (Possani *et al.*, 1981*a,b*, 1982).

Lethality tests

The mouse lethality of various protein fractions of toxins was observed after intraperitoneal injection of different amounts of protein in 0.1–0.3 ml of saline or ammonium acetate buffer into 18–20 g adult albino mice (Mexican strain). In the present paper, 'lethal' means that the component at the dose injected was enough to kill the tested mouse within 20 h of injection, 'toxic' denotes that the mouse showed one or more of the symptoms excitability, salivation, temporary paralysis of rear limbs, dyspnoea, and diarrhea, but that it recovered within 20 h after injection, and 'non-toxic' indicates that the mouse showed normal behaviour similar to that seen on injection of a saline control.

Purification of the toxins

The soluble part of the dried venom from *C. noxius* or *T. serrulatus* was separated by gel filtration on Sephadex G-50 followed by ion-exchange chromatography on CM-cellulose (CM-32; Whatman, Clifton, NJ, U.S.A.) as previously described (Possani *et al.*, 1981*a,b*). During chromatography, protein concentrations were measured spectrophotometrically by assuming that $A_{280}^{0.1\%} = 1$. For sequence determination, chemical modification and enzymic digestion, the protein content was determined by amino acid analysis.

Reduction, alkylation and amino acid analysis

Reduction of the toxins by dithiothreitol and alkylation with iodo[2-¹⁴C]acetic acid or with 4-

vinylpyridine was performed as described previously (Possani *et al.*, 1981*a,b*). Reduced and carboxymethylated (RCM)-toxin or pyridylethylated (PE)-toxin was chromatographed on a Sephadex G-50 (fine grade) column (2.5 cm × 120 cm) with 20% (v/v) acetic acid as eluent. The contents of the protein-containing tubes were pooled and freeze-dried. Sample treatment and amino acid analysis with a Durrum D-500 analyser have been described previously (Possani *et al.*, 1981*a,b*).

Sequence determination

Automated Edman degradation (Edman & Begg, 1967) of the toxins was performed by using a Beckman model 890C or 890M liquid-phase sequencer. Our procedures for the 890C instrument have been described previously (Possani *et al.*, 1981*a*). The 890M instrument, which can handle samples as small as 0.2 nmol, automatically converts the thiazoline derivative of each amino acid into the phenylthiohydantoin derivative.

The following procedures were used with the 890M instrument. Sequencer-grade reagents were purchased from Beckman Instruments (Palo Alto, CA, U.S.A.). Retention of the peptide film in the spinning cup was facilitated by adding Polybrene (3 mg) in water (0.10 ml) at the beginning of every sequencing run. The Polybrene was freed of impurities by prior treatment in the cup by using two Polybrene precycles (Beckman programme 1) and three peptide Quadrol cycles (programme 3). Peptide samples were applied to the cup in 30% (v/v) acetic acid, dried by running the start-up subroutine (no. 347339) and conditioned by running the Polybrene precycle. Repetitive Edman degradation was carried out with programme 3, with benzene for solvent S1 and ethyl acetate containing aldehyde-free acetic acid (1 ml/l) and dithiothreitol (10 mg/l) for solvent S2.

The resulting amino acid phenylthiohydantoin derivatives were identified by reverse-phase chromatography of about 40% of each sample on an octadecylsilane Ultrasphere column (0.46 cm × 25 cm) from Altex-Beckman, with a Hewlett-Packard model 1084B h.p.l.c. system equipped with a variable-wavelength u.v. detector set at 269 nm. The phenylthiohydantoin derivatives of serine and threonine were initially identified by their retention times and those of several characteristic breakdown products. These assignments were confirmed by identification at 320 nm of their dehydro derivatives (Hawke *et al.*, 1982). A mixture of amino acid phenylthiohydantoin derivatives (Pierce Chemical Co, Rockford, IL, U.S.A.) containing diethyl phthalate as internal standard was run at the 100 pmol level for standardization. With the buffer system of Hawke

et al. (1982), this mixture was completely separated in less than 13 min. Although the background was sufficiently low that 20 pmol of an amino acid phenylthiohydantoin derivative could be identified, more than 1 nmol of each toxin was sequenced.

Enzymic digestion of toxin Ts γ

Toxin Ts γ was digested with *Staphylococcus aureus* proteinase V8 (Miles Laboratories, Stoke Poges, Bucks., U.K.) at two different pH values. First, freeze-dried PE-toxin (180 nmol) was dissolved in 1.5 ml of 0.1 M-ammonium acetate buffer containing 2 mM-EDTA and adjusted to pH 4.0. Proteinase V8 (25 μ g) was added, and the toxin solution was kept at 40°C for 6 h, with gentle shaking every 30 min. The solution was filtered into a Sephadex G-50 (fine grade) column (1.8 cm $\times$ 85 cm) equilibrated with 20% (v/v) acetic acid. The contents of the tubes corresponding to the same fraction obtained after chromatography were pooled and freeze-dried.

Secondly, RCM-toxin TS γ (200 nmol), prepared by reduction and carboxymethylation with iodo-[¹⁴C]acetic acid (Amersham International, Amersham, Bucks., U.K.), was dissolved in 1.2 ml of 0.05 M-sodium phosphate buffer containing 2 mM-EDTA and adjusted to pH 7.8. After proteinase V8 (25 μ g) was added, the solution was kept at 40°C for 12 h, heated at 80°C for 15 min to denature the enzyme, and freeze-dried. As needed, small portions of the residual solid were dissolved in aq. 0.1% trifluoroacetic acid and separated by reverse-phase chromatography on an octadecylsilane column with a linear gradient of 0.1% trifluoroacetic acid in water to 0.1% trifluoroacetic acid in acetonitrile.

CNBr treatment of toxins Cn II-14 and Ts γ

A solution of PE-toxin Cn II-14 (0.5 μ mol) and CNBr (10 mg/mg of toxin) in aq. 70% (w/v) formic acid (1.8 ml) was kept overnight at 25°C and evaporated under a stream of N₂. The residue was separated by reverse-phase chromatography on octadecylsilane with the linear gradient described above.

Toxin Ts γ (2.6 mg) was dissolved in 70% formic acid (2.0 ml) preflushed with N₂ and concentrated to 1.8 ml. Part (0.15 ml) was kept as a control. The rest was treated with solid CNBr (25 mg) for 24 h at 25°C as described previously (Biedermann *et al.*, 1980). The cleavage reaction was practically complete. The resulting two peptides were separated on a Sephadex G-25 (superfine grade) column (0.9 cm $\times$ 200 cm) equilibrated with 10% (v/v) acetic acid. The larger peptide was collected, freeze-dried and chromatographed on CM-cellulose under the same conditions used for purifi-

cation of the toxins (loaded in 0.02 M-ammonium acetate buffer, pH 4.7, and eluted with a linear gradient of 0–0.55 M-NaCl). This 55-residue peptide was eluted as a single symmetrical peak in tubes 135–145 instead of tubes 178–190, which is where native toxin Ts γ would have been eluted. This peptide showed only one band when analysed by polyacrylamide-gel electrophoresis in the system of Reisfeld *et al.* (1962). About 5% of the peptide was bioassayed for toxicity, and the remainder was subjected to 18 cycles of automated Edman degradation. The control sample treated with 70% formic acid alone and chromatographed on Sephadex G-50 under the same conditions was as toxic as native toxin Ts γ .

Sequence comparison by metric analysis

Amino acid segments from 18 scorpion toxins were compared with one another by metric analysis. First, algorithm UU (Sellers, 1980; Erickson *et al.*, 1982a) and our set of genetic distances between pairs of amino acids (Erickson & Sellers, 1983) were used to align segments 1–50 of toxin Ts γ and Cn II-14 with the available sequences of 16 other toxins. In this way, the segments of the latter toxins that correspond to the 50-residue segments of toxin Ts γ and Cn II-14 were identified. Secondly, algorithm SS (Sellers, 1980; Erickson & Sellers, 1983) was used to calculate the 153 distances between pairs of these 18 scorpion toxin segments. The algorithms SS and UU were implemented as computer programs written in the C programming language and run on a Digital Equipment Corp. model VAX 11-780 superminicomputer running under the UNIX operating system. Finally, these 18 segments were clustered by using these distances to construct the minimum spanning tree linking these segments. Thus, starting with the smallest distance and continuing until all 18 segments were connected to the tree, the next larger distance was added to the tree only if it did not complete a closed circuit.

Results

Sequence determination

C. noxius venom has been separated by gel-permeation chromatography on Sephadex G-50 followed by ion-exchange chromatography on CM-cellulose (Possani *et al.*, 1981b). As shown in Fig. 1, fraction II-14 was resubmitted to ion-exchange chromatography to remove two minor impurities. Fraction II-14-3 contained pure toxin (Cn II-14) and was used for toxicity assay and reduction and pyridylethylation. Only 2% of the initial soluble venom corresponds to the fraction II-14-3. Edman degradation of PE-toxin Cn II-14 (50 nmol) in a model 890C sequencer provided the

Fig. 1. Final purification of toxin Cn II-14

The fraction II-14 obtained by ion-exchange chromatography, as previously described (Possani *et al.*, 1981b), was applied to a CM-cellulose ion-exchange column (0.9 cm $\times$ 30 cm) after dialysis against 50 mM-sodium phosphate buffer, pH 6.0. A total of 53.7 absorbance units at 280 nm in 44 ml was loaded on to the column and eluted with a linear gradient (■) of 0–0.5 M-NaCl in the same buffer (250 ml of each solution). Fractions 1–4 were pooled according to A_{280} (○) as indicated by the horizontal bars. The final recovery was 95.4% of the material applied to the column, and fractions II-14.1 to II-14.4 correspond respectively to 17.7%, 9.3%, 70.4% and 2.6% of the material recovered. Toxin Cn II-14 corresponds to 2% of the initial soluble venom. Fractions II-14.1 and II-14.4 were 'non-toxic', whereas fractions II-14.2 and II-14.3 were 'lethal' to mice. L indicates loading and G indicates the point where the gradient started. The inset shows the electrophoretic pattern of toxin II-14.3 in polyacrylamide gel in the urea/ β -alanine system described by Reisfeld *et al.* (1962). A is the application point, T is the tracking dye.

first 56 residues (Fig. 2a) with a repetitive yield of 95% per cycle. CNBr cleavage of this PE-toxin followed by reverse-phase chromatography gave a major 34-residue peptide. Part (2.5 nmol) of this peptide was sequenced with a model 890M sequencer to identify residues 31–64, with a repetitive yield of 95% per cycle. The amino acid composition of this peptide after acid hydrolysis agreed well with the composition determined by sequencing: hydrolysis values (sequencing values) Asp 3.6 (4), Thr 2.3 (2), Ser 3.0 (3), Glu 1.1 (1), Pro 2.5 (2), Gly 6.1 (6), Leu 2.2 (2), Tyr 3.9 (4), Phe 0.6 (1), Lys 3.2 (3), His 0.7 (1) and PE-Cys + Arg 4.8 (5). Since no residues were seen in the next two cycles, Cys-64 was the last residue. The complete amino acid sequence of toxin Cn II-14 is shown in Fig. 2(a).

Soluble *T. serrulatus* venom has been separated by gel-permeation and ion-exchange chromatography to give pure toxins Ts γ , III-8, and IV-5 (Possani *et al.*, 1981a), which correspond respectively to 11%, 3.7% and 1.5% of the initial soluble venom. Direct sequencing of PE-toxin γ (90 nmol) in a model 890C sequencer for 46 steps gave only one amino acid residue at each step (Fig. 2b), with a repetitive yield of 95% per cycle. Residues 7–24

were confirmed by direct sequencing of the 55-residue peptide 7–61 obtained by CNBr cleavage of native toxin γ . PE-toxin γ (1.5 mg) was digested with *S. aureus* proteinase V8 at pH 4 (Fig. 3a). Automated Edman degradation of component 2 (60 nmol) for 31 cycles proceeded with a repetitive yield of 96% per cycle and gave an unequivocal sequence through to Ala-57. The amino acid composition of this peptide after acid hydrolysis is comparable with the composition determined by sequencing; Asp 3.0 (3), Thr 1.3 (1), Ser 2.7 (2), Glu 0.9 (0), Pro 2.2 (2), Gly 4.5 (4), Ala 3.1 (3), Val 2.0 (2), Met 0.0 (0), Ile 1.1 (1), Leu 1.5 (1), Tyr 2.8 (3), Phe 0.0 (0), His 0.5 (0), Lys 4.4 (4) and Arg 1.3 (1).

Radiolabelled RCM-toxin γ (2 mg) was digested with *S. aureus* proteinase V8 at pH 7.8 (Fig. 3b). Automated Edman degradation of component 2 (3.5 nmol) gave an unequivocal sequence for six cycles. The position of the cysteine residue was confirmed by counting the radioactivity of a sample of each amino acid phenylthiohydantoin derivative, which gave less than 40 c.p.m. for cycles 1–5 and 1960 c.p.m. for cycle 6. The amino acid composition of this peptide after acid hydrolysis agreed well with the composition determined

Fig. 2. Complete primary structure of scorpion toxins

(a) Toxin Cn II-14 from *C. noxius*. Residues were determined by sequential Edman degradation directly (D) and after cleavage with CNBr. (b) Toxin Ts γ from *T. serrulatus*. Residues were determined as in (a), and also by proteinase V8 cleavage at pH 4.0 (V8A) and pH 7.8 (V8N).

by sequencing: carboxymethyl-Cys 0.7 (1), Asp 1.0 (1), Thr 0.8 (1), Ala 1.0 (1), Lys 0.6 (1) and Arg 0.6 (1). Since no residues were seen in the next two cycles, Cys-61 was the last residue. The complete amino acid sequence of toxin Ts γ is shown in Fig. 2(b). It is consistent with the amino acid composition and the total of 61 amino acid residues previously determined for toxin Ts γ (Possani *et al.*, 1981a).

Pure samples of PE-toxins Ts III-8 and Ts IV-5 were rechromatographed on Sephadex G-50. PE-toxin III-8 (100 nmol) was sequenced for 50 cycles to identify new residues 42–50 as Ala-Cys-Tyr-Cys-Tyr-Gly-Val-Pro-Asn (Figs. 5 and 8). PE-toxin IV-5 (50 nmol) was sequenced for 51 cycles to identify new residues 42–51 as His-Ile-Leu-Cys-Tyr-Cys-Tyr-Gly-Leu-Pro (Figs. 5 and 8). The repetitive yield for both runs was 96% per cycle.

Toxicity of peptide Ts γ 7–61

CNBr treatment of toxin Ts γ removed the first six residues to provide the 55-residue peptide 7–61, which contained all four disulphide bridges. This

peptide was not toxic when injected into mice at a dose (0.094 mg) greater than 3 times the lethal dose of native toxin Ts γ . Thus the toxicity of toxin Ts γ is greatly diminished or possibly even abolished when the *N*-terminal hexapeptide is removed.

Metric analysis of scorpion toxin segments

These studies provided the complete amino acid sequences of the Mexican scorpion toxin Cn II-14 and of the Brazilian scorpion toxin Ts γ and extensive *N*-terminal sequence data for the Brazilian scorpion toxins Ts III-8 and Ts IV-5. These sequences were compared with one another and with 14 other sequences of representative scorpion toxins by the mathematical methods of metric analysis (Sellers, 1980; Erickson *et al.*, 1982a; Erickson & Sellers, 1983). For example, Fig. 4 shows the three pairwise comparisons of complete amino acid sequences of three scorpion toxins: Ts γ , Cn II-14 and CsE v3. The third toxin is variant 3 from *Centruroides sculpturatus* Ewing (Fontecilla-Camps *et al.*, 1980), a scorpion whose range includes the Arizonan desert of the United States

Fig. 3. Separation of the enzymically cleaved peptides from toxin *Ts. γ* (a) Gel filtration on Sephadex G-50 of PE-toxin *Ts. γ* (1.3 mg) after digestion for 6 h with *S. aureus* proteinase V8 (25 μg) at pH 4.0. The column (1.8 cm × 85 cm) was equilibrated with 20% acetic acid. Horizontal bars indicate the pooling of the 2.05 ml fractions. Components 1–7 were freeze-dried and recovered for amino acid analysis. (b) Toxin *Ts. γ* after digestion with proteinase V8 at pH 7.8 was applied to an h.p.l.c. chromatograph as described in the Materials and methods section. For this experiment 5 nmol portions in 50 μl of solution were applied to an octadecylsilane Ultrasphere column and eluted at flow rate of 1 ml/min. The gradient (----) was generated by mixing solution A (0.1% trifluoroacetic acid in water) and solution B (0.1% trifluoroacetic acid in acetonitrile). The chromatogram was complete in 40 min. Fractions 1–10 were collected by hand and used for amino acid analysis and radioactivity measurements. Peptide 2 was used for sequence determination. After 24 min no appreciable amounts of material absorbing at 254 nm were recovered. The recovery from the column was practically quantitative.

and the Sonoran desert of Mexico. The Figure legend describes the interpretation of these results.

The distance between the last two toxins (36 mbc) is much smaller than the distances between them and the Brazilian toxin (54 and 66 mbc). These results are consistent with both the range (North America versus Brazil) and genus (*Centruroides* versus *Tityus*) of the three scorpions. More than one-third of the residues of toxin CsE v3 are also present at the same alignment positions as those of the other two toxins, including all eight cysteine residues. In other words, these three sequences have 23 amino acid residues in common, as shown at the bottom of Fig. 4. Also shown are the regions of α-helical structure (a) and β-sheet structure (b) observed in the three-dimensional structure of toxin CsE v3 by X-ray crystallography (Fontecilla-Camps *et al.*, 1980). Interestingly, four of the five patches of common residues are located at these four regions of secondary structure. Except for the terminal cysteine residue, none of the last 13 residues are shared by all three sequences.

When these comparisons were extended to include other scorpion toxins from North America and North Africa, the latter were seen to be closer to Brazilian scorpion toxin *Ts. γ* than to any of the North American scorpion toxins. Since this result was based upon just one Brazilian scorpion toxin

sequence, this comparative study was extended to include our *N*-terminal sequence data for two other Brazilian scorpion toxins, *Ts. III-8* and *Ts. IV-5*. In order to compensate for the unavailable C-terminal sequence data, the last 11 residues of toxin *Ts. γ* were ignored. Thus segment 1–50 of toxin *Ts. γ* was selected as a model segment containing 22 of the 23 common residues and all four regions of secondary structure shown in Fig. 4.

The 18 toxins of the present study include the three major toxins (*Ts. γ*, *Ts. III-8* and *Ts. IV-5*) from the Brazilian scorpion, five toxins [Cn II-14, CsE I and CsE v3 (Babin *et al.*, 1975; Fontecilla-Camps *et al.*, 1980), and Csx II and Csx III (Rochat *et al.*, 1979)] from North American scorpions, nine toxins [AaH I, AaH II and AaH III (Rochat *et al.*, 1979), Bot I, Bot II and Bot III (Vargas *et al.*, 1982), Lqq IV (Kopeyan *et al.*, 1982), and Lqq V and Amm V (Rosso & Rochat, 1982)] from North African scorpions and one toxin [Be M10 (Ovchinnikov & Grishin, 1982)] from a Central Asian scorpion. Algorithm UU (Sellers, 1980; Erickson *et al.*, 1982a) was used to identify the *N*-terminal segments from 17 other scorpion toxin sequences that correspond to the 50-residue model segment of toxin *Ts. γ*. Then the pairwise distances among these 18 related segments were calculated by using algorithm SS (Sellers, 1980; Erickson & Sellers,

Fig. 4. Comparison by metric analysis of the complete sequence for Brazilian scorpion toxin *Ts γ* and the Mexican scorpion toxins *Cn II-14* and *CsE v3*

The distance d between each pair of sequences is given at the right and one metric alignment corresponding to that distance is shown at the left. The distance between two amino acid sequences is defined as the minimum number of base changes and deletions needed to convert two DNA sequences that could code for the amino acid sequences into a common DNA sequence. It is measured in units of minimum base changes (mbc) and is calculated from (A) the number of identical amino acid residues, (B) the number of aligned residue pairs that differ by 1mbc, (C) the number of aligned residue pairs that differ by 2mbc, and (D) the number of residues aligned with a null (a 'missing' residue, represented by a dash, '-') by using the formula $d = B + 2C + 3D$. The values of A, B, C and D are given below the distance in the form [A/B/C/D]. For each of the standard alignments shown in this Figure, the similarity line between the pair of sequences shows A identical residues, contains B plus signs ('+') for the B aligned residue pairs differing by 1mbc, and contains C + D blanks (' ') for the C aligned residue pairs differing by 2mbc and the D residues aligned with a null. When more than one metric alignment exists, the valence line is given above the first sequence and the number of metric alignments is indicated to its right. The valence line contains a dash ('-') for each univalent alignment position, which is common to all of the metric alignments, and contains a field consisting of a number N followed by a series of colons (e.g. '4:::') for each multivalent alignment region, which can be aligned in N equivalent ways. The number of metric alignments is the product of the multivalency values N.

1983). Fig. 5 shows the results of 17 such calculations and indicates the distances and corresponding metric alignments following the format of Fig. 4. Fig. 6 tabulates all 153 pairwise distances among these 18 scorpion toxin segments.

Fig. 7 shows the minimum spanning tree that connects all 18 segments by the shortest possible distances without creating closed loops. The 17 distances that comprise this tree are calculated in Fig. 5 and are enclosed in heavy boxes in Fig. 6. The minimum spanning tree shows the North American scorpion toxin segments at the top, the North African scorpion toxin segments at the bottom, and the South American scorpion toxin segments lying in between. All five North American scorpion toxin segments are closer to one another than to any of the South American or North African scorpion toxin segments. Similarly, all nine North African scorpion toxin segments are closer to one another than to any of the North American or South American scorpion toxin segments.

The 18 segments were divided into five clusters

at the four distances greater than 30mbc (wavy lines in Fig. 7). The five North American scorpion toxin segments form cluster A, two Brazilian scorpion toxin segments (*Ts γ* and *Ts III-8*) form cluster B, the third Brazilian scorpion toxin segment comprises cluster C, seven of the North African scorpion toxin segments and the Central Asian scorpion toxin segment constitute cluster D, and the last two North African scorpion toxin segments (*AaH I* and *AaH III*) make up cluster E. The minimum spanning tree connecting these clusters is thus A-B-C-D-E. This tree places the Brazilian scorpion toxin segments in clusters B and C between the North American scorpion toxin cluster and the pair of North African scorpion toxin clusters. This result is relatively robust, because it is not changed by deleting one or two of the Brazilian scorpion toxin segments from the segment set and recalculating the minimum spanning tree. When the *Ts γ* toxin segment is deleted, cluster C (*Ts IV-5*) remains in the middle, even though cluster B (*Ts III-8*) merges with cluster A: (A,B)-C-D-E.

Toxin	Metric alignments					Distance
	1	10	20	30	40	50 54
Css II	KEGYLVSKSTGCKYECLKLGDN	DYCLRECKQQYGKSSGGYCYA	F	ACWCTHLYEQ		8 mbc
	KEGYLVSKSTGCKYECLKLGDN	DYCLRECKQQYGKSSGGYCYA	F	ACWC L ++		[49/2/3/0]
Css III	KEGYLVSKSTGCKYECLKLGDN	DYCLRECKQQYGKSSGGYCYA	F	ACWCEALPDH		
	1	10	20	30	40	50 54
Css III	KEGYLVSKSTGCKYECLKLGDN	DYCLRECKQQYGKSSGGYCYA	F	ACWCEALPDH		27 mbc
	KEGYLV KS GCKY+CLKLG+N+ C	+ECK ++ +S GYCYA	F	ACWCE+LP+		[36/9/9/0]
CsE v3	KEGYLVKKS	DGCKYGC	KLGENEGCDTECKAKNQGGSYGYCYA	F	ACWCEGLPES	
	1	10	20	30	40	50 54
CsE v3	KEGYLVKKS	DGCKYGC	KLGENEGCDTECKAKNQGGSYGYCYA	F	ACWCEGLPES	27 mbc
	K+GYLV+K+ GCK C	KLGEN+ C++ECK K+ GGSYGYCY+	F	+C CEGLP+S		[36/10/7/1]
CsE I	KDGYLV	KEKTKCYKLGENDFCNRECKWKHIGGSYGYCYG	F	GCYCEGLPDS		
	1	10	20	30	40	50 53
CsE I	KDGYLV	KEKTKCYKLGENDFCNRECKWKHIGGSYGYCYG	F	GCYCEGLPDS		12 mbc
	KDGYLV+ +GCKK+CYKLG+ND+CNREC+	KH+GGSYGYCYG	F	GCYCEGL+SDS		[43/8/2/0]
Cn II-14	KDGYLVDA	KGCKKNCYKLGNDYCNRECKMKHRRGGSYGYCYG	F	GCYCEGLSDS		
	1	10	20	30	40	50 53
Cn II-14	-----4:::-----3:::-----	-----	-	-----		12 alignments
	KDGYLVDA	KGCKKNCYKLGNDYCNRECKMKHRRGGSYGYCYG	F	GCYCEGLSDS		41 mbc
	K+GYLV+D +GCK +C+++ ++YC REC++K +G+S GYC +			+CYC GL++		[25/18/7/3]
Ts γ	KEGYLMDH	EGCKLSCFIRP-SGYCGRECGIK-KGSS-GYCAW	P	ACYCYGLPNW		
	1	10	20	30	40	50
Ts γ	KEGYLMDH	EGCKLSCFIR	PSGYCGRECGI-KK GSSGYCAW	P	ACYCYGLPNW	2 alignments
	KEGY MDH	EGCK+SCFIR	P+G+C++ C + K +SSGYCAW	P	ACYCYG+PN	21 mbc
Ts III-8	KEGYAMDH	EGCKFSCFIR	PAGFC	DGYCKTHLK ASSGYCAW	P	[37/8/5/1]
	1	10	20	30	40	51
	1	10	20	30	40	50
Ts γ	KEGYLMDH	EGCKLSCFIR	PSGYCGRECGIK K GSSGYCAW- P	2:-----		2 alignments
	K+GY++++ + C +C +	++YC++ C K K + SGYC W +		-ACYCYGLPNW		48 mbc
Ts IV-5	KDGPVPEY	DNCAYICWNY	DNAYCDKLCKDK K ADSGYCYWV H	ILCYCYGLPXX		[22/14/14/2]
	2	10	20	30	40	50 53
Ts IV-5	-----3: :-----2:-----	-----	-	-----		6 alignments
	KDGPVPEYDN	CAYICWNY	DNAYCDKLC KDKK	ADSGYCYWV-H--ILCYCYGLPXX		49 mbc
	KDGY V++++ C+Y+C+	NAYC++ C + K ++SGYC W+ + + CYCYLP				[26/14/10/5]
AaH II	KDGYIVDDVN	CTYFCG--	RNAYCNEEC TKLK	GESGYCQWASPYGNACYCYKLPDH		
	2	10	20	30	40	50 54
AaH II	KDGYIVDDVN	CTYFC	GRNAYCNEEC TKLK	GESGYCQWASPYGNACYCYKLPDH		3 mbc
	KDGYIVDD N	CTYFC	GRNAYCNEEC TKLK	GESGYCQWASPYGNACYCYK+PDH		[51/1/1/0]
Bot III	KDGYIVDDRN	CTYFC	GRNAYCNEEC TKLK	GESGYCQWASPYGNACYCYKVPDH		
	2	10	20	30	40	50 54
Bot III	KDGYIVDDRN	CTYFC	GRNAYCNEEC TKLK	GESGYCQWASPYGNACYCYKVPDH		12 mbc
	KDGYIVDD+N	CT+FC	GRNAYC+N+EC +K	GESGYCQWASPYGNAC CYK+PD+		[44/6/3/0]
Lqq V	KDGYIVDDKN	CTFFC	GRNAYCNDEC KKKG	GESGYCQWASPYGNACWCYKLPDR		
	2	10	20	30	40	50 54
Lqq V	KDGYIVDDKN	CTFFC	GRNAYCNDEC KK	KGGESGYCQWASPYGNACWCYKLPDR		4 mbc
	KDGYI+DD N	CTFFC	GRNAYC+DEC KK	KGGESGYCQWASPYGNACWCYKLPDR		[50/2/1/0]
Amm V	KDGYIIDD LN	CTFFC	GRNAYCDDEC KK	KGGESGYCQWASPYGNACWCYKLPDR		
	2	10	20	30	40	50 54
	2	10	20	30	40	50 54
	-----3: :-----2:-----	-----	-	-----		3 alignments
Lqq V	KDGYIVDDKN	CTFFC	GRNAYCNDEC KK	KGGESGYCQWASPYGNACWCYKLPDR		20 mbc
	+DGYI+DDK+ C++FC		GRNAYC++EC KK	G+ESG C A++YGNACWCYKLPD+		[38/11/3/1]
Be M10	RDGYIADDKD	CAYFC	GRNAYCDEEC KK	-GAESGKCWYAGQYGNACWCYKLPDW		
	2	10	20	30	40	50 53
	2	10	20	30	40	50 54
Lqq V	KDGYIVDDKN	CTFFC	GRNAYCNDEC KK	KGGESGYCQWASPYGNACWCYKLPDR		24 mbc
	+D+YI+DDKN C + C		G+N+YCN EC +K	G+ESGYCQW + YGNACWC KLPD+		[36/10/7/0]
Lqq IV	RDAYIADDKN	CVYTC	GSNSYCNTEC TK	DGAESGYCQWLGYGNACWCIKLPDK		
	3	10	20	30	40	50 55
Lqq IV	RDAYIADDKN	CVYTC	GSNSYCNTEC TK	DGAESGYCQWLGYGNACWCIKLPDK		20 mbc
	RDAYIA +N	CVY C	+ NSYCN C TK	DGA SGYCQWLGYGNACWC+ LPD+		[41/4/8/0]
Bot I	RDAYIAQPEN	CVYEC	AQNSYCNDEC TK	DGATSGYCQWLGYGNACWCCKDLPDN		
	2	10	20	30	40	50 54
Bot I	RDAYIAQPEN	CVYEC	AQNSYCNDEC TK	DGATSGYCQWLGYGNACWCCKDLPDN		10 mbc
	RDAYIAQPEN	CVYEC	A+NSYCNDEC TK	+GA+SGYCQWL G+ GNAC C+DLPD+		[45/6/2/0]
Bot II	RDAYIAQPEN	CVYEC	AKNSYCNDEC TK	NGAKSGYCQWLGRWGNACYCIDLPDK		
	2	10	20	30	40	50 54

Toxin	Metric alignments										Distance
	2	10	20	30	40	50	54				
Bot I	RDAYIAQPEN	CVYEC	AQNSYCNDLC	TK	DGATSGYCQWL	GKYGNACWCKD	LPDN				35 mbc [29/15/7/2]
	RD+YI+ P N	CVY C	++ + C++LC	+K	+G++SG++	L+ +G ACWCKD	LPDN				
AaH I	RDGYIYVPPN	CVYHC	VP-P-CDGLC	KK	NGGSSGSSCFL	VPSGLACWCKD	LPDN				
	2	10	20	30	40	50	52				
AaH I	RDGYIYVPPN	CVYHC	VP P CDGLC	KK	NGGSSGSSCFL	VPSGLACWCKD	LPDN				12 mbc [41/8/2/0]
	RDGYIV+++N	CVYHC	VP P CDGLC	KK	NG+ SGS++FL+P	SGLACWC +LPDN					
AaH III	RDGYIVNSKN	CVYHC	VP P CDGLC	KK	NGAKSGSCGFL	IPSGLACWCVAL	PDN				
	2	10	20	30	40	50	52				

Fig. 5. Comparison by metric analysis of the 17 closest pairs of amino acid segments from 18 scorpion toxins. The segments are the first 50 residues of toxin Ts γ and the related segments at or near the N-termini of two other toxins from the Brazilian scorpion (Ts), five toxins from North American scorpions (Cn, CsE and Css), nine toxins from North African scorpions (AaH, Amm, Bot and Lqq), and one toxin from Central Asia (Be). The metric analysis notations used in this Figure are defined in detail in the legend to Fig. 4. Since the segments of sequences used in the present figure contain 50–54 residues and the alignment spans 59 amino acid positions, each segment contains five to nine blank positions lacking amino acid residues (see Discussion).

Fig. 6. Pairwise distances between 18 scorpion toxin segments. Calculation of the 17 distances shown in heavy boxes is shown in Fig. 5.

Discussion

Toxins from *C. noxius* and *T. serrulatus* venom

Toxin Cn II-14 is the first long polypeptide toxin from *Centruroides noxius* venom for which the complete primary structure is known (Fig. 2a). We have previously reported (Possani *et al.*, 1982) the complete sequence of noxiustoxin, a 39-residue toxin in this venom that blocks the K⁺ channels of the squid giant axon (Carbone *et al.*, 1982). Both of these toxins contain a methionine residue, which is relatively rare in scorpion toxin sequences. We have also purified and partially sequenced two other major toxins (Cn II-9 and Cn II-10) and two minor toxins (Cn II-12 and Cn II-13) from this

venom (Possani *et al.*, 1981b). The terms ‘major’ and ‘minor’ refer to the relative concentrations in the venom rather than to relative toxicity.

Several small proteins toxic to mice have been isolated from the venom of the Brazilian scorpion *Tityus serrulatus* Lutz and Mello. Gomez & Diniz (1966) obtained two toxic fractions from which a single protein called tityustoxin was isolated. Toledo & Neves (1976) purified and partially characterized another toxin from this venom. Initially we obtained six different toxic fractions from this venom (Possani *et al.*, 1977). More recently, we have purified and characterized four major toxins from ten fractions toxic to mice (Possani *et al.*, 1981a). On the basis of identical

Fig. 7. Minimum spanning tree between the 18 scorpion toxin segments

The boxes correspond to the segments shown in Figs. 5 and 8 and the lengths of the connecting lines are proportional to the distance (in mbc) between the segments, as calculated in Fig. 5. The four wavy lines separate five toxin clusters.

amino acid compositions and *N*-terminal sequences through to Ala-41, the toxin initially designated as toxin γ (Possani *et al.*, 1977) and later isolated separately as toxins Ts II-11 and Ts III-10 (Possani *et al.*, 1981a) are considered to be the same toxic protein and will be called toxin γ in the future. Thus *T. serrulatus* venom contains three major toxins (Ts γ , Ts III-8 and Ts IV-5) and at least three minor toxins (Ts II-9A, Ts II-10 and Ts IV-6).

The amino acid sequence of toxin Ts γ is only the second complete primary structure reported for a polypeptide toxin from a South American scorpion (Fig. 2b). An important part of this structure determination was cleavage of the toxin after Glu-26 with *S. aureus* proteinase V8 at pH 4 followed by

direct Edman sequencing of residues 27–57. Assignment of the *C*-terminal four residues was accomplished by cleavage after Asp-55 with *S. aureus* proteinase V8 at pH 7.8 followed by direct Edman sequencing of residues 56–61.

Cleavage of toxin Ts γ with CNBr in 70% formic acid removed the *N*-terminal hexapeptide. The resulting 55-residue peptide 7–61 was not toxic in mice at a dose 3-fold higher than the lethal dose for the native toxin. This loss in toxicity is not due to the acidic cleavage conditions because the control sample of toxin Ts γ that had been treated with 70% formic acid alone was just as lethal as the native toxin. The X-ray-crystallographic study of toxin CsE v3 (Fontecilla-Camps *et al.*, 1980) showed that the four *N*-terminal residues form one of three strands of β -sheet (Fig. 4). The method of Chou & Fasman (1978) predicts that the first six *N*-terminal residues of toxin Ts γ probably form part of an α -helix. If either of these secondary structural features is present at the *N*-terminus of toxin Ts γ , the absence of residues 1–6 would probably prevent peptide 7–61 from being similar to native toxin Ts γ in secondary structure or toxicity.

While we were revising our manuscript, the complete sequence of toxin VII, a β -toxin from *T. serrulatus*, was reported by Bechis *et al.* (1984). Although it was suggested that toxin VII and our toxin Ts γ were identical, we have not found the *C*-terminal cysteine residue to be amidated; this is the single difference between these two toxins.

Sequence comparison of 18 scorpion toxins by metric analysis

The relationships of the three major toxins from *T. serrulatus* to the other scorpion toxins whose sequences are substantially completely known were investigated with the rigorous sequence comparison methods of metric analysis. These procedures have been used previously to compare the amino acid sequences of plant lectins (Cunningham *et al.*, 1979), calmodulin domains (Erickson *et al.*, 1980) and immunoglobulin variable domains (Erickson *et al.*, 1982b; Litman *et al.*, 1982), and the DNA sequences of influenza-virus genes (Shaw *et al.*, 1982), immunoglobulin variable genes (Litman *et al.*, 1983) and human interferon (Erickson *et al.*, 1984).

In order to include toxins Ts III-8 and Ts IV-5 in the comparison, it was necessary to compare the *N*-terminal segments of these sequences rather than the complete sequences. The first 50 residues of toxin Ts γ were chosen as representative of the entire sequence because this segment includes seven of the eight cysteine residues and all of the regions of α -helix and β -sheet in the three-dimensional structure of toxin variant 3 from *C. sculpturatus* Ewing.

X-ray	bbbb		aaaaaaaaa	bbbb	bbbb
	1	10	20	30	40
Css II	KEGYLVSKSTGCKYECLKLGDN	NDYCLRECKQY	GKSSGGYCYA	F	ACWCTHLYEQ
Css III	KEGYLVSKSTGCKYECLKLGDN	NDYCLRECKQY	GKSSGGYCYA	F	ACWCEALPHD
CsE v3	KEGYLVSKSDGCKYGLKLGNE	GCDTECKAKNQ	GGSYGYCYA	F	ACWCEGLPDS
CsE I	KDGYLVEK	TGCKKTCYKLGENDFCN	RECKWKHIGGSYGYCYG	F	GCYCEGLPDS
Cn II-14	KDGYLVDA	KCKKNCYKLGKNDYCN	RECRMKHRGGSYGYCYG	F	GCYCEGLSDS
	1	10	20	30	40
Ts III-10	KEGYLMDH	EGCKLSCFIR	PSGYCGRECGIK	K	GSSGYCAW
Ts III-8	KEGYAMDH	EGCKFSCFIR	PAGFCDGYCKTHLK	ASSGYCAW	P
	1	10	20	30	40
Ts IV-5	KDGYPVEYDN	CAYICWNY	DNAYCDKLCKDK	K	ADSGYCYWV
	1	10	20	30	40
AaH II	KDGYIVDDVN	CTYFC	GRNAYCNEEC	TKLK	GESGYCQWASPYGNACYCYKLPDH
Bot III	KDGYIVDDRN	CTYFC	GRNAYCNEEC	TKLK	GESGYCQWASPYGNACYCYKVPDH
Lqq V	KDGYIVDDKN	CTFFC	GRNAYCNDEC	KK	KGGESGYCQWASPYGNACWCYKLPDR
Amm V	KDGYIIDLN	CTFFC	GRNAYCDEEC	KK	KGGESGYCQWASPYGNACWCYKLPDR
Be M10	RDGYIADDKD	CAYFC	GRNAYCDEEC	KK	GAESGKCWYAGQYGNACWCYKLPDW
Lqq IV	RDAYIADDKN	CVYTC	GSNSYCNTCC	TK	DGAESGYCQWLKGYGNACWCYKLPDK
Bot I	RDAYIAQPEN	CVYEC	AQNSYCNCLC	TK	DGATSGYCQWLKGYGNACWCYKLPDN
Bot II	RDAYIAQPEN	CVYEC	AKNSYCNCLC	TK	NGAKSGYCQWLGRWGNACYCIDLPDK
	1	10	20	30	40
AaH II	RDGYIYVPPN	CVYHC	VP P	CDGLC	KK
AaH III	RDGYIVNSKN	CVYHC	VP P	CDGLC	KK
	1	10	20	30	40
Common		C C	C C	G	C C

Fig. 8. Composite alignment of the first 50 residues of toxin Ts γ and the related sequences of 17 other scorpion toxins. The numbering from 1 to 59 refers to this alignment and not the sequence numbering. The X-ray-crystallography results shown at the top indicate the residues of toxin CsE v3 that exist in the α -helix (aaa) or the three strands of β -sheet (bbb) (cf. Fontecilla-Camps *et al.*, 1980). The seven common residues shown at the bottom are present at the same alignment position in all 18 sequences. Toxin Ts III-10 is the same as toxin Ts γ .

A composite alignment of these scorpion toxin segments is shown in Fig. 8. This result generally follows the metric alignments shown in Fig. 5. In both of these Figures, breaks in the alignments are present in order to maximize the vertical alignment of amino acids having similar codon sets. The diversity of amino acid residues at each alignment position is striking. At only seven positions do all 18 segments have a common residue, namely cysteine at alignment positions 12, 16, 25, 29, 51 and 53 and glycine at position 39.

Since the segments contain only 50–54 residues and the alignment spans 59 amino acid positions, each segment contains five to nine blank positions lacking amino acid residues. The positions of these blanks, which are determined by the metric-analysis results, are often conserved between similar segments. All five North American scorpion toxin segments have only one residue at the six alignment positions 44–49. In contrast, all of the North African scorpion toxin segments have six residues at these positions but lack residues at the five alignment position 11, 17–19 and 30. In addition, seven of the nine North African scorpion toxin segments lack a residue at alignment position 33.

The segments from the three major Brazilian scorpion toxins resemble the North American scorpion toxin segments by lacking residues at alignment positions 45, 47 and 48, resemble the North African scorpion toxin segments by lacking in two cases a residue at position 33, and resemble

none of the other segments by lacking a residue at alignment position 20. In addition, two of the Brazilian scorpion toxin segments resemble two of the North American scorpion toxin segments by lacking a residue at position 9. Similarly, all three of the Brazilian scorpion toxin segments resemble two of the North African scorpion toxin segments by lacking a residue at alignment position 35. Thus the segments from the South American scorpion toxins resemble the segments from the North American and North African scorpion toxins more than do segments from the North American toxins resemble those from the North African toxins.

Finally, the 18 segments in Fig. 8 form five clusters, as illustrated in Fig. 7. The three major toxins from *T. serrulatus* fall into two clusters distinct from the North American scorpion toxin cluster and the two North African scorpion toxin clusters, contrary to the suggestion by Bechis *et al.* (1984). It should be interesting to see whether or not the major toxins from other *Tityus* species found in Brazil, Trinidad and Venezuela also fall into the South American scorpion toxin clusters. The fact that the Brazilian scorpion toxin segments are intermediate between the North American and North African scorpion toxin clusters is consistent with the hypothesis that the South American and African continents were joined by a land connection in the distant past.

This work was supported in part by Grant PCCBBNA 001370 from the Mexican National Council of Science

and Technology (CONACyT) and by Grant GM 32622 from the National Institute of General Medical Sciences, U.S. Public Health Service, and by Dirección General de Asuntos de Personal Académico from Universidad Nacional Autónoma de México. We are indebted to Dr. Bruno Soerensen Cardozo, Director of the Instituto Butantan (São Paulo, Brazil), and to the Conselho de Pesquisas of the same Institute for approving the project that granted the supply of scorpion venom used in this work. We thank Professor Ottesen of the Carlsberg Laboratory for the use of sequencer facilities. The technical assistance of Mrs. Lone Sørensen, Mrs Bodil Corneliussen and Mrs Denise Merkle-Lehman is gratefully acknowledged.

References

- Babin, D. R., Watt, D. D., Goos, S. M. & Mlejnek, R. V. (1975) *Arch. Biochem. Biophys.* **166**, 125–134
- Bechis, G., Sampieri, F., Yuan, P. M., Brando, T., Martin, M. F., Diniz, C. R. & Rochat, H. (1984) *Biochem. Biophys. Res. Commun.* **122**, 1146–1153
- Biedermann, K., Montali, U., Martin, B. M., Svendsen, I. & Ottesen, M. (1980) *Carlsberg Res. Commun.* **45**, 225–235
- Cahalan, M. D. (1975) *J. Physiol. (London)* **244**, 511–534
- Carbone, E., Wanke, E., Prestipino, G., Possani, L. D. & Maeticke, A. (1982) *Nature (London)* **296**, 90–91
- Catterall, W. A. (1980) *Annu. Rev. Pharmacol. Toxicol.* **20**, 15–43
- Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276
- Cunningham, B. A., Hemperly, J. J., Hopp, T. & Edelman, G. M. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3218–3222
- Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
- Erickson, B. W. & Sellers, P. H. (1983) in *Time Warps, String Edits, and Macromolecules* (Sankoff, D. & Kruskal, J. D., eds.), pp. 55–91, Addison-Wesley Publishing Co., Reading
- Erickson, B. W., Watterson, D. M. & Marshak, D. R. (1980) *Ann. N. Y. Acad. Sci.* **356**, 378–379
- Erickson, B. W., Sekulski, J. M. & Sellers, P. H. (1982a) in *Methods in Protein Sequence Analysis* (Elzinga, M., ed.), pp. 553–554, Humana Press, Clifton
- Erickson, B. W., Gerber-Jensen, B., Wang, A. C. & Litman, G. W. (1982b) *Mol. Immunol.* **19**, 347–355
- Erickson, B. W., May, L. T. & Sehgal, P. B. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7171–7175
- Fletcher, M. D. & Fletcher, P. L. (1984) *Abstr. Digest. Dis. Sci.* **19**, 948
- Fontecilla-Camps, J. C., Alamssy, R. J., Suddath, F. L., Watt, D. D. & Bugg, C. E. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6496–6500
- Gomez, M. V. & Diniz, C. R. (1966) *Mem. Inst. Butantan* **33**, 899–902
- Gomez, M. V., Diniz, C. R. & Barbosa, T. S. (1975) *J. Neurochem.* **24**, 331–336
- Hawke, D., Yuan, P. & Shively, J. E. (1982) *Anal. Biochem.* **120**, 302–311
- Kopeyan, C., Martinez, G., Lissitzky, S., Miranda, F. & Rochat, H. (1974) *Eur. J. Biochem.* **47**, 483–489
- Kopeyan, C., Martinez, G. & Rochat, H. (1982) *Toxicon* **20**, 71 (abstr.)
- Lima, E. G. & Freire Maya, L. (1977) *Toxicon* **15**, 225–234
- Litman, G. W., Erickson, B. W., Lederman, L. & Mäkelä, O. (1982) *Mol. Cell. Biol.* **45**, 49–57
- Litman, G. W., Berger, L., Murphy, K., Litman, R., Hinds, K., Jahn, C. L. & Erickson, B. W. (1983) *Nature (London)* **303**, 349–352
- Machado, J. C. & Silveira-Filho, J. F. (1976) *Mem. Inst. Butantan* **40**, 1–9
- Narahashi, T., Shapiro, D. I., Deguchi, T., Scuka, M. & Wang, C. M. (1972) *Am. J. Physiol.* **222**, 850–857
- Norman, R. I., Schmid, A., Lombet, A., Barhanin, J. & Lazdunski, M. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4164–4168
- Ovchinnikov, Y. A. & Grishin, E. V. (1982) *Trends Biochem. Sci.* **7**, 26–28
- Possani, L. D., Alagón, A. C., Fletcher, P. L., Jr. & Erickson, B. W. (1977) *Arch. Biochem. Biophys.* **180**, 394–403
- Possani, L. D., Martin, B. M., Mochca-Morales, J. & Svendsen, I. (1981a) *Carlsberg Res. Commun.* **46**, 195–205
- Possani, L. D., Dent, M. A. R., Martin, B. M., Maeticke, A. & Svendsen, I. (1981b) *Carlsberg Res. Commun.* **46**, 207–214
- Possani, L. D., Martin, B. M. & Svendsen, I. (1982) *Carlsberg Res. Commun.* **47**, 285–289
- Reisfeld, R. A., Lewis, U. J. & Williams, D. E. (1962) *Nature (London)* **195**, 281–283
- Rochat, H., Bernard, P. & Couraud, F. (1979) *Adv. Cytopharmacol.* **3**, 325–334
- Rosso, P. & Rochat, H. (1982) *Toxicon* **20**, 76–77 (abstr.)
- Sampaio, S. V., Laure, C. J. & Giglio, J. R. (1983) *Toxicon* **21**, 265–277
- Sellers, P. H. (1980) *J. Algorithms* **1**, 359–373
- Shaw, M. W., Lamb, R. A., Briedif, D. J., Erickson, B. W. & Choppin, P. W. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6817–6821
- Sitges, M., Bayón, A. & Possani, L. D. (1984) *Abstr. Annu. Meet. Am. Soc. Neurosci.* 198
- Toledo, D. & Neves, A. G. A. (1976) *Comp. Biochem. Physiol. B* **55**, 249–253
- Vargas, O., Gregoire, J., Martin, M. F., Bechis, G. & Rochat, H. (1982) *Toxicon* **20**, 79 (abstr.)
- Zhdanova, L. N., Adamovich, T. B., Nasimov, I. V., Grishin, E. V. & Ovchinnikov, Y. A. (1977) *Bioorg. Khim.* **3**, 485–493