

# The disulfide bridges of toxin 2 from the scorpion *Centruroides noxius* Hoffmann and its three-dimensional structure calculated using the coordinates of variant 3 from *Centruroides sculpturatus*

G.B. Gurrola, G. Moreno-Hagelsieb, F.Z. Zamudio, M. García, X. Soberon, L.D. Possani\*

Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apartado Postal 510-3, Avenida Universidad 2001, Cuernavaca 62271, Mexico

Received 7 May 1994

## Abstract

The disulfide bridges of toxin 2 from the venom of the scorpion *Centruroides noxius* Hoffmann were found by amino acid sequence determination of fragments of native toxin, produced by enzymatic cleavage and separated by high-performance liquid chromatography (HPLC). They are: Cys<sup>12</sup>-Cys<sup>65</sup>, Cys<sup>16</sup>-Cys<sup>41</sup>, Cys<sup>25</sup>-Cys<sup>46</sup> and Cys<sup>29</sup>-Cys<sup>48</sup>. The coordinates of the X-ray diffraction structure of toxin variant 3 of *C. sculpturatus* [(1980) Proc. Natl. Acad. Sci. USA 77, 6496–6500] were used to construct a three-dimensional model of toxin 2. All the amino acid replacements were easily accommodated, and the modeled structure reveals a clustered pattern of sequence variation, which may help to identify residues responsible for functional differences among toxins of mammals and insects.

**Key words:** Scorpion toxin; Disulfide bridge; Homology modeling; *Centruroides noxius*; *Centruroides sculpturatus*

## 1. Introduction

To elucidate the structure–function relationships of scorpion toxins and their targeted ion channels [2] more information is needed, not only on their primary structure, of which many are known [3,4], but also on their three-dimensional folding [1,5]. One of the structural features of the Na<sup>+</sup> channel-directed toxins is the high content of disulfide bridges, but few disulfide pairs have been directly determined [1,5,6]. Here we describe the experimental determination of the disulfide bridges of toxin 2, a mammalian-specific toxin. We also present a model of its three-dimensional structure, based on the crystal structure of its insect-specific homologue, variant 3 of *C. sculpturatus* [1]. Insights into possible structure–function relationships are discussed.

## 2. Material and methods

Toxin 2 was purified as described by Zamudio et al. [7]. This peptide (1.2 mg, equivalent to 170 nmol) was hydrolysed by chymotrypsin and trypsin, and separated by HPLC, following the procedure of Sugg et al. [8]. Heterodimeric peptides containing the native disulfide bridges were identified by microsequencing in a Millipore ProSequencer [7]. Since the primary structure of this toxin is known it was easy to assign its disulfide pairing.

Computational analysis made use of the coordinates of variant 3 of the related scorpion *C. sculpturatus* [1]. The suit of programs for homology modeling from Biosym Technologies (San Diego, USA) was used for model generation (including minimization, and molecular dynamics at 400 K for 5 ps and then 5 ps at 300 K). A water shell was included.

## 3. Results and discussion

The few experimentally determined disulfide bridge patterns of Na<sup>+</sup> channel-specific scorpion toxins are those of variant 3 from *C. sculpturatus* [1], toxin M9 from *Buthus eupeus* [5], toxin III and toxin II [5,6], from *Androctonus australis*, a typical  $\alpha$ -scorpion toxin [9]. The disulfide bridges of these toxins all correspond to equivalent positions on the primary structure (Cys<sup>12</sup>-Cys<sup>65</sup>, Cys<sup>16</sup>-Cys<sup>41</sup>, Cys<sup>25</sup>-Cys<sup>46</sup> and Cys<sup>29</sup>-Cys<sup>48</sup>), except for the insect toxin I from *A. australis* [10], which has the pairing Cys<sup>16</sup>-Cys<sup>37</sup>, Cys<sup>22</sup>-Cys<sup>42</sup>, Cys<sup>26</sup>-Cys<sup>44</sup>, and Cys<sup>38</sup>-Cys<sup>64</sup>. Toxin 2 from *C. noxius* [7], initially called component II-9.2.2 [11], was shown to recognize Na<sup>+</sup> channels of several tissues [12–14]. This toxin is a typical  $\beta$ -scorpion toxin [15].

Since the species specificity and the fine mechanism of action of insect, and mammalian  $\alpha$ - and  $\beta$ -scorpion toxins are different, we decided to determine the relative positions of the disulfides of toxin 2 from *C. noxius*. Separation of enzymatically cleaved peptides of native toxin 2 gave rise to approximately 30 components when applied to a C<sub>18</sub> reverse-phase column on HPLC (data not shown). Under sequence analysis, peptides that eluted at 47.15 min (step-gradient 9.82 min), 59.28 (8.51), 64.36 (8.28) and 68.76 min gave two amino acids each at every step of the microsequencing process. The values in parenthesis after the time (min) mean that the same component before sequencing was further applied to HPLC and eluted at the time indicated, when using a continuous gradient mode (step-gradient). Since the primary structure of toxin 2 was known [7], this allowed us to assign the the disulfide pairing unequivocally (Table 1). The

\*Corresponding author. Fax: (52) (73) 172 388.

E-mail: possani@pbr322.ceingebi.unam.mx

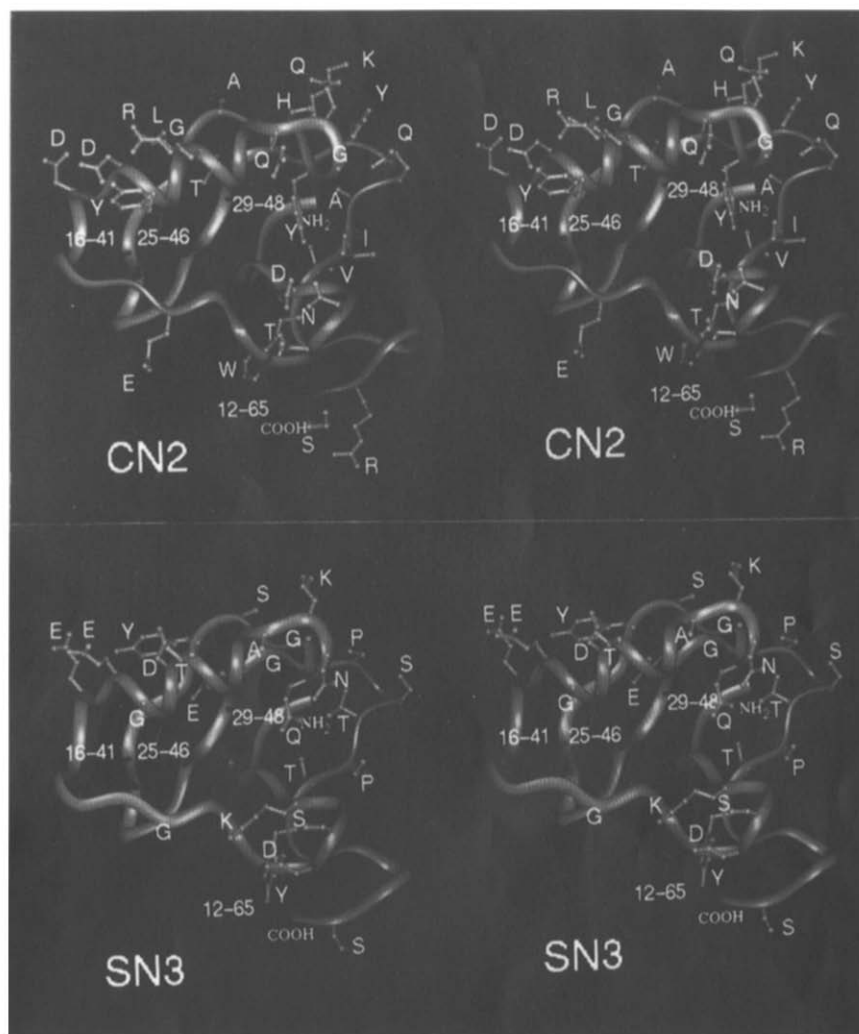


Fig. 1. Stereo-image of the three dimensional models of toxin 2 from *C. noxius* (CN2), and variant 3 of *C. sculpturatus* (SN3). Coordinates for variant 3 of *C. sculpturatus* (Brookhaven data bank 1sn3) were used for modeling the CN2 structure, using a Silicon-Graphics IRIS 4D/35 and a Cray Y-MP4/464; with programs from Biosym Technologies Inc. Energy minimization, and molecular dynamics at 400 K for 5 ps and then 5 ps at 300 K were employed to refine the model; a water shell 5 Å thick was included. NH<sub>2</sub>- and COOH-termini are indicated. Variant amino acids (one letter code) are represented as ball and stick side chains, protruding from the main chain represented as a solid ribbon. Disulfide bridge-forming atoms, used as a reference, are represented by dark sticks, and labeled with numbers corresponding to the cysteines involved. Other identical amino acids are not shown. The complete primary structure of these two proteins is shown in Table 1.

relative positions of the disulfide bridges of toxin 2 correspond to equivalent positions in variant 3 of *C. sculpturatus* [1] and toxin II from *A. australis* [6].

We next modelled the structure of toxin 2 of *C. noxius*, using the available coordinates of the three-dimensional structure of variant 3 (Brookhaven entry 1SN3, from the original reference of Fontecilla-Camps et. al. [1]). The alignment of the two sequences was straightforward and the amino acid residues of toxin 2 were substituted on the SN3 structure. A simple energy minimization protocol revealed very good compatibility of the toxin 2 sequence with that of SN3 structure. In order to have a slightly better sampling of conformational space we also conducted limited molecular dynamics simulations which further adjusted the model to a sterically plausible one.

Inspection of the model in Fig. 1 shows the preservation of a face rich in aromatic residues (corresponding to the front of the 'hand model' proposed by Fontecilla-Camps et al. [1], shown here facing away from the viewer). The figure uses a ribbon representation, in which only the side chain corresponding to variable amino acids are shown as stick representations. Numerous amino acid differences between the two toxins occur mainly towards the opposite face (back of the hand model, facing towards the viewer in the figure). For example, in CN2 there are charged amino acids, such as glutamic acid at position 15 (Glu<sup>15</sup>), and Arg<sup>64</sup>, facing this side, which replace, respectively, Gly<sup>15</sup> and Ser<sup>64</sup>, present in SN3. Also noteworthy is the substitution of Asp<sup>7</sup> in CN2 for Lys<sup>7</sup> of SN3, which faces the front of the hand, at the

Table 1

Amino acid sequence of heterodimeric peptides corresponding to the disulphide bridges of toxin 2

Elution time HPLC	Amino Acid Sequence						Corresponding disulfide
47.15 (9.82) min	1	2	3	4	5	6	Cys12-Cys65
	Asn-Thr-Gly-CYS-Lys-Tyr						
	Xxx-Ser						
59.28 (8.51) min	1	2	3	4	5		Cys29-Cys48
	Glu-CYS-Lys						
	Xxx-Thr-His-Leu-Tyr						
64.36 (8.28) min	1	2	3				Cys16-Cys41
	Glu-CYS-Leu						
	Xxx-Tyr						
68.67 min	1	2	3				Cys25-Cys46
	Xxx-Leu-Arg						
	Ala-CYS-Trp						
	1	10	20	30	40	50	60
CN2	KEGYLV	DKNTG	CKYECL	KLGDND	YCLRECK	QQGYKG	AGGYYCYAFACWCTHLYEQAI
SN3	KEGYLV	VKKS	SDGCKY	GCLKL	GENEG	CDTECKA	KNQGGSGYGYCYAFACWCEGLPESTPTTYPLPNKSC

Amino acid sequences were obtained by microsequencing peptides purified by HPLC after cleavage with chymotrypsin and trypsin. Numbers on top of the amino acid mean the positions in the sequence; Xxx, blank position corresponding to a 1/2 Cys residue (cysteine); CYS, position of a cysteine residue (usually in our machine we can identify a component that absorbs at 330 nm which corresponds to this amino acid). The two bottom lines contain the complete primary structure of CN2, as determined by Zamudio et al. [7], and SN3, as corrected by Fontecilla-Camps et al. [1]. Cysteinyl residues are printed in bold.

bottom center of Fig. 1. From observation of Fig. 1 it seems that the portion facing forward at the back of the hand of these two molecules may play an important role in defining species specificity. On the other hand, the analysis made earlier by Kobayashi et al. [16] shows another important structural feature of the scorpion toxins. The spatial arrangement of these toxins seems to follow the same pattern: they all have the same general motif, around which many variants have been constructed by natural selection [17]; the only highly conserved residues are the cysteines [16]. Thus the substitutions of amino acids in key positions of the toxin, at the protruding J- and B-loops, located, respectively, at relative positions 15–25 and 40–45 [18], and at the confluent N- and C-terminal regions of the molecules [1,5,10], seem to be important structural determinants of the species specificity and/or fine mechanism of action of these toxins (eg.  $\alpha$ - vs.  $\beta$ -scorpion toxins), as already noted [4]. Loret et al. [19] have proposed that the species specificity might reside in the C-terminal region of the molecule. Toxins having proline-rich segments at the C-terminus are thought to be more rigid, so they would not fit well into the binding sites of the Na<sup>+</sup> channels of mammals, while those having a more flexible C-terminus, like  $\gamma$ -toxin (toxin VII) from *Tityus serrulatus* [19], can recognize the Na<sup>+</sup> channel of either mammals or insects.

Although chemical modification of lysine at position 56 of toxin II from *A. australis* [20] showed it to be crucial for toxicity, the study of the structure–function relationship of Na<sup>+</sup> channel scorpion toxins is less ad-

vanced than in other toxins, such as charybdotoxin, a K<sup>+</sup> channel toxin [21]. Park and Miller [21] showed by genetic engineering that a single lysine residue at position 27 was fundamental to toxin activity. Also, Gurrola et al. [22] and Vaca et al. [23] have found that the N-terminal segment of noxiustoxin, another K<sup>+</sup> channel scorpion toxin, was important for its toxicity. The latter authors showed that a synthetically prepared nonapeptide of noxiustoxin was sufficient for channel recognition and modulation.

Thus, concerning the species specificity of Na<sup>+</sup> channel scorpion toxins, chemical or genetic modifications of the residues discussed above should still be conducted before a clear-cut picture can emerge. The construction and expression of artificially mutated toxins, by genetic engineering of the cloned cDNA genes [24–26], should allow the hypotheses discussed above to be tested.

**Acknowledgements:** This research was partially supported by a Howard Hughes Medical Institute award (75191–527104), DGAPA-UNAM Grant IN 205893; and Mexican CONACYT Grant 0018-N9105 to L.D.P. The technical assistance of Timoteo Olamendi and Fredy Coronas is kindly appreciated.

## References

- [1] Fontecilla-Camps, J.C., Almassy, R.J., Suddath, F.L., Watt, D. D. and Bugg, C.E. (1980) Proc. Natl. Acad. Sci. USA 77, 6496–6500. See also recent paper: Zhao, B., Carson, M., Ealick, S.E. and Bugg, C.E. (1992) J. Mol. 227, 239–252.
- [2] Catterall, W. (1980) Annu. Rev. Pharmacol. Toxicol. 20, 15–43.

- [3] Rochat, H., Bernard, P. and Couraud, F. (1979) in: *Advances in Cytopharmacology*, vol. 3 (Ceccarelli, B. and Clementi, F. eds.) pp. 325–334, Raven Press, NY.
- [4] Possani, L.D. (1984) in: *Handbook of Natural Toxins*, vol.2 (Tu, A.T. ed.) pp. 513–550, Marcel Dekker, NY.
- [5] Fontecilla-Camps, J.C., Habersetzer-Rochat, C. and Rochat, H. (1988) *Proc. Natl. Acad. Sci. USA* 85, 7443–7447. See also: Pashkov, V.S., Maiorov, V.N., Bystrov, V.F., Hoang, A.N., Volkova, T.M. and Grishin, E.V. (1988) *Biophys. Chem.* 31, 121–131. See also: Mikou, A., LaPlante, S.R., Guittet, E., Lallemand, J.Y., Martin-Eau Claire, M.F. and Rochat, H. (1992) *J. Biomol. NMR*, 2, 57–70.
- [6] Kopeyan, C., Martinez, G., Lissitzky, S., Miranda, F. and Rochat, H. (1974) *Eur. J. Biochem.* 47, 483–489.
- [7] Zamudio, F., Saavedra, R., Martin, B.M., Gurrola, G.B., Herion, P. and Possani, L.D. (1992) *Eur. J. Biochem.* 204, 281–292.
- [8] Sugg, E.E., Garcia, M.L., Reuben, J.P., Patchett, A.A. and Kacsorowski, G.J. (1990) *J. Biol. Chem.* 265, 18745–18748.
- [9] Jover, E., Couraud, F. and Rochat, H. (1980) *Biochem. Biophys. Res. Commun.* 95, 1607–1624. See also: Couraud, F., Jover, E., Dubois, J.M. and Rochat, H. (1982) *Toxicon*, 20, 9–16.
- [10] Darbon, H., Rochat, H., Kopeyan, C., Van Rietschoten, J. (1982) *Toxicon* 20, 64 (Abstract). See also Darbon, H., Zlotkin, E., Kopeyan, C., Van Rietschoten, J. and Rochat, H. (1982) *Int. J. Pept. Protein Res.* 20, 320–330.
- [11] Possani, L.D., Steinmetz, W.E., Dent, M.A.R., Alagon, A.C. and Wuthrich, K. (1981) *Biochim. Biophys. Acta* 669, 183–192.
- [12] arbone, E., Wanke, E., Prestipino, G., Possani, L.D. and Maelicke, A. (1982) *Nature* 296, 90–91.
- [13] Sitges, M., Possani, L.D. and Bayon, A. (1987) *J. Neurochem.* 48, 1745–1752.
- [14] Yatani, A., Kirsch, G.E., Possani, L.D. and Brown, A.M. (1988) *Am. J. Physiol.* 254 (Heart Circ. Physiol. 23) H443–H451.
- [15] Kirsch, G.E., Skattebol, A., Possani, L.D. and Brown, A.M. (1989) *J. Gen. Physiol.* 93, 67–83.
- [16] Kobayashi, J., Takashima, H., Tamaoki, H., Kyogoku, Y., Lambert, P., Kuroda, H., Chino, N., Watanabe, T.X., Kimura, T., Sakakibara, S., Moroder, L., (1991) *Biopolymers* 31, 1213–1220.
- [17] Menez, A., Bontems, F., Roumestand, C., Gilquin, B. and Toma, F. (1992) *Proc. Royal Soc. Edinburg* 99B, 83–103.
- [18] Meves, H., Simard, J.M. and Watt, D.D. (1984) *J. Physiol., Paris*, 79, 185–191.
- [19] Loret, E.P., Martin-Eauclaire, M.F., Mansuelle, P., Sampieri, F., Granier, C. and Rochat, H. (1991) *Biochemistry* 30, 633–640.
- [20] Habersetzer-Rochat, C. and Sampieri, F. (1976) *Biochemistry* 15, 2254–2261.
- [21] Park, C.S. and Miller, C. (1992) *Biochemistry* 31, 7749–7755.
- [22] Gurrola, G.B., Molinar-Rode, R., Sitges, M., Bayon, A. and Possani, L.D. (1989) *J. Neural Transmission* 77, 11–20.
- [23] Vaca, L., Gurrola, G.B., Possani, L.D. and Kunze, D.L. (1993) *J. Membrane Biol.* 134, 123–129.
- [24] Bougis, P.E., Rochat, H. and Smith, L.A. (1989) *J. Biol. Chem.* 264, 19259–19265.
- [25] Becerril, B., Vazquez, A., Garcia, C., Corona, M., Bolivar, F. and Possani, L.D. (1993) *Gene* 128, 165–171.
- [26] Becerril, B., Corona, M., Mejia, B.M., Martin, B.M., Lucas, S., Bolivar, F. and Possani, L.D. (1993) *FEBS Lett.* 335, 6–8.