The three-dimensional structure of the toxic peptide Cl13 from the scorpion Centruroides limpidus

Andrea Estefanía López-Giraldo, Timoteo Olamendi-Portugal, Lidia Riaño-Umbarila, Baltazar Becerril, Lourival D. Possani, Muriel Delepierre, Federico del Río-Portilla

Title: **The three-dimensional structure of the toxic peptide Cl13 from the scorpion *Centruroides limpidus***

Name of authors: Andrea Estefanía López-Giraldo[1], Timoteo Olamendi-Portugal[2], Lidia Riaño-Umbarila[2,3], Baltazar Becerril[2], Lourival D. Possani[2], Muriel Delepierre[4*] and Federico del Río-Portilla[1*]

Institutions:

1. Instituto de Química, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, CdMx, 04510, México.

2. Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad, 2001 Colonia Chamilpa, Cuernavaca, Morelos 62210 México.

3. Cátedra CONACYT - Instituto de Biotecnología, Universidad Nacional Autónoma de Mexico, Av. Universidad, 2001 Colonia Chamilpa, Cuernavaca, Morelos 62210 México

4. CNRS UMR3528, Department of Structural Biology and Chemistry, Institut Pasteur, Paris, France.

Abstract

Cl13 is a toxin purified previously from the venom of the Mexican scorpion *Centruroides limpidus*. This toxin affects the function of voltage gated Na^+-channels, human subtypes Nav1.4, Nav1.5 and Nav1.6 in a similar manner as other known β-toxins from scorpion venoms. Here, we report a correction of the primary structure of Cl13, previously published. The peptide does contain 66 amino acids, but residue 58 is a tryptophan and the last C-terminal amino acid is an amidated lysine, instead of arginine. The main contribution of this communication is the determination of the 3D-structure of Cl13, by solution NMR, showing that Cl13 has the classical cysteine-stabilized α/β (CSα/β) folding. It has a triple stranded antiparallel beta sheet commonly present in scorpion sodium channel a-toxins. In addition, we report and discuss a comparison of Cl13 structure with two other toxins (Cn2 and Css2) from scorpions of the same genus *Centruroides*, which shows important surface

* To whom correspondence should be addressed

similarities with the structure reported here. Finally, the lack of neutralization of Cl13 toxin by two single-chain antibody fragments (scFvs), named LR and 10FG2, which are capable of neutralizing various toxins from Mexican scorpions, is revised. In particular, 10FG2 is capable of neutralizing toxins Cll1 and Cll2 of the same scorpion *C. limpidus*. The reasons why LR and 10FG2 are unable of neutralizing Cl13 toxin are discussed.

Keywords: NMR solution structure; scorpion toxins; sodium channels; toxin Cl13.

1. Introduction

Mexico is the country where a very high number of human stings by scorpions occurs (Chippaux and Goyffon, 2008).It is estimated to be close to 300,000 people stung per year (Secretaria de Salud, 2015). It also harbors the highest biodiversity of scorpion species in the world, with 289 species described until now (Santibáñez-López et al., 2015). From this number of species, only 21 are assumed to be dangerous to humans (González-Santillán and Possani, 2018). *Centruroides limpidus,* here thereafter abbreviated *C. limpidus,* is possibly the most important species in this regard, because it is extensively distributed in the states of Morelos, Guerrero, Mexico and Michoacan. The venom of this species has been extensively studied (Alagón et al., 1988; Cid-Uribe et al., 2019; Dehesa-Dávila et al., 1996; Lebreton et al., 1994; Riaño-Umbarila et al., 2013). Initially, two main toxic peptides were described Cll1 (Ramírez et al., 1994) and Cll2 (Dehesa-Dávila et al., 1996) which are capable of impairing proper function of sodium ion-channels (Possani et al., 1999). Our group is interested in developing human antibody fragments capable of protection against toxins of this scorpion and others of the genus *Centruroides* of Mexico (Riaño-Umbarila et al., 2019, 2016). Previous work by our group showed that a single-chain antibody (LR) directed against toxin Cn2 of another species of scorpion (*C. noxius*) was capable of neutralizing the entire soluble venom of this species (Riaño-Umbarila et al., 2011). Similarly, a single-chain antibody of human origin (10FG2) showed to be adequate for complete neutralization of the above mentioned toxins Cll1 and Cll2 (Riaño-Umbarila et al., 2019). Besides, a mixture of these antibody fragments was not capable of fully protecting against the entire soluble venom of *C. limpidus*, contrary to what it was shown with toxin Cn2 and the corresponding whole venom as well as some other venoms. The experimental mice injected with these two single-chain antibodies survived the challenge,

but still showed symptoms of intoxication. This prompted us to conduct further characterization of possible minor toxic components present in the venom of *C. limpidus*, which were not protected by our newly developed antibodies. In fact, a minor component, denominated Cl13 was found. It is very toxic to mice. The peptide was isolated, sequenced, physiologically characterized and published (Olamendi-Portugal et al., 2017). For the scFv 10FG2 that recognizes Cll1 and Cll2, the cognate epitope on the toxins have been located. Since the amino acid sequence similarity of Cl13 compared to that of Cll1 and Cll2 is over 70% identical, it suggests the need for further knowledge of their 3D-structures. For this reason the determination of the three-dimensional structure of Cl13 was undertaken. With this information, we hoped conducting comparative studies with the various scorpion toxins of the genus *Centruroides*, for which the structure and function are known. The NMR analysis of Cl13 indicated that a discrepancy in the reported primary structure of peptide was found. The sequence analysis by Edman degradation of the peptide was repeated with more material and the discrepancy was elucidated. The conduction of this work was very important, as we describe below, because the NMR results helped solving an ambiguity found in the primary structure of Cl13, as earlier published. Position 58 corresponds to a tryptophan and position 66 was identified as an amidated lysine, instead of an arginine. In addition, the determination of Cl13 3D-structure allowed to conduct a comparative analysis with the other known toxic peptides, which is the main contribution of this manuscript.

2. Material and Methods

2.1 Source of material

Scorpions of the species *C. limpidus* were collected in Iguala (latitude 17.850, longitude 100.367) with official permit of SEMARNAT (number 004474/18), brought to the laboratory and milked for venom extraction. The soluble venom was fractionated, as earlier described by our group (Olamendi-Portugal et al., 2017) for obtaining purified Cl13. The pure peptide was used for NMR studies and *de novo* peptide sequencing.

2.2 Amino acid sequence

A PPSQ-31A Protein Sequencer from Shimadzu Scientific Instruments, Inc. (Columbia, MD, USA) was used for automatic Edman degradation. The peptide Cl13 was used in native and alkylated format for sequence determination. In order to obtain overlapping

segments of the Cl13 the peptide was enzymatically digested and processed as earlier described (Olamendi-Portugal et al., 2017).

2.3 Mass spectrometry

The molecular weights of toxins and peptides after digestion were obtained by mass spectrometry measurements using a LCQFleet apparatus, from Thermo Fisher Inc, (San Diego, CA, USA).

2.4 NMR procedures

NMR experiments were recorded on an either 800 MHz Avance NEO (Bruker Biospin, Billerica, MA, USA) spectrometer with an 18.1 Tesla magnetic field or a 600MHz Avance III HD spectrometer with a 14.1 Tesla magnetic field. The two spectrometers were equipped with a cryogenically cooled triple resonance 1H[13C/15N] probe. Spectra were recorded using TOPSPIN 4.05 (Bruker Biospin). The lyophilized toxin (about 4 mg) was dissolved in 180 µl of 10% D2O 20 mM acetate pH 5 buffer (Euriso-Top, Saclay, France) to produce a 2.8 mM (7846 MW) sample that was studied by NMR using a 3 mm Shigemi tube (Shigemi, Alison Park, PA, USA). Experiments were run at 283, 288, 298, 303, 313 and 318 K with either a 9615 Hz or 7194 Hz sweep widths. Spectra were referenced to external DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid). The 2D proton NMR spectra were collected with 4 K data points in the t2 dimension and 512 or 1024 t1 increments, with typically 32-64 scans per increment. For spin system identification and sequential assignment, the following 2D NMR homonuclear experiments were recorded: (a) total correlation spectroscopy or Clean TOCSY (Griesinger et al., 1988) using a MLEV-17 pulse scheme with 80 ms or 40 ms isotropic mixing period (Bax and Davis, 1985; Levitt et al., 1982), (b) double-quantum spectroscopy (DQCOSY) (Wagner and Zuiderweg, 1983), (c) double-quantum filtered COSY (DQFCOSY) (Piantini et al., 1982; Rance et al., 1983) and (d) nuclear Overhauser effect spectra (NOESY) (States et al., 1982). NOESY spectra were acquired with mixing times of 100, 150 and 250 ms. Water suppression was achieved either by pre-saturation during the recycle delay or, in the NOESY experiment, by double pulsed-field gradient spin echoes (DPFGSE) (Dalvit et al., 1991; Hwang and Shaka, 1995). Assignment of signals to peptide protons was achieved by the standard method developed

by Wüthrich (Wüthrich, 1986). HN-Hα coupling constants were measured using the modified J-doubling method in the frequency domain (Delrioportilla et al., 1994; Garza-García et al., 2001) from 2D-TOCSY at 303 K traces. To satisfy the J-doubling method requirement, TOCSY experiment was re-transformed using 1024 X 65,536 complex points to have a resolution of 0.20 Hz / point for each trace used.

In addition, two dimensional heteronuclear experiments were run at 308K on the 800 MHz, namely ^{15}N HSQC, Multiplicity edited ^{13}C HSQC, ^{13}C HSQC-TOCSY and ^{13}C HMBC, to back up the proton assignments (Medvedeva et al., 1993). ^{13}C and ^{15}N chemical shift assignments were used to obtain Phi and Psi angles using the TalosN web server (Shen and Bax, 2013).

2.5 Structural calculation

A homonuclear ^{1}H-^{1}H NOESY spectrum acquired at 308 K with a 150 ms mixing time was used to obtain distance constraints for structure calculations. The experimental results obtained by NMR were processed with the software nmrPipe (Delaglio et al., 1995). CARA 1.5 (Keller, 2004) software was used for spin system identification, peak picking and NOE derived distance constraints assignments. The structural calculation was run using CYANA 2.1 (Downing and Güntert, 2004) with chemical shift tolerance between 0.02 and 0.015 ppm. The 20 protein´s models with the lowest energy values and without violations were selected between 500 structures. Additionally, 20 models with the lowest energy values were minimized and refined using molecular dynamic calculations with AMBER 9, the protocol applied was as described by Gurrola (Gurrola et al., 2012). The refined structural model of the toxin Cl13 was validated and deposited in the Protein Data Bank (PDB: 6VXW). Structures were analyzed and visualized with UCSF Chimera (Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, 2004).

3. Results

3.1 Extraction and sequence

Due to the ambiguity found between the earlier published amino acid sequence regarding position 58 of the primary structure of Cl13, it was decided to purify a greater

amount of toxin and repeat the determination of the Edman degradation. For the sequence reported in Olamendi-Portugal (Olamendi-Portugal et al., 2017), we used 75 µg of purified Cl13; for the present communication we used 250 µg. The C-terminal fragment obtained by digesting the peptide with Asp-N endopeptidase allowed determining the full sequence from residue number 53 to 66. To our surprise the last residue was lysine, instead of arginine, as earlier surmised based on mass spectrometry results. The first time the sequence was obtained by Edman degradation did not allow the identification of residue in position 66. The molecular mass obtained from the purified peptides of the C-terminal segments Glu53 to Lys66 was 1723.5 Da (see Olamendi-Portugal et al., 2017) and under the repeated experiment reported here is 1722.9 Da, which within the experimental error of our mass spectrometer means that it corresponds to the same molecular weight. These results agree with the expected molecular mass of the peptide found. It is important to mention that this C-terminal peptide had the cysteine reduced and alkylated. Knowing that the last residue was an amidated Lys66 and that in position 58 there was Trp, as shown on the NMR results, the fully corrected amino acid sequence of Cl13 was obtained. The molecular mass experimentally observed is (7846.6) and the expected molecular mass for the newly determined sequence is (7846.0), thus confirming the right amino acid sequence of Cl13.

3.2 NMR assignments and structure calculations

The sequence specific assignment of Cl13 were acquired according to a standard procedure (Wüthrich, 1986) using spectra recorded at different temperatures ranging from 283 to 313 K to solve ambiguities due to overlapping signals. Although, linewidths of amide resonances are quite broad at 283 K for all the amide protons those sharpened up differentially when increasing the temperature the ultimate case being Phe44 amide proton only detectable at 313 K. Briefly, spin systems of all amino acid residues were identified via their through-bond connectivities observed in DQFCOSY, TOCSY and DQCOSY whereas sequential assignments were obtained using the through space connectivities (NH_i-NH_{i+1}, $H\alpha_i$-NH_{i+1} and $H\beta_i$-NH_{i+1}) observed between neighboring residues in the NOESY spectra. While fairly strong NOEs sequential interactions were observed between the Hα proton of Leu60 and the Hδ of Pro61 allowing to assign the *trans* conformation for the Leu60-Pro61 peptide bond. Although unambiguous NOE sequential interaction could not

be stated between the Hα protons of Trp58 and Pro59 as a result of unfortunate overlap between Phe44 Hα and Hβ protons with Trp58Hα and Pro59 Hα protons respectively, the up-field shifted value of the Pro59 Hα chemical shift, 3.62 ppm together with the lack of NOE interaction between Trp58 Hα and Pro59 Hδ protons are in favor of a Trp58-Pro59 peptide bond in the *cis* conformation (Dyson et al., 1988).

To complete the proton assignments 2D ^{13}C-^{1}H HSQC, HSQC-TOCSY and HMBC were acquired using the natural unlabeled protein. The carbon chemical shifts displayed characteristic ranges for each amino acid providing sufficient information to back up the proton assignments. The ^{13}C-HSQC with multiplicity editing was of significant importance because it allowed direct identification of Gly Hα protons, Val, Ser and Thr Hβ protons. The ^{13}C-^{1}H HSQC-TOCSY was quite helpful to complete side chain assignments of Pro, Lys, Arg whereas the HMBC was used to complete carbon signals of aromatic side chains (Medvedeva et al., 1993).

The Cl13 HN-Hα coupling constants, proton, carbon and nitrogen chemical shifts are deposited in the Biological Magnetic Resonance Bank (BMRB: 30727). The 150 ms mixing time NOESY and 80 ms TOCSY at 308 K spectra showed a reasonable signal dispersion which allowed identifying the 66 spin systems of all amino acid residues; however, the dispersion was not good enough to measure Hα-Hβ coupling constants. The ^{1}H, ^{13}C and ^{15}N completeness assignment are 93%, 66% and 72% respectively at natural abundance sample used. The solution structure of Cl13 (Figure 1) was obtained from 2,102 NOEs with off diagonal assignment; 253 of them are long-range NOEs |i-j |>= 5, 132 medium-range NOEs 1 <|i-j|< 5 and 1,717 short-range NOEs |i-j|<= 1 (Figure 2), 67 angle restrictions from TalosN and 48 HN-Hα coupling constants. Figure 1A shows the Cl13 NMR solution structures, which comprise an α-helix (Tyr24-Arg32), a triple stranded antiparallel β-sheet (Lys1-Tyr4, Ala45-Leu51 and His38-Tyr42), three loops (β1-α1, α1- β2 and β2-β3) and four disulfide bridges. Figure 1B shows the superposition of the 20 lowest energy structures. The final average backbone RMSD to mean was 0.66 and considering the heavy atoms RMSD to mean was 1.04. A summary of the experimental constraints and structure calculation statistics are shown in Supplementary Material Table S1. Temperature coefficients of amide proton as a function of temperature were measured from 10°C to

45°C (1D spectra are shown in the Supplementary Material Figure S1). The sample in H_2O was freeze dried and dissolved in D_2O to assign the remaining amide assuming they are hydrogen bonded.

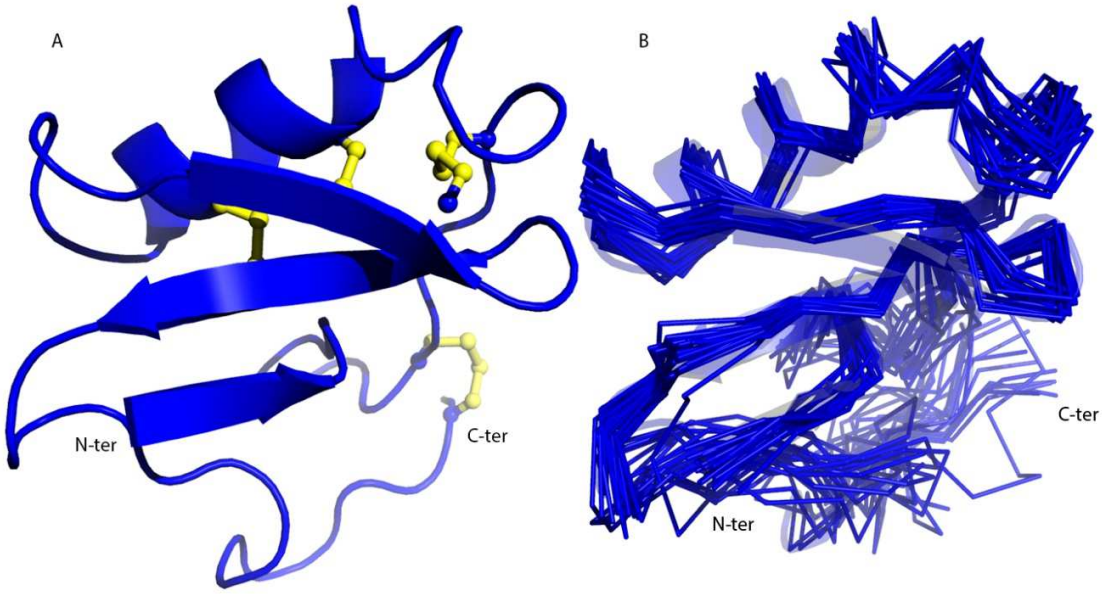


Figure 1. A) NMR solution structure of Cl13 with the lowest energy. The disulfides are painted in yellow. B) Backbone superposition of the 20 lowest energy structures of Cl13. The C-terminus is amidated. Structure calculation without disulfide bonds and superposition of the Cl13 with and without disulfide bond restrictions are in Supplementary Material Figure S2-3 respectively.

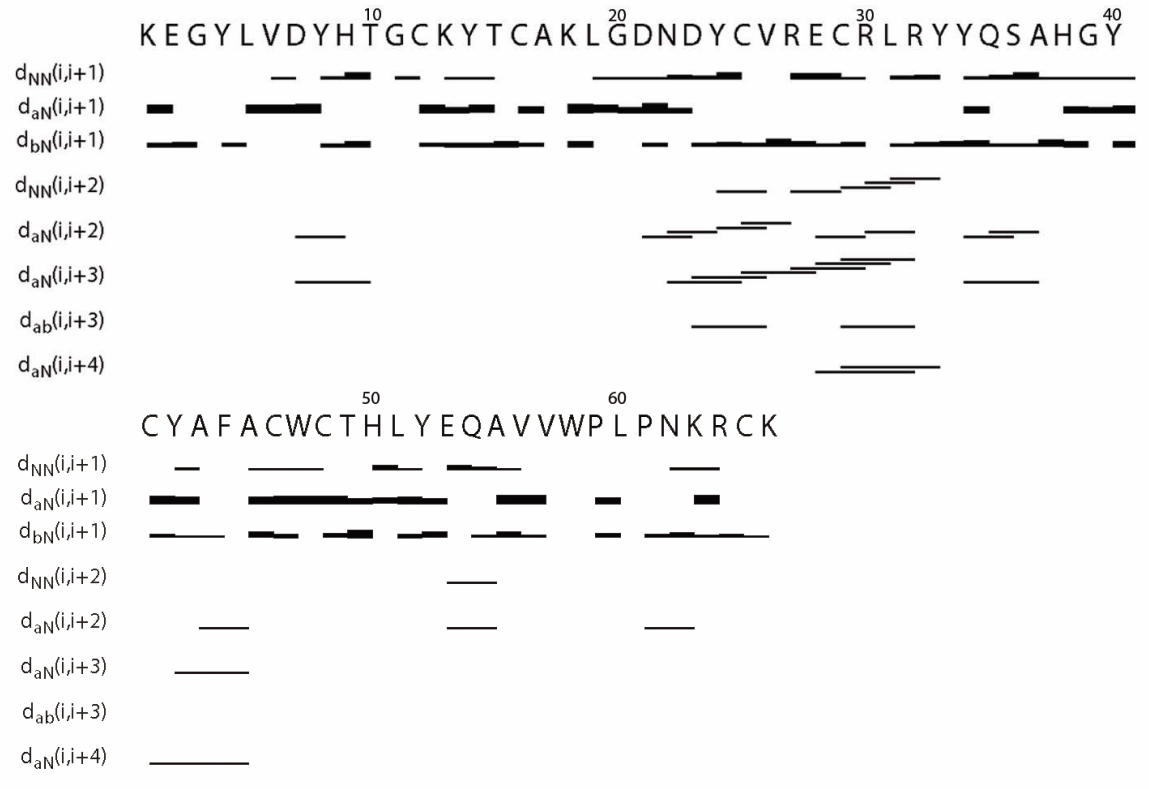


Figure 2. NOE's used in the structure calculations. Thick, medium and thin lines correspond to short, medium and long range NOE's. It could be seen clearly the α-helix in between residues Asp23-Tyr33, a double β-stranded antiparallel β-sheet (His38-Tyr42 and Ala45-Leu49. The β-strand between Glu2-Leu5 is hardly seen; however, the angle measurement and proximity found with the second β-strand shows its existence. NMR statistics and Ramachandran plot are in Supplementary Material Figure S4.

We performed two different structure calculations to define the disulfide bonds: the first one, without S–S bond constraint; and, the second one with S–S bonds (C1–C8, C2–C5, C3–C6, and C4–C7) relative to the primary structure. The two calculations converged to very similar structures, see Figure S-2. The S-S bonds used for the second calculation were those observed by NMR. Indeed, NOE's were clearly observed between $H\beta_i$-$H\beta_j$ and $H\alpha_i$-$H\beta_j$ of the cysteines present in each disulfide bond. We present the Cys29Hβ and-Cys48 Hβ NOE correlation in Figure 3. It can be seen that no overlap can hinder the interpretation. These calculations and the NOE's indicate that the specific cysteines are close by and therefore should be forming a bond between them. Also, the other Hβ regions (Cys16-Cys41, Cys25-Cys46 and Cys65-Cys12) can be observed in Figure S-3. Hα/Hα, Hα/Hβ NOEs between C41 and C46, which are not covalently linked each other, were also

observed as a result of beta strand formation. Considering these two results and the fact that this disulfide bond arrangement is similar to all other β-toxin whose structures are known, we added those restrictions to the calculation.

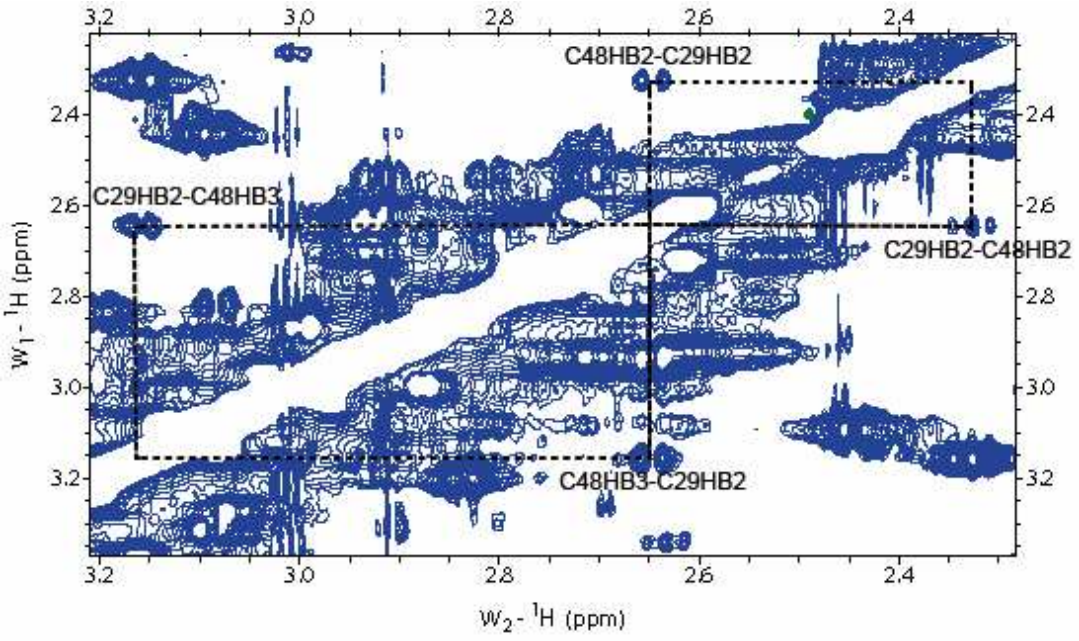


Figure 3. Hβ-Hβ region of Cys29-Cys48 2D NOESY disulfide correlation obtained for Cl13 experiment at 150 ms mixing time, 308 K. pH5 and 2.8 mM toxin concentration; assignments are labeled in black. The whole NOE correlation regions of cysteines forming disulfide bonds are in the Supplementary Material Figure S3. It can be seen that those correlations are well defined. This figure was made with NMRFAM-SPARKY software (Lee et al., 2015).

Figure 4 shows the sequence alignment of Cll1, Cll2, Cl13, Cn2, Css2 and CsEv2 toxins which have high identity percentage and tertiary structures reported in the PDB for Cl13, Cn2, Css2 and CsEv2 toxins only. The amino acids that form β-strands are highlighted in blue meanwhile those that form α-helices are highlighted in red. The differences found in the secondary structure elements are highlighted in Figure 4. Additionally, Cl13 shows a well defined first β-strand. Figure 5 shows Cn2, Css2 and Cl13 backbone superposition of their structures. All of these toxins have a CSα/β folding motif.

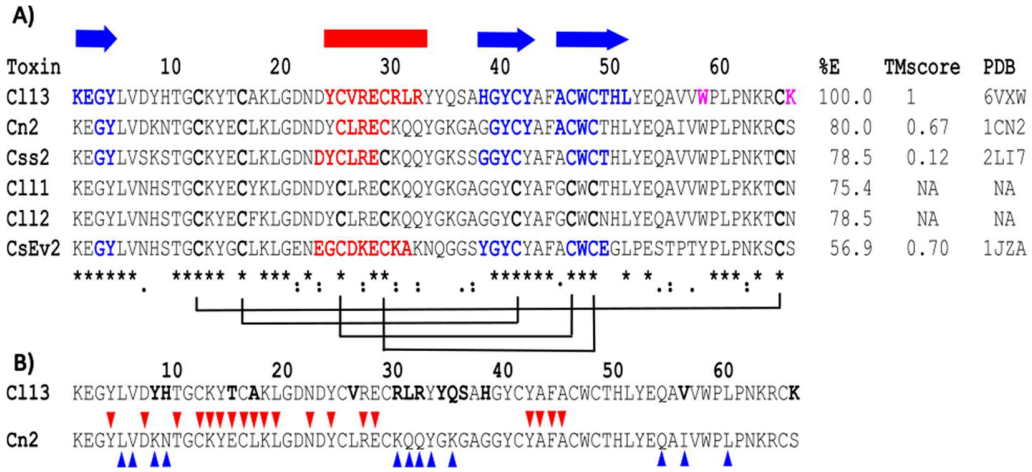


Figure 4. A) Sequence alignment of beta toxins Cl13, Cn2, Css2, Cll1, Cll2 and CsEv2. The two amino acids corrected of Cl13, Trp58 and Lys66, are colored in pink. The β-sheet region is colored in blue and the α-helix in red. Disulfide bonds are marked as C1-C8, C2-C5, C3-C6 and C4-C7. The corrected residues in toxin Cl13 are colored in magenta. Secondary structures were determined via Chimera software. Cll1 and Cll2 do not have their 3D structures reported yet. The asterisk (*), colon (:), and dot (.) indicate identical amino acid residues, conserved substitution, and semi-conserved substitutions in all sequences used in the alignment, respectively. Below, it is shown the disulfide bonds connectivity. This alignment was made with the server T-coffee (Notredame C, Higgins DG, 2000). %E is the identity percentage value considering exact match between amino acids. The TM score (Xu and Zhang, 2010) compares 3D structures, it is convenient to consider not only the sequences alignments, but also 3D structures to see their similarities. TM score makes this comparison and produces a number between 0 and 1. The value is 0 when they do not match. The value is 1 if they match perfectly. B) Sequence alignment of Cn2 and Cl13 highlights the amino acid residues recognized in Cn2 by scFv LR in red triangles and the ones recognized by scFv 10FG2 in blue triangles. Differences between aminoacids C113 and Cn2 are highlighted in bold.

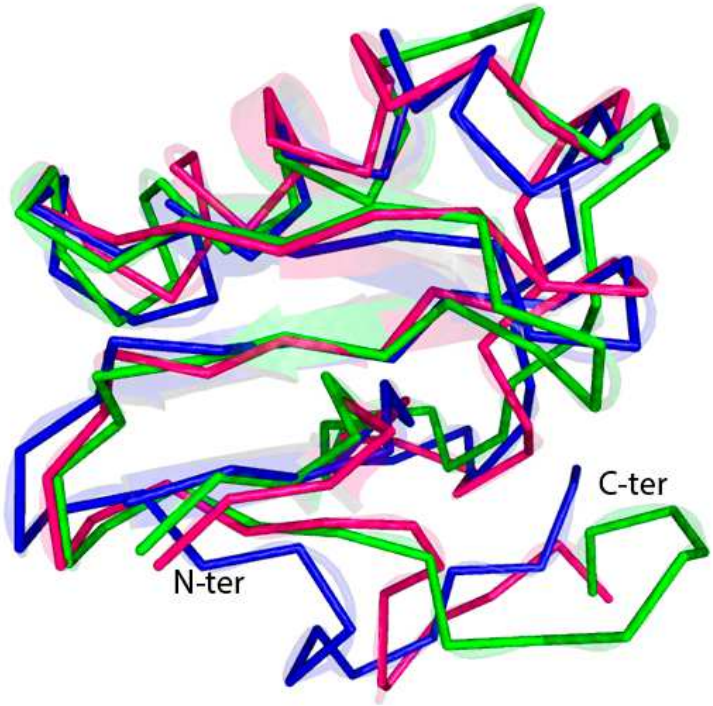


Figure 5. Structure comparison between Cn2 (pink), Css2 (green) and Cl13 (blue) where can be observed that all these toxins have similar structure. We have eliminated the His tag for Css2. Differences in activity should be due to specific amino acid positions.

The structures being established, we did calculate the hydrophobic (figure 6) and the coulombic (figure 7) regions of Cl13, as well as those of Css2 of *Centruroides suffusus suffusus* (Saucedo et al., 2012) and Cn2 of *Centruroides noxius Hoffmann* (Pintar et al., 1999), as these are highly similar in activity and have their 3D structure solved in order to contrast those regions which have been relevant for the toxin function.

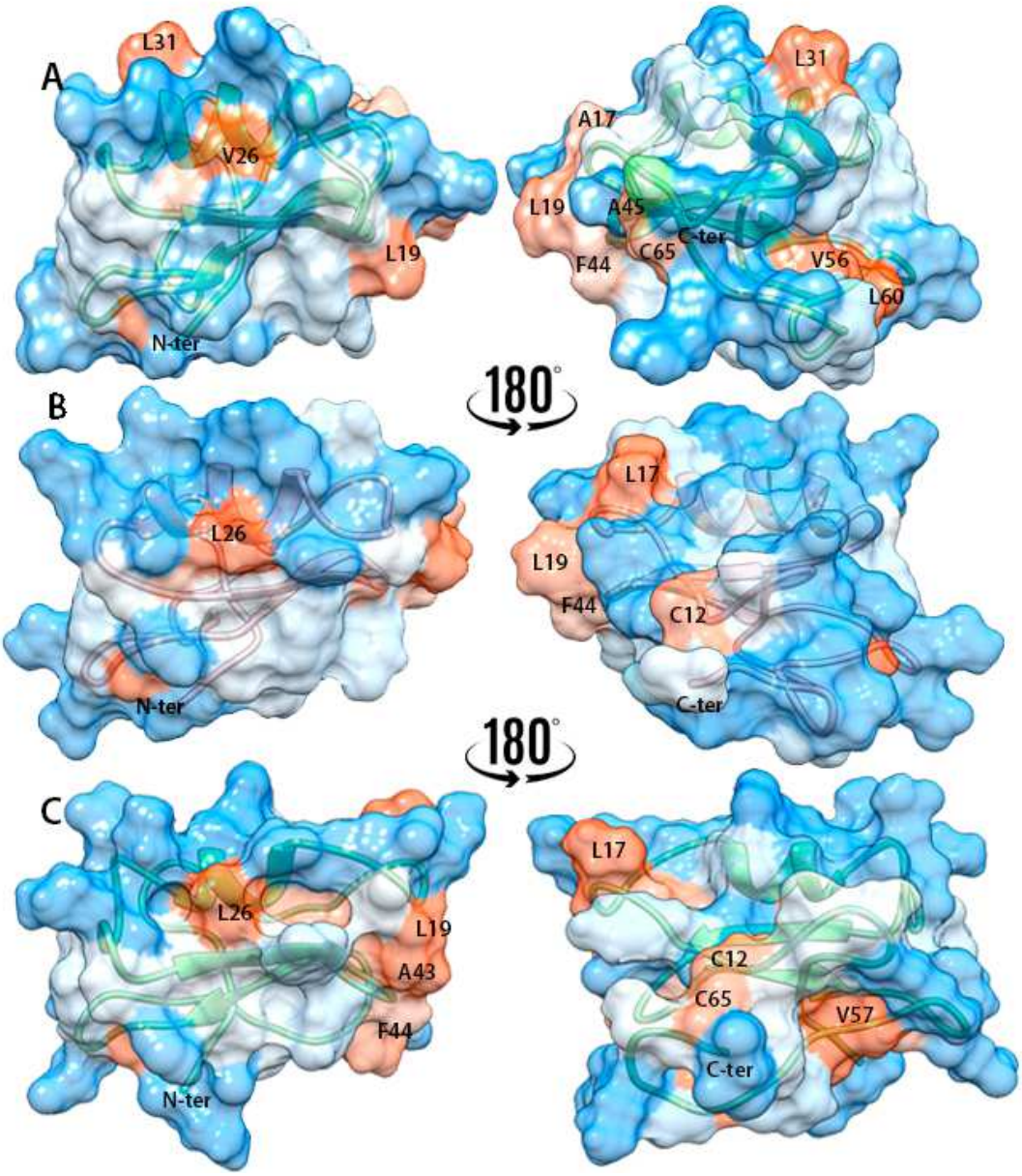
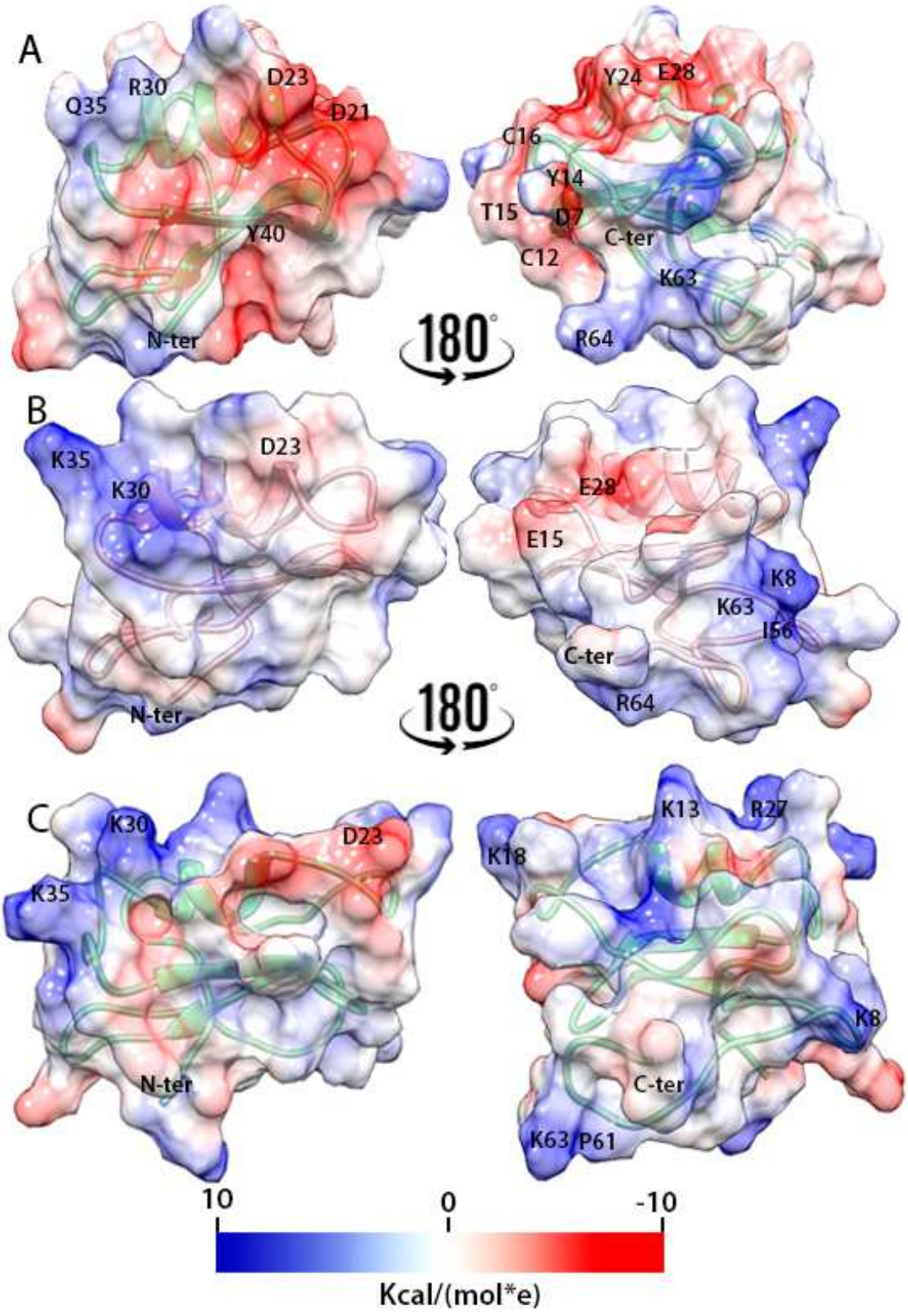


Figure 6. The hydrophobic surfaces were calculated using UCSF Chimera (Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, 2004) of the toxins A) Cl13, B) Css2 and C) Cn2. The surfaces were colored according to a scale of three colors where blue indicates a polar region, white a neutral region and orange a hydrophobic region.


A
R30
D23
Q35
D21
Y40
N-ter
Y24
E28
C16
Y14
T15
D7
C-ter
C12
K63
R64
180°
B
K35
D23
K30
E28
E15
K8
K63
I56
C-ter
R64
N-ter
180°
C
K30
D23
K35
N-ter
K13
R27
K18
K8
C-ter
K63
P61
10
0
-10
Kcal/(mol*e)

1

Figure 7. Comparison between the coulombic surfaces of A) Cl13 and B) Cn2 and C) Css2. Positively charged regions are in blue, negatively charged regions are in red. We use UCSF Chimera software (Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, 2004) to create this figure.

4. Discussion

Cid-Uribe et al. (Cid-Uribe et al., 2019) in their recent published transcriptomic analysis of venom components from *C. limpidus* reported the presence of a gene that possibly codes for the toxin Cl13. The authors report the identification of 59 transcripts with similar amino acid sequence as those described for known Na^+-channel specific toxins. One of them (CLiNaTbet37, C0HK69) was exactly the same as the sequence now corrected for Cl13.

Taking into account the high sequence homology with two other beta-scorpion toxins Css2 of *Centruroides suffusus suffusus* (Saucedo et al., 2012) and Cn2 of *Centruroides noxius Hoffmann* (Pintar et al., 1999) whose structures have been solved by our teams, we used the two NMR experimental beta-scorpion toxin structures in order to compare them with the resulting structure of Cl13 (Figure 1A). Indeed, all of them are toxic to mammals; however, while Cl13 and Css2 have a strong activity over hNav1.5 channel (Olamendi-Portugal et al., 2017; Schiavon et al., 2006), Cn2 has a strong activity over hNav1.6 channel. An intrinsic property of toxin Cl13 is that although it also affects hNav 1.6 channels it displays some interesting and unusual effects on the inactivation process.

The sequence identity of the toxins shown in Figure 4 is higher than 56% (78% with Cn2 and 56.9% with Css2). The 3D structures differed in limited regions, see Figures 3-6. Indeed, sequence alignment of Cl13, Cn2, Css2 and CsEv2 toxins shows a high identity percentage which correlates with their conserved tertiary structures reported in the PDB sharing the typical CSα/β folding. The TM score is 0.67 for Cn2 and Cl13, indicating that there is a good match between these two structures (see Figure 4). The differences between these toxins are only in the length of amino acid segments which form the secondary structure elements. However, the alignment between Cl13 and Css2 produces very different results because there is some extra amino acids present in Css2 which give a small TM score value; nevertheless, these two structures are very similar considering visual comparison (Figure 5). It is interesting to note that the first β-strand in Cl13 involving

residues 2 to 4 while for Cn2 and Css2 only two amino acids are involved. We are thus expecting that binding differences are due to exposed amino acids which results in the lack of antibody recognition. Glu15 in Cn2 was replaced with Thr15 in Cl13 which is exposed (Figure 4.B, 6 and 7). These key differences between Cn2 and Cl13 could help to explain the lack of interaction between scFv LR and toxin Cl13.

Lior Cohen et al (Cohen et al., 2005) while describing the common features in the functional surface of scorpion β-toxins highlighted the role of some hydrophobic amino acids of the Css4 toxin in the interaction with the sodium channel receptor site 4. Some of these residues are conserved in the hydrophobic surfaces of Cl13, Cn2 and Css2 as shown Figure 6. The largest hydrophobic patch is formed by the amino acids Ala17, Leu19 and Phe44, which are conserved in the three toxins. This hydrophobic patch is situated in the loop preceding the helix. The two amino acids Ala17 and Leu19 are also conserved within a hydrophobic patch in Css4 which was demonstrated, introducing some point mutations, that this toxin segment is important for the Css4 function. Another small hydrophobic patch is formed by the amino acid Val26 (Leu26 for Cn2 and Css2) which is conserved in the three toxins and is located on one side of the helix. Finally, there is a third hydrophobic patch in Cl13 formed by Ala45 and Cys65, which is also observed in Cn2 and Css2, but in these toxins it is formed by Cys12 and Cys65.

Relevance of positive residues in the biological activity of the cysteine-stabilized proteins for their binding to channels has been widely studied (Clairfeuille et al., 2019; Estrada et al., 2011; Goldstein et al., 1994; Hidalgo and MacKinnon, 1995; Krezel et al., 1995; Loret et al., 1994; MacHado et al., 2018; Stampe et al., 1994), using different methods like chemical modifications and mutational analysis (Hidalgo and MacKinnon, 1995; Kharrat et al., 1990). For that reason, we compared the coulombic surface of Cl13 with that of other beta toxins which are highly similar in structure and activity namely Cn2 and Css2 (Figure 7). Several positive patches were observed. The first patch in the C-terminal region includes Lys63 for the three toxins, Arg64 for Cn2 and Cl13 and Lys66 for Cl13.

Previously, mutagenesis studies and a neutralizing monoclonal antibody have shown the relevance of His64 in Nav channel modulation which help to guide the α-toxin into a stable binding pose (Clairfeuille et al., 2019). For Cl13 and Cn2 toxins, in this position, it is filled

with an Arg64 which forms a positive patch in C-terminal region with Lys63. In Cn2, this patch is less conserved and is formed by Lys63 and Pro61.

Then, there is a second positive patch formed by the amino acids 30 and 35 that is more or less conserved in the three toxins considering that there is difference in the amino acid composition, this patch is situated in the loop α1-β2. Finally, the amino acid Lys 8 which has been shown to play an important role for the interaction with the channel (Clairfeuille et al., 2019; Fabrichny et al., 2012) in the case of α-toxin, is conserved in Cn2 and Css2, whereas it is replaced by a tyrosine followed by a histidine in Cl13. This leads to the presence of a third positive patch that is not present in Cl13 as clearly observed in Figure 4. It clearly appears that Cl13 displays the largest negative surface as compared to Cn2 and Css2.

It is important to remember the fact that only 10FG2 single-chain antibody fragment of human origin is able to completely neutralize the two β-toxins Cll1 and Cll2, (Riaño-Umbarila et al., 2019) and that a mixture LR and 10FG2 antibody fragments are not capable to fully protect against the entire soluble venom of *C. limpidus*, contrary to what it was shown with toxin Cn2 of *C. noxius*. For this reason the minor component Cl13 was isolated, sequenced and physiologically characterized (Olamendi-Portugal et al., 2017). It is very toxic to mice. The residue Glu15 known to have a crucial role in the action of β-toxins in term of toxicity (Izhar et al., 2004; Schiavon et al., 2012) is replaced here by a threonine. Cl13 is one of the most toxic components that we have studied together with Cn2, which suggests that there are other residues in the toxin that contribute to the binding and alteration of the functioning of the above mentioned sodium channels.

The study described here shows that the three-dimensional structure of Cl13 is very similar to that of Cn2 and Css2 and therefore is likely to be similar to that of Cll1 and Cll2 taking into account the high sequence similarity among these toxins, although the 3D structures of these two last toxins are not known yet (Figure 4). As reported earlier (Olamendi-Portugal et al., 2017), the three stretches namely positions 7-9, 30-38 and 62-66, that have been reported to be different for Cl13 as compared to the two toxins Cll1 and Cll2, might be responsible for the lack of its interaction with the single chain neutralizing antibodies. In previous studies of the 3D structure of the ternary complex of the scFv LR, Cn2 and the

predecessor scFv of 10FG2 (RU1), it was found that the core of the binding site of scFv LR is located near the N-terminus of the α-helix's region of toxin Cn2 whereas that of scFv RU1 is located near the corresponding C-terminus(Riaño-Umbarila et al., 2016).

The alignment shown in figure 4B highlights the amino acid residues recognized in Cn2 by scFv LR in red and the ones recognized by scFv 10FG2 in blue. In Cl13 toxin, the amino acid residues that differ between Cn2 and Cl13 are highlighted in bold with respect to the sequence of the Cn2 toxin, Figure 4B. These differences coincide with important anchor sites of the LR scFv in the Cn2 toxin, since the different contacts formed with Glu15 (a saline bridge with His35 of the VH domain of the scFv, hydrogen bonds with Ala33, Gly102, van der Waals contacts with A33H, G99H, V101H in LR) (Canul-Tec et al., 2011). In the same way, scFv 10FG2 has a very important anchorage with residues Gln31 and Glu32 of Cn2 (Riaño-Umbarila et al., 2019), which are different in Cl13 (Leu31 and Arg31) indicating that these interactions could not be established with toxin Cl13.

5. Concluding remarks

Having assigned all protons chemical shifts and also most of carbon and nitrogen chemical shifts for Cl13 at natural abundance, its NMR solution structure was solved showing a cysteine-stabilized α/β (CSα/β) motif commonly found in sodium channel β-toxins. The structural comparison analysis shows that the three-scorpion sodium β toxins active on mammals share a similar hydrophobic surface and some conserved positive regions, but also a large negative region for Cl13 which suggest that the activity of these disulfide-stabilized proteins is regulated by different interaction interfaces. Indeed, the three stretches already pointed out as well as the nature of residue at position 15, which in Cl13 is a Thr whereas in most of the others is Glu, which might be responsible for the lack of interaction of Cl13 toxin with LR. The changes in positions 30, 31 and 32 do not permit the interaction with 10FG2 and as a consequence, the null neutralization of Cl13 and the deficient neutralization of the *Centruroides limpidus* whole venom.

6. Declaration of interest

The authors declare that they have not known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. Acknowledgements

The authors acknowledge Dr. Fernando Z. Zamudio for the service of determining the Edman degradation and mass spectrometry data of this manuscript. The work in Paris was supported by fund from the Pasteur Institute and the *Centre National de la Recherche Scientifique* (UMR 3528). The 800 MHz NMR spectrometer of the Institut Pasteur was partially funded by the *Région Ile de France* (SESAME 2014 NMRCHR grant no 4014526). Partially supported by grant FORDECYT 303045 from *Consejo Nacional de Ciencia y Tecnología* to BB and LDP; also, *Dirección General del Personal Académico* UNAM IN210319 and *Dirección General de Cómputo y de Tecnologías de la Información y Comunicación*, Supercomputing Project LANCAD-UNAM-DGTIC-145 given to FRP.

8. Bibliography

Alagón, A.C., Guzmán, H.S., Martin, B.M., Ramírez, A.N., Carbone, E., Possani, L.D., 1988. Isolation and characterization of two toxins from the Mexican scorpion Centruroides limpidus limpidus Karsch. Comp. Biochem. Physiol. B. 89, 153–61. https://doi.org/10.1016/0305-0491(88)90277-5

Bax, A., Davis, D.G., 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. J. Magn. Reson. 65, 355–360. https://doi.org/10.1016/0022-2364(85)90018-6

Canul-Tec, J.C., Riaño-Umbarila, L., Rudiño-Piñera, E., Becerril, B., Possani, L.D., Torres-Larios, A., 2011. Structural basis of neutralization of the major toxic component from the scorpion Centruroides noxius Hoffmann by a human-derived single-chain antibody fragment. J. Biol. Chem. 286, 20892–20900. https://doi.org/10.1074/jbc.M111.238410

Chippaux, J.P., Goyffon, M., 2008. Epidemiology of scorpionism: A global appraisal. Acta Trop. https://doi.org/10.1016/j.actatropica.2008.05.021

Cid-Uribe, J.I., Meneses, E.P., Batista, C.V.F., Ortiz, E., Possani, L.D., 2019. Dissecting toxicity: The venom gland transcriptome and the venom proteome of the highly venomous scorpion centruroides limpidus (Karsch, 1879). Toxins (Basel). 11. https://doi.org/10.3390/toxins11050247

Clairfeuille, T., Cloake, A., Infield, D.T., Llongueras, J.P., Arthur, C.P., Li, Z.R., Jian, Y., Martin-Eauclaire, M.F., Bougis, P.E., Ciferri, C., Ahern, C.A., Bosmans, F., Hackos, D.H., Rohou, A., Payandeh, J., 2019. Structural basis of a-scorpion toxin action on Na v channels. Science (80-.). 363, 1–25. https://doi.org/10.1126/science.aav8573

Cohen, L., Karbat, I., Gilles, N., Ilan, N., Benveniste, M., Gordon, D., Gurevitz, M., 2005. Common features in the functional surface of scorpion β-toxins and elements that confer specificity for insect and mammalian voltage-gated sodium channels. J. Biol. Chem. 280, 5045–5053. https://doi.org/10.1074/jbc.M408427200

Dalvit, C., Shapiro, G., Bohlen, J., Parella, T., 1991. Technical aspects of an efficient multiple solvent suppression pulse sequence. Magn. Reson. Chem. 37, 7–14.

Dehesa-Dávila, M., Ramírez, A.N., Zamudio, F.Z., Gurrola-Briones, G., Liévano, A., Darszon, A., Possani, L.D., 1996. Structural and functional comparison of toxins from the venom of the scorpions Centruroides infamatus infamatus, Centruroides limpidus limpidus and Centruroides noxius. Comp. Biochem. Physiol. - B Biochem. Mol. Biol. 113, 331–339. https://doi.org/10.1016/0305-0491(95)02031-4

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., Bax, A., 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293. https://doi.org/10.1007/BF00197809

Delrioportilla, F., Blechta, V., Freeman, R., 1994. Measurement of Poorly Resolved Splittings by J Doubling in the Frequency Domain. J. Magn. Reson. Ser. A. https://doi.org/10.1006/jmra.1994.1238

Downing, A.K., Güntert, P., 2004. Automated NMR Structure Calculation With CYANA. Protein NMR Tech. 278, 353–378. https://doi.org/10.1385/1-59259-809-9:353

Dyson, H.J., Rance, M., Houghten, R.A., Lerner, R.A., Wright, P.E., 1988. Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of a reverse turn. J. Mol. Biol. 201, 161–200. https://doi.org/10.1016/0022-2836(88)90446-9

Estrada, G., Restano-Cassulini, R., Ortiz, E., Possani, L.D., Corzo, G., 2011. Addition of positive charges at the C-terminal peptide region of CssII, a mammalian scorpion peptide toxin, improves its affinity for sodium channels Nav1.6. Peptides 32, 75–79. https://doi.org/10.1016/j.peptides.2010.11.001

Fabrichny, I.P., Mondielli, G., Conrod, S., Martin-Eauclaire, M.F., Bourne, Y., Marchot, P., 2012. Structural insights into antibody sequestering and neutralizing of Na + channel α-type modulator from old world scorpion venom. J. Biol. Chem. 287, 14136–14148. https://doi.org/10.1074/jbc.M111.315382

Garza-García, A., Ponzanelli-Velázquez, G., del Río-Portilla F, 2001. Deconvolution and measurement of spin-spin splittings by modified J doubling in the frequency domain. J. Magn. Reson. 148, 214–219. https://doi.org/10.1006/jmre.2000.2235

Goldstein, S.A.N., Pheasant, D.J., Miller, C., 1994. The charybdotoxin receptor of a Shaker K+ channel: Peptide and channel residues mediating molecular recognition. Neuron 12, 1377–1388. https://doi.org/10.1016/0896-6273(94)90452-9

González-Santillán, E., Possani, L.D., 2018. North American scorpion species of public health importance with a reappraisal of historical epidemiology. Acta Trop. 187, 264–274. https://doi.org/10.1016/j.actatropica.2018.08.002

Griesinger, C., Otting, G., Wüthrich, K., Ernst, R.R., 1988. Clean tocsy for 1H spin system identification in macromolecules. J. Am. Chem. Soc. 110, 7870–7872. https://doi.org/10.1021/ja00231a044

Gurrola, G.B., Hernández-López, R.A., Rodríguez De La Vega, R.C., Varga, Z., Batista, C.V.F., Salas-

Castillo, S.P., Panyi, G., Del Río-Portilla, F., Possani, L.D., 2012. Structure, function, and chemical synthesis of Vaejovis mexicanus peptide 24: A novel potent blocker of Kv1.3 potassium channels of human T lymphocytes. Biochemistry 51, 4049–4061.

Hidalgo, P., MacKinnon, R., 1995. Revealing the architecture of a K+ channel pore through mutant cycles with a peptide inhibitor. Science (80-.). 268, 307–310. https://doi.org/10.1126/science.7716527

Hwang, T.L., Shaka, A.J., 1995. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. J. Magn. Reson. - Ser. A 112, 275–279. https://doi.org/10.1006/jmra.1995.1047

Izhar, K., Lior, C., Nicolas, G., Dalia, G., Michael, G., 2004. Conversion of a scorpion toxin agonist into an antagonist highlights an acidic residue involved in voltage sensor trapping during activation of neuronal Na + channels. FASEB J. 18, 683–689. https://doi.org/10.1096/fj.03-0733com

Keller, R., 2004. The computer aided resonance assignment tutorial, Goldau, Switzerland: Cantina Verlag.

Kharrat, R., Darbon, H., Granier, C., Rochat, H., 1990. Structure-activity relationships of scorpion α-neurotoxins: Contribution of arginine residues. Toxicon 28, 509–523. https://doi.org/10.1016/0041-0101(90)90295-I

Krezel, A.M., Kasibhatla, C., Hidalgo, P., Mackinnon, R., Wagner, G., 1995. Solution structure of the potassium channel inhibitor agitoxin 2: Caliper for probing channel geometry. Protein Sci. 4, 1478–1489. https://doi.org/10.1002/pro.5560040805

Lebreton, F., Delepierre, M., Ramírez, A.N., Balderas, C., Possani, L.D., 1994. Primary and NMR Three-Dimensional Structure Determination of a Novel Crustacean Toxin from the Venom of the Scorpion Centruroides limpidus limpidus Karsch. Biochemistry 33, 11135–11149. https://doi.org/10.1021/bi00203a010

Lee, W., Tonelli, M., Markley, J.L., 2015. NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31, 1325–1327. https://doi.org/10.1093/bioinformatics/btu830

Levitt, M.H., Freeman, R., Frenkiel, T., 1982. Broadband heteronuclear decoupling. J. Magn. Reson. 47, 328–330. https://doi.org/10.1016/0022-2364(82)90124-X

Loret, E.P., Soto Del Valle, R.M., Mansuelle, P., Sampieri, F., Rochat, H., 1994. Positively charged amino acid residues located similarly in sea anemone and scorpion toxins. J. Biol. Chem. 269, 16785–16788.

MacHado, L.E.S.F., De Paula, V.S., Pustovalova, Y., Bezsonova, I., Valente, A.P., Korzhnev, D.M., Almeida, F.C.L., 2018. Conformational Dynamics of a Cysteine-Stabilized Plant Defensin Reveals an Evolutionary Mechanism to Expose Hydrophobic Residues. Biochemistry 57, 5797–5806. https://doi.org/10.1021/acs.biochem.8b00753

Medvedeva, S., Simorre, J.P., Brutscher, B., Guerlesquin, F., Marion, D., 1993. Extensive 1H NMR resonance assignment of proteins using natural abundance gradient-enhanced 13C-1H

correlation spectroscopy. FEBS Lett. 333, 251–256. https://doi.org/10.1016/0014-5793(93)80664-G

Notredame C, Higgins DG, H.J., 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 17, 302:205.

Olamendi-Portugal, T., Restano-Cassulini, R., Riaño-Umbarila, L., Becerril, B., Possani, L.D., 2017. Functional and immuno-reactive characterization of a previously undescribed peptide from the venom of the scorpion Centruroides limpidus. Peptides 87, 34–40. https://doi.org/10.1016/j.peptides.2016.11.008

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, F.T.J., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. Comput Chem. 25, 1605-1612.

Piantini, U., Sørensen, O.W., Ernst, R.R., 1982. Multiple Quantum Filters for Elucidating NMR Coupling Networks. J. Am. Chem. Soc. 104, 6800–6801. https://doi.org/10.1021/ja00388a062

Pintar, A., Possani, L.D., Delepierre, M., 1999. Solution structure of toxin 2 from Centruroides noxius Hoffmann, a β-scorpion neurotoxin acting on sodium channels. J. Mol. Biol. 287, 359–367. https://doi.org/10.1006/jmbi.1999.2611

Possani, L.D., Becerril, B., Delepierre, M., Tytgat, J., 1999. Scorpion toxins specific for Na+-channels. Eur. J. Biochem. https://doi.org/10.1046/j.1432-1327.1999.00625.x

Ramírez, A.N., Martin, B.M., Gurrola, G.B., Possani, L.D., 1994. Isolation and characterization of a novel toxin from the venom of the scorpion Centruroides limpidus limpidus Karsch. Toxicon 32, 479–490. https://doi.org/10.1016/0041-0101(94)90300-X

Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R., Wüthrich, K., 1983. Improved spectral resolution in COSY 1H NMR spectra of proteins via double quantum filtering. Biochem. Biophys. Res. Commun. 117, 479–485. https://doi.org/10.1016/0006-291X(83)91225-1

Riaño-Umbarila, L., Contreras-Ferrat, G., Olamendi-Portugal, T., Morelos-Juárez, C., Corzo, G., Possani, L.D., Becerril, B., 2011. Exploiting cross-reactivity to neutralize two different scorpion venoms with one single chain antibody fragment. J. Biol. Chem. 286, 6143–6151. https://doi.org/10.1074/jbc.M110.189175

Riaño-Umbarila, L., Gómez-Ramírez, I. V., Ledezma-Candanoza, L.M., Olamendi-Portugal, T., Rodríguez-Rodríguez, E.R., Fernández-Taboada, G., Possani, L.D., Becerril, B., 2019. Generation of a broadly cross-neutralizing antibody fragment against several mexican scorpion venoms. Toxins (Basel). 11. https://doi.org/10.3390/toxins11010032

Riaño-Umbarila, L., Ledezma-Candanoza, L.M., Serrano-Posada, H., Fernández-Taboada, G., Olamendi-Portugal, T., Rojas-Trejo, S., Gómez-Ramírez, I. V., Rudiño-Piñera, E., Possani, L.D., Becerril, B., 2016. Optimal neutralization of Centruroides noxius venom is understood through a structural complex between two antibody fragments and the Cn2 toxin. J. Biol. Chem. 291, 1619–1630. https://doi.org/10.1074/jbc.M115.685297

Riaño-Umbarila, L., Olamendi-Portugal, T., Morelos-Juárez, C., Gurrola, G.B., Possani, L.D., Becerril,

B., 2013. A novel human recombinant antibody fragment capable of neutralizing Mexican scorpion toxins. Toxicon 76, 370–6. https://doi.org/10.1016/j.toxicon.2013.09.016

Santibáñez-López, C.E., Francke, O.F., Ureta, C., Possani, L.D., 2015. Scorpions from Mexico: From species diversity to venom complexity. Toxins (Basel). https://doi.org/10.3390/toxins8010002

Saucedo, A.L., Del Rio-Portilla, F., Picco, C., Estrada, G., Prestipino, G., Possani, L.D., Delepierre, M., Corzo, G., 2012. Solution structure of native and recombinant expressed toxin CssII from the venom of the scorpion Centruroides suffusus suffusus, and their effects on Nav1.5 Sodium channels. Biochim. Biophys. Acta - Proteins Proteomics 1824, 478–487.

Schiavon, E., Pedraza-Escalona, M., Gurrola, G.B., Olamendi-Portugal, T., Corzo, G., Wanke, E., Possani, L.D., 2012. Negative-shift activation, current reduction and resurgent currents induced by β-toxins from Centruroides scorpions in sodium channels. Toxicon 59, 283–293. https://doi.org/10.1016/j.toxicon.2011.12.003

Schiavon, E., Sacco, T., Cassulini, R.R., Gurrola, G., Tempia, F., Possani, L.D., Wanke, E., 2006. Resurgent current and voltage sensor trapping enhanced activation by a β-scorpion toxin solely in Na v1.6 channel: Significance in mice Purkinje neurons. J. Biol. Chem. 281, 20326–20337. https://doi.org/10.1074/jbc.M600565200

Secretaria de Salud, 2015. Casos por entidad federativa de enfermedades no transmisibles hasta la semana epidemiológica, in: V.D.B.Epidemiologia (Ed.), B. d. V. Epidemiologica. Mexico, p. 52.

Shen, Y., Bax, A., 2013. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J. Biomol. NMR 56, 227–241. https://doi.org/10.1007/s10858-013-9741-y

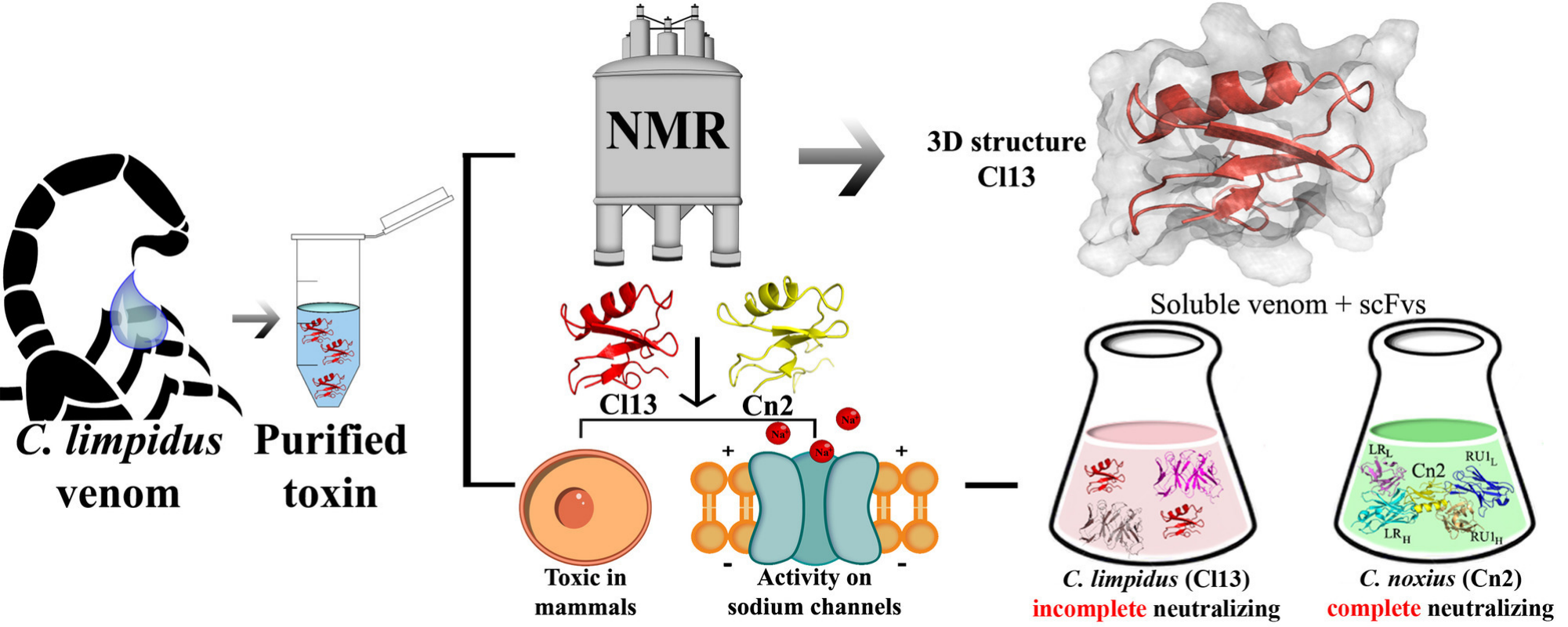
Stampe, P., Kolmakova-Partensky, L., Miller, C., 1994. Intimations of K+ Channel Structure from a Complete Functional Map of the Molecular Surface of Charybdotoxin. Biochemistry 33, 443–450. https://doi.org/10.1021/bi00168a008

States, D.J., Haberkorn, R.A., Ruben, D.J., 1982. A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. J. Magn. Reson. 48, 286–292. https://doi.org/10.1016/0022-2364(82)90279-7

Wagner, G., Zuiderweg, E.R., 1983. Two-dimensional double quantum 1H NMR spectroscopy of proteins. Biochem. Biophys. Res. Commun. 113, 854–60. https://doi.org/10.1016/0006-291x(83)91077-x

Wüthrich, K., 1986. NMR with Proteins and Nucleic Acids. Europhys. News. https://doi.org/10.1051/epn/19861701011

Xu, J., Zhang, Y., 2010. How significant is a protein structure similarity with TM-score = 0.5? Bioinformatics 26, 889–895. https://doi.org/10.1093/bioinformatics/btq066


C. limpidus
venom
Purified
toxin
NMR
3D structure
Cl13
Cl13
Cn2
Toxic in
mammals
Activity on
sodium channels
Soluble venom + scFvs
C. limpidus (Cl13)
incomplete neutralizing
C. noxius (Cn2)
complete neutralizing
Cn2