ELSEVIER

Journal of Molecular Graphics and Modelling 20 (2002) 389–398

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

A proposed 3D structure for crotamine based on homology building, molecular simulations and circular dichroism

A.M. Siqueira [a], N.F. Martins [b], M.E. De Lima [a], C.R. Diniz [c], A. Cartier [d], D. Brown [e], B. Maigret [d,f,*,1]

[a] *Department of Biochemistry and Immunology, Federal University of Minas Gerais (UFMG), Caixa Postal 486, Belo Horizonte, MG 31270-901, Brazil*
[b] *Department of Celular Biology, Biology Institute, Brasilia University (UnB), Brasília, DF 70862-030, Brazil*
[c] *Fundação Ezequiel Dias, Belo Horizonte, MG, Brazil*
[d] *Laboratory of Theoretical Chemistry, UMR CNRS/UHP 756, Henri Poincaré University, BP 239, 54506 Vandoeuvre les Nancy Cedex, France*
[e] *Laboratory of Organic Materials with Specific Properties, UMR CNRS 5041, University of Savoie, Savoie Technolac, 73376 Bourget du Lac, France*
[f] *Institute of Chemistry, Brasilia University (UnB), Brasilia, DF 70862-030, Brazil*

Received 2 April 2001; received in revised form 30 October 2001; accepted 10 November 2001

Abstract

Crotamine, isolated from the venom of the South American rattlesnake *Crotalus durissus terrificus* is a strongly basic 42-aminoacid polypeptide belonging to the small basic myotoxin family. As no tridimensional structure is available for this myotoxin subfamily, despite its important pharmacological interest, we propose in this paper a theoretical 3D model for crotamine. Starting from a homology modelling procedure, followed by intensive molecular dynamics (MD) simulations in water and complementary CD experiments, the designed 3D model is the first example of a tridimensional structure in this family of small basic myotoxins. Crotamine, therefore, belongs to a newly identified structural family presenting a common fold also found in β-defensin and antopleurine-B. The proposed 3D model will be used for future calculations about crotamine aggregation and interaction with membranes.

Keywords: Homology building; Molecular simulations; Circular dichroism

1. Introduction

Snake venoms contain sets of proteins that are responsible for their neurotoxic, cardiotoxic, hemorrhagic and myotoxic activities. Among these toxins, myotoxins are of particular interest as they exhibit a broad spectrum of interesting biological responses. The myotoxin family comprises several toxin subclasses which may be placed in two main groups: myotoxins such as cardiotoxin, melittin and phospholipase A2, involved in rapid lysis of sarcolemma, myofibril clumping and hypercontraction of sarcomeres; and small basic myotoxins such as crotamine, which bind strongly to excitable membranes leading to the contraction of skeletal muscles [1,2]. Despite a considerable amount of work, there is still no clear understanding of the structure–function relationships of myotoxins that might account for their biological properties.

In this work, we have focused on crotamine because of (i) its potential use for pharmaceutical applications, (ii) the lack of structural data for this myotoxin subfamily; and (iii) because the exact mechanism of their interaction with biological membranes is still subject of much debate.

Crotamine isoforms can be isolated from the venom of the South American rattlesnake *Crotalus durissus terrificus* [3]. It is a strongly basic (pI = 9.51–9.73) polypeptide with a molecular weight of 4.89 kDa, composed of 42 amino acids, containing three disulfide bridges [4]. The general biological action of crotamine is the depolarization of cell membranes. Crotamine acts on the voltage sensitive sodium channels of the skeletal muscle sarcolemma, inducing a sodium influx by the opening of the ryanocine receptor. The toxin also seems to alter the Ca^{2+} ion influx in the sarcoplasmic reticulum [5–7]. This modification of ion fluxes is responsible for the depolarization and asynchronous tetanic contractions of skeletal muscles and also for the vacuolisation of the sarcoplasmic reticulum. Crotamine was also seen to be a histamine releaser in mast cells. The analgesic activity of crotamine was recently investigated and found to exhibit time- and dose-dependence effects that suggests an opioid

* Corresponding author. Tel.: +33-383-912528; fax: +33-383-912530.
E-mail address: bernard.maigret@lctn.uhp-nancy.fr (B. Maigret).
[1] On leave from UMR UHP/CNRS 7565 during year 2000.

1093-3263/02/$ – see front matter
PII: S1093-3263(01)00139-5

action mechanism, 30-fold more potent than that of morphine [8,9].

To understand the relationship between the structure of crotamine and its multiple biological properties, it would be necessary to have a good knowledge of its three-dimensional structure (3D structure). Experimental studies on the structure–function relationship of crotamine with chemically produced fragments have already revealed that the myotoxic activity is achieved with either the N- or the C-terminal regions of its primary sequence [10,11]. Studies on the optical rotatory dispersion, circular dichroism, laser Raman and proton NMR of crotamine have been reported [12,13]. Among these, NMR data indicates that the protein might contain β-strands and α-helices as secondary structures, and suggest at least two structural folding states for the monomer in solution. Moreover, crotamine conformational heterogeneity may occur at different pH [14]. Crotamine, although being usually a monomer, is able to aggregate in different ways, and to form dimers with two identical cross-linked, inter-chain disulfide bonds. The self-association of crotamine is condition dependent [15,16]. But, according to the recent discovery that *Crotalus durissus terrificus* venom shows variations in the composition with respect to crotamine, indicating the presence of several polymorphic variants [3,4], experiments may give different results according to the purification procedure and observation conditions used. Therefore, all the structural NMR and CD analysis reported long ago [12–16] might be difficult to correlate with the behaviour of a pure crotamine molecule in the physiological conditions.

The lack of precise structural data of crotamine plus its possible conformational heterogeneity in solution makes it difficult to obtain a clear understanding of its behavior and, therefore, its function. Among the physico–chemical techniques usually used to solve protein 3D structures, computer modeling is now considered to be a useful complementary method, providing valuable information when difficulties in its purification and/or crystallization make it unpractical to use NMR or X-ray crystallography. In the case of crotamine, because crystallization has proved very difficult and no published NMR structure yet exists, we have undertaken its homology modeling in order to contribute to the understanding of the molecular behaviour of crotamine in solution.

2. Methodology

The protocol used to derive the crotamine 3D model can be divided into three different phases: (i) sequence alignment and analyses, (ii) model building and stability testing and (iii) model evaluation.

2.1. Sequence alignment and sequence analysis

Homology modeling implies the possibility of generating sequence alignments of the target protein to one or several templates having 3D structures already solved. We generated sequence alignments using the algorithms Fasta and Blast, both in sequence databases and in the Protein Data Bank (PDB) structure database [17], to search for a crotamine family signature and to identify possible 3D templates.

An extensive search for pattern similarities was then performed. For that purpose, we first analyzed possibilities of secondary structure of crotamine and related myotoxins, looking for a consensus secondary structure pattern. To get information about the crotamine myotoxin family, a ProtoMap analysis [18] was performed in addition to similar analyses provided by the Predict Protein [19], the NPS [20], the FOLD [21] and the NCBI [22] servers which are designed to dissect similarities within protein families.

2.2. $3D^0$ model building

The first step in our procedure was to obtain a reasonable starting point for the 3D structure of crotamine, by homology modeling. Next, we carried out several rounds of molecular dynamics (MD) calculations. Homology modeling is now widely used to produce reasonable structures, and when this technique is handled with care one can be confident about its final results, at least as a starting point for further and improved calculations.

A preliminary 3D model was first obtained using the MSI InsightII software [23] and its Homology and Modeler modules using the alignment retained in the previous step. Next, the connections between these pre-built pieces were done using the loop search procedure, that checks for similar pieces in the PDB database. The position of the disulphide bonds between the Cys residues was achieved using the aligned Cys residues and the corresponding disulphide bridges in the β-defensin-12 molecule. The complete model structure in this way obtained was refined by several energy minimization rounds (1000 steps of steepest descents followed by conjugate gradients until convergence). The side-chain positions were first optimized keeping the full backbone fixed. This constraint was then removed until the conjugate gradient algorithm converged. The MSI Discover module was used for that purpose with a distance-dependent dielectric constant and no cut-off. No charged groups (N- and C-terminal groups, Asp, Glu, Arg and Lys side-chains) were considered at this point, to avoid unexpected ionic intra-molecular interactions. The MSI CVFF force field was used. This $3D^0$ model was used as the entry point for MD simulations.

2.3. Molecular simulations

After the crude energy minimized $3D^0$ model for crotamine was obtained, it was validated by means of a protocol based on MD simulations. In this procedure, the structural stability of the model, obtained by homology modeling, is tested by means of MD simulations. We would consider a model to be acceptable if its overall average structure were conserved in time. This step is very CPU demanding as the

only valid testing of the stability of the crotamine model should be to simulate its behavior in a box of solvent. We, therefore, performed such calculations in four steps.

- The first step is a 1-ns MD simulation of the protein in vacuum in order to test its intrinsic stability. Here, just like the $3D^0$ procedure, a distance-dependent dielectric constant was used and all amino acids were considered uncharged. No cut-off distance value was used. One structure was recorded every 10 ps, so that a sample of 100 structures was obtained for further analysis. The most stable structure emerging from this procedure was energy-minimized, becoming the $3D^1$ model.
- The $3D^1$ model retained after the completion of the previous step, obtained as an intrinsically stable crotamine 3D structure, was placed in a cubic box of 1843 water molecules (the default CVFF water representation is a mixture of TIPS3 and SPC water). The side length of the cubic periodic box was 40 Å. At this point, all the partial charges of ionizable groups were introduced and the dielectric constant was set to 1. In order to maintain electroneutrality in the whole system, 8 chlorine ions were added at random in the water bath. This $3D^{1W}$ starting model was stabilized by an energy minimization step (conjugate gradients) followed by a short constant-volume controlled-temperature (NVT) MD at 300 K, and by another round of minimization. Here, only 100 ps of MD were performed, as the goal was only to force the system to reach a stable initial state, especially considering the water molecules arrangement around the ionic solute moieties. The Coulombic interactions were calculated using the Ewald summation [24] to ensure the consistency with the subsequent ddgmq calculations (see later). All the calculations here (minimization and MD) were performed using the MSI Discover3 software and the CVFF forcefield, yielding our $3D^2$ model.
- The achievement of a stable model structure yielded the entry point for a longer 1 ns MD simulation of crotamine in water for which the ddgmq parallel MD software [25–27] was used with the same CVFF forcefield used for the Discover calculations. The Coulombic interactions were again calculated using the Ewald summation. The simulation was carried out under controlled-pressure, controlled-temperature, NPT conditions using loose-coupling techniques [28], where the shape of the box was kept cubic and the average temperature and pressure were maintained at 300 K and 1 bar, respectively. At the completion of this step, our final crotamine model was obtained.
- In order to evaluate the thermodynamic stability of the resulting model according to the choice described above of the CVFF MSI force field, another 0.5 ns NVT simulation was recorded. This was done at 300 K using the software Discover and a 10-Å cut-off, under the CFF91 MSI forcefield. The choice of the simulation length will be discussed in Section 3.

All the Discover and Discover3 calculations were performed on an IBM SP2 supercomputer at CINES [29], while the ddgmq simulations were carried out on a Cray T3E supercomputer at IDRIS [30].

2.4. CD spectra

The *Crotalus durissus terrificus* venom was kindly danated by the Fundação Ezequiel Dias (Belo Horizonte, MG, Brazil). The crotamine sample was obtained in a very pure form according to the protocol described in [4]. The data were collected with a Jobin–Yvon spectropolarimeter. Samples were prepared at a concentration of 250 μM in 1.6 mM phosphate buffer at pH 7.0. No pH variations were performed to look for possible conformational modifications at other pHs. Since CD spectra between 250 and 200 nm were needed for the prediction of the secondary structure (α-helix, parallel and anti-parallel β-sheet, turn and others), far UV spectra were generated. The data were analyzed with the circular dichroism analysis software DICROPROT [31].

3. Results

3.1. Sequence alignments and analysis

Using the Fasta and Blast searches in the sequence databases, we found that outside of the myotoxin family only proteins with low sequence identity are found. The sequence homology scores drop from around 90% for myotoxins to around 40% and below for several other proteins, with no intermediates values.

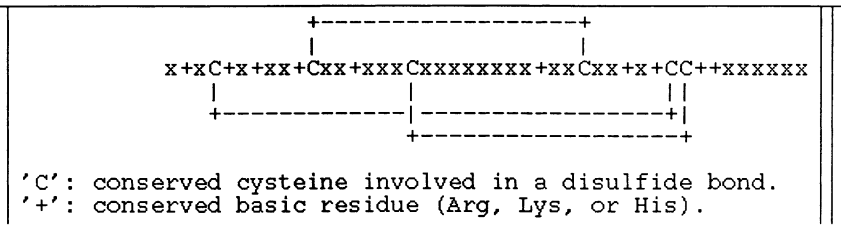
The analysis of common patterns for this myotoxin family (Clustalw, Multalin, Prosite, and ProDom analysis) provides a myotoxin signature consisting of three very conserved disulfide bonds and basic residues (Table 1). This signature was useful for choosing next the 3D template for crotamine, which should respect its signature.

The secondary structure analysis performed using different methods (PHD at the PredictProtein server, Consensus obtained from the NPS server) reveals a total absence of well-established helix and sheet patterns. The crotamine is expected to be around 90% coiled, with a very short sheet of two residues around residues 16–17. The GLOBE method, therefore, predicts crotamine to be a small, spherical globule without any secondary structure organization.

Because no 3D template was detected coming from the myotoxin family itself, the only possible 3D templates had to be found in the 30% range of sequence identity. The best proteins ranked from a Fasta search in the PDB database belong to the defensin family, with around 36% for a β-defensin-12, the structure of which was solved by NMR [32] (PDB code 1BNB [33]). Because it is better to check the possibility of using other appropriate templates, another search was performed in the FOLD database. Considering the myotoxin signature defined above, it appears that the

Table 1
The short-basic myotoxins signature

O57540/23-65	YKQCHKKGGHCFPKEKICIPPSSDFGKMDCRWRWKCCKKGSGK	O57540
O73799/22-64	YKQCHKKGGHCFPKEKICIPPSSDFGKMDCRWRWKCCKKGSGK	O73799
MYX4_CRODU/9-51	YKQCHKKGGHCFPKEKICIPPSSDFGKMDCRWRWKCCKKRSGK	P24334
MYXC_CRODU/1-42	YKQCHKKGGHCFPKEKICLPPSSDFGKMDCRWRWKCCKKGSG-	P01475
MYX1_CROVV/1-42	YKQCHKKGGHCFPKEKICIPPSSDLGKMDCRWKWKCCKKGSG-	P01476
MYX2_CROVC/1-43	YKRCHKKGGHCFPKEKICTPPSSDFGKMDCRWKWKCCKKGSVN	P12029
MYX2_CROVV/1-43	YKRCHKKEGHCFPKTVICLPPSSDFGKMDCRWKWKCCKKGSVN	P19861
MYX1_CROVC/1-43	YKRCHKKEGHCFPKTVICLPPSSDFGKMDCRWKWKCCKKGSVN	P12028
MYXC_CROVH/1-43	YKRCHKKGGHCFPKTVICLPPSSDFGKMDCRWKWKCCKKGSVN	P01477
MYX_CROAD/1-43	YKRCHKKGGHCFPKTVICLPPSSDFGKMDCRWRWKCCKKGSVN	P24330
MYX2_CRODU/22-64	YKRCHIKGGHCFPKEKICIPPSSDFGKMDCPWRRKSLKKGSGK	P24332
MYX1_CRODU/23-65	YKRCHIKGGHCFPKEKICIPPSSDFGKMDCPWRRKCCKKGSGK	P24331
MYX3_CRODU/23-65	YKRCHIKGGHCFPKGKICIPPSSDFGKMDCPWRRKCCKKGSGK	P24333



only template that respects the best this signature is the 1BNB β-defensin-12 protein, as proposed before from the PDB search. The alignment between crotamine and 1BNB sequences depicted on Fig. 1 shows that the 1BNB template, while presenting a perfect alignement with the pattern of cysteine residues as present in the myotoxin signature, allows only 3 of the 10 basic amino acids of this signature to be perfectly aligned. This occurs mainly in the C-terminal sequences. Nevertheless, three other Arg residues are found in 1BNB which could more or less corresponds to similar basic residues in crotamine.

In some cases, a low sequence identity may nevertheless corresponds to a significant 3D conservation (the Pleckstrin domain is a good example of a very conserved 3D pattern while sequence identities below 20% [34]), we checked for such possible relationship enlarging the common sequence searches. For that purpose, we submitted several queries, such as “sea anemones”, or “sodium channels” to both sequences and PDB databases, and have considered the dendrograms obtained between the identified sequences. The result of such a search allows us to identify several other proteins as possible templates, despite their weak sequence identity with crotamine (the dendrogram obtained with the keyword “sea anemones” shows for example the close connection between crotamine and anthopleurin-B, PDB code 1APF).

Because the sequence-identity searches gave the 1BNB PDB structure as a possible template, and the dendrogram-searches produced the 1APF as another possibility, among others, it was questionable if these two proteins may share common structural patterns, despite their very low sequence identity. For that purpose, we have chosen one structure of each NMR set and have looked for possible pieces of superimposed moieties. The result shows that the two proteins share a common 3D pattern made of two loops connected by two disulfide bridges. Latter we will call these loops “L1” and “L2”, “L3” being the undisplayed loop. Therefore, associating this common pattern to the myotoxin II signature, it appears that the 1BNB structure is the most appropriate template for the homology modeling as it shares both the 3D common pattern identified as above and the myotoxin signature. The alignment that we have used between crotamine and this 1BNB sequence is depicted in Fig. 1. This alignment is slightly different from the one found by Fasta in order to include the necessary gaps in loop regions but not in the signature core itself.

3.2. *Stability of the $3D^0$ starting model: the $3D^1$ model*

The 3D coordinates of the crotamine amino acids aligned with those of the 1BNB template where first obtained. Because the 1BNB structure which was solved by NMR technique presents 20 possible conformations, a preliminary study of this NMR sample was firstly undertaken in order to choose the unique 3D template. Displaying the dispersion of the 1BNB backbone conformations shows a high conservation of the core of the protein, while the N-terminal region is much more flexible. We, thus, decided to take the average (using all heavy atoms) of the 20 NMR conformation as the 3D template for modeling crotamine. Inspection of the average structure compared to the 20 individual ones reveals that the protein β-sheet core was always preserved. Using, therefore, the average β-defensin-12 structure instead of a particular individual one will not introduce a strong bias in the homology modelling procedure.

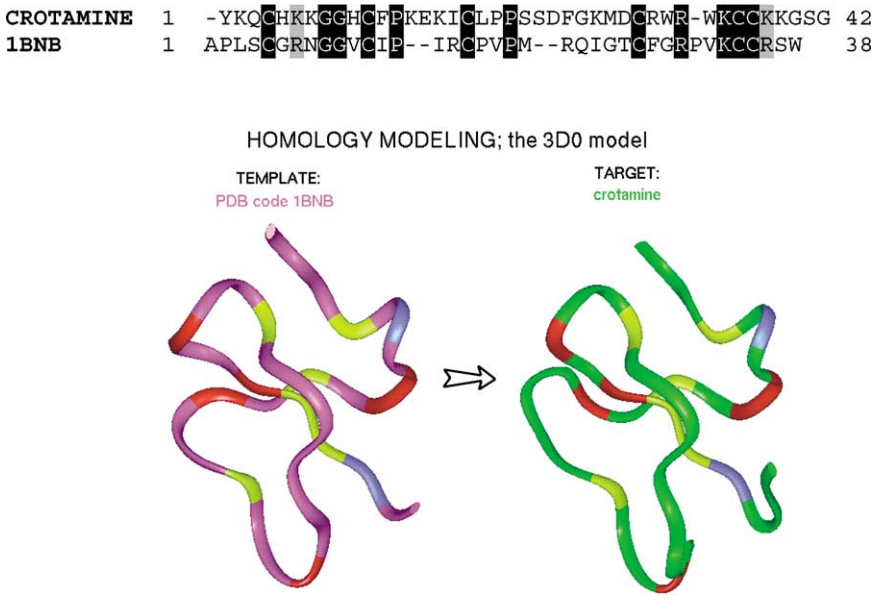


Fig. 1. The proposed alignment between crotamine and the best scored PDB structure 1BNB. The C^{α} ribbons are drawn for both molecules, and colored green for the crotamine $3D^0$ model, and purple for the 1BNB average NMR structure. The Cys common residues are colored yellow, the other common residues (Gly, Pro, Arg, Lys) in both sequences are colored red, and the similar Arg/Lys ones in blue.

The energy-minimized model obtained directly after the homology step was first checked for its quality with the Procheck algorithm [35]. The RMSD (calculated on the backbone C^{α} atoms) between this minimized structure and the original one coming immediately after the homology modeling was 1.8 Å. This structure was considered as the $3D^0$ starting point for the 1-ns MD simulation of the system in vacuum.

The analysis of the RMSD map obtained between each recorded structure during the 1 ns MD simulation in vacuum shows that, after a small rearrangement from the initial conformation (RMSD of 3.2 Å between the starting conformation and the first stabilized one after 100 ps MD), it can be seen that the structure is very stable during the whole MD, and that the RMSD among all the C^{α} atoms of the protein does not deviate by more than 1 Å. A short transition occurs in the middle of the simulation (between 450–490 ps) but the structure quickly returns to its previous state until the end of the simulation (the RMSD between the conformation obtained at 200 and 900 ps is 1.2 Å).

This weakly fluctuating structure was thus submitted to an energy minimization enhancement, taking the last conformation recorded during the MD as the entry point of the minimization procedure. We have considered this last energy-minimized conformation of the protein as the $3D^1$ model. This model was also submitted to the Procheck

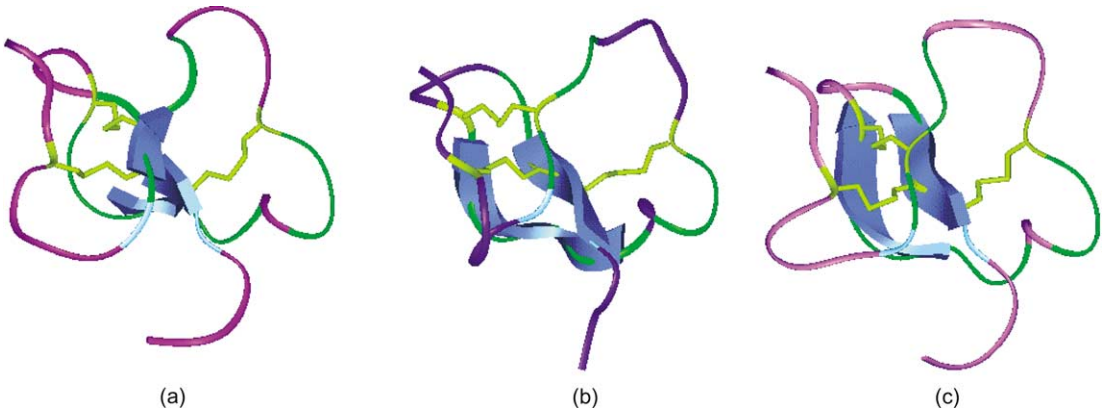


Fig. 2. Comparison between (a) the starting $3D^2$ conformation for the MD in solution, (b) the final ones after the ddgmq 1 ns MD and (c) the Discover 0.5 ns MD showing the secondary structures observed (colored in blue). The Cys bridges are depicted and colored yellow. The green parts are the ones structurally conserved within this protein family (see Fig. 5). One can see that only two β-strands are obtained after the dynamics in solution (b and c) conversely to the starting conformation (a).

program in order to verify all its geometrical data. Many hydrogen bonds are observed in this conformation, especially H-bonds that are stabilized between the C-terminal carbonyl of Gly42 and both the charged side-chains of Lys27 and Lys38, respectively. Other H-bond interactions are observed between side-chain moieties: between the Asp29 side-chain and those of Lys7 and Arg33 respectively; between the Glu15 side-chain and those of Arg31 and Lys16 respectively; between the Asp24 side-chain and that of Lys27.

3.3. Stability of the $3D^1$ solvated model: the $3D^2$ model

The $3D^1$ model obtained as described above was placed at the center of a cubic box of solvent. After stabilizing this model (minimization + 100 ps MD + minimization) the overall crotamine conformation does not change significantly at least at the backbone level (the RMSD fluctuates a little between the generated conformers during this step). The only changes occur for the ionic groups finding their best position according to the surrounding water molecules. No strong salt bridges were observed between Lys/Arg residues and the Glu/Asp ones. Procheck results were satisfactory for this $3D^2$ model. Considering the evolution of the RMSD between the recorded structures during this step it can be seen that the model is already quite stable at both the backbone and side-chain levels.

3.4. Stability of the $3D^2$ solvated model: the final model

Starting with the $3D^2$ model obtained at the above step, considered as a stable starting point for MD in water, the evolution of the conformational behavior of crotamine during the 1-ns ddgmq simulation was checked by recording one conformational state each 5 ps and then using this sample of 200 conformers for more detailed analyses. It can be seen for the RMSD evolution that the structure reached a stable plateau after 200 ps of simulation, and next fluctuated around a stable core. The radius of gyration, as well as the moment of inertia of the conformers remain stable, indicating an almost spherical stable shape. The global dipole moment of the protein increased during the MD, due to rearrangement of some polar charged groups, and decreased after 800 ps possibly returning to its value after 200 ps of simulation.

Looking to the evolution of the distances between the charged groups, it appears that a few intra-molecular salt bridges formed during the MD simulation in water, most of these charged groups preferring to be solvent-exposed and to maintain a stable shell of hydration around them. Such salt bridges, nevertheless, are not completely absent, and two were formed during the first 100 ps involving Asp29: with Lys7 on one hand and with Arg33 on the other. Others are rare events, such as between Glu15 and Lys14. The variations in the distances between Asp24 and the positive side-chains in its vicinity (Lys27, Lys35, Lys38 and Lys39) are mainly due to the larger fluctuations of the C-terminal tail and between this tail and the $L1_b$ loop. The charged triad Lys14–Glu15–Lys16 is always protruding at the top of the $L1_a$ loop, but without any direct interactions between them. The $L1_a$ loop, which encompasses these residues, is quite mobile, yielding motions of large amplitude for this triad. The molecule, therefore, appears as a charged sphere, surrounded mostly by a shell of positively charged residues.

3.5. Stability of the $3D^2$ solvated model: the influence of the simulating conditions

In order to validate the previous results, we have submitted the $3D^2$ model to another round of MD simulation in water. These steps were considered necessary:

- to check the usefulness of the Ewald treatment for long-range electrostatics as many papers discussed the validity of the cut-off approach versus more sophisticated, but more costly models (see for example [36]);
- to check the consistency of the CVFF results for proteins MD simulations. As these calculations were done, we have recently encountered some problems with the CVFF force field working on another peptide (work to be published) and this situation forced us to start another MD simulation, with the same starting point as used in the ddgmq MD, but using now the CFF91 forcefield;
- the only way to check the convergence of our MD sampling should be to perform another simulation, starting with different conditions and to compare the statistical behavior of the system until it will be similar to the one obtained after the first simulation, if possible.

Considering the above reasons, we have chosen to use the Discover program for performing this comparison (see methods), and to stop the calculation when a common behavior was found between the two trajectories. This situation was achieved after 0.5 ns of simulation, when the averaged radius of gyration, moment of inertia and global dipole moment gave values similar to those obtained previously with ddgmq. We present in Fig. 2, the final conformations obtained during this step, compared to the initial $3D^2$ one and to the final one obtained after the 1-ns ddgmq simulation. It can be seen that the flexibility of the N- and C-chains is mainly responsible for the RMSD values obtained between the C^α traces of the three conformers. Meanwhile, the general organization of the core around the L1 and L2 loops is quite stable, a result in agreement with the analysis of the 1-ns ddgmq MD. The last Discover conformation, just like the last ddgmq one, presents a well-established central β-strand connecting the L2 loop. This strand was observed in most of the conformers after 200 ps of ddgmq or Discover MDs, giving us some belief of its reality in the real system. The position of the strand in the ddgmq model is very similar to the corresponding strand in β-defensin-12.

To summerize, it appears that the most probable crotamine 3D structure is built from a stable core organized around the

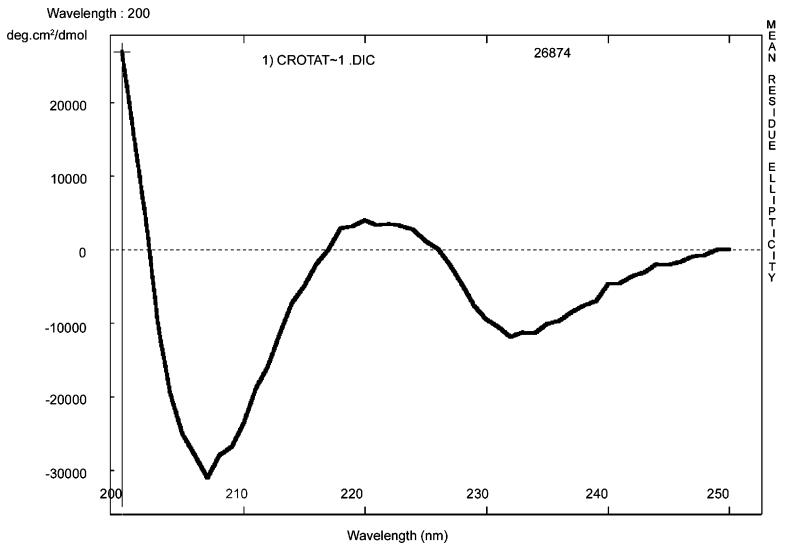


Fig. 3. The recorded CD spectrum for crotamine.

L1 and L2 loops and maintained by a β-strand stretching between residues 34 and 37 on one side and between residues 26 and 30 on the other side. A weak helical tendency is visible in the $L1_b$ loop just before the connection to the strand.

3.6. Circular dichroism spectra

The analysis of the crotamine spectrum (Fig. 3) by DICROPROT, with several tools, showed that the toxin has a very poor percentage of secondary structure. From the ellipticity, the α-helix is estimated to be 0.186%. From the Bolotina method, as well as the Chang, Fasman and Yang algorithms, the percentages of helix were calculated to be around 2% suggesting a small tendency to form helix in solution. The Varselec analysis of the spectrum also showed a low percentage of secondary structure, around 2% of a helix, but no strand and a high probability of coils.

This result confirms the theoretical model, except for the β-strand, which is observed during the last MD simulations. This may be due to the size of this strand which is quite small and could not be recognized by the CD analysis algorithms, which are trained on larger protein structures in which β-strands are usually longer.

4. Discussion

From the results presented above, a quite clear picture of the crotamine conformational behavior emerges, making it possible to use this model for further work such as site-directed mutagenesis. The protein is a small, stable globule presenting a large positively charged envelope, as seen from the electrostatic potential maps obtained with the Delphi Insight module and depicted in Fig. 4. One side of this globule presents a strong positive electrostatic potential and could be considered as the membrane-anchoring part (N-terminus, Lys2, Lys6, Lys7, Arg31 and Arg33, while the other side presents a large depression in which the negative electrostatic potential surface appears due to the Glu15, Asp24 and C-terminus carboxylates. Several protruding positive areas corresponding to residues Lys14, Lys16, Lys27 and Lys38 surround this depression. Lys35 has a particular position, participating in the positive face on the side of the depression.

In the proposed final model, only two β-strands are found compared to the initial conformation presenting the three strands coming from the 1BNB template. This is mainly the result of a different balance between intra- versus inter-atomic interactions in crotamine compared to the starting β-defensin-12 structure due to sequence differences inside this region: the stability of the β-sheets seems very sensitive to solvation and, therefore, to the balance between intra- and inter-molecular interactions. The third strand of the β-sheet disappeared during the MD because of preferred interactions, at least for some time, between the involved amino acid side-chains and the solvent molecules, compared to the backbone–backbone interactions stabilising the sheet. This highlights clearly why doing long MD in explicit solvent seems inevitable to reach stable and plausible models.

Another important point reveals from the present study if that comparison between experimental and computation results should be done warily. In the present case of crotamine, the available NMR data, in favor of a conformational heterogeneity, were apparently in contradiction with the MD results showing a stable conformational behaviour. In fact the conformational heterogeneity, as observed from the NMR studies [13], seems to be due to the experimental conditions used by these authors. It has been now demonstrated

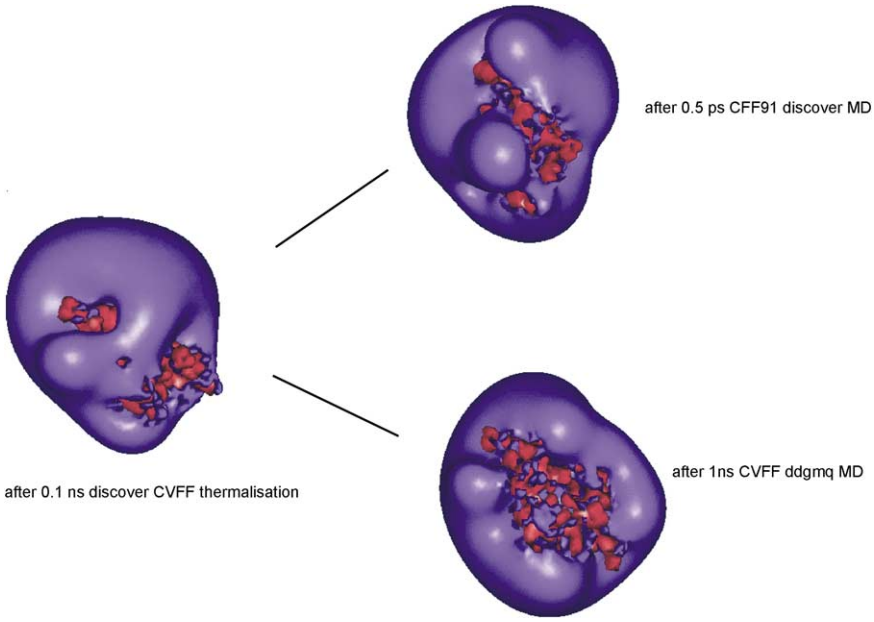


Fig. 4. Comparison between the molecular electrostatic potentials obtained for (a) the starting $3D^2$ conformation for the MD in solution and the final ones after (b) the ddgmq 1 ns MD and (c) the Discover 0.5 ns MD. The positive lobes are colored in blue and the negative ones in red, showing the high "positive ion"-like electrostatic behavior of the protein.

```
CROTAMINE  1  -----YKQCHK----------KGGHCFPK-EKICLPPSSDFGKMDCRWRWK-CCKKGSG  42
1APF       1  -----GVPCLCDSDGPRPRGNTLSGILWFYPS-GCPSG--WHNCKAHGPNIGWCCKK    49
1BNB       1  ----APLSCGR----------NGGVCIPI---RCPVP--MRQIGTCFGRPVKCCRSW   38
1B8W       1  FVQHRPRDCES----------INGVCRHKDTVNCREI----FLADCYNDGQKCCRK    42
1D6B       1  IMFFEMQACWS----------HSGVCRDKSERNCKPM----AWTYCENRNQKCCEY    42
```

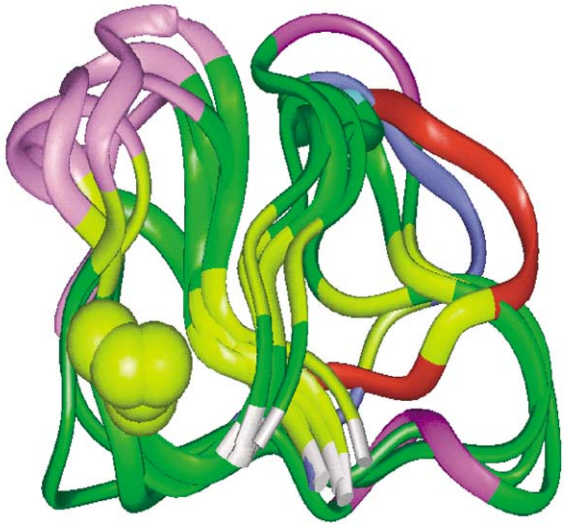


Fig. 5. Comparison between the obtained crotamine model and the NMR structures of anthopleurine-B (1APF PDB code), β-defensin-12 (1BNB PDB code), DLP-1 (1B8W PDB code) and DLP-2 (1D6B PDB code) proteins: sequence alignment and superposition of the C^{α} trace ribbons. Only the superposed parts of each structure between the common Gly residue (in white) and the common Cys–Cys–X N-terminal sequences (X colored white) are depicted in the ribbons. The C^{α} atoms of the common Cys residues found on the N-terminal side are drawn as CPK spheres. The Cys residues are colored in yellow, the common identified folds are colored in green, the common loops in light-purple. The loops appearing different in 1B8W and 1D-B are colored in blue and red respectively. The "extended" folds found in crotamine is colored dark-purple.

[3,4] that *Crotalus durissus terrificus* venom shows variations in the composition with respect to crotamine, indicating the presence of several polymorphic variants. These multiple variants showed 5-6 amino acid differences with the original sequence of Laure [10], a rather high substitution rate for a peptide of 42 amino acids. The protocol used by Endo et al. [13] in 1989 to isolate crotamine from the venom extract was not sufficient to obtain pure crotamine isoforms. This, added to the pH at which a conformational transition is proposed by these authors (acidic pH), may explain that their results do not correspond to the behavior of a pure crotamine molecule at the physiological conditions the simulation is supposed to mimic. These structural data were, therefore, questionable ones. On the contrary, our CD experiment, performed on a pure material, agrees perfectly with the simulations.

Recently, the 3D structures of two small proteins isolated from the venom of platypus (*Ornithorhynchus anatinus*) were published, being considered to be defensin-like peptides [37,38]. We used these NMR structures, available in the PDB (PDB codes 1B8W and 1D6B) for superposition to our model. The results depicted in Fig. 5 clearly show the conservation of well identified 3D structural patterns between crotamine, anthopleurine-B, β-defensin-12, and the defensin-like peptides DLP-1 and DLP-2: the central core connecting the L1 and L2 loops is well conserved, despite a weak sequence identity. The dispersion of the L2 loop observed in Fig. 5 corresponds roughly to the fluctuation of this loop in the theoretical crotamine model. The $L1_b$ loop is much shorter in the DLP peptides compared to the others, but the cysteine residues remain almost at the same place. An interesting feature is the conservation of the 3D position of the first cysteine residue in all these structures, therefore giving a stable anchoring point for the fluctuating and less conserved N-terminal moieties. The C-terminal lysine or arginine residues remain at the same position in space, and considering the role given to such residues in anthopleurine-B (Lys48, Lys49 in APF) from site-directed mutagenesis [39], it can be postulated that Lys38 and Lys39 in crotamine could also play a role in crotamine expression and biological activity. Lys37 in APF was postulated to play a role on channel interaction. By similarity, we could propose the same role for Lys27 in crotamine. Trp33 of APF was proposed to be involved in a hydrophobic contact with the receptor [40]. No such Trp residue was found in either crotamine, β-defensin-12 or DLP peptides presenting, according to our "structural" alignment, a Lys (Lys35 in crotamine) residue which therefore could be involved in an electrostatic interaction with a negatively charged group on the receptor. In our crotamine model, the Trp34 side-chain has a position very similar to those found for Trp45 in APF suggesting a similar hydrophobic interaction with the receptor. Site-directed mutagenesis of these residues in crotamine should give us additional insights on the credibility of the model here proposed.

5. Conclusions

We have obtained a quite robust and convincing 3D-structural model for crotamine. This model is the first example for the 3D structure of this family of small basic myotoxins (the model will be deposited in the PDB). This model will be the template for further studies both at theoretical and experimental levels, such as the study of crotamine dimerisation, site-directed mutagenesis, etc. Our calculations clearly demonstrate that we are faced with a new structural family of small proteins sharing a common structural fold, distinct from what has been observed up to now. This observation is in perfect agreement with the conclusions of Torres et al. about DLP and β-defensin-12 [40]. We have extended this family to crotamine and have shown that the anthopleurin-B proteins also share this distinct fold. The specificity of all the proteins belonging to this family should be found in the nature of the sidechains, which is likely to play a crucial role in defining their biological functions. We propose that key residues should be found in what we call the L1, L2 and L3 loops, the role of which could serve as a particular behaviour, such as membrane anchoring, receptor anchoring, receptor isoform selection and receptor inactivation. Further work is in progress to better identify the concerned amino acids.

References

[1] A.C. Matavel, D.L. Ferreira-Alves, P.S. Beirão, J.S. Cruz, Tension generation and increase in voltage-activated Na^+ current by crotamine, Eur. J. Pharmacol. 348 (1998) 167–173.
[2] M.L. Santoro, M.C. Sousa-e-Silva, L.R. Gonçalves, S.M. Almeida-Santos, D.F. Cardoso, I.L. Laporta-Ferreira, M. Saiki, C.A. Peres, I.S. Sano-Martins, Comparison of the biological activities in venoms from three subspecies of the South American rattlesnake, Comp. Biochem. Physiol. C. Pharmacol. Toxicol. Endocrinol. 122 (1999) 61–73.
[3] M.H. Toyama, E.M. Carneiro, S. Marangoni, R.L. Barbosa, G. Corso, A.C. Boschero, Biochemical characterization of two crotamine isoforms isolated by a single step RP-HPLC from *Crotalus durissus terrificus* (South American rattlesnake) venom and their action on insulin secretion by pancreatic islets, Biochim. Biophys. Acta 1474 (2000) 56–60.
[4] G. Radis-Baptista, N. Oguiura, M.A. Hayashi, M.E. Camargo, K.F. Grego, E.B. Oliveira, T. Yamane, Nucleotide sequence of crotamine isoform precursors from a single South American rattlesnake (*Crotalus durissus terrificus*), Toxiconology 37 (1999) 973–984.
[5] J.E. Fletcher, M. Hubert, S.J. Wieland, Q.H. Gong, M.S. Jiang, Similarities and differences in mechanisms of cardiotoxins, melittin and other myotoxins, Toxiconology 34 (1996) 1301–1311.
[6] S.J. Hong, C.C. Chang, Novel inhibition of contractility by wortmannin in skeletal muscle, Br. J. Pharmacol. 124 (1998) 849–856.
[7] C.L. Owby, Structure, function and biophysical aspects of the myotoxins from snake venoms, J. Toxicology 17 (1998) 213–238.
[8] A.C. Mancin, A.M. Soares, C.A. Giglio, S.H. Andriao-Escarso, C.A. Vieira, J.R. Giglio, The histamine releasers crotamine, protamine and compound 48/80 activate specific proteases and phospholipases A2, Biochem. Mol. Biol. Int. 42 (1997) 1171–1177.
[9] A.C. Mancin, A.M. Soares, S.H. Andriao-Escarso, V.M. Faca, L.J. Greene, S. Zuccolotto, I.R. Pela, J.R. Giglio, The analgesic activity of crotamine, a neurotoxin from *Crotalus durissus terrificus* (South

American rattlesnake) venom: a biochemical and pharmacological study, Toxiconology 36 (1998) 1927–1937.
[10] C.J. Laure, The primary structure of crotamine, Hoppe Seylers Z. Physiol. Chem. 356 (1975) 213–215.
[11] B. Baker, A.T. Tu, J.L. Middlebrook, Binding of myotoxin a to cultured muscle cells, Toxiconology 31 (1993) 271–284.
[12] D.L. Cameron, A.T. Tu, Chemical and functional homology of myotoxin a from prairie rattlesnake venom and crotamine from South American rattlesnake venom, Biochim. Biophys. Acta 532 (1978) 147–154.
[13] T. Endo, M. Oya, H. Ozawa, Y. Kawano, J.R. Giglio, T. Miyazawa, A proton nuclear magnetic resonance study on the solution structure of crotamine, J. Protein Chem. 8 (1989) 807–815.
[14] O.G. Hampe, M.M. Vozari-Hampe, J.M. Goncalves, Crotamine conformation: effect of pH and temperature, Toxiconology 16 (1978) 453–460.
[15] J.R. Beltran, Y.P. Mascarenhas, A.F. Craievich, C.J. Laure, Saxs study of structure and conformational changes of crotamine, Biophys. J. 47 (1985) 33–35.
[16] O.G. Hampe, Model studies of crotamine self-association, Braz. J. Med. Biol. Res. 22 (1989) 17–24.
[17] Research Collaboratory for Structural Bioinformatics, Protein Data Bank, PDB: http://www.rcsb.org/.
[18] ProtoMap, An Automatic Hierarchical Classification of all SwissProt Proteins, Hebrew University, Israel: http://www.protomap.cs.huji.ac.il/.
[19] The PredictProtein server, European Molecular Biology Laboratory, Heidelberg, DE: http://www.embl-heidelberg.de/predictprotein/.
[20] Network Protein Sequence Analysis, NPS, Lyon, France: http://pbil.ibcp.fr/.
[21] UCLA-DOE Fold-Recognition server, FOLD, Los Angeles, CA, USA: http://fold.doe-mbi.ucla.edu/Home/.
[22] National Center for Biotechnology Information, NCBI, National Library of Medicine, Bethesda, MD, USA: http://www.ncbi.nlm.nih.gov/.
[23] Molecular Simulation, Inc., San Diego, CA, USA: http://www.msi.com/.
[24] A. Smith, A replicated data molecular dynamics strategy for the parallel Ewald sum, Comput. Phys. Com. 67 (1992) 392–406.
[25] D. Brown, M. Minoux, Maigret B, A domain decomposition parallel processing algorithm for molecular dynamics simulations of systems of arbitrary connectivity, Comp. Phys. Comm. 103 (1997) 170–186.
[26] D. Brown, B. Maigret, Large scale molecular dynamics simulations using the domain decomposition approach, in: Proceedings of the 24th SpeedUp worshop, 24–25 September 1998, Berne, Switzeland.
[27] ddgmq: http://www.univ-savoie.fr/labos/lmpc/gmq.html.
[28] D. Brown, J.H.R. Clarke, A loose coupling constant pressure molecular dynamics algorithm for use in the modelling of polymer materials, Comp. Phys. Comm. 62 (1991) 360–369.
[29] CINES, Centre Informatique National de l'Enseignement Supérieur, Montpellier, France: http://www.cnusc.fr/.
[30] IDRIS, Institut du Developpement et des Ressources en Informatique Scientifique, France: http://www.idris.fr/.
[31] DICROPROT: http://pbil.ibpc.fr/DICROPROT/.
[32] G.R. Zimmermann, P. Legault, M.E. Selsted, A. Pardi, Solution structure of bovine neutrophil beta-defensin-12: the peptide fold of the beta-defensins is identical to that of the classical defensins, Biochemistry 34 (1995) 13663–13671.
[33] H.M. Berman, J. Westbrook, Z. Feng, G. Gililand, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein Data Bank, Nucl. Acid Res. 28 (2000) 235–242.
[34] Pleckstrin Homology (PH) domain: http://www.embl-heidelberg.de/~wilmanns/structures/ph.html.
[35] Procheck: http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html.
[36] T. Fox, P.A. Kollman, The application of different solvation and electrostatics models in molecular dynamics simulations of ubiquitin: how well is the X-ray structure "maintained", Proteins 25 (1996) 315–334.
[37] A.M. Torres, X. Wang, J.I. Fletcher, D. Alewood, P.F. Alewood, R. Smith, R.J. Simpson, G.M. Nicholson, S.K. Sutherland, C.H. Gallagher, G.F. King, P.W. Kuchel, Solution structure of a defensin-like peptide from platypus venom, Biochem. J. 341 (1999) 785–794.
[38] A.M. Torres, G.M. de Plater, M. Doverskog, L.C. Birinyi-Strachan, G.M. Nicholson, C.H. Gallagher, P.W. Kuchel, Defensin-like peptide-2 from platypus venom: member of a class of peptides with a distinct structural fold, Biochem. J. 348 (3) (2000) 649–656.
[39] P.K. Kera, G.R. Benzinger, G. Lipkind, C.L. Drum, D.A. Hanck, K.M. Blumenthal, Multiple cationic residues of anthopleurin-B that determine high affinity and channel isoform discrimination, Biochemistry 34 (1995) 8533–8541.
[40] B.L. Dias-Kadambi, K.A. Combs, C.L. Drum, D.A. Hanck, K.M. Blumenthal, The role of exposed tryptophan residues in the activity of the cardiotonic polypeptide anthopleurin-B, J. Biol. Chem. 271 (1996) 23828–23835.