

Tertiary Structure of Conotoxin GIIIA in Aqueous Solution[†]Jean-Marc Lancelin,^{‡§||} Daisuke Kohda,[‡] Shin-ichi Tate,[‡] Yuchio Yanagawa,[‡] Teruo Abe,[‡] Mei Satake,[‡] and Fuyuhiko Inagaki^{*‡}*Department of Molecular Physiology, The Tokyo Metropolitan Institute of Medical Science, 3-18-22 Honkomagome, Bunkyo-ku, Tokyo 113, Japan, The Institute of Physical and Chemical Research, Wako, Saitama 351-01, Japan, and Brain Research Institute, Niigata University, Niigata 951, Japan**Received January 22, 1991; Revised Manuscript Received April 19, 1991*

ABSTRACT: The three-dimensional structure of conotoxin GIIIA, an important constituent of the venom from the marine hunting snail *Conus geographus* L., was determined in aqueous solution by two-dimensional proton nuclear magnetic resonance and simulated annealing based methods. On the basis of 162 assigned nuclear Overhauser effect (NOE) connectivities obtained at the medium field strength frequency of 400 MHz, 74 final distance constraints of sequential and tertiary ones were derived and used together with 18 torsion angle (ϕ , χ_1) constraints and 9 distance constraints derived from disulfide bridges. A total of 32 converged structures were obtained from 200 runs of calculations. The atomic root-mean-square (RMS) difference about the mean coordinate positions (excluding the terminal residues 1 and 22) is 0.8 Å for backbone atoms (N, C $^\alpha$, C). Conotoxin GIIIA is characterized by a particular folding of the 22 amino acid peptidic chain, which is stabilized by three disulfide bridges arranged in cage at the center of a discoidal structure of approximately 20-Å diameter. The seven cationic side chains of lysine and arginine residues project radially into the solvent and form potential sites of interaction with the skeletal muscle sodium channel for which the toxin is a strong inhibitor. The present results provide a molecular basis to elucidate the remarkable physiological properties of this neurotoxin.

Conotoxin GIIIA belongs to the family of neurotoxic peptides that was isolated from the venoms of the fish-hunter snails of the marine *Conus* genus. Conotoxins exhibit a number of physiological properties with striking selectivities from a receptor subtype to another [for reviews, see Olivera et al. (1985, 1990)]. Among this family of neurotoxins, only conotoxin GI, an acetylcholine receptor inhibitor, was structurally reported in three dimensions (Kobayashi et al., 1989; Pardi et al., 1989). Conotoxin GIIIA was isolated (Nakamura et al., 1983; Cruz et al., 1985) from the venom of *Conus geographus* L. and demonstrated to bind to the voltage-gated sodium channel of skeletal muscle cell but not to those of heart, brain, or axonal origin (Cruz et al., 1985; Moczydlowski et al., 1986). The three-dimensional structural study of such a molecule is consequently useful to probe the unknown molecular arrangement of the muscle sodium channel, which plays a central role in the electrical excitability of the cells.

The toxin was sequenced as a 22 amino acid peptide (Sato et al., 1983; Cruz et al., 1985) containing three disulfide bridges (Figure 1). The locations of these disulfide bridges were very recently established by using a pertinent combination of isotopic labeling and mass spectrometry (Hidaka et al., 1990). Due to the presence of these three disulfide bridges, conotoxin GIIIA is anticipated to have a rather rigid molecular structure in solution in spite of its small number of residues. In addition, with the lack of suitable crystalline derivatives, no X-ray analyses have been reported. These make conotoxin

GIIIA an attractive object for structure determination by NMR. In the present paper, we now report the three-dimensional structure of conotoxin GIIIA in aqueous solution elucidated by means of ¹H NMR.

MATERIALS AND METHODS

NMR Measurements. Conotoxin GIIIA was purified from the venom of *C. geographus*, L. as previously described (Yanagawa et al., 1988). The sample was dissolved at 13 mM concentration in ²H₂O or H₂O (10% ²H₂O). The pH was adjusted to an uncorrected glass electrode reading (pH*) of 2.5 by using a Radiometer PHM 84 and an Ingold thin-glass electrode. All the NMR spectra were recorded on a JEOL JNM-GX spectrometer operating at 400 MHz for the proton frequency. Chemical shifts were quoted relative to 2,2-dimethyl-2-silapentane-5-sulfonate used as an internal reference. DQF-COSY¹ (double-quantum filtered correlation spectroscopy) (Rance et al., 1983), PE-COSY (primitive exclusive COSY) (Müller, 1987), HOHAHA (homonuclear Hartmann-Hahn spectroscopy) (Bax & Davies, 1985), and NOESY (Jeener et al., 1979; Macura et al., 1981) spectra were recorded in the phase-sensitive mode (States et al., 1982). The water resonance was suppressed by means of the DANTE pulse (Zuiderweg et al., 1986) during the relaxation period. Two-dimensional spectra were recorded at 10 and 28 °C with 512 (*t*₁) × 2048 (*t*₂) data points (512 × 4096 for PE-COSY) with a spectral width of 4000 Hz. HOHAHA spectra were recorded in H₂O and ²H₂O with a mixing time of 48 ms. NOESY spectra were recorded with mixing times of 250 ms in H₂O and 65, 125, 250, 500 ms in ²H₂O. After zero filling once in the *t*₁ dimension, phase-shifted sine-bell functions were

[†] This work was supported in part by Grants-in Aid from the Ministry of Education, Science and Culture of Japan and Special Coordination Funds for Promoting Science and Technology from the Science and Technology Agency of Japan.

* Author to whom correspondence should be addressed.

‡ The Tokyo Metropolitan Institute of Medical Science.

§ Present address: Laboratoire de Biochimie Structurale associé au CNRS, Université d'Orléans, F45067 Orléans, France.

|| The Institute of Physical and Chemical Research.

‡ Brain Research Institute, Niigata University.

¹ Abbreviations: DQF-COSY, double-quantum filtered correlated spectroscopy; PE-COSY, primitive exclusive COSY; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; NOE, nuclear Overhauser effect; DANTE, delays alternating with nutation for tailored excitation; RMS, root mean square; RMSD, RMS difference; Hyp, hydroxyproline.

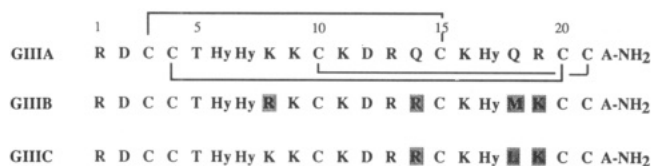


FIGURE 1: Sequence of conotoxin GIIIA compared with the minor conotoxins GIIIB and GIIIC isolated from the venom of *C. geographus* L. where Hy means *trans*-4-hydroxyproline (Cruz et al., 1985). Other minor conotoxins were also identified with replacement of Hy6 and Hy7 by a proline residue. Conotoxin GIIIA and GIIIB are identical with geographotoxins I and II (Hidaka et al., 1990; Sato et al., 1983). The amino acid replacements indicated by shaded boxes are conservative except for Gln-14.

applied prior to Fourier transformation.

Computing Procedure. All calculations were achieved on a Silicon Graphics Iris 4D/70G workstation. An initial peptide structure was built in an extended conformation with QUANTA/CHARMM (version 3.0, Polygen Corp., Waltham, MA). Structures were calculated in two steps using successively two XPLOR programs (Brünger et al., 1987). The first one is based on the variable target function protocol with a soft asymptotic NOE potential (Nilges et al., 1988a). The random sampling of structures in the conformation space was achieved by assigning random initial velocities according to a Maxwell distribution at 1000 K. At this stage, only distance constraints were introduced. Methylene protons without stereospecific assignments were treated as pseudoatoms, and the correction factors were added to the upper and lower distance constraints (Wüthrich et al., 1983). The roughly annealed structures obtained were then used as starting points for the second stage based on the dynamical simulated annealing protocol (Nilges et al., 1988b). The velocities were randomly reassigned at 1000 K to introduce another randomness into the calculation. Dihedral angle constraints from *J* couplings were introduced at this level. The interproton distances involving methylene protons without stereospecific assignment were referred to single $((r^{-6})^{-1/6})$ average distances so that no corrections for pseudoatom treatments were made at this stage. After achievement of the simulated annealing process at 1000 K, the disulfide bridges were covalently connected before the final cooling stage. As extensively described previously (Nilges et al., 1988b), the nonbonded energetic component is described by a simple van der Waals repulsion term, F_{repe} :

$$F_{\text{repe}} = 0 \quad \text{if } r \geq sr_{\text{min}} \\ = k_{\text{vdw}}(s^2r_{\text{min}}^2 - r^2)^2 \quad \text{if } r < sr_{\text{min}} \quad (1)$$

where k_{vdw} is the force constant; r_{min} , the standard values for the sum of the van der Waals radii of two neighboring atoms; and s , the van der Waals radius scale factor that was set to 0.825 in the variable target function protocol step and 0.80 in the dynamical simulated annealing step. The interproton distance and torsional angle constraints are described by the classical square-well potentials, F_{NOE} and F_{tor} , respectively, (Clare et al., 1986a), given below for clarity:

$$F_{\text{NOE}} = k_{\text{NOE}}(r_{ij} - r_{ij}^u)^2 \quad \text{if } r_{ij} > r_{ij}^u \\ = 0 \quad \text{if } r_{ij}^l \leq r_{ij} \leq r_{ij}^u \\ = k_{\text{NOE}}(r_{ij} - r_{ij}^l)^2 \quad \text{if } r_{ij} < r_{ij}^l \quad (2a)$$

and

$$F_{\text{tor}} = k_{\text{tor}}(\phi_i - \phi_i^u)^2 \quad \text{if } \phi_i > \phi_i^u \\ = 0 \quad \text{if } \phi_i^l \leq \phi_i \leq \phi_i^u \\ = k_{\text{tor}}(\phi_i - \phi_i^l)^2 \quad \text{if } \phi_i < \phi_i^l \quad (2b)$$

where k_{NOE} and k_{tor} are the force constants, r_{ij} the distance

between atoms i and j , and ϕ_i the dihedral angle considered. Neither electrostatic, 6–12 Lennard-Jones terms, nor hydrogen bonds were taken into account. The C-terminal amide group was excluded from the calculations. The target functions used for maintaining the correct bond lengths, angles, and improper torsions and peptide-bond dihedral angles were not modified relative to the original method (Nilges et al., 1988a).

Analysis of the Calculated Structures. For quantitative assessment of the convergence of the calculations, minimum root-mean-square (RMS) NOE violations, number of NOE violations, and the values of F_{repe} , F_{NOE} , and F_{tor} were examined. A NOE violation is defined as a deviation of the calculated interproton distance from the experimental distance constraint. For quantitative comparisons of the different structures, minimum RMS differences (RMSDs) were calculated for the specified atoms for the specified range of residues. A mean structure was obtained by averaging the coordinates of the structures that were first superimposed to the best converged structure for a minimum pairwise RMSD of the backbone atoms (N, C $^{\alpha}$, C) for residues 2–21. Because such a structure was distorted in valence geometry, it was subjected to a restrained Powell minimization that comprised 200 cycles with the soft van der Waals radii reduced by a factor of $s = 0.25$, 200 cycles with $s = 0.5$, and 800 cycles with $s = 0.8$ (Clare et al., 1986b). The final coordinates were obtained by fitting to the mean structure for minimum RMSDs for the backbone atoms of residues 2–21. Structural statistics are presented essentially in a comparable way with recent reports where similar XPLOR protocols were used (Kraulis et al., 1989; Omichinski et al., 1990; Clare et al., 1990).

RESULTS

Sequential Assignments and Secondary Structure. Sequence-specific resonance assignments of conotoxin GIIIA were achieved according to the conventional method developed by Wüthrich (1986). First, the NMR resonances were assigned to the spin systems of the specific amino acid types by using DQF-COSY and HOHAHA. Figure 2 displays a representative part of the HOHAHA spectrum taken in H₂O. All expected cross peaks between amide and C $^{\alpha}$ H protons were observed. For amide protons with small $^3J_{\text{HN}\alpha}$ coupling constants, the magnetization transfer was inefficient and no magnetization transfer was observed beyond the C $^{\alpha}$ H protons. This was the case for Arg-13 ($^3J_{\text{HN}\alpha} = 1$ Hz), Lys-16 ($^3J_{\text{HN}\alpha} = 1.5$ Hz), and Arg-19 ($^3J_{\text{HN}\alpha} = 1$ Hz). The chemical shift correlations between amide and C $^{\alpha}$ H protons were further extended to the side-chain protons with the HOHAHA experiment in ²H₂O (data not shown). Due to the acidic pH condition, the exchange rate of N⁵H proton of Lys and N⁴H proton of Arg with the solvent were slow. These protons gave additional cross peaks in the HOHAHA spectrum in H₂O. The magnetization transfer starting from N⁵H of Lys and N⁴H of Arg confirms the assignments made in ²H₂O for the side-chain protons. The N-terminal Arg-1 was identified from the remaining spin system after complete analysis. The three spin systems of hydroxyprolines (Hyp) were easily recognized in the HOHAHA spectrum in ²H₂O by their typical pattern formed by the C $^{\beta}$ H, C $^{\beta}$ H', C $^{\gamma}$ H, C $^{\delta}$ H, and C $^{\delta}$ H' protons.

Since all observable protons were assigned to specific types of amino acid residues, the second step was to align the spin systems according to the primary structure of conotoxin GIIIA with distance connectivities by NOESY experiment. This was achieved by using the NOE observed between an amino acid i and $i+1$ with the following pairs of protons: C $^{\alpha}$ H(i)-NH($i+1$) ($d_{\alpha\text{N}}$), C $^{\beta}$ H(i)-NH($i+1$) ($d_{\beta\text{N}}$) and NH(i)-NH($i+1$)

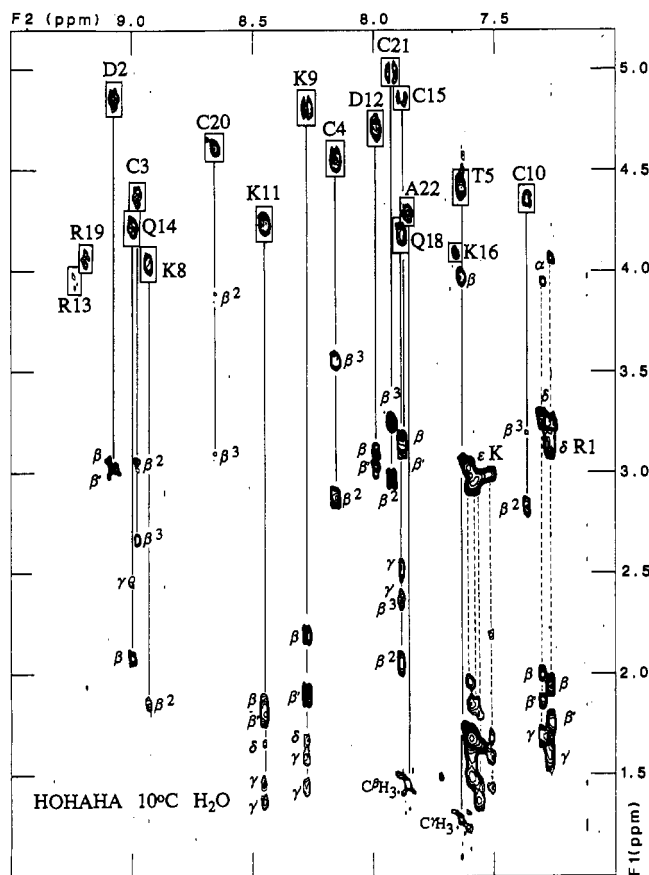


FIGURE 2: NH (F_2 axis)–aliphatic (F_1 axis) region of the HOHAHA spectrum recorded at 400 MHz (48 ms of mixing time) of conotoxin GIIIA at 10 °C in 90% H_2O /10% 2H_2O at pH 2.5. Spin systems are connected from the (NH, $C^\alpha H$) cross peaks (shown in the boxes). β^2 and β^3 protons, when indicated, refer to the stereospecifically assigned $C^\beta H$ protons. Dotted lines indicate magnetization transfer from N^H of Lys-8, -16, -11, and -9 and N^H of Arg-13 and -19 from left to right. Arg-1 is indicated apart.

(d_{NN}). Due to the small size of the peptide, it was found convenient to conduct experiments at 10 °C in addition to those at 28 °C. Lower temperature increased the viscosity of the solvent and appreciably increased the overall intensity of the NOESY cross peaks. Moreover, two missing amide to $C^\alpha H$ proton correlations of Lys-9 and Asp-12 at 28 °C were identified at 10 °C due to the shift of the water resonance. In addition, we noticed a sharpening of the amide proton of the Cys-10 at 10 °C, which gave rise to observable NOE cross peaks for this residue that completely vanished at 28 °C.

Since conotoxin GIIIA contains only one Thr and one Ala, Thr-5 and Ala-22 could be used as starting points for the sequential assignment process. Figure 4 shows the part of the NOESY spectrum recorded in H_2O at 10 °C for the $C^\alpha H$ –(i)–NH($i+1$) connectivities. From Thr-5, the assignment proceeded via $d_{\alpha N}$ until the N-terminal Arg-1. Similarly, starting from Ala-22 at the C-terminus led to the sequential assignment of the 18–22 segment. The connectivity observed in the NOESY spectra recorded at 10 °C in 2H_2O (data not shown), between the $C^\alpha H$ of Thr-5 and the $C^\beta H$ protons of one Hyp, identified the Hyp-6 and characterized a trans peptidic bond between Thr-5 and Hyp-6. The central segment from Hyp-7 to Hyp-17 contains only one Lys–Lys dipeptide unit, Lys-8 and Lys-9. This dipeptide unit was identified in the NOESY spectrum via a strong $d_{\alpha N}$. One of the two Lys gave an amide to $C^\alpha H$ correlation with one of the two remaining Hyp spin systems. This led us to the assignment of the Hyp-7. From the tripeptide sequence Hyp-7, Lys-8, and

Table I: 1H Resonance Assignments of Conotoxin GIIIA at 28 °C and pH 2.5

residue	chemical shift (ppm) ^a			
	NH	$C^\alpha H$	$C^\beta H$	others
R1		4.07	1.93, 1.93	$C^\gamma H$ 1.63, 1.63; $C^\delta H$ 3.17, 3.17; N^H 7.22
D2	9.00	4.86	3.07, 3.01	
C3	8.89	4.38	3.05, 2.69*	
C4	8.11	4.57	3.56*, 2.89	
T5	7.58	4.44	3.98	$C^\gamma H$, 1.38
HY6		4.79	2.39, 2.02	$C^\gamma H$ 4.68; $C^\delta H$ 4.08, 3.86
Hy7		4.80	2.43, 2.34	$C^\gamma H$ 4.62; $C^\delta H$ 3.72, 3.54
K8	8.83	4.05	1.96*, 1.86	$C^\gamma H$ 1.30, 1.30; $C^\delta H$ 1.49, 1.49; $C^\epsilon H$ 3.03, 3.03; N^H 7.56
K9	8.20	4.70	2.19, 1.90	$C^\gamma H$ 1.58, 1.45; $C^\delta H$ 1.68, 1.68; $C^\epsilon H$ 3.00, 3.00; N^H 7.44
C10	7.32	4.35	3.21*, 2.84	
K11	8.32	4.23	1.86, 1.79	$C^\gamma H$ 1.46, 1.38; $C^\delta H$ 1.66, 1.66; $C^\epsilon H$ 2.99, 2.99; N^H 7.53
D12	7.93	4.70	3.11, 3.00	
R13	9.15	3.94	1.99, 1.89	$C^\gamma H$ 1.71, 1.71; $C^\delta H$ 3.26, 3.26; N^H 7.23
Q14	8.89	4.22	2.08, 2.08	$C^\gamma H$ 2.46, 2.46; NH_2 7.69; ^b NH_2 6.99 ^b
C15	7.86	4.87	3.16, 3.12	
K16	7.60	4.09	1.96, 1.87	$C^\gamma H$ 1.58, 1.50; $C^\delta H$ 1.70, 1.70; $C^\epsilon H$ 2.95, 2.95; N^H 7.50
Hy17		4.67	2.41, 1.96	$C^\gamma H$ 4.51; $C^\delta H$ 3.81, 3.23
Q18	7.84	4.19	2.08, 2.36*	$C^\gamma H$ 2.36, 2.55; NH_2 7.61; ^b NH_2 7.08 ^b
R19	9.06	4.06	1.94, 1.94	$C^\gamma H$ 1.77, 1.77; $C^\delta H$ 3.25; N^H 7.21
C20	8.48	4.62	3.87, 3.09*	
C21	7.90	4.97	3.25*, 2.99	
A22	7.82	4.30	$C^\beta H$, 1.45	
CONH ₂	7.65 ^b			
CONH _E	7.13 ^b			

^a Accuracy of the chemical shifts measures ± 0.02 ppm. An asterisk indicates the H^β proton when the $C^\beta H$ methylene protons have been stereospecifically assigned. ^b Assignment and chemical shift measured at 10 °C. H₂ and H_E refer to NH protons cis and trans to C $^\gamma$ (C $^\alpha$ for Ala-22), respectively.

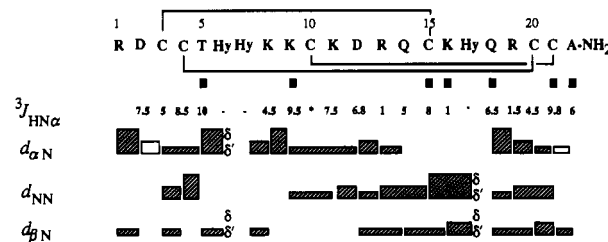


FIGURE 3: Summary of the sequential NOE connectivities involving the NH, $C^\alpha H$, and $C^\beta H$ protons measured at 10 °C and 250-ms mixing time. NOEs are classified into strong, medium, and weak according to the height of the shaded bar underneath the sequence. Open bars indicate bleached NOESY cross peaks at 10 °C subsequently measured at 28 °C. For hydroxyproline, sequential NOEs involving $C^\beta H$ are indicated. The accuracy of $^3J_{HN\alpha}$ measures ± 0.5 Hz on a one-dimensional resolution-enhanced spectrum at 28 °C. The asterisk indicates a broad NH resonance for which no $^3J_{HN\alpha}$ have been measured. Filled squares indicate slowly exchanging NH protons.

Lys-9, the sequential assignment proceeded via $d_{\alpha N}$ until Asn-14, which was connected to Cys-15 via d_{NN} and then Lys-16. The NOE observed between both the amide and $C^\beta H$ of Lys-16 to the $C^\delta H$ protons of the remaining Hyp confirmed the assignment of Hyp-17. This completed the sequential assignment of conotoxin GIIIA (Table I). The distance connectivities are summarized in the scheme of Figure 3. In addition to short-range NOEs, Figure 3 provides a summary of slowly exchanging amide protons as well as the $^3J_{HN\alpha}$ coupling constants. These coupling constants were measured on a resolution-enhanced one-dimensional spectrum in H_2O

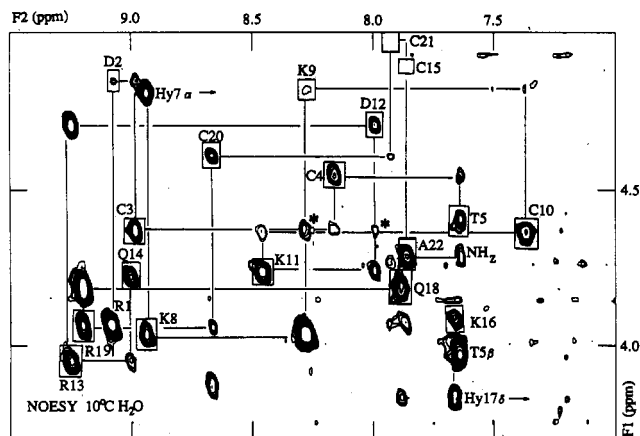


FIGURE 4: Part of the NH (F_2 axis)— C^α H (F_1 axis) of the NOESY spectrum recorded in 90% H_2O /10% 2H_2O at 400 MHz, 10 °C, and 250 ms of mixing time. NH to C^α H cross peaks of Cys-15 and Cys-21 are attenuated or vanished due to water suppression. These cross peaks are best observed at 28 °C. Asterisks indicate long-range NOEs, NH (K9)— C^α H (C3) and NH (D12)— C^α H (C10). The latter was only assigned on the basis of the structures obtained and not used in the calculations.

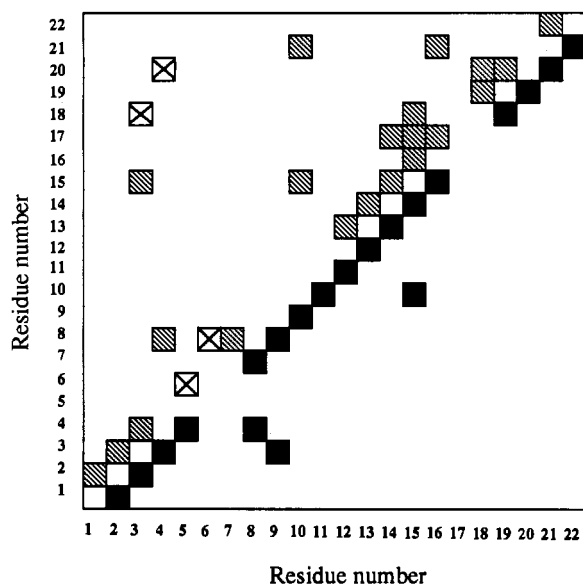


FIGURE 5: Diagonal plot representation of NOEs found for conotoxin GIIIA. The two axes are scaled with the primary sequence. The squares connect the pairs of the amino acids for which NOEs were observed. The filled squares in the lower right part indicate backbone to backbone NOEs. The hatched squares and crossed squares in the upper left part indicate backbone to side chain and side chain to side chain NOEs, respectively. Whenever both NOEs are observed, only hatched squares are displayed.

at 28 °C. Despite concomitant line-broadening, no sizeable differences were noticed at 10 °C for these couplings according to the appearance of the NH to C^α H HOHAHA cross peaks (Figure 2). The slowly exchanging amide protons had slow exchange rates with 2H_2O at 10 °C, enough to be observable in an overnight NOESY experiment that confirmed the assignment made on one-dimensional spectra. The rest of NOEs, besides the intraresidual ones, were mainly observed between α -proton and aliphatic proton regions of the NOESY spectrum in 2H_2O at 10 °C. The connectivities are shown in the diagonal plot of Figure 5. This diagonal plot, together with Figure 3, provides the basis for a qualitative interpretation of the secondary structure (Wüthrich, 1986). Very little secondary structure could be identified from these data. Only a β -turn formed by the segment 2–5 was recognized on the basis of the

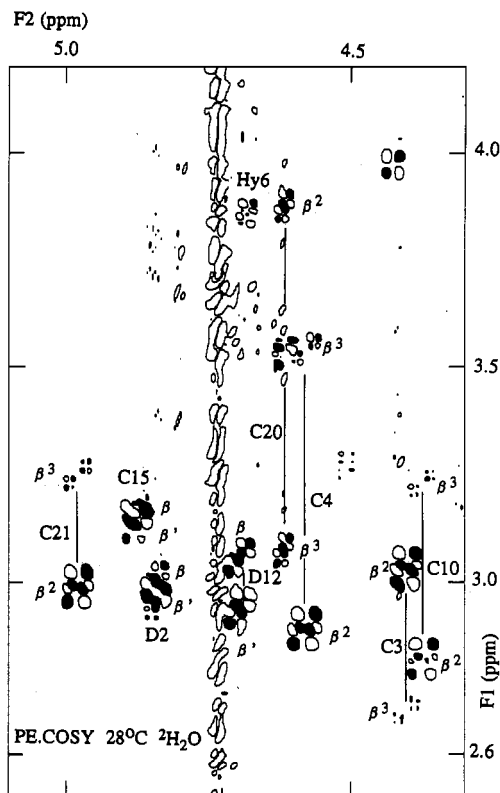


FIGURE 6: Part of the C^α H (F_2 axis)— C^β H (F_1 axis) of the PE-COSY spectrum recorded at 28 °C in 2H_2O and 400 MHz. Positive and negative contours are distinguished by filled or open contours. The t_1 noise strip is due to the residual 2HOH signal. Whenever possible, stereospecific assignments are specified.

Table II: Stereospecific Assignments of Conotoxin GIIIA

	C3	C4	K8	C10	Q18	C20	C21
$^3J_{\alpha\beta^2}$ (Hz) ^a	11.5	12.0	10.5	12.0	12.0	4.0	11.5
$^3J_{\alpha\beta^3}$ (Hz) ^a	2.0	3.0	3.0	2.0	2.5	3.5	2.0
NOE ^b NH- β^2	S	S	S	M	S	M	S
NOE ^b NH- β^3	W	W	W	W	M	S	W
conformer	t^2g^3	t^2g^3	t^2g^3	t^2g^3	t^2g^3	g^2g^3	t^2g^3

^a Measured with an accuracy of ± 0.5 Hz at 28 °C from the PE-COSY spectrum in D_2O . ^b Relative intensities observed in a NOESY experiment at 10 °C and 250 ms of mixing time, as strong (S), medium (M), and weak (W).

characteristic NOE pattern and $^3J_{HN\alpha}$ coupling constants. The slowly exchanging amide proton of Thr-5 confirmed the presence of a tight turn which could be classified in a type I class (Richardson, 1981). No other structural elements could be securely established at this stage.

Stereospecific Assignments and Input Data for the Three-Dimensional Structure Calculations. Figure 6 shows a part of the PE-COSY spectrum taken in 2H_2O at 28 °C. $^3J_{\alpha\beta}$ coupling constants were measured from the in-phase components of the cross peaks between C^α H and C^β H protons on the best resolved F_2 dimension. These coupling constants were combined with the intraresidual NH to C^β H NOEs in order to achieve the stereospecific assignments of the diastereotopic C^β H protons (Wagner et al., 1987; Hyberts et al., 1987). Using only these two sets of parameters, we obtained unambiguous stereospecific assignments of C^β H protons for Cys-3, Cys-4, Lys-8, Cys-10, Gln-18, Cys-20, and Cys-21 as described in Table II.

As input data for structure calculations, we used 62 sequential NOEs plus 24 long-range NOEs, which were classified into three distance constraints: ≤ 2.5 , ≤ 3.0 , and ≤ 4.0 Å corresponding to the strong, medium, and weak intensity

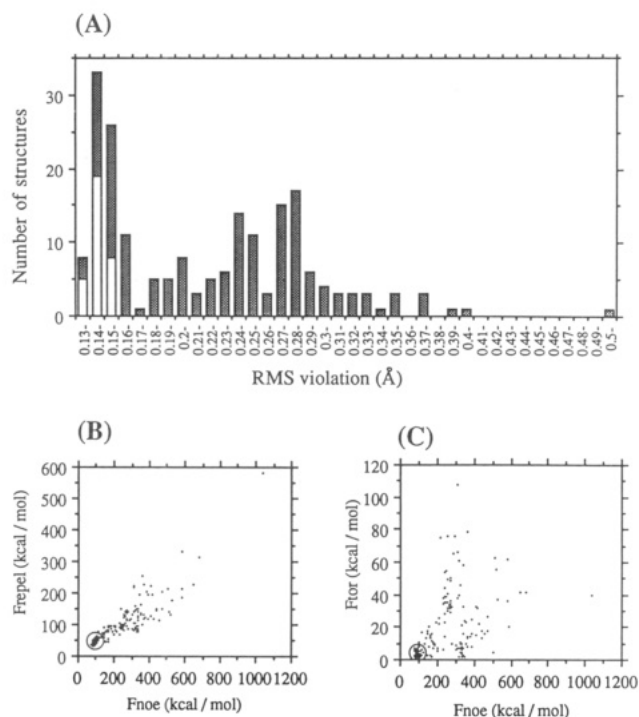


FIGURE 7: (A) Statistical distributions of calculated 200 structures of conotoxin GIIIA according to the RMS NOE violation where the open bars correspond to the 32 converged structures, (B) correlation diagram of the square-well potentials F_{NOE} versus F_{repel} , and (C) correlation diagram of the square-well potentials F_{NOE} versus F_{tor} . F_{repel} , F_{NOE} , and F_{tor} are defined according to eq 1, 2a, and 2b, respectively (see the text). All of the 32 final structures are localized inside the circles. Note that the F_{NOE} term is connected to the RMS NOE violation according to $F_{\text{NOE}} = k_{\text{NOE}}(\text{number of NOE})(\text{RMS NOE violation})^2$.

of NOESY cross peaks at 10 °C and 250 ms of mixing time. The NOE data were then converted to 74 final distance constraints (the complete lists of NOE constraints are available from the supplementary materials). A total of 11 constraints on the angle ϕ and the 7 on the angle χ_1 were also introduced according to the following rules (Wagner et al., 1987): a $^3J_{\text{HN}\alpha}$ greater than 8 Hz constrains the ϕ angle in the range of $-120 \pm 40^\circ$ and a $^3J_{\text{HN}\alpha}$ less than 5 Hz in the range of $-65 \pm 25^\circ$. A t^2g^3 side-chain conformation constrained the χ_1 angle in the $-60 \pm 60^\circ$ range and a g^2g^3 in the $+60 \pm 40^\circ$ range. We explicitly included the disulfide bridges between Cys-3 and Cys-15, Cys-4 and Cys-20, Cys-10 and Cys-21 (Hidaka et al., 1990). This was formulated as distance constraints between sulfur atoms S(*i*)–S(*j*) ($2.02 \pm 0.02 \text{ \AA}$) and between sulfur and C $^\beta$ atoms S(*i*)–C(*j*) and S(*j*)–C(*i*) ($2.99 \pm 0.50 \text{ \AA}$) across the bridges. No additional distance constraints were introduced such as slowly exchanging amide proton information (hydrogen bond) or intraresidual NOEs.

Structure Calculations. At first, we computed 25 solution structures in order to evaluate the folding of the peptidic chain and the overall coherence of our input data set. Among the structures with a reasonable RMS NOE violation (less than 0.2 Å), one distance constraint between C $^\alpha$ H of Hyp-7 and C $^\alpha$ H of Lys-8 had a systematic violation of more than 1 Å. After reexamination of the NOESY spectrum with the help of the preliminary folding, this cross peak was reassigned to C $^\alpha$ H of Hyp-6 instead of Hyp-7 ($\Delta\delta < 0.02$ ppm). The NOE between C $^\alpha$ H of Thr-5 and C $^\gamma$ H of Hyp-7 systematically violated by more than 0.5 Å was also relaxed. With these two modifications, we computed 200 solution structures. This required about 2 weeks on our Iris workstation. Figure 7 gives the graphical representations of the distribution of the 200

Table III: Structural Statistics^a

	32 structures	mean structure
RMS deviations from experimental distance restraints (Å) (83) ^b	0.145 ± 0.006	0.140
RMS deviations from experimental dihedral restraints (deg) (18)	1.6 ± 0.4	1.6
F_{NOE} (kcal·mol ⁻¹) ^c	87.9 ± 7.4	81.7
F_{tor} (kcal·mol ⁻¹) ^c	3.09 ± 1.47	2.94
F_{repel} (kcal·mol ⁻¹) ^c	44.6 ± 6.6	54.9
$E_{\text{L-J}}$ (kcal·mol ⁻¹) ^d	-13.0 ± 7.1	-8.8
RMS deviations from idealized geometry		
bonds (Å) (350)	0.010 ± 0.0006	0.011
angles (deg) (649)	1.972 ± 0.016	4.120
impropers (deg) (83) ^e	0.776 ± 0.116	0.599
dihedral (deg) (21) ^f	4.295 ± 0.324	4.379

^aThe "32 structures" refers to the final set of dynamical simulated annealing structures; the "mean structure" refers to the restrained minimized structure obtained by restrained minimization of the averaged coordinate of the individual 32 structures; the number of specified constraints are given in parentheses. ^bNo structures have distance violations greater than 0.5 Å. ^cThe values of the force constants used for the calculation of the square-well potentials are the original values of 50 kcal·mol⁻¹·Å⁻¹ and 200 kcal·mol⁻¹·rad⁻², for NOE and torsion angle potentials, respectively (eqs 2a and 2b). The value of the van der Waals repulsion term (eq 1) is calculated with the original force constant of 4 kcal·mol⁻¹·Å⁻⁴ with the van der Waals radii scaled by a factor 0.8 of the standard value used in the CHARMM empirical function (Brooks et al., 1983). ^d $E_{\text{L-J}}$ is the Lennard-Jones van der Waals energy calculated with the CHARMM empirical energy function (Brooks et al., 1983), which was not included in the simulated annealing process. ^eThe improper torsion terms are used to maintain the planar geometries as well as the chiralities. ^fDihedral angle term only used for restraining the peptidic ω angle. The conformation around the peptide bond is maintained as trans. Hyp-6, Hyp-7, and Hyp-17 are allowed a cis–trans interconversion with a 40 times reduced force constant of the ω angle dihedral potential term.

Table IV: Atomic RMS Differences of 32 Structures versus Mean Structure^a

	RMSDs (Å)	
	residues 2–21	residues 2–10, 14–21
backbone		
(N, C $^\alpha$, C)	0.85 ± 0.28	0.55 ± 0.30
(N, C $^\alpha$, C, O, C $^\beta$)	1.02 ± 0.31	0.71 ± 0.32
all heavy atoms	1.81 ± 0.34	1.33 ± 0.32

^aThe terms "32 structures" and "mean structure" refer to the same as described in Table III.

calculated structures in the different classes of RMS NOE violation and the correlation diagrams of 200 calculations for two different energetic potentials, (F_{NOE} , F_{repel}) and (F_{NOE} , F_{tor}). A sharp peak occurs at the 0.13–0.17 Å interval for RMS NOE violation and around 100, 50, and 6 kcal·mol⁻¹ for F_{NOE} , F_{repel} , and F_{tor} , respectively (shown with circles in Figure 7B,C). This clearly shows the convergence of the structures distinguished from the rest of those trapped in local minima occasionally with different foldings. Thus, we selected the 67 structures of the 0.13–0.17 Å interval of RMS NOE violations from which the final set of 32 structures was retrieved with no violations greater than 0.5 Å (shown in white-washed bar in Figure 7A). Structural statistics for the converged structures and atomic RMSDs are given in the Table III and IV, respectively (individual correlation data between the converged structures are available in the supplementary materials). Figure 8 shows the best-fit overlays of the 32 converged structures, and the mean structure is shown in Figure 9. Figure 10 gives the variations of the number of distance constraints and RMSDs with the sequence of conotoxin GIIIA.

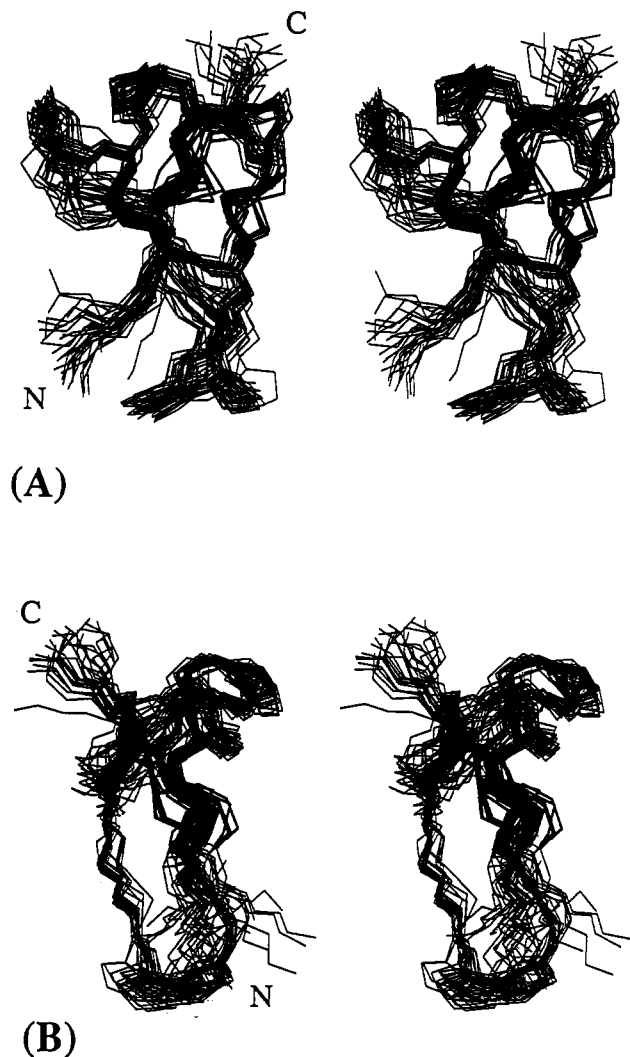


FIGURE 8: Stereopairs of the best-fit superposition of the 32 final converged structures of conotoxin GIIIA. In addition to the (N, C α , C) backbone atoms, C β , C γ , and C δ of Hyp are displayed with C β and sulfur atoms of Cys in bold lines. The stereopair displayed in (B) is a 90° left side view of (A).

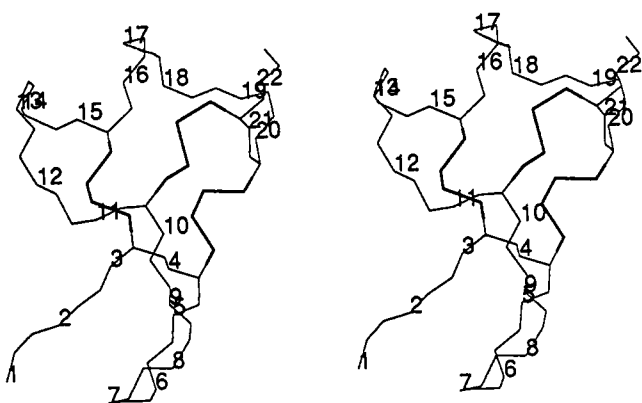


FIGURE 9: Stereoview of backbone atoms (N, C α , C) trace of the mean structure of conotoxin GIIIA obtained by averaging and minimizing the 32 converged coordinates displayed in Figure 8. The sulfur and C β of Cys are represented in bold lines in addition to the backbone atoms. The numbering refers to the primary sequence of conotoxin GIIIA.

In order to evaluate the contribution of the information on the disulfide bridges to the success of folding of the peptidic chain, we have made additional calculations where the associated constraints were removed from the data set. The folding

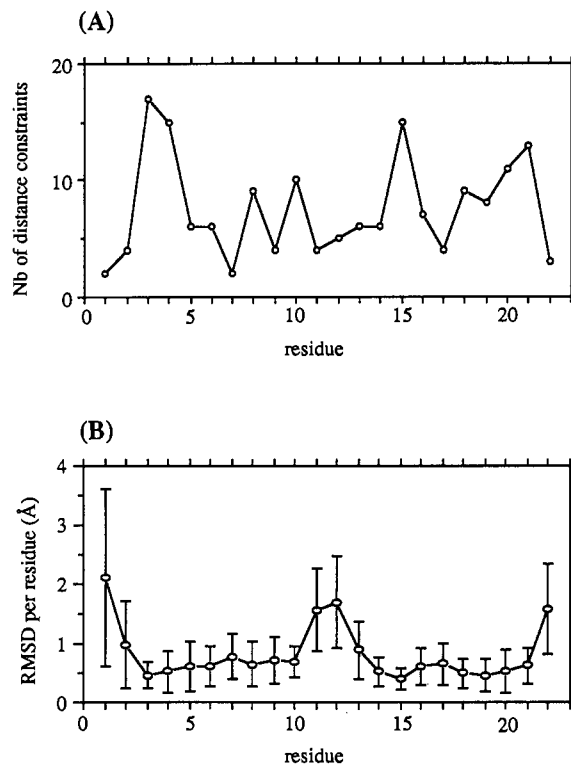


FIGURE 10: Distribution of (A) the number of experimental distance constraints and (B) the RMS difference from the mean structure per residue on the primary sequence of conotoxin GIIIA.

Table V: Dihedral Angles of Disulfide Bridges and Their Standard Deviations for the "32 Converged Structures" of Conotoxin GIIIA

disulfide bridges	dihedral angle (deg) ^a		
	χ_2	χ_3	χ_2'
C3-C15	-97 ± 9 (-91)	-68 ± 4 (-69)	-50 ± 5 (-48)
C4-C20 ^b	-69 ± 79 (-97)	-54 ± 28 (-47)	-119 ± 80 (-148)
C10-C21	-101 ± 26 (-123)	-117 ± 27 (-118)	-4 ± 36 (-10)

^aThe values for the mean structure, as defined in Table III, are given in parentheses. ^bSeven out of the "32 converged structures" containing mirror images of C4-C20 disulfide bridges were omitted from the analysis.

found for the less violated structures (ca. RMS NOE violation of 0.16 Å) was similar in every respects to the previous one. However, due to the exceptional array of the sulfur atoms, it would have been difficult to assign the paired cysteines without further refinements.

Description of the Structure. The general folding of conotoxin GIIIA is composed of a successive pair of tight turns from Asp-2 to Thr-5 and from Thr-5 to Lys-8 followed by an almost extended part from Lys-8 to Asp-12. A loop from Asp-12 to Cys-15 is followed by a small right-handed helix with only one turn until Gln-18. Finally, the tail exits from the core by a final loop placing the C-terminus in an almost opposite direction relative to the N-terminus. The conspicuous arrangement of the three disulfide bridges, which are successfully defined in conformation, forms a cage of sulfur atoms at the center of the molecule. Only the disulfide bridge between Cys-4 and Cys-20 gives some mirror images (7 out of 32). Table V gives the values of dihedral angles of the disulfide bridges among the converged structures and for the mean structure. If we exclude the structures with the mirror image forms, the disulfide bridges have a left-handed helicity with χ_2 and χ_3 torsion angles in the range of the most commonly encountered conformation for disulfide bridges ($\chi_1 \sim -60^\circ$ and $\chi_2 \sim \chi_3 \sim -90^\circ$) (Richardson, 1981). Only Cys-20 takes

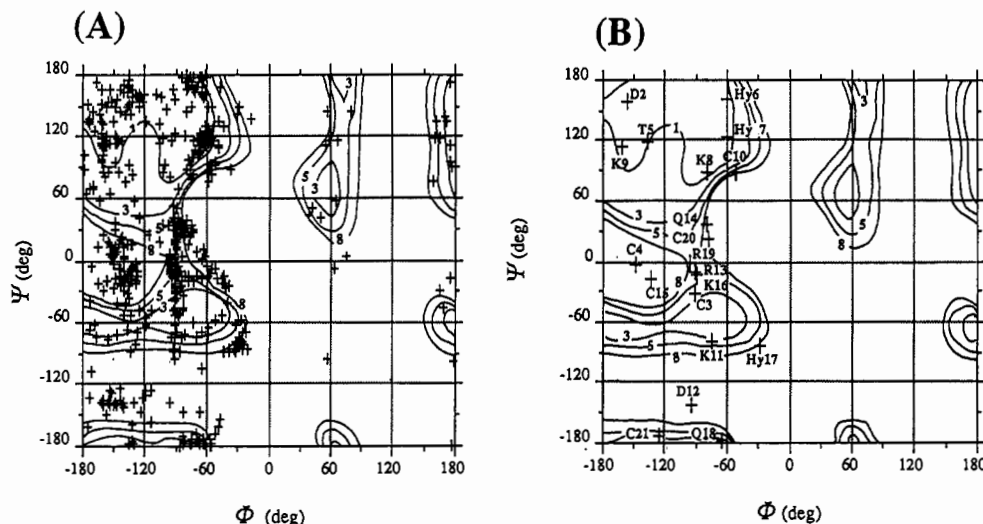


FIGURE 11: Ramachandran plots of the backbone conformational (ϕ, ψ) angles of conotoxin GIIIA superimposed on the contours of the calculated energies (given in kcal·mol⁻¹) for poly(L-Ala) (Brant & Schimmel, 1967). (A) Plot for the 32 converged structures and (B) plot for the mean structure as defined in Table III.

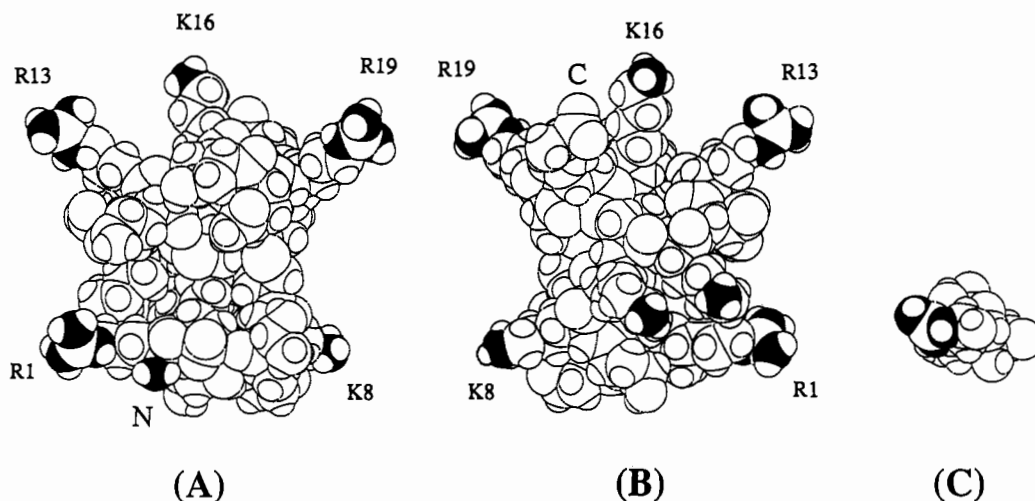


FIGURE 12: Space-filling representation of the mean structure of conotoxin GIIIA. Black spheres indicate the side-chain nitrogen atoms of Arg, Lys, and the N-terminal group. Pictures A and B represent the two opposite faces of conotoxin GIIIA. For comparison, tetradotoxin is displayed in picture C at the same scale. In B, Lys-9 (left) and Lys-11 (right) face with each other.

a g^2g^3 conformation with a χ_1 angle of $+60^\circ$. Figure 11 provides plots of dihedral backbone angles in the (ϕ, ψ) conformation space (Ramachandran plots) for the 32 converged structures (A) and the mean structure of conotoxin GIIIA (B). These Ramachandran plots are good indicators of the overall quality of the results.

The specific backbone folding of conotoxin GIIIA determines the orientation of the basic side chains of Arg-1, Lys-8, Arg-13, Lys-16, and Arg-19 in a radial orientation relative to the center defined by the sulfur cage. The side chains of Lys-9 and Lys-11 are located on the seam side of the structure and face each other. This is illustrated in Figure 12 by the space-filling model of the mean structure. The maximum distances between each cationic site range from 20 to 25 Å.

DISCUSSION

We have determined the three-dimensional structure of conotoxin GIIIA on the basis of NMR information by a two-step simulated annealing method. The first simulated annealing step used a variable target function in real space (Nilges et al, 1988a). In this step, the random sampling in the conformational space was achieved by assigning random initial velocities for each atom according to a Maxwell dis-

tribution at 1000 K, starting from an extended structure. In order to confirm the random sampling of the structures in the conformational space, six representative trajectories (given in the supplementary material) were examined according to Nilges et al. (1988a). Pairwise RMSDs starting from 0 Å at the initial point diverged to as much as 6 Å during the first step of the structure calculation using the variable target function protocol and then regularly decreased until the convergence point. Because backbone RMSDs obtained for 10 conotoxin GIIIA structures randomized in the dihedral (ϕ, ψ) angle space were calculated to be 6.7 ± 1.3 Å, we can conclude that the initial stage of the simulation samples a wide range of the conformational space. The statistical analysis of the 200 resulted structures shows that the structures with less RMS NOE violations also correspond to those with minimum energy values of each energy potential term (Figure 7B,C). Thus, we found that the criterion used for the structure selection was sufficient although it was based solely on the NOE terms. The nonselected structures (RMS NOE violation >0.18 Å in Figure 7A) were all trapped in local minima. They all present larger RMSDs from the converged mean structure (data provided in the supplementary material).

Considering the 32 converged structures of conotoxin GIIIA

displayed in Figure 8, it appears to be conformationally well-defined. However, from Figure 8 and Figure 10, one can realize that the less defined part of the structure lies in the Lys-11 to Arg-13 portion. This may be explained by the lack of dihedral ϕ angle constraints for Lys-11 and Asp-12 as well as a minimum number of distance constraints for this part of the molecule (Figure 10A). Moreover, the striking behavior of the amide proton of Cys-10, which sharpened at low temperature relative to the ambient, is also a good indicator of a particular dynamic of this segment. The inspection of the structures shows that the conformation of the peptide bonds of Hyp-6, Hyp-7, and Hyp-17 are defined as trans, cis, and trans, respectively. This is in accordance with the NOEs found between C^αH of Thr-5 and C^βH of Hyp-6 and those found between Lys-16 and Hyp-17. Unfortunately, the chemical shift degeneracy of the C^αH of Hyp-6 and Hyp-7 (vide supra) did not allow the observation of a NOE between these two C^αH protons, so that cis peptidic bond conformation for Hyp-7 was not experimentally supported. Besides the NH proton of Thr-5 engaged in the type I β -turn, some other slowly exchanging protons were observed, which were expected to be involved in the hydrogen bond formation. From the inspection of the mean structure of conotoxin GIIIA (Figure 9), N–O distances and N–H–O angles were evaluated for the candidates of hydrogen bonds. We used a N–O distance of less than 3.3 Å and a N–H–O angle of greater than 110° (Richardson, 1981) as a criterion of hydrogen-bond formation. We found hydrogen bonds between Lys-9 NH and Hyp-7 C=O (3.0 Å, 117°) and between Cys-15 NH and Arg-13 C=O (2.75 Å, 139°). For Lys-16, Gln-18, and Cys-21, the slowly exchanging amide protons could be rationalized by hydrogen-bond formation with the carbonyl oxygens of Cys-15 (3.4 Å, 108.5°), Arg-13 (3.5 Å, 160°), and Gln-18 (3.7 Å, 133°), respectively, with some distortion on angles or distances. These protons are located in the turn and the small helical part formed by Cys-15, Lys-16, Hyp-17, and Gln-18. No explanation can be provided for the slowly exchanging NH proton of the C-terminal Ala-22.

Small and highly constrained cysteine-rich peptides have already been reported. Representative examples can be taken from the recently introduced knottins mini-protein family (Le-Nguyen et al., 1990). Some of them were determined at high resolution either by NMR and X-ray studies (Holak et al., 1989a,b) or NMR alone (Heitz et al., 1989; Gariépy et al., 1986). Contrary to knottins, no disulfide bridges interpenetrate a macrocycle formed by other two ones in conotoxin GIIIA. From Figure 8, it is easy to realize that conotoxin GIIIA is a relatively flat and discoidal molecule (ca. 6 Å thick \times 15 Å in diameter for the backbone) with the cationic side chains of Arg and Lys distributed at the periphery (Figure 12). One of the discoidal surfaces is characterized by Lys-9 and Lys-11, which face with each other. The chemical shift degeneracy of a number of C^γH, C^δH of Arg and C^αH of Lys is in accordance with the full solvent exposure and the flexibility of these side chains. Our results indicate that, except for the Lys-11 to Arg-13 segment, conotoxin GIIIA is based on a rigid core that defines the overall radial orientations of flexible cationic side chains. In terms of molecular interaction, the mobility of the side chains of Arg and Lys residues should be understood as an advantage for a good fit to a receptor by local optimized interactions of those side chains.

Modified conotoxins at the amino-terminal position or the N^δH position of Lys could be cross-linked covalently to the skeletal muscle sodium channel (Cruz et al., 1989). Consequently, it can be expected that the side chains of all these basic

amino acids interact with anionic sites located within the macromolecular arrangement of the channel. Actually, these anionic sites have been detected through number of electrophysiological experiments (Begenisich, 1987). They are expected to be located in the extracellular "funnel" part of the channel and would surround a specific carboxylate implicated in the binding of tetrodotoxin and saxitoxin. Recently, the presence of such a carboxylate was confirmed by an experiment of a single point mutation of a highly conserved glutamic acid (Noda et al., 1989). Tetrodotoxin and saxitoxin are two water-soluble guanidinium cations of small size compared to conotoxin GIIIA (ca. 300 vs 4000 Å³, see also Figure 12). They are known to block the different subtypes of voltage-gated sodium channel except for the heart one. On the other hand, conotoxin GIIIA was demonstrated (Cruz et al., 1985) to bind competitively with tetrodotoxin and saxitoxin on the skeletal muscle subtype. Consequently, it is possible that one of the guanidinium or ammonium cations of conotoxin GIIIA particularly competes with the binding site of tetrodotoxin and saxitoxin.

The complex structures of the voltage-gated ion channels associated to the transmembrane environment make difficult the comprehensive understanding of the molecular basis of their actions. The structure of conotoxin GIIIA provides an entryway-fingerprint of the sodium channel of the skeletal muscle subtype and highlights the difference between those of the brain and other subtypes. Its molecular architecture based on a rigid core and flexible side chains is most probably a general feature of a number of peptidic inhibitors.

ACKNOWLEDGMENTS

J.-M.L. thanks the Science and Technology Agency of the Government of Japan for a postdoctoral award 1990-1991 and the Centre National de la Recherche Scientifique of France for financial support.

SUPPLEMENTARY MATERIAL AVAILABLE

Four tables listing of NOEs and XPLOR input files of constraints and three figures giving statistical data for the calculated structures (8 pages). Ordering information is given on any current masthead page.

Registry No. Conotoxin GIIIA, 86394-16-3.

REFERENCES

- Bax, A., & Davis, D. G. (1985) *J. Magn. Reson.* 65, 393–402.
- Begenisich, T. (1987) *Annu. Rev. Biophys. Biophys. Chem.* 16, 247–263.
- Brant, D. A., & Schimmel, P. R. (1967) *Proc. Natl. Acad. Sci. U.S.A.* 58, 428–435.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem.* 4, 187–217.
- Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science* 235, 458–460.
- Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986a) *EMBO J.* 5, 2729–2735.
- Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986b) *J. Mol. Biol.* 191, 523–551.
- Clore, G. M., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. T. (1990) *J. Mol. Biol.* 214, 811–817.
- Cruz, L. J., Gray, W. R., Olivera, B. M., Zeikus, R. D., Kerr, L., Yoshikami, D., & Moczydlowski, E. (1985) *J. Biol. Chem.* 260, 9280–9288.

- Cruz, L. J., Kupryszewski, G., LeCheminant, G. W., Gray, R. W., Olivera, B. M., & Rivier, J. (1989) *Biochemistry* 28, 3437-3442.
- Gariépy, J., Lane, A., Frayman, F., Wilbur, D., Robien, W., Schoolnik, G. K., & Jardetzky, O. (1986) *Biochemistry* 25, 7854-7866.
- Heitz, A., Chiche, L., Le-Nguyen, D., & Castro, B. (1989) *Biochemistry* 28, 2392-2398.
- Hidaka, Y., Sato, K., Nakamura, H., Kobayashi, J., Ohizumi, Y., & Shimonichi, Y. (1990) *FEBS Lett.* 264, 29-32.
- Holak, T. A., Gondol, D., Otlewski, J., & Wilusz, T. (1989a) *J. Mol. Biol.* 210, 635-648.
- Holak, T. A., Bode, W., Huber, R., Otlewski, J., & Wilusz, T. (1989b) *J. Mol. Biol.* 210, 649-654.
- Hyberts, S. G., Marki, W., & Wagner, G. (1987) *Eur. J. Biochem.* 164, 625-635.
- Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys.* 71, 4546-4553.
- Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Nishiuchi, Y., Sakakibara, S., Braun, W., & Go, N. (1989) *Biochemistry* 28, 4853-4860.
- Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry* 28, 7241-7257.
- Le-Nguyen, D., Heitz, A., Chiche, L., Castro, B., Boigegrain, R. A., Favel, A., & Colletti-Previero, M. A. (1990) *Biochimie* 72, 431-435.
- Macura, S., Hyang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson.* 43, 259-281.
- Moczydlowski, E., Olivera, B. M., Gray, W. R., & Strichartz, G. R. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 5321-5325.
- Mueller, L. (1987) *J. Magn. Reson.* 72, 191-196.
- Nakamura, H., Kobayashi, J., Ohizumi, Y., & Hirata, Y. (1983) *Experientia* 39, 590-591.
- Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988a) *Protein Eng.* 2, 27-38.
- Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988b) *FEBS Lett.* 229, 317-324.
- Noda, M., Suzuki, H., Numa, S., & Stuhmer, W. (1989) *FEBS Lett.* 259, 213-216.
- Olivera, B. M., Gray, W. R., Zeikus, R., McIntosh, J. M., Varga, J., Rivier, J., de Santos, V., & Cruz, L. J. (1985) *Science* 230, 1338-1343.
- Olivera, B. M., Rivier, J., Clark, C., Ramillo, C. A., Corpuz, G. P., Abogadie, F. C., Mena, E. E., Woodward, S. R., Hillyard, D. R., & Cruz, L. J. (1990) *Science* 249, 257-263.
- Omichinski, J. G., Clore, G. M., Appella, E., Sakaguchi, K., & Gronenborn, A. M. (1990) *Biochemistry* 29, 9324-9334.
- Pardi, A., Galdes, A., Florance, J., & Maniconte, D. (1989) *Biochemistry* 28, 5494-5501.
- Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 117, 479-485.
- Richardson, J. S. (1981) *Adv. Protein Chem.* 34, 97-102.
- Sato, S., Nakamura, H., Ohizumi, Y., Kobayashi, J., & Hirata, Y. (1983) *FEBS Lett.* 155, 277-280.
- States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson.* 48, 286-292.
- Trimmer, J. S., & Agnew, W. S. (1989) *Annu. Rev. Physiol.* 51, 401-418.
- Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol.* 196, 611-639.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
- Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol.* 169, 949-961.
- Yanagawa, Y., Abe, T., Satake, M., Odani, S., Suzuki, J., & Ishikawa, K. (1988) *Biochemistry* 27, 6256-6262.
- Zuiderweg, E. R. P., Hallenga, K., & Olejniczak, E. T. (1986) *J. Magn. Reson.* 70, 336-343.