Solution Structures of α-Conotoxin G1 Determined by Two-Dimensional NMR Spectroscopy†

Arthur Pardi,*,‡ Alphonse Galdes,§ James Florance,§ and Duane Maniconte‖

Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, Colorado 80309-0215, Department of Chemistry, Rutgers, The State University of New Jersey, New Brunswick, New Jersey 08903, and Boc Group Technical Center, Health Care, Murray Hill, New Jersey 07974

Received November 18, 1988; Revised Manuscript Received March 17, 1989

ABSTRACT: Two-dimensional NMR data have been used to generate solution structures of α-conotoxin G1, a potent peptide antagonist of the acetylcholine receptor. Structural information was obtained in the form of proton–proton internuclear distance constraints, and initial structures were produced with a distance geometry algorithm. Energetically more favorable structures were generated by using the distance geometry structures as input for a constrained energy minimization program. The results of both of these calculations indicate that the overall backbone conformation of the molecule is well-defined by the NMR data whereas the side-chain conformations are generally less well-defined. The main structural features derived from the NMR data were the presence of tight turns centered on residues Pro^5 and Arg^9. The solution structures are compared with previous proposed models of conotoxin G1, and the NMR data are interpreted in conjunction with chemical modification studies and structural properties of other antagonists of the acetylcholine receptor to gain insight into structure–activity relationships in these peptide toxins.

Conotoxins are small basic peptides isolated from marine snails of the genus *Conus* (Gray et al., 1981; Olivera et al., 1985; Cruz et al., 1985). These snails immobilize their prey by a highly specialized venom apparatus where the venom is injected into the prey with a disposable hollow tooth that functions as a harpoon (Kohn, 1958; McMichael, 1971). The deadliest of the *Conus* species is the *Conus geographus*, which has reportedly caused human fatalities (Kohn, 1958; McMichael, 1971). The α-conotoxins, such as G1 from *Conus geographus*, are lethal due to muscular paralysis, especially of the respiratory muscles. The mechanism of action is by antagonism of acetylcholine binding to the postsynaptic acetylcholine nicotinic receptor at the neuromuscular junction (McManus et al., 1981). This is the same mechanism of action as the larger snake toxins such as α-bungarotoxin (Low, 1979).

The α-conotoxins are small peptides, 13–15 amino acids in length, and the primary structure of conotoxin G1, determined by Gray et al. (1981), is shown in Table I. Using Chou–Fasman techniques, and comparing sequence similarities with α-bungarotoxin, Gray and co-workers proposed a tertiary structure for G1 that includes two β turns, one being the putative "binding loop" from Gly^8 to Tyr^{11} and the other turn centered on Pro^5 (Gray et al., 1985). This structure allows for two cationic groups (an arginine and the α-amino terminus on Glu^1) to have a separation of approximately 11 Å. This charge separation has been observed in X-ray crystal structures of rigid antagonists of the nicotinic acetylcholine receptor, including curare, and is believed necessary for activity. An alternate tertiary structure for conotoxin G1 has been proposed by Hider (1985) based on circular dichroism data and Chou–Fasman analysis, with this model predicting a 50% α-helical structure. Both models are based on limited structural data, and so in this study, we present the structure of conotoxin G1 derived from two-dimensional (2D)[1] nuclear magnetic resonance data and computational techniques.

Present methodology makes it possible to determine the three-dimensional structures of peptides and small proteins in solution from NMR data [see Wüthrich (1986) for a review]. The NMR structural data consist mainly of proton–proton internuclear distances derived from two-dimensional nuclear Overhauser effect experiments. Structures consistent with these distance constraints are then generated by computational methods such as a distance geometry algorithm, constrained molecular dynamics, constrained molecular mechanics, or constrained Monte Carlo techniques (Crippen, 1977; Kuntz et al., 1979; Havel et al., 1979; Braun & Go, 1985; Havel & Wüthrich, 1985; Clore et al., 1985, 1986; Bruenger et al., 1986; Holak et al., 1987; Bassolino et al., 1988).

Preliminary NMR studies of α-conotoxin G1 have been performed by Kobayashi and co-workers (Kobayashi et al., 1984), and solution structures were generated from NOE distance data with a minimization algorithm working in torsion angle space (Kobayashi et al., 1986). Since no proton resonance assignments or NOE data were reported by these workers, we initiated an independent structure determination of α-conotoxin G1. We made essentially complete proton resonance assignments on G1, and structures were generated with a distance geometry program. The results of the distance

† This work was supported in part by grants from the Searle Scholars Program of the Chicago Community Trust (85-C110) and NIH AI27026 to A.P. The 400-MHz spectrometer was purchased with partial support from NSG Grant CHEM-8300444. This work is dedicated to the memory of Duane Maniconte.

* To whom correspondence should be addressed.
‡ University of Colorado at Boulder.
§ Boc Group Technical Center.
‖ Rutgers, The State University of New Jersey.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two dimensional; COSY, two-dimensional correlation spectroscopy; NOE, nuclear Overhauser effect; DQF-COSY, two-dimensional double-quantum-filtered correlation spectroscopy; TOCSY, two-dimensional total correlation spectroscopy; RELAYED-COSY, two-dimensional relayed coherence-transfer spectroscopy; ppm, parts per million; AMX, weakly coupled three-spin system; G1, α-conotoxin G1; FID, free induction decay.

0006-2960/89/0428-5494$01.50/0

geometry calculations show that the backbone conformation of conotoxin G1 is quite well-defined from these data. These structures were then subjected to constrained molecular mechanics techniques. The solution structures will be compared with the previously proposed models for conotoxin G1 and with data on other antagonists of the acetylcholine receptor.

Materials and Methods

NMR Data. Conotoxin G1 was purchased from the Peptide Institute Inc., Osaka, Japan, and was used without further purification. Samples were prepared at a concentration of approximately 7 mM in either 100% ^2H_2O (Stohler Isotopes) or 90% H_2O/10% ^2H_2O at pH 4.8 in a manner previously described (Bach et al., 1987). All NMR spectra were run on a Varian XL-400 NMR spectrometer operating a 400 MHz with a sample temperature of 5.0 ± 0.5 °C. 2D NMR spectra were obtained in the phase-sensitive absorption mode with quadrature detection in both dimensions (Mueller & Ernst, 1978). Methods used for data acquisition and data processing of the 2D NMR experiments were similar to that described elsewhere (Bach et al., 1987). The following 2D NMR experiments were used to make resonance assignments and obtain structural information: 2D NOE, COSY, and RELAYED-COSY experiments in 90% H_2O/10% ^2H_2O and 2D NOE, DQF-COSY, TOCSY, and RELAYED-COSY experiments in 100% ^2H_2O. The 2D NOE spectra were obtained with mixing times of 100, 200, and 300 ms. All NMR data were transferred to a microVAX II computer, and data processing was performed by using the FORTRAN program FTNMR (Hare Research, Inc.).

Distance Geometry Structures. The distance geometry program DSPACE (Hare Research, Inc.) was used to calculate structures for conotoxin G1 on a microVAX II computer. DSPACE uses a metric–matrix embedding technique to generate trial structures that are subsequently refined by using nonlinear optimization algorithms to minimize the differences between the structures and the input distance constraints (Hare et al., 1986; Nerdal et al., 1988). The input for the distance geometry program consists of upper and lower distance bounds estimated from the 2D NOE spectra. All protons were explicitly defined in the distance geometry calculations except for the methyl protons on Ala^6 where a pseudoatom was substituted for this methyl group (Wüthrich et al., 1983). For any experimental distances involving the Ala^6 methyl group, 1 Å was added to these upper distance bounds. Most distances were entered with a precision of ±0.6 Å except for distances involving protons on prochiral centers and the aromatic protons on the tyrosine ring that were handled as previously described (Pardi et al., 1988). In addition to the experimental distance constraints, a set of distances that define the covalent structure of the peptide, including the two disulfide linkages Cys^2–Cys^7 and Cys^3–Cys^{13}, were included in the input for the distance geometry algorithm. L-Amino acids and trans-peptide bonds were assumed for all amino acid residues.

Constrained Energy Minimization. Energy minimization of the distance geometry structures was performed on a VAX 8530/8350/11-780 computer cluster with the AMBER molecular mechanics software package (Weiner & Kollman, 1981; Weiner et al., 1984). The "all atom" force field was used, in which hydrogen atoms are explicitly included. A (positively) charged N-terminus was used, and the net atomic charges for the C-terminal amide were set to −0.68 (N) and +0.34 (H). A distance-dependent dielectric constant of $\epsilon = r_{ij}$ was used; all 1–4 nonbonded and electrostatic interactions were scaled by 0.5, and no nonbonded cutoff was used. A minimization was judged to have converged when the final rms energy gradient was less than 0.1 kcal/mol. Typically, this required approximately 2000 iterations for each structure.

The initial coordinates for energy minimization were taken directly from the distance geometry structure. Coordinates for the undefined atoms (such as the methyl hydrogens on Ala^6) were generated from the heavy-atom coordinates. In addition, the psuedoatom lone pairs on the cystine sulfur atoms, which are required by the AMBER force field, were added such that the torsion angle τ(LP–S–S–C$^\beta$) = τ(C$^\beta$–S–S–C$^\beta$) ± 100°. Unless this was done, the initial energy was very high (>6000 kcal/mol), and many more iterations were required for convergence.

In the constrained energy minimizations, an extra harmonic term of the form $K(r - r_0)^2$ was added to constrain the internuclear proton–proton distances (r) to the NMR-determined distance (r_0). Various values in the range of 1–20 kcal/(mol·Å^2) were used for the force constant K, with K = 10 kcal/(mol·Å^2) giving the best results. Since stereospecific assignments of protons on prochiral centers were not made, the pair of protons that best met the NOE-determined distance constraint was selected.

Results

Resonance Assignment of G1. The tyrosine aromatic, the histidine aromatic, and the alanine methyl proton resonances of conotoxin G1 were readily identified by comparison with the chemical shifts of "random coil" peptides (Wüthrich, 1986; Bundi & Wüthrich, 1979). The sequential resonance assignment procedure developed by Wüthrich and co-workers was used to assign the rest of the proton spectrum (Wüthrich et al., 1983). This procedure basically consists of employing a variety of 2D NMR experiments to map out both spin–spin J coupling connectivities and nuclear Overhauser effect connectivities. The J coupling type experiments group proton resonances into individual spin systems which often allows identification of each spin system to specific amino acid type(s). Sequential resonance assignment is completed by application of 2D NOE spectra that map out connectivities between protons on neighboring amino acid residues. The interproton distances used in sequential assignment involve connectivities between an amide proton on one amino acid residue, to the amide, C$^\alpha$, and/or C$^\beta$ protons on the previous residue. The following notation is commonly used to describe these distances (Wüthrich et al., 1984):

$$d_{\alpha N} = d(C^\alpha H_i, NH_{i+1}) \qquad d_{NN} = d(NH_i, NH_{i+1})$$
$$d_{\beta N} = d(C^\beta H_i, NH_{i+1})$$

Figure 1 shows a portion of the double-quantum-filtered COSY (Piantini et al., 1982; Shaka & Freeman, 1983) experiment on conotoxin G1 where a number of amino acid spin systems are illustrated. By combining this spectrum with the TOCSY (Braunschweiler & Ernst, 1983; Davis & Bax, 1985) spectrum (data not shown), it was possible to unambiguously identify the nonexchangeable proton resonances of Glu^1, Pro^5, Ala^6, Gly^8, and Arg^9 because these residues have unique spin systems in the peptide. The remaining nonexchangeable protons were grouped into individual "AMX" spin systems (Wüthrich, 1986) but could not be assigned to specific amino acid residues at this point. The next step in the assignment procedure was to attach the exchangeable amide protons to individual spin systems. Figure 2 shows the NH–C$^\alpha$H "fingerprint" region of the COSY spectrum in H_2O where the predicted number of NH–C$^\alpha$H cross-peaks for the amino acid sequence of conotoxin G1 are identified. Many of the amide protons could be unambiguously attached to the previously

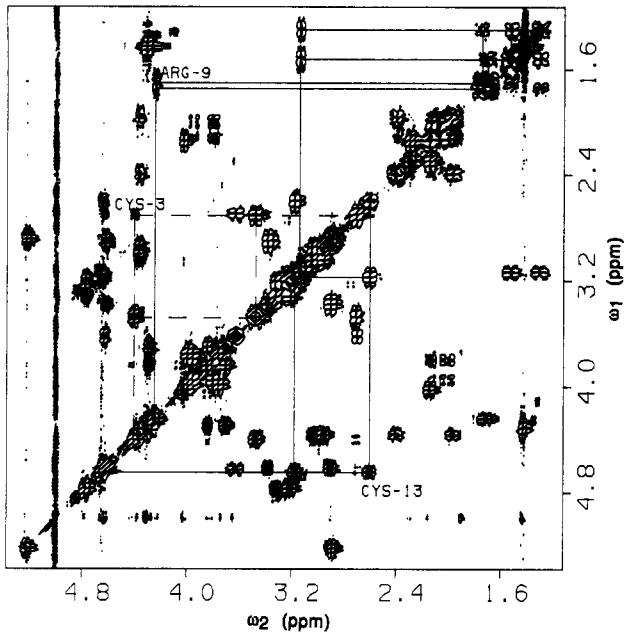


FIGURE 1: Contour plot of a portion of the 400-MHz DQF-COSY spectrum of conotoxin G1 at 5 °C, 7 mM peptide, p^{2}H 4.8, in ^2H_2O. Spin–spin J coupling connectivities are shown for Cys^3, Cys^{13}, and Arg^9. Positive and negative contours are plotted in the figure. This spectrum was acquired with sweep widths of 4200 Hz in both dimensions with the carrier set to the water frequency. The spectrum was acquired with 1024 complex points in t_2, 450 complex FIDs in t_1, and 32 transients for each FID.

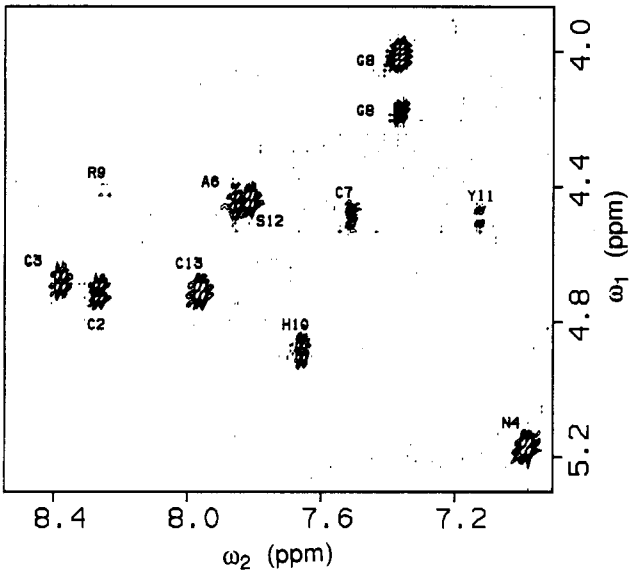


FIGURE 2: Contour plot of the C$^\alpha$H–NH fingerprint region of a COSY spectrum of conotoxin G1 in H_2O. Conditions and data acquisition parameters are similar to those given in Figure 1. The C$^\alpha$H–NH cross-peaks are labeled with standard single-letter amino acid abbreviations.

identified spin systems from this COSY spectrum; however, for C$^\alpha$ protons with degenerate chemical shifts, the RELAYED-COSY spectrum (data not shown) was needed to resolve any ambiguities.

Once protons were grouped into individual amino acid residues, a NOESY (Jeener et al., 1979; Macura & Ernst, 1980) spectrum in H_2O was used to "walk" from one amino acid residue to its neighboring residues. Figure 3 shows the "fingerprint" region of the NOESY spectrum in H_2O of conotoxin G1 with the $d_{\alpha N}$ connectivities illustrated. The $d_{\alpha N}$ connectivities were started from Glu^1 since its C$^\alpha$ proton had been previously identified. A pattern of Glu-AMX-AMX-

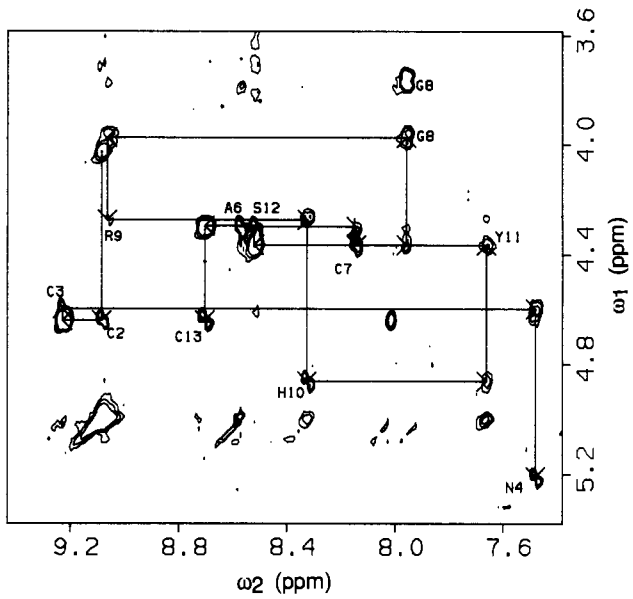
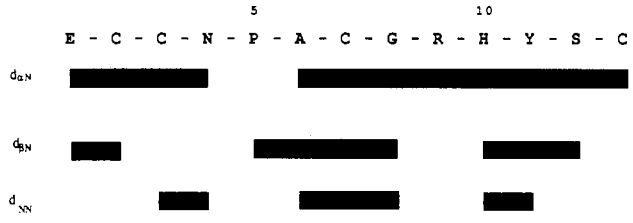


FIGURE 3: Contour plot of the fingerprint region of a 2D NOE spectrum of conotoxin G1 in H_2O. The spectrum was acquired with a mixing time of 300 ms under conditions similar to those given in Figure 1 except that 64 transients were collected for each FID. The $d_{\alpha N}$ connectivities used to make sequential resonance assignment are illustrated. The cross-peaks corresponding to the intraresidue C$^\alpha$H–NH NOEs are labeled with their single-letter amino acid abbreviations.

Table I: Sequential Connectivity Diagram for Conotoxin G1



AMX $d_{\alpha N}$ connectivities was observed in these spectra corresponding to amino acid residues Glu^1-Cys^2-Cys^3-Asn^4. The $d_{\alpha N}$ connectivities end at this point because Pro^5 does not have an amide proton. Since the Ala^6 residue had been previously identified, the $d_{\alpha N}$ connectivities were initiated from this point and a pattern of Ala-AMX-Gly-Arg-AMX-AMX-AMX-AMX corresponding to amino acid residues Ala^6-Cys^7-Gly^8-Arg^9-His^{10}-Tyr^{11}-Ser^{12}-Cys^{13}. Thus, the complete sequential resonance assignment for conotoxin G1 could be made solely from $d_{\alpha N}$ connectivities. The additional $d_{\beta N}$ and d_{NN} connectivities observed in conotoxin G1 are summarized in Table I and were used to confirm the sequential resonance assignment made from the $d_{\alpha N}$ connectivities. There are no $d_{\alpha N}$, $d_{\beta N}$, or d_{NN} connectivities from residue 4 to 5, but sequential connectivities were observed from C$^\alpha$H on Asn^4 to the C$^\delta$Hs on Pro^5, which again confirms the resonance assignments for these residues.

Assignments for the aromatic side-chain protons of His^{10} and Tyr^{11} were confirmed from the 2D NOE spectrum in ^2H_2O by observation of NOEs from the C$^\beta$ protons to the aromatic ring protons. Conotoxin G1 has an amidated C-terminus, and the exchangeable imide protons were assigned from 2D NOE and COSY spectra in H_2O as were the NH_2 and N$^\epsilon$H side-chain protons of Asn^4 and Arg^9, respectively. The proton chemical shifts for conotoxin G1 at 5 °C, pH 4.8, are listed in Table II.

Input Distance Constraints for the Distance Geometry and Energy Minimization Calculations. Estimates for the interproton distance constraints used in the distance geometry

Table II: Chemical Shifts, δ, for the Assigned Proton Resonances of Conotoxin G1 at 5 °C, pH 4.8[a]

amino acid residue	NH	$C^{\alpha}H$	$C^{\beta}H$	others
Glu-1		4.02	2.15, 2.15	$C^{\gamma}H$ 2.30, 2.30
Cys-2	9.08	4.64	3.62, 2.70	
Cys-3	9.22	4.61	3.38, 2.90	
Asn-4	7.48	5.22	2.89, 2.87	$N^{\delta}H$ 7.92, 7.00
Pro-5		4.36	2.40, 1.96	$C^{\gamma}H$ 2.12, 2.01; C^{δ} 3.95, 3.80
Ala-6	8.57	4.31	1.43	
Cys-7	8.14	4.36	3.48, 2.71	
Gly-8	7.96	3.76, 3.98		
Arg-9	9.06	4.26	1.76, 1.67	$C^{\gamma}H$ 1.52, 1.30; $C^{\delta}H$ 3.13 $N^{\epsilon}H$ 7.23
His-10	8.32	4.81	3.33, 3.21	$C^{\epsilon}H$ 8.62; $C^{\delta}H$ 7.36
Tyr-11	7.66	4.36	3.03, 2.96	$C^{\delta}H$ 6.99, 6.99; $C^{\epsilon}H$ 6.68, 6.68
Ser-12	8.51	4.30	3.83, 3.70	
Cys-13	8.69	4.64	3.18, 2.60	NH^{b} 8.00, 7.28

[a] The chemical shifts have errors of ±0.02 ppm and are relative to the HDO or H_2O signal that resonates at 5.00 ppm under these conditions. Where no numbers are given, these resonances were not observed or not assigned. [b] G1 has an amidated C-terminus.

calculations were made from 2D NOE spectra in H_2O and ^2H_2O taken with 200–300-ms mixing times. Since conotoxin G1 is a relatively small peptide, no NOEs were observed at room temperature, and the sample temperature had to be lowered to 5 °C in order to observe negative NOEs. The effects of spin diffusion (Kalk & Berendsen, 1976) should be negligible for a peptide of this size under the experimental conditions, and so were ignored in the analysis. Volume integrals were measured for individual cross-peaks in the 2D NOE spectra, and distances were estimated by standard methods (Hare et al., 1986). For the spectra in ^2H_2O, the geminal C^{δ} protons of Pro^5 were used as a calibration with a distance of 1.8 Å, and for spectra in H_2O, the largest $d_{\alpha N}$ NOE was used as a calibration with a distance of 2.3 Å. The distances determined by this analysis were estimated to have a precision of better than ±0.6 Å, and the distance constraints used in the distance geometry calculations are listed in Table III. For resolved protons on prochiral centers, no additional terms were added to the upper bounds, but instead the recently described procedure of relaxing chirality constraints in the distance geometry calculations was used to account for the lack of stereospecific assignments (Pardi et al., 1988). For unresolved resonances on prochiral protons, and the C^{δ} and C^{ϵ} ring proton resonances of tyrosine, additional terms (given in Table III) were added to the upper distance bounds to account for the lack of stereospecific assignments of these protons (Wüthrich et al., 1983; Pardi et al., 1988).

The distance constraints for the energy-minimized structures were the same as those described above except for interactions involving the Ala^6 methyl protons, the Tyr^{11} aromatic ring protons, and protons on prochiral centers. For distance constraints involving these protons, the additional terms given in Table III were not added to the upper bounds, but the particular pair of protons that gave the best fit to distance constraints in the distance geometry structure was selected as the constraint for the energy-minimized structures. This leads to a much tighter set of distance constraints for the energy-minimized structures than for the distance geometry structures.

Distance Geometry Structures. Eleven structures were generated with the DSPACE distance geometry program. One structure formed a "pseudo mirror image" structure of the others, as judged by superposition of the structures, and so was ignored in the rest of the data analysis. The rms deviations among the backbone N, C, O, and C^{α} atoms in these structures are given in Table IV (the mirror image structure had a deviation about a factor of 2 larger than any of the others). Refinement for each structure was judged complete when there were few or no distance violations greater than 0.05 Å, and

Table III: Input Distance Constraints (in Å) for the Distance Geometry Calculations on Conotoxin G1[a]

Glu-1			
$C^{\delta}H_2$	Cys-2	NH	2.7–5.7 b
Cys-2			
NH	Cys-2	$C^{\beta 1}H$	2.7–3.9
NH	Cys-2	$C^{\beta 2}H$	2.8–4.0
NH	Ser-12	NH	2.7–3.9
$C^{\beta 1}H$	His-10	$C^{\delta}H$	3.2–4.4
$C^{\beta 2}H$	Ser-12	NH	2.9–4.1
Cys-3			
NH	Cys-3	$C^{\beta 1}H$	2.3–3.5
NH	Cys-3	$C^{\beta 2}H$	2.7–3.9
NH	Asn-4	NH	2.3–3.5
$C^{\alpha}H$	Asn-4	NH	2.3–3.5
$C^{\alpha}H$	Tyr-11	$C^{\delta 1,2}H$	2.9–8.1 r
$C^{\alpha}H$	Tyr-11	$C^{\epsilon 1,2}H$	2.6–7.8 r
$C^{\alpha}H$	Cys-13	$C^{\beta 1}H$	2.2–3.4
$C^{\alpha}H$	Cys-13	$C^{\beta 2}H$	2.7–3.9
$C^{\beta 1}$	Cys-13	$C^{\alpha}H$	2.2–3.4
Asn-4			
NH	Asn-4	$C^{\beta 2}H$	2.4–3.6
$C^{\alpha}H$	Pro-5	$C^{\delta 1}H$	2.0–3.2
$C^{\alpha}H$	Pro-5	$C^{\delta 2}H$	2.0–2.9
$C^{\alpha}H$	Tyr-11	$C^{\epsilon 1,2}H$	3.1–8.3 r
$C^{\beta 2}H$	Pro-5	$C^{\delta 1}H$	2.2–3.4
$C^{\beta 2}H$	Pro-5	$C^{\delta 2}H$	2.5–3.7
Pro-5			
$C^{\beta 1}H$	Ala-6	NH	2.8–4.0
$C^{\delta 1}H$	Ala-6	NH	2.6–3.8
Ala-6			
NH	Cys-7	NH	2.1–3.3
$C^{\beta}H_3$	Cys-7	NH	2.7–4.9 m
Cys-7			
NH	Cys-7	$C^{\beta 1}H$	2.9–4.1
NH	Cys-7	$C^{\beta 2}H$	2.3–3.5
NH	Gly-8	NH	2.4–3.6
$C^{\alpha}H$	Gly-8	NH	2.7–3.9
$C^{\beta 2}H$	Tyr-11	$C^{\delta 1,2}H$	2.4–7.6 r
$C^{\beta 2}H$	Gly-8	NH	3.3–4.5
Gly-8			
$C^{\alpha 1}H$	Arg-9	NH	2.8–4.0
Arg-9			
$C^{\alpha}H$	His-10	NH	2.4–3.6
NH	His-10	NH	3.0–4.2
His-10			
NH	His-10	$C^{\beta 2}H$	2.5–3.7
NH	Tyr-11	NH	2.3–3.5
$C^{\alpha}H$	Tyr-11	NH	2.3–3.5
$C^{\delta}H$	His-10	$C^{\beta 2}H$	2.5–3.7
$C^{\beta 2}H$	Tyr-11	NH	3.0–4.5
Tyr-11			
$C^{\alpha}H$	Tyr-11	$C^{\delta 1,2}H$	2.2–7.4 r
$C^{\alpha}H$	Tyr-11	$C^{\epsilon 1,2}H$	3.0–8.2 r
$C^{\beta 2}H$	Ser-12	NH	3.0–4.2
$C^{\epsilon 1,2}H$	Cys-13	$C^{\beta 2}H$	3.1–8.3 r
$C^{\epsilon 1,2}H$	Cys-13	$C^{\beta 1}H$	3.1–8.3 r
Ser-12			
NH	Ser-12	$C^{\beta 1}H$	2.7–3.9
NH	Ser-12	$C^{\beta 2}H$	2.9–4.1
$C^{\alpha}H$	Cys-13	NH	2.0–2.9
Cys-13			
NH	Cys-13	$C^{\beta 1}H$	2.4–3.6
NH	Cys-13	$C^{\beta 2}H$	2.5–3.7

[a] Lower and upper distance bounds are given for each constraint. The letters b, m, and r after the upper distance bounds indicate that additional factors of 1.8, 1.0, and 4.0 Å, respectively, were added to the upper distance bands. The additional terms corresponding to the letters b, m, and r account for lack of stereospecific assignments for methylene, methyl, and tyrosine ring protons, respectively. Distances for some of the sequential connectivities given in Table I were not used in the distance geometry algorithm because accurate distances could not be measured due to spectral overlap.

FIGURE 4: Stereoview of the superposition of the backbone atoms for seven distance geometry structures of conotoxin G1. The structures were superimposed as the best fit between the C, N, O, C^{α}, and C^{β} atoms.

Table IV: Statistics for the Distance Geometry Structure of G1[a]

av total residual NMR distance violations (Å)	av no. of distance violations >0.1 Å	av no. of distance violations >0.05 Å	av rms deviation for backbone atoms[b] (Å)	av rms deviation for non-hydrogen atoms (Å)
0.09 ± 0.07	0.3 ± 0.7	3.2 ± 6.0	1.74 ± 0.28	2.73 ± 0.52

[a] Values are listed with their standard deviations. [b] The rms deviation was calculated by using the C, N, O, and C^{α} atoms.

Table IV also summarizes the extent of refinement for the G1 structures. Figure 4 shows a superposition of the peptide backbone of seven distance geometry structures for conotoxin G1.

Energy Minimization of the Distance Geometry Structures. Various strategies were used to refine the 10 distance geometry structures by energy minimization. Unconstrained minimization leads to low-energy structures with the final AMBER total energies ranging from −136 to −110 kcal/mol. However, these energy-minimized structures differed significantly from the initial distance geometry structures and, more importantly, were less consistent with the NMR data. The average rms deviation between the calculated and experimental distances was 0.24 Å for the energy-minimized structures vs 0.18 Å for the distance geometry structures (note that these calculations were made by using the tighter set of distance constraints discussed above). In contrast, constrained energy minimization gave structures which were consistent with the NMR data but which had higher final energies. Thus, after constrained energy minimization with a force constant $K = 10$ kcal/(mol·Å^2), the average deviation from the experimental proton–proton distances was 0.04 Å, but the energies were −104 to −40 kcal/mol, or well above those obtained from unconstrained minimizations.

When the structures obtained from constrained minimization were subjected to a final unconstrained energy minimization, there was a dramatic decrease in energy ($E = -152$ to −110 kcal/mol), and the resultant structures were consistent with the NMR data (average rms deviation from experimental distances was 0.14 Å). In particular, the lowest energy structure obtained in this fashion is 16 kcal/mol more stable than the lowest energy structure obtained through unconstrained minimization.

The results obtained for all 10 structures through this procedure are shown in Table V. As can be seen from this table, the final structures are generally consistent with the NMR data. The total violation is much greater than those observed in the distance geometry structures because of much tighter constraints for the energy-minimized structures. The observed violations arise mainly from the ring protons of His^{10} and Tyr^{11}. The position of the side-chain atoms differs somewhat from that of the distance geometry structures (average rms deviation 1.62 Å for the non-hydrogen atoms), reflecting movement to energetically more favored positions. In contrast, the position of the corresponding backbone atoms shows a smaller rms deviation of 0.98 Å, indicating that the NMR constraints uniquely define the basic backbone folding of conotoxin G1. Indeed, as with the distance geometry structures, all 10 energy-minimized structures possess a very similar backbone folding pattern, and the average rms deviation for the backbone among all 10 structures is 1.53 Å.

Discussion

Two-dimensional NMR spectroscopy has been used to make essentially complete resonance assignments of the proton spectrum of α-conotoxin G1. Analysis of the chemical shifts in Table III gives the first indication that conotoxin G1 does not have a "random coil" conformation in solution. The NH chemical shifts in G1 range from 9.22 to 7.66 ppm which is much larger than the range of shifts observed in random-coil peptides (Bundi & Wüthrich, 1979). A similar trend is seen for the C^{α} proton chemical shifts. Another indication that conotoxin G1 forms a non-random-coil structure in solution can be seen in the chemical shifts of prochiral methylene protons in the peptide. There are 17 methylene carbon centers in conotoxin G1, and for 14 of these, the 2 prochiral methylene proton resonances have resolvable chemical shifts indicating nonequivalent chemical environments (see Table III). The

Table V: Relative Energies and Extent of Refinement of the Constrained Energy-Minimized Structures

structure	rel energy[a] (kcal/mol)	total residual NMR distance violations (Å)	total no. of distance violations >0.5 Å	av rms deviation for backbone atoms[b] (Å)	av rms deviation for all non-hydrogen atoms (Å)
A	44.3	3.29	1	0.87	1.37
B	26.2	3.02	1	0.79	1.22
C	0.0	4.36	2	1.03	1.80
D	33.4	6.64	3	1.02	1.73
E	24.2	6.39	5	1.20	1.90
F	8.8	6.99	4	0.96	1.69
G	33.1	1.16	0	0.57	1.35
H	45.5	1.99	2	0.87	1.48
I	20.7	3.20	1	1.20	1.74
J	27.6	8.81	5	1.27	1.91

[a] Total energy from AMBER, relative to that of the lowest energy structure C. The absolute energy of C was −152 kcal/mol. [b] The rms deviation was calculated by using the C, N, O, and C^{α} atoms.

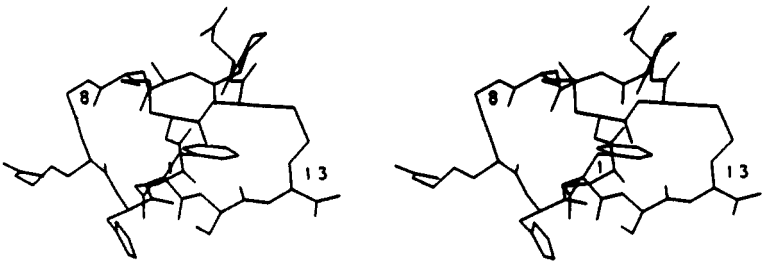


FIGURE 5: Stereoview of C, which is the lowest energy structure of the 10 constrained energy-minimized structures of conotoxin G1. All non-hydrogen atoms are shown. The coordinates of the energy-minimized structures of G1 are available from the authors upon request.

three exceptions are the C^{α} proton and the C^{γ} proton resonances on Glu^1 and the C^{δ} proton resonances on Arg^9. These protons are in regions (the N-terminal residue and the end of a long side chain) that might be expected to be some of the more flexible parts of the molecule.

Structures of Conotoxin G1 Produced from NMR Distance Data. Kobayashi et al. (1985) have reported preliminary results on the structures of conotoxin G1 produced from 2D NMR data. However, to our knowledge, the resonance assignments and the NMR-derived distances used in the structure calculations have not been reported, thereby making it impossible to directly compare their results with ours. Kobayashi et al. show a superposition of the backbone atoms for their 10 structures [Figure 2 in Kobayashi et al. (1985)] which can be compared with our structures in Figure 4. It is clear the Kobayashi's structures are much less well-defined and have much larger rms deviations than those reported here. Unfortunately, in the absence of detailed data, it is impossible to say why their structures are less well-defined, and so no further comparisons will be made to their structures.

Intramolecular hydrogen-bonding interactions are important for stabilizing secondary and tertiary structures in proteins (Richardson, 1981; Creighton, 1985). Since NMR distance data come mainly from proton–proton internuclear distances, it is not possible to directly observe hydrogen-bonding partners in ^1H NMR experiments. However, hydrogen-bonding partners can often be deduced from proximity, geometric, and energetic considerations. The energy-minimized structures were analyzed for potential hydrogen-bonding interactions in order to identify tight turns in the molecule. All 10 structures show 2 reverse turns, the first between Asn^4 and Cys^7 and the second between Gly^8 and Tyr^{11}. In most of the structures, the first turn is a regular β-turn, and the second is a β-turn or one or two γ-turns. The evidence for the 1,4 β-turn is that the Asn^4 carbonyl is generally within hydrogen-bonding distance, ≈2 Å, of the Cys^7 amide proton in the 10 energy-minimized structures. The γ-turns seen in our structures were unexpected. γ-Turns are rare but have been reported for a few proteins including thermolysin (Mathews, 1972). Mainly on the basis of potential hydrogen-bonding interactions, some of the energy-minimized structures had two adjacent γ-turns centered on Arg^9. Adjacent γ-turns have been previously proposed in peptide models related to HC toxin, a cyclic tetrapeptide fungal toxin which can adopt conformations locked by three γ-turns (Heitz et al., 1987), and in another toxin tetrapeptide, chlamydocin, which has a proposed conformation with two adjacent γ-turns (Rich et al., 1987).

The positions of the turns in the energy-minimized structures are in general agreement with the model proposed by Gray et al. (1981), though their nature is somewhat different. The lowest energy structure C, shown in Figure 5, illustrates this point. The first turn in this structure is a nonclassical β-turn characterized by Pro^5 ($\phi = -50°$, $\psi = -40°$), while the second turn is composed of two subsequent γ-turns, the first between Cys^7 and Arg^9 (Gly^8 $\phi = 72°$, $\psi = -78°$) and the second between Gly^8 and Tyr^{11} (His^{10} $\phi = 66°$, $\psi = -84°$). Two short hydrogen bonds stabilize the γ-turns: $CO(Cys^7)$–$HN(Arg^9)$ = 1.96 Å and $CO(Arg^9)$–$HN(Tyr^{11})$ = 2.01 Å. The net effect is that the arginyl side chain is pinched off between these two turns. These general features are seen in all 10 energy-minimized structures. It is difficult to say which of the 10 structures is "best" since, as seen in Table V, the structure with the lowest energy (structure C) is not the structure with the smallest distance violations (structure G); therefore, for this level of refinement and analysis, all structures are considered equally viable.

In contrast to the Gray model, none of the energy-minimized structures possess a salt bridge between Glu^1 and Arg^9. The average distance between the arginyl guanidinium group and the glutamate carboxylate is ≈17 Å (ranging from 11.7 to 20.4 Å). In our structures, both groups are on different faces of the molecule, and there are no NOEs observed between these two residues, so there is no experimental evidence for their proximity. Another general feature concerning the conformation of the energy-minimized structures is that Arg^9 is at the tip of a sharp turn. This arrangement is different from the snake toxins which contain an arginine-containing loop, called the "active tip", that is proposed to be critical for bioactivity.

Hider (1985) has proposed a model for G1 that contains 50% α-helical structure based on CD data. The structures described in our study clearly do not contain such a helix. The highly compact structure of G1 with a β-turn, another β-turn or one to two γ-turns, and two disulfide linkages may give rise to a CD spectrum that has features similar to that for an α-helix.

Structure–Activity Relations in α-Conotoxins. One structural requirement for the paralytic properties of curare-like acetylcholine antagonists usually is the presence of two "acetylcholine-like" units about 11 Å apart (Stenlake, 1981). Each unit is composed of a cationic center (quaternary nitrogen) separated by approximately 5 Å from an electronegative group (carbonyl oxygen). Also, many of these antagonists have the two acetylcholine units separated by hydrophobic regions. By analogy, two positive charges are thought to be required for the paralytic activity of conotoxin G1 (Gray et al., 1985). These charges must originate from the Arg^9 guanidyl side chain and the amino-terminus NH_3^+ group. The average distance between the positive charge on Glu^1 and Arg^9 is 15.5 Å (ranging from 13.6 to 17.8 Å) in the energy-minimized structures. This distance is larger than the 11-Å distance found in many rigid acetylcholine antagonists. However, we note there was no NMR distance data connecting these two residues and small rotations of the torsional angles on the Arg^9 side chain can easily move the cationic sites closer to the ideal distance without a large change in energy. In addition, the binding of the toxin to the receptor could facilitate such rotations.

Our lowest energy structure had a two adjacent γ-turn conformation centered on Arg^9. Chemical modification experiments on conotoxin G1 with arginine-specific reagents such as cyclohexane-1,2-dione indicated that the modified peptide had very reduced affinity for the acetylcholine receptor, pointing to the importance of Arg^9 for activity (Florance et al., 1986). Similar results have been found with the larger snake toxins (Yang et al., 1974), and so a cationic residue in this region of the peptide would appear to be important for the bioactivity of this class of toxins.

Conclusions

Two-dimensional NMR data have been used to generate solution structures for the α-conotoxin G1. Initial structures were generated with a distance geometry algorithm and were then refined with a constrained energy minimization program. The results indicate that the backbone conformation of G1 is well-defined by the NMR data (with an rms deviation of <1.0 Å) but the side-chain conformations are generally less well-defined. The peptide backbone contains two tight turns centered on residues Pro^5 and Arg^9, and the overall folding is quite similar to the model for G1 proposed by Gray et al. (1985). Gray's model consisted of two β-turns whereas the NMR structures are generally more consistent with a β-turn for Pro^5 and one or two γ-turns centered on Arg^9. The NMR data contain no evidence for an α-helix in the G1 structure. It is difficult to show that a molecule exists in a single stable conformation from NMR data; however, Hider (1985) has shown that the circular dichroism spectrum of G1 is constant over a pH range of 1–9 and also does not change in the presence of trifluoroethanol. These data are consitent with G1 existing in a single conformation in solution.

Structures of rigid antagonists of the nicotinic acetylcholine receptor, such as curare, possess two positive charges separated by approximately 11 Å. Although the distance between the positive charges in the NMR structures of G1 was generally larger than this, it appears that a separation of 11 Å is not inconsistent with the NMR data. To help probe the structure–activity relations in G1, additional NMR data and computational methods are being used to obtain better defined structures of G1 (especially for the side-chain conformations). High-resolution COSY-type experiments are being used to measure $^3J_{HN\alpha}$ and $^3J_{H\alpha\beta}$ spin–spin coupling constants to help restrict the conformation of the ϕ and χ torsion angles, respectively. Constrained molecular dynamics simulations, in both the absence and the presence of solvent, have been performed on G1, and preliminary results indicate that these structures are very similar to those presented here (K. S. Kim and A. Galdes, unpublished results). NMR structural studies are also being performed on the homologous 15 amino acid residue α-conotoxin M1. This combination of further refinements of the G1 structure and structures of the homologous conotoxin M1 along with results from modification and activity studies on the α-conotoxins should lead to better understanding of the possible structure–activity relations in these neurologically active peptide toxins.

Acknowledgments

We thank Elizabeth Cheng-Kwock for help in data analysis and Dr. D. H. Hare for providing us with copies of his FTNMR and DSPACE software and for helpful discussions.

References

Abu Khaled, M., Urry, D. W., & Okamoto, K. (1976) *Biochem. Biophys. Res. Commun. 72*, 162.

Bach, A. C., II, Selsted, M. E., & Pardi, A. (1987) *Biochemistry 26*, 4389.

Bassolino, D. A., Hirata, F., Kitchen, D. B., Kominos, D., Pardi, A., & Levy, R. M. (1988) *Int. J. Supercomput. App. 2*, 41.

Berendson, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981) in *Intermolecular Forces* (Pullman, B., Ed.) Reidel, Dortrecht, Holland.

Braun, W., & Go, N. (1985) *J. Mol. Biol. 186*, 611.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521.

Bruenger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801.

Bundi, A., & Wuethrich, K. (1979) *Biopolymers 18*, 285.

Clore, G. M., Bruenger, A. T., Karplus, M., & Gronenborn, A. M. (1985) *J. Mol. Biol. 186*, 435.

Clore, G. M., Bruenger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *J. Mol. Biol. 191*, 523.

Crippen, G. M. (1977) *J. Comp. Physiol. 24*, 96.

Cruz, L. J., Gray, W. R., Yoshikama, D., & Olivera, B. M. (1985) *J. Toxicol., Toxin Rev. 4*, 107.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820.

Florance, J., Galdes, A., Konteatis, Z., Gold, B., & Lenz, G. (1986) *Int. Symp. HPLC Pept. Proteins Oligonucleotides 739*, 34.

Gray, W. R., Luque, A., & Olivera, B. M. (1981) *J. Biol. Chem 256*, 4734.

Gray, W. R., Luque, F. A., Galyean, R., Shepard, R. C., Stone, B. L., Reyes, A., Alford, J., McIntosh, M., Olivera, B. M., Cruz, L. J., & Rivier, J. (1984) *Biochemistry 23*, 2796.

Gray, W. R., Middlemas, D. M., Zeikus, R., Olivera, B. M., & Cruz, L. J. (1985) in *Peptides, Structure and Function. Proceedings of the 9th American Peptide Symposium*, p 823, Pierce Chemical Co., Rockford, IL.

Hagler, A. T. (1974) *J. Am. Chem. Soc. 96*, 5319.

Hare, D., Shapiro, L., & Patel, D. J. (1986) *Biochemistry 25*, 7445.

Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 281.

Havel, T. F., Crippen, G. M., & Kuntz, I. D. (1979) *Biopolymers 18*, 73.

Heitz, F., Kaddari, F., Raniriseheno, H., & Lazaro, R. (1987) *Int. J. Pept. Protein Res. 30*, 801.

Hider, R. C. (1985) *FEBS Lett. 184*, 181.

Holak, T. A., Prestegard, J. H., & Forman, J. D. (1987) *Biochemistry 26*, 4652.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546.

Kalk, A., & Berendsen, H. J. C. (1976) *J. Magn. Reson. 24*, 343.

Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & Van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179.

Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Nishiuchi, Y., & Sakakibara, S. (1984) in *Peptide Chemistry 1983*, p 163, Protein Research Foundation, Osaka, Japan.

Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Nishiuchi, Y., Sakakibara, S., Braun, W., & Go, N. (1987) in *Peptide Chemistry 1986*, p 127, Protein Research Foundation, Osaka, Japan.

Kohn, A. J. (1958) *Hawaii Med. J. 17*, 528.

Kuntz, I. D., Crippen, G. M., & Kollman, P. A. (1979) *Biopolymers 18*, 939.

Low, B. (1979) *Handb. Exp. Pharmacol. 52*, 57.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95.

Matthews, B. W. (1972) *Macromolecules 5*, 818.
McManus, O. B., Musick, J. R., & Gonzalez, C. (1981) *Neurosci. Lett. 24*, 57.
Nerdal, W., Hare, D. R., & Reid, B. R. (1988) *J. Mol. Biol. 201*, 717.
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.
Pardi, A., Hare, D. R., Selsted, M. E., Morrison, R. D., Bassolino, D. A., & Bach, A. C., II (1988) *J. Mol. Biol. 201*, 625.
Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800.
Rich, D. H., Kawai, M., & Jasensky, R. D. (1983) *Int. J. Pept. Protein Res. 21*, 35.
Richardson, J. (1981) *Adv. Protein Chem. 34*, 167.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169.
Stenlake, J. B. (1981) in *Burger's Medicinal Chemistry* (Wolff, M. B., Ed.) p 431, Wiley, New York.
Van Gunsteren, W. F., & Berendson, H. J. C. (1977) *Mol. Phys. 34*, 1311.
Weiner, P., & Kollman, P. A. (1981) *J. Comput. Chem. 2*, 287.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, M. C., Alagona, G., Profeta, S., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715.
Yang, C. C., Chang, C. C., & Liou, I. F. (1974) *Biochim. Biophys. Acta 365*, 1.

Characterization of Lens α-Crystallin Tryptophan Microenvironments by Room Temperature Phosphorescence Spectroscopy†

Jeffrey W. Berger and Jane M. Vanderkooi*

Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received September 23, 1988; Revised Manuscript Received March 14, 1989

ABSTRACT: Room temperature phosphorescence techniques were used to study the structural and dynamic features of the tryptophan residues in bovine α-crystallin. Upon excitation at 290 nm, the characteristic signature of tryptophan phosphorescence was observed with an emission maximum at 442 ± 2 nm. The phosphorescence intensity decay was biphasic with lifetimes of 5.4 ms (71%) and 42 ms (29%). Phosphorescence quenching measurements strongly suggest that each component corresponds to one class of tryptophans with the more buried residues having the longer emission lifetime. Three small-molecule quenchers were surveyed, and in order of increasing quenching efficiency: iodide < nitrite < acrylamide. A heavy-atom effect was observed in iodide solutions, and an upper limit of 5% was placed on the quantum yield of triplet formation in iodide-free solutions, while the phosphorescence quantum yield was estimated to be $\sim 3.2 \times 10^{-4}$. The temperature dependence of the phosphorescence lifetime was measured between 5 and 40 °C. Arrhenius plots exhibited discontinuities at 26 and 29 °C for the short- and long-lived components, respectively, corresponding to abrupt transitions in segmental flexibility. Denaturation studies revealed conformational transitions between 1 and 2 M guanidine hydrochloride, and 4 and 6 M urea. Long-lived phosphorescence lifetimes of 3 and 7 ms were measured in 6 M guanidine hydrochloride and 8 M urea, respectively, suggesting that some structural features are preserved even at very high concentrations of denaturant. Our studies demonstrate the sensitivity of room temperature phosphorescence spectroscopy to the structure of α-crystallin, and the applicability of this technique for monitoring conformational changes in lens crystallin proteins.

Eye lens fibers contain a highly concentrated solution of lens-specific crystallin proteins, and the short-range ordering of the crystallins is essential for the maintenance of lens transparency and for proper focusing of light on the retina (Tardieu & Delaye, 1988). Cataracts are characterized by aggregation and insolubilization of crystallin proteins, and it is believed that protein conformational changes precede the onset of opacification. A variety of biochemical and spectroscopic techniques have been used to probe the structure of lens crystallins, and to monitor the conformational changes associated with aging (Lerman & Borkman, 1978; Ozaki et al., 1983; Libondi et al., 1986; Liang, 1987) and cataract (Harding, 1972, 1981; McNamara & Augusteyn, 1984).

α-Crystallin is an aggregate composed of A and B subunits each with a molecular weight of ~20000 (Bloemendal, 1981). α-A_2 and α-B_2 are the primary gene products, while α-A_1 and α-B_1 arise via posttranslational adenosine cyclic 3′,5′-phosphate (cAMP)[1]-dependent phosphorylation (Chiesa et al., 1988).

† This work was supported by an NIH Cell and Molecular Biology training grant (J.W.B.) and by NIH Grant GM 21487 (J.M.V.).
* To whom correspondence should be addressed.

[1] Abbreviations: cAMP, adenosine cyclic 3′,5′-phosphate; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; Gdn-HCl, guanidine hydrochloride.

0006-2960/89/0428-5501$01.50/0