Three-Dimensional Structure of Phoratoxin in Solution: Combined Use of Nuclear Magnetic Resonance, Distance Geometry, and Restrained Molecular Dynamics†

G. Marius Clore,* Dinesh K. Sukumaran, Michael Nilges, and Angela M. Gronenborn

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG

Received September 5, 1986; Revised Manuscript Received November 10, 1986

ABSTRACT: The solution conformation of phoratoxin, a 46-residue plant protein, has been investigated by ^{1}H nuclear magnetic resonance (NMR) spectroscopy. The spectrum is assigned in a sequential manner by a combination of two-dimensional NMR techniques to demonstrate through-bond and through-space (<5 Å) connectivities. A set of 331 approximate interproton distance restraints and six ϕ backbone torsion angle restraints is derived from the two-dimensional nuclear Overhauser enhancement and double quantum filtered homonuclear correlated spectra, respectively. These restraints are used as the basis of a structure determination with a metric matrix distance geometry algorithm. A total of eight structures are computed in this manner and subjected to refinement by restrained molecular dynamics in which the experimental restraints are incorporated into the total energy function of the system in the form of square well effective potentials. The average atomic root mean square difference between the final eight structures and the mean structure obtained by averaging their coordinates is 1.7 ± 0.5 Å for the backbone atoms and 2.1 ± 0.5 Å for all atoms. The overall shape of phoratoxin is that of the capital letter L, similar to that of crambin and α_1-purothionin, with the longer arm comprising two α-helices at an angle of ~140° to each other and the shorter arm a mini-antiparallel β-sheet and a loop made up of two turns and a strand.

Phoratoxin is a member of a group of low molecular weight water-soluble protein toxins that are ubiquitous throughout the plant kingdom and include the purothionins and viscotoxins (Samuelsson & Ekblad, 1967; Samuelsson et al., 1968; Samuelsson & Patterson, 1971; Killstrand & Samuelsson, 1973; Mak & Jones, 1978). Pharmacological experiments have shown that these toxins are haemolytic, lyse a wide variety of mammalian cells, cause vascoconstriction of vessels in skin and skeletal muscle, produce hypotension and bradydycardia, and have a negative inotropic effect on heart muscle (Rosell & Samuelsson; Anderson & Johannson, 1973; Okada & Yoshizumi, 1973; Carraco et al., 1982).

The amino acid sequence homology between these toxins ranges from 30% to 50%. In addition, they display 30–50% homology with respect to crambin, a hydrophobic, water-insoluble plant protein of no yet known biological activity (Van Etten et al., 1965) whose crystal structure has been solved to very high resolution (Hendrickson & Teeter, 1981).

Recently, we presented a ^{1}H NMR[1] study on α_1-purothionin (Clore et al., 1987) and derived a set of 310 approximate interproton distance restraints from NOE measurements, which were subsequently used to determine its three-dimensional structure by the combined use of distance geometry and restrained molecular dynamics calculations (Clore et al., 1986a). The three-dimensional structure of α_1-purothionin is indeed similar to that of crambin, in agreement with the observed 33% amino acid sequence homology. Phoratoxin displays 39% and 45% sequence homology with respect to crambin and α_1-purothionin, respectively (see Figure 1). One would therefore expect the three-dimensional structure of phoratoxin to be similar to that of crambin and α_1-purothionin.

To test this hypothesis, we have carried out a ^{1}H NMR study on the solution structure of phoratoxin. First, the ^{1}H NMR spectrum is assigned in a sequential manner by a combination of two-dimensional NMR techniques to demonstrate through-bond and through-space (<5 Å) connectivities (Wüthrich et al., 1982). A set of 331 approximate interproton distance restraints and six ϕ backbone torsion angle restraints is then derived from the NOESY and DQF-COSY spectra, respectively. This is followed by metric matrix distance geometry calculations (Crippen & Havel, 1978; Havel et al., 1984; Havel & Wüthrich, 1984, 1985; Sippl & Scheraga, 1986) to generate a set of structures that are consistent with the experimental restraints. Finally, these structures are refined by restrained molecular dynamics calculations (Brünger et al., 1986; Clore et al., 1985, 1986a,b; Kaptein et al., 1985; Nilsson et al., 1986).

Experimental Procedures

Sample. Phoratoxin was a gift of Prof. G. Samuelsson and was purified from the mistletoe *Phorandondron tomentosum* subsp. *macrophyllum* as described previously (Mellstrand & Samuelsson, 1973). This material is in fact a mixture of 85% phoratoxin A and 15% phoratoxin B (G. Samuelsson, personal communication). These two proteins differ in only a single amino acid residue at position 25 with phoratoxin A having a valine and phoratoxin B an isoleucine at this position (Mellstrand & Samuelsson, 1974; Thunberg, 1983).

The samples for NMR contained 8.6 mM phoratoxin in either 90% H_2O/10% D_2O or 99.96% D_2O at pH 3.1. All experiments were carried out at 25 °C.

NMR Spectroscopy. NMR spectra were recorded on a Bruker AM500 spectrometer equipped with digital phase

†This work was supported by the Max-Planck Gesellschaft and Grant Cl 86/1-1 of the Deutsche Forschungsgemeinschaft (G.M.C. and A.M.G.).

[1] Abbreviations: NMR, nuclear magnetic resonance spectroscopy; NOE, nuclear Overhauser effect or enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; DQF-COSY, double quantum filtered homonuclear correlated spectroscopy; 1 cal = 4.183 J; rms, root mean square.

0006-2960/87/0426-1732$01.50/0

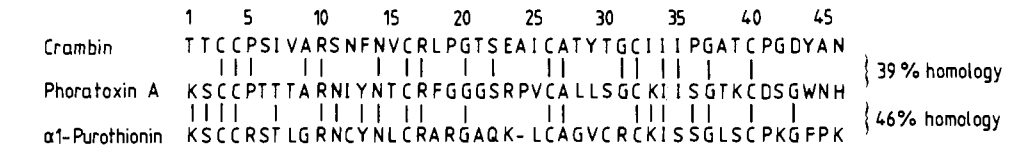

FIGURE 1: Comparison of the amino acid sequences of phoratoxin A (Mellstrand & Samuelson, 1974), α_1-purothionin (Jones & Mak, 1977), and crambin (Teeter et al., 1981). Phoratoxin B differs from phoratoxin A in having an isoleucine instead of a valine at position 25 (Thunberg, 1983). The numbering of the residues is that of phoratoxin and crambin, and the alignment with α_1-purothionin, which has one residue less, is that which gives maximum homology (Teeter et al., 1981).

shifters and an ASPECT 3000 computer. All two-dimensional spectra were recorded in the pure-phase absorption mode with the time proportional phase incrementation method (Redfield & Kuntz, 1975, Bodenhausen et al., 1980) as described by Marion and Wüthrich (1983). The following spectra were recorded in D_2O and H_2O: NOESY (Jeener et al., 1979; Macura et al., 1981), DQF-COSY (Rance et al., 1983), and MLEV17 HOHAHA (Davis & Bax, 1985; Bax & Davis, 1985; Braunschweiler & Ernst, 1985) spectra. NOESY spectra were recorded at mixing times of 100, 150, and 200 ms. HOHAHA spectra were recorded at several mixing times ranging from 15 to 70 ms in order to demonstrate successively direct, single, and multiple relayed through-bond connectivities (Davis & Bax, 1985; Bax & Davis, 1985). For measurements in H_2O, the H_2O resonance was suppressed by selective irradiation during the relaxation delay and in the case of the NOESY spectra during the mixing time as well (Wider et al., 1984). An additional set of NOESY spectra in H_2O was also recorded without solvent irradiation by replacing the last 90° pulse in the sequence by a semiselective jump-return (90_x–t–90_{-x}) pulse with the carrier placed at the position of the solvent (Plateau & Gueron, 1982). All spectra were recorded with sweep widths of 7042 Hz. The digital resolution employed for the NOESY and HOHAHA spectra was 6.88 Hz/point in both dimensions, and this was achieved by appropriate zero filling in the t_1 dimension only. In the case of the DQF–COSY spectra, the digital resolution was 1.72 Hz/point in the t_2 dimension and 6.88 Hz/point in the t_1 dimension. Typically, 64–160 transients were collected for each of 512 increments with a relaxation delay of 1–1.2 s between successive transients.

Calculations. Metric matrix distance geometry calculations were carried out with the program DISGEO (Havel & Wüthrich, 1984, 1985; Havel, 1986). All energy minimization and restrained molecular dynamics calculations were carried out as described by Clore et al. (1986b) and Brünger et al. (1986) on a CRAY-XMP using a CRAY version (A. T. Brünger, unpublished data) of the program CHARMM (Brooks et al., 1983). Analysis of the structures and molecular dynamics trajectories was carried out with a modified version of the function network of FRODO (Jones, 1978) interfaced with CHARMM on an Evans & Sutherland PS330 color graphics system.

Empirical energy potentials for the energy minimization and molecular dynamics calculations were taken from Brooks et al. (1983) and modified in order to treat all hydrogen atoms explicitely (D. States and M. Karplus, unpublished data). Solvent molecules were not included explicitly in the simulations, but the effect of solvent was approximated by multiplying the electrostatic energy term by a $1/r$ screening function (Brooks et al., 1983). The nonbonded interactions were switched off by a cubic switching function, between 6.5 and 7.5 Å, with pairs up to 8 Å included in the nonbonded list. Integration of the equations of motion was performed by a Verlet integrator algorithm (Verlet, 1967) with initial velocities assigned to a Maxwellian distribution at the appropriate temperature. The time step of the integrator was 0.001 ps, and the nonbonded interaction lists were updated every 0.02 ps.

Results and Discussion

Assignment of the ^1H NMR Spectrum. Sequence-specific resonance assignment is the prerequisite for the determination of a protein structure in solution (Wüthrich et al., 1982; Billeter et al., 1982). This was relatively straightforward and was carried out by first identifying amino acid spin systems by means of direct and relayed through-bond connectivities, followed by the sequential assignment of resonances by means of short (<5 Å) through-space connectivities involving the NH, $C^\alpha H$, and $C^\beta H$ protons as well as the $C^\delta H$ protons of proline (Wagner & Wüthrich, 1982; Zuiderweg et al., 1983; Strop et al., 1983; Weber et al., 1985; Kline & Wüthrich, 1985; Clore et al., 1985, 1986c, 1987; Wagner et al., 1986; Zarbock et al., 1986). The former was achieved with the HOHAHA spectra, examples of which are shown in Figures 2 and 3 for the NH-aliphatic and $C^\alpha H$-aliphatic regions, respectively, while the latter was carried out with the NOESY spectra, examples of which are shown in Figures 4 and 5 for the NH–NH and NH-aromatic/NH-aliphatic regions, respectively. A plot of the $C^\alpha H$-aliphatic region of a 150-ms mixing time NOESY spectrum is available as supplementary material (see paragraph at end of paper regarding supplementary material).

One feature of the assignment is noteworthy. This concerns the microheterogeneity at residue 25, which is a valine in phoratoxin A and an isoleucine in phoratoxin B, phoratoxin A and B being present in a ratio of ~5:1 (Mellstrand & Samuelson, 1974; Thunberg, 1983; G. Samuelson, personal communication). The resonances of the valine and isoleucine at this position were easily identified (see the HOHAHA spectra in Figures 2 and 3). The only effect of this microheterogeneity on other residues is to cause a small difference in the chemical shift of the NH protons of residues 26–30 in phoratoxin B relative to phoratoxin A. This is best seen in the NOESY spectrum of the NH–NH region (Figure 4). The observed patterns of NOEs for these residues, however, are identical in the two phoratoxins.

A summary of the observed short range ($|i - j| \leq 5$) NOEs involving the NH, $C^\alpha H$, and $C^\beta H$ protons as well as the $C^\delta H$ protons of proline is given in Figure 6. The complete list of resonance assignments is available as supplementary material.

Secondary Structure. The regular secondary structure elements present in phoratoxin can be identified from the data in Figure 6 (Wüthrich et al., 1984). There are two helices extending from residues 7 to 19 and 23 to 29 that are characterized by a stretch of NH(i)–NH(i + 1), $C^\alpha H(i)$–NH(i + 3), and $C^\alpha H(i)$–$C^\beta H(i + 3)$ NOEs, as well as the presence of slowly exchanging NH protons. There are two small β-strands from residues 2 to 5 and 32 to 35 characterized by a stretch of $C^\alpha H(i)$–NH(i + 1) NOEs and in the case of Pro-5 by a $C^\alpha H(4)$–$C^\delta H(5)$ NOE and apparent values of >9 Hz for the $^3J_{HN\alpha}$ coupling constants indicative of ϕ backbone torsion angles in the range −80° to −180° (Pardi et al., 1984). These two strands form a mini-antiparallel β-sheet characterized by

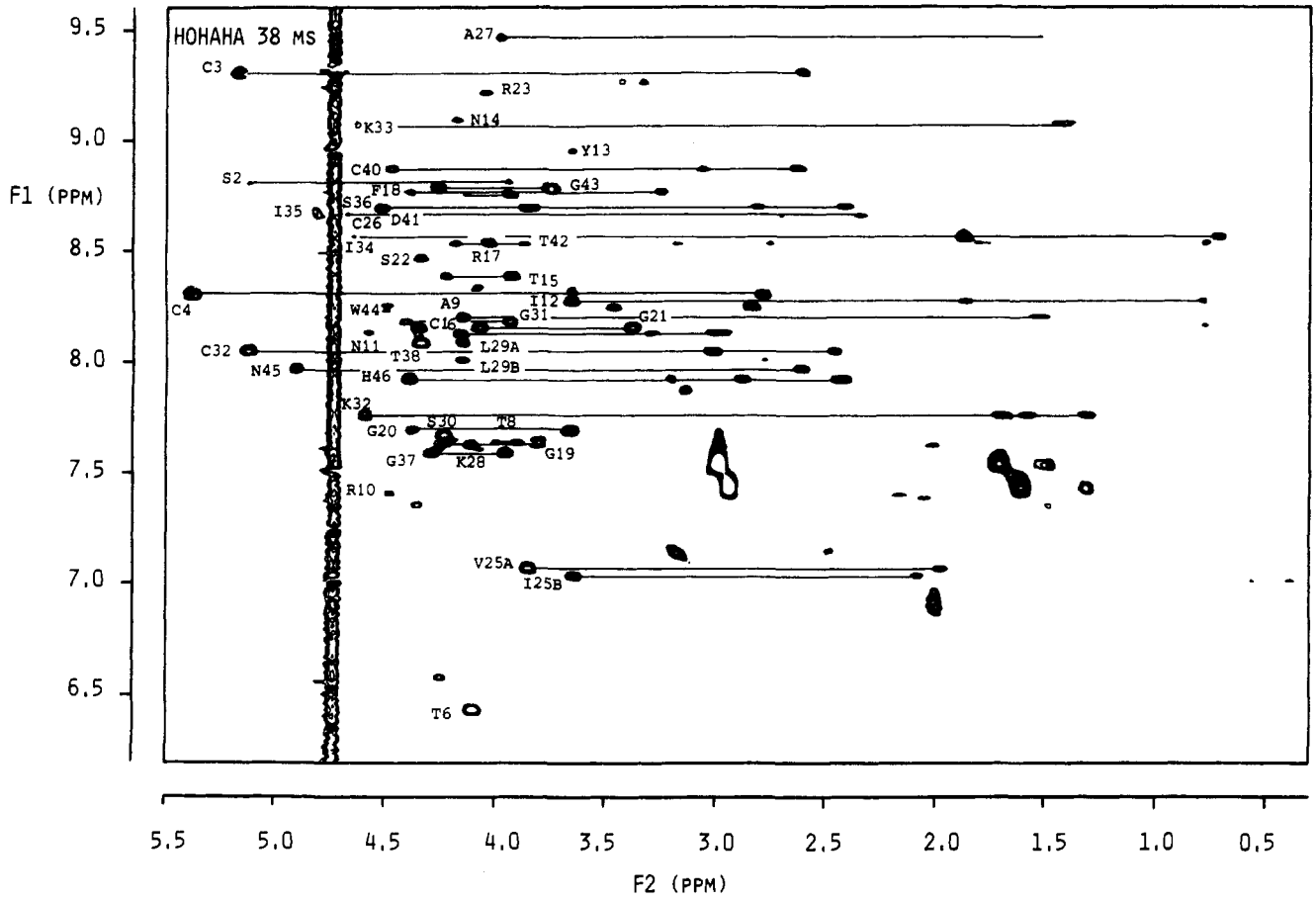

FIGURE 2: NH (F1 axis)–aliphatic (F2 axis) region of the HOHAHA spectrum of phoratoxin in H_2O. Some relayed connectivities are indicated by continuous lines, and the labels are at the positions of the direct NH–C$^{\alpha}$H cross-peaks. The spectrum is unsymmetrized.

interstrand NOEs between the C$^{\alpha}$H protons of Ser-2 and Ile-34 and of Cys-4 and Cys-32 and between the NH protons of Cys-3 and Lys-33, from the CαH protons of Ser-2 and Cys-4 to the NH protons of Ile-35 and Lys-33, respectively, and from the C$^{\alpha}$H proton of Cys-32 to the C$^{\delta}$H protons of Pro-5. This set of interstrand NOEs enables one to assign two interstrand backbone hydrogen bonds between the NH proton of Cys-3 and the carbonyl oxygen atom of Lys-33 and between the NH proton of Lys-33 and the carbonyl oxygen atom of Cys-3. The presence of these two hydrogen bonds is further supported by the observation of slow H → D exchange for the NH protons of Cys-3 and Lys-33. In addition to these regular structure elements, a number of turns can be identified from the short-range NOEs (Wagner et al., 1986). These comprise residues 4–6, 19–22, 30–31, 36–39, and 41–45.

Interproton Distance and Dihedral Restraints. At short mixing times, the intensities of the NOESY cross-peaks are approximately proportional to r^{-6}, where r is the distance between two protons (Wagner & Wüthrich, 1979; Kumar et al., 1981; Dobson et al., 1982; Keepers & James, 1984; Clore & Gronenborn, 1985). We used the NOESY spectra with mixing times of 150 and 200 ms to assign all NOESY cross-peaks and the spectra recorded with mixing times of 100 and 150 ms to assign the NOESY cross-peak intensities. This classification procedure was carried out essentially as described previously (Braun et al., 1983, 1986; Williamson et al., 1985; Kline et al., 1986; Clore et al., 1985, 1986a,b). The NOESY cross-peak intensities were divided into three classes, strong, medium, and weak, corresponding to the distance ranges of 1.8–2.8, 1.8–3.3, and 2.5–5 Å, respectively, used in the tertiary structure computations. The complete list of all the NOEs used in the computations is available as supplementary material. This consists of a total of 331 NOE restraints comprising 135 short-range ($|i - j| \leq 5$) and 60 long-range ($|i - j| > 5$) interresidue and 136 intraresidue interproton distance ranges. The number and type of NOE restraints are similar to those used in the determination of the solution structure of α_1-purothionin (Clore et al., 1986a). The number of interresidue distances, both short and long range, is also similar to that used in our model studies on crambin (Clore et al., 1986b; Brünger et al., 1986), but the number of intraresidue distances used here is significantly larger.

In addition to the NOE restraints, three other groups of restraints were used in the computations. (i) Six ϕ backbone torsion angles restrained to a range of −80 to −180° on the basis of values of $^3J_{HN\alpha} > 9$ Hz (Pardi et al., 1984). These involve Cys-3, Cys-4, Thr-6, Cys-32, Lys-33, and Thr-38. (ii) Four distance restraints corresponding to the two backbone hydrogen bonds between Cys-3 and Lys-33 identified as described in the previous section. For each hydrogen bond the N–O and NH–O distances are restrained to ranges of 2.3–3.3 and 1.3–2.3 Å, respectively. (iii) Nine distance restraints for the three disulfide bridges between Cys-3 and Cys-40, Cys-4 and Cys-32, and Cys-16 and Cys-26. For each disulfide bridge there are three distance restraints, S_i–S_j, S_i–C^{β}_j, and S_j–C^{β}_i, which are constrained to values of 2.02 ± 0.05, 2.99 ± 0.05, and 2.99 ± 0.05 Å, respectively.

Tertiary Structure Computation. The strategy we have employed here for the computation of the tertiary structure of phoratoxin follows the one that we previously used for α_1-purothionin (Clore et al., 1986a). This involves a two stage procedure: (i) the generation of a set of structures that approximately satisfies all experimental and covalent restraints using the DISGEO distance geometry program (Havel &

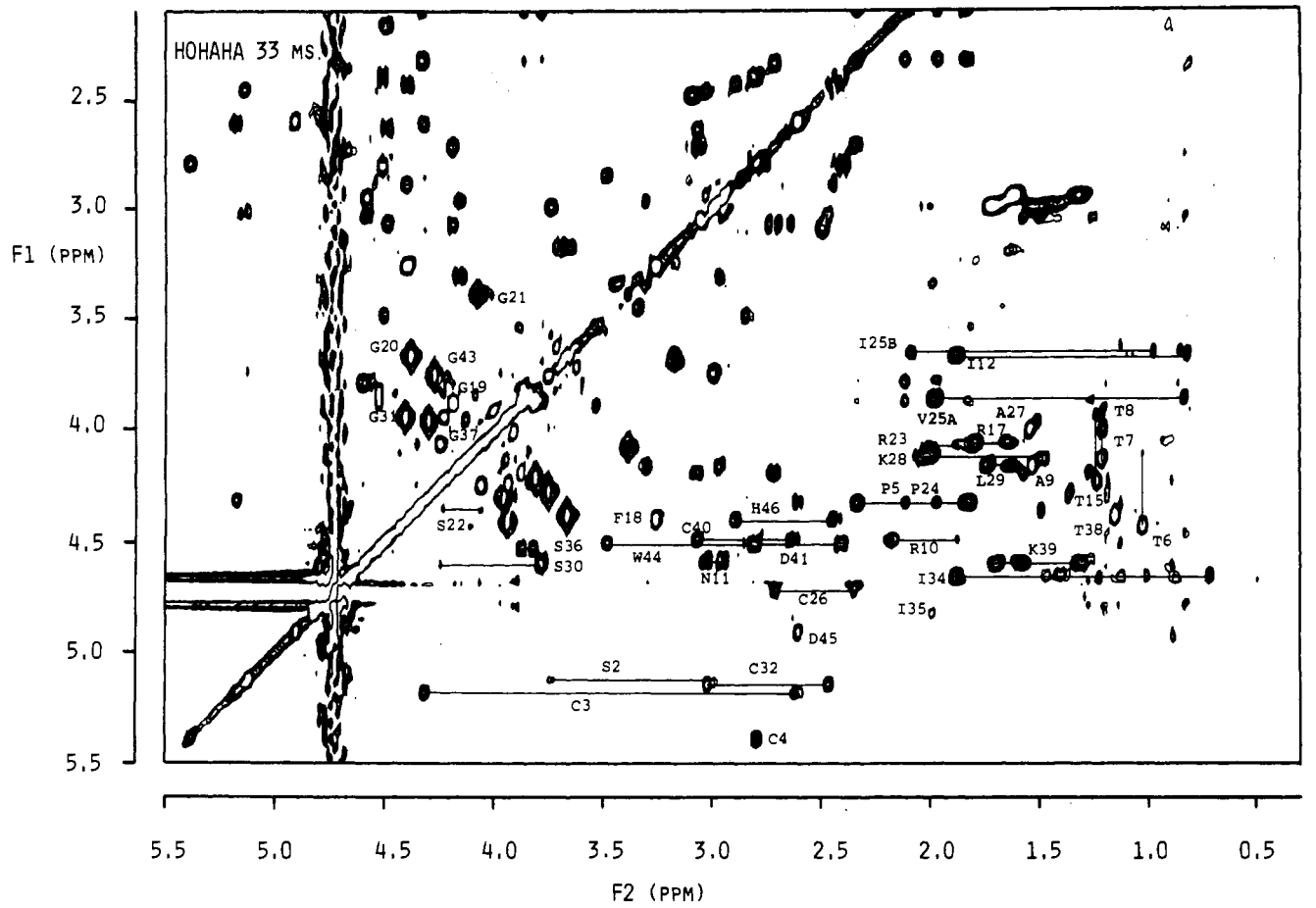

FIGURE 3: A portion of the $C^{\alpha}H$ (F1 axis)–aliphatic (F2 axis) region of the HOHAHA spectrum of phoratoxin in D_2O. Direct and relayed connectivities are present, and some spin networks originating from the $C^{\alpha}H$ protons are indicated by continuous lines. The spectrum is unsymmetrized.

Wüthrich, 1984, 1985; Havel, 1986) based on the metric matrix (Crippen & Havel, 1978; Havel et al., 1984); (ii) the refinement of these structures using a combination of restrained energy minimization and restrained molecular dynamics (Clore et al., 1985, 1986b; Brünger et al., 1986). In our experience this dual approach is the most efficient in terms of computational time although we note that restrained molecular dynamics is also capable of determining the tertiary structure starting from unfolded structures (Brünger et al., 1986; Clore et al., 1986b).

In the distance geometry calculations, the distance ranges for protons for which stereospecific assignments could not be made were appropriately corrected for the pseudoatom representation used by DISGEO as described by Wüthrich et al. (1983). The distance geometry calculations then proceeded in four phases: (i) the determination of a complete set of upper and lower limits for the distances between all atoms in the structure by triangulation from the covalent restraints (viz., bond lengths, bond angles, and planes), the van der Waals restraints (which set a minimum distance between any two nonbonded atoms), and the experimental distance and dihedral angle restraints; (ii) the embedding of a set of substructures comprising the C, C^{α}, N, and $C^{\alpha}H$ backbone atoms and the nonterminal C^{β} and C^{γ} atoms which are consistent with the distance bounds between these atoms; (iii) the computation of a set of structures that approximately fits all the distance data by a procedure known as metrization in n-dimensional distance space (Havel et al., 1984); (iv) restrained least-squares refinement (1500 cycles) in Cartesian coordinate space of these structures with respect to all the distances. To complete the structure determination stage, the pseudoatoms were replaced by real atoms and all hydrogen atoms were built on to generate a set of structures known as DG(i).

The refinement stage involves a combination of restrained energy minimization and restrained molecular dynamics in which the NOE interproton distance and ϕ backbone torsion angle restraints are incorporated into the total energy function of the system in the form of square well effective potentials (Clore et al., 1986a):

$$E_{\text{NOE or }\phi} = \begin{cases} c(x_i - x_i^{u})^2 & \text{if } x_i > x_i^{u} \\ 0 & \text{if } x_i^{l} \leq x_i \leq x_i^{u} \\ c(x_i - x_i^{l})^2 & \text{if } x_i < x_i^{l} \end{cases} \quad (1)$$

where x_i is the calculated value of a particular NOE interproton distance or ϕ backbone torsion angle and x_i^{u} and x_i^{l} are the upper and lower limits of the target ranges of this NOE interproton distance or ϕ backbone torsion angle. The force constant c was set to 40 kcal mol^{-1} Å^{-2} for the NOE restraints and to 40 kcal mol^{-1} rad^{-2} for the ϕ backbone torsion angle restraints. In contrast to the distance geometry calculations, which use pseudoatoms with appropriate distance corrections for protons that could not be stereospecifically assigned, the energy calculations employ a single $(\langle r^{-6} \rangle)^{-1/6}$ mean distance (Clore et al., 1985, 1986a,b; Brünger et al., 1986). The explicit distance restraints used for the two assigned backbone hydrogen bonds and the three disulfide bridges in the distance geometry calculations were not included as effective restraints potentials. This is because the hydrogen-bonding restraints are taken over by the hydrogen-bonding and electrostatic potential terms and the disulfide bridge restraints by the appropriate bond, angle, and dihedral potential terms.

The refinement proceeded in two phases: (i) 500 cycles of

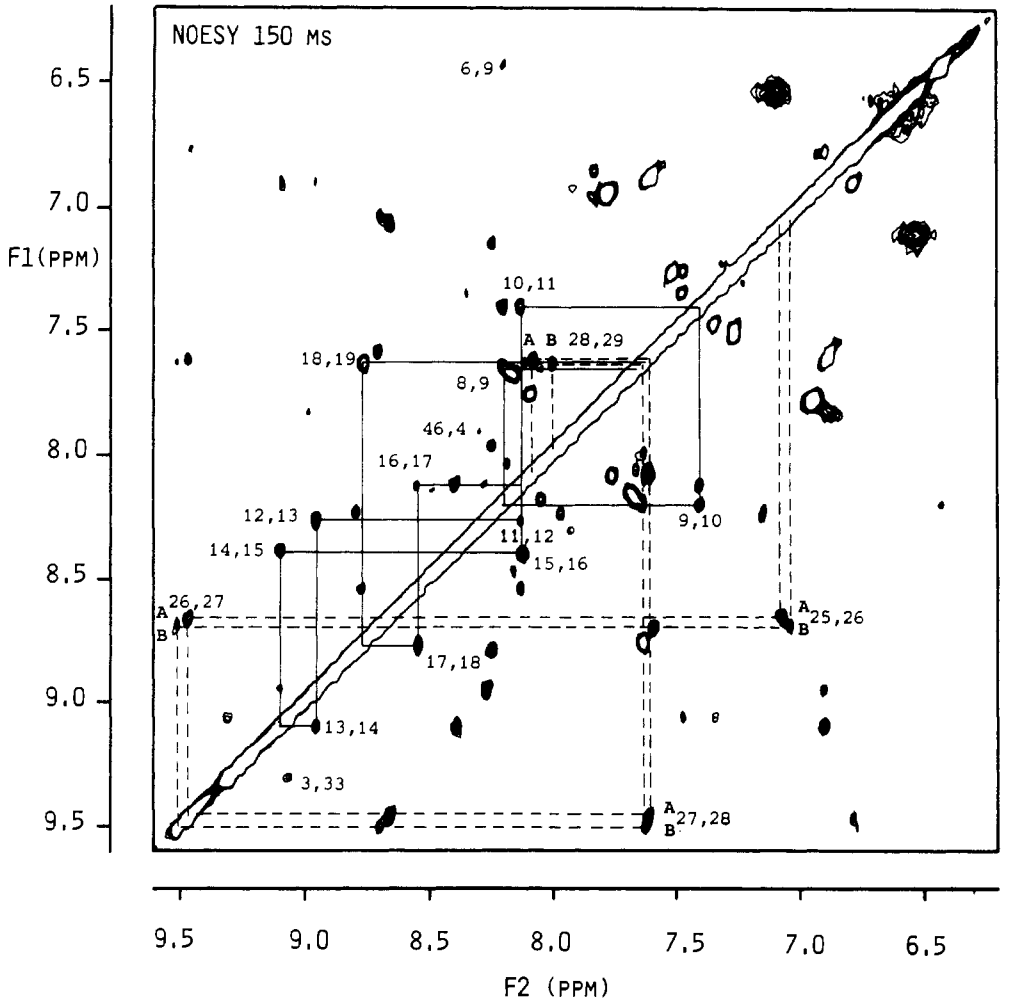

FIGURE 4: NH (F1 axis)–NH (F2 axis) region of the 150-ms NOESY spectrum of phoratoxin in H_2O. The sequence of $d_{NN}(i, i+1)$ connectivities extending from residues 8 to 19 and 25 to 29 are indicated by continuous (—) and dashed (– –) lines, respectively. Note that there are two sets of NH peaks for residues 25–29 corresponding to phoratoxin A and B, which are present in a ratio of approximately 5:1. The spectrum is unsymmetrized.

restrained energy minimization to generate structures DGm(i); (ii) 2 ps of equilibration and thermalization (Brooks et al., 1983) followed by 10 ps of restrained molecular dynamics at 400 K. The final restrained dynamics structures, RDDG(i), were obtained by averaging the coordinates of the last 7 ps of the trajectory, followed by 500 cycles of restrained energy minimization, additionally constrained by weak harmonic constraints (Bruccoleri & Karplus, 1986) to correct for minor distortions in bond lengths and angles introduced by the averaging procedure. (Note that this last restrained energy minimization step results in only very small atomic rms shifts of <0.2 Å).

Three mean structures, $\overline{\text{DG}}$, $\overline{\text{DGm}}$, and $\overline{\text{RDDG}}$, were computed by averaging the coordinates of the individual DG(i), DGm(i), and RDDG(i) structures, respectively. Because these mean structures are poor in stereochemical terms and exhibit very bad nonbonded contacts, they were subjected to a multiple-step restrained energy minimization procedure (Brünger et al., 1986; Clore et al., 1986b) in which the van der Waals radii were slowly increased from one-fourth of their full values in the first 200 cycles, to half of their full values in the next 200 cycles, and finally to their full values in the final 400 cycles. This resulted in structures ($\overline{\text{DG}}$)m, ($\overline{\text{DGm}}$)m, and ($\overline{\text{RDDG}}$)m.

Converged Structures. Eight converged DG structures were computed and subjected to further refinement to yield the various other structures described in the previous section. The results are summarized in Figures 7–11 and Tables I–IV. The best fit superposition of the backbone atoms of the DG, DGm, and RDDG structures is shown in Figure 7, the superposition of the NOE restraints on the backbone atoms of the (RDDG)m structure in Figure 8, and the best fit superposition of the backbone atoms of the mean $\overline{\text{DG}}$, $\overline{\text{DGm}}$, and $\overline{\text{RDDG}}$ structures and of all the atoms (except hydrogens) of the restrained energy minimized mean structures ($\overline{\text{DG}}$)m, ($\overline{\text{DGm}}$)m, and ($\overline{\text{RDDG}}$)m in Figure 9. The atomic rms distributions of the DG, DGm, and RDDG structures about their respective mean structures is shown in Figure 10, as a function of residue number. The atomic rms shifts between each set of DG, DGm, and RDDG structures and the atomic rms differences between the mean structures, $\overline{\text{DG}}$, $\overline{\text{DGm}}$, and $\overline{\text{RDDG}}$, and between the restrained energy minimized mean structures, ($\overline{\text{DG}}$)m, ($\overline{\text{DGm}}$)m, and ($\overline{\text{RDDG}}$)m, plotted as a function of residue number, are available as supplementary material. The variation of the ϕ and ψ backbone torsion angles of the restrained energy minimized mean structures, ($\overline{\text{DG}}$)m, ($\overline{\text{DGm}}$)m, and ($\overline{\text{RDDG}}$)m, as a function of residue number is shown in Figure 11. (Similar plots for all the other structures are available as supplementary material.) The values of the atomic rms distributions and shifts between the various structures are given in Table I, the ϕ and ψ angular rms

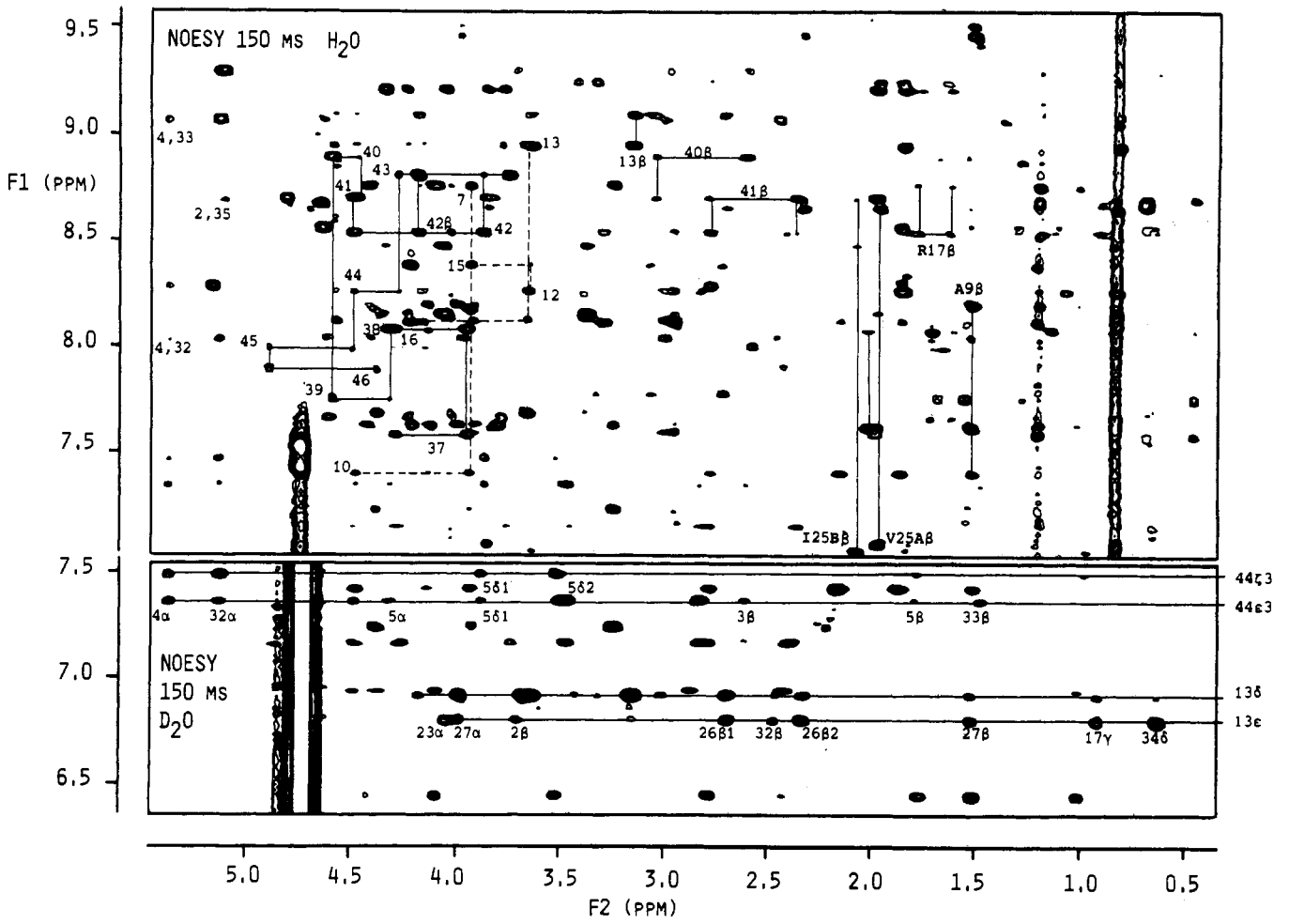

FIGURE 5: NH/aromatic (F1 axis)–aliphatic (F2 axis) region of the 150-ms NOESY spectrum of phoratoxin. The upper spectrum was recorded in H_2O and the lower one, mainly showing aromatic to aliphatic connectivities, in D_2O. Some $d_{\alpha N}(i, i + 1)$ and $d_{\alpha N}(i, i + 3)$ connectivities in the H_2O spectrum are indicated by continuous (—) and dashed (– –) lines, respectively, and the peaks are labeled by residue number at the positions of the NH(i)–C$^{\alpha}$H(i) intraresidue cross-peaks. Also indicated in the H_2O spectrum are some $d_{\beta N}(i, i + 1)$ connectivities with the peaks labeled by residue number followed by the letter β at the position of the NH(i)–C$^{\beta}$H(i) intraresidue cross-peaks. Long-range interresidue NOEs from the C$^{\zeta 3}$H and C$^{\epsilon 3}$H protons of Trp-44 and the C$^{\delta}$H and C$^{\epsilon}$H protons of Tyr-13 are indicated in the D_2O spectrum. The spectra are unsymmetrized, and the spectrum in H_2O was recorded with the jump-return sequence for the last 90° pulse to suppress the water resonance.

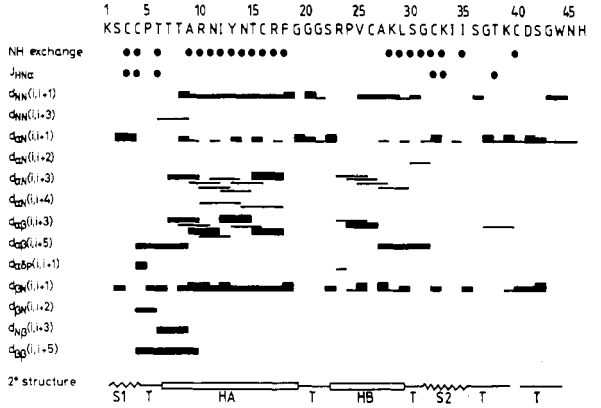

FIGURE 6: Sequence of phoratoxin A together with a summary of the observed short-range ($|i - j| \leq 5$) NOEs involving the NH, C$^{\alpha}$H, and C$^{\beta}$H protons as well as the C$^{\delta}$H protons of the proline residues. The NOEs are classified into strong, medium, and weak by the thickness of the line. NH protons that are still present after 24 h of dissolving the protein in D_2O are indicated by closed circles (●). Apparent values of $^3J_{HN\alpha} > 9$ Hz as measured from the DQF-COSY spectrum are also indicated by closed circles (●). Apart from residue 25, only the chemical shifts of the NH protons of residues 25–29 differ slightly in phoratoxin A and B; for residue 25 (Val in phoratoxin A and Ile in phoratoxin B), the chemical shifts of the NH protons differ slightly (by 0.04 ppm) and the chemical shifts of the other resonances somewhat more (see Table I). The pattern and relative intensities of the NOEs observed for phoratoxin A and B, however, are identical.

differences in Table II, the NOE interproton distance deviations and radii of gyration in Table III, and the energies of the structures in Table IV.

Examination of the above data clearly shows that the distance geometry calculations generate a set of similar structures all of which satisfy the experimental restraints. [Note that the comparatively large values of the NOE interproton distance deviations in Table III and of the NOE restraints energy in Table IV for the DG structures are principally due to the fact that for distances involving protons that could not be stereospecifically assigned these values are calculated with a single $(\langle r^{-6} \rangle)^{-1/6}$ average distance rather than a corrected distance range to a pseudoatom.] The average atomic rms difference between any pair of DG structures is 2.1 ± 0.3 Å for the backbone atoms and 2.8 ± 0.4 Å for all atoms, and that between the DG structures and the mean DG structure about which they are distributed is 1.4 ± 0.3 Å for the backbone atoms and 1.8 ± 0.3 Å for all atoms. These values are approximately the same for the DGm structures but interestingly slightly larger for the RDDG structures (Table I). This is probably due to the fact that a larger range of bond lengths, bond angles, and ω peptide bond torsion angles is sampled during the course of the restrained molecular dynamics trajectory, which in turn allows a wider region of conformational space to be sampled. The ϕ and ψ angular rms differences

Table I: Atomic rms Differences[a]

	atomic rms difference (Å)	
	backbone atoms	all atoms
(A) Distributions		
⟨DG⟩ vs. ⟨DG⟩	2.1 ± 0.3	2.8 ± 0.4
⟨DGm⟩ vs. ⟨DGm⟩	2.1 ± 0.3	2.7 ± 0.4
⟨RDDG⟩ vs. ⟨RDDG⟩	2.6 ± 0.7	3.2 ± 0.6
⟨DG⟩ vs. $\overline{DG}$	1.4 ± 0.2	1.8 ± 0.3
⟨DGm⟩ vs. $\overline{DGm}$	1.4 ± 0.3	1.9 ± 0.3
⟨RDDG⟩ vs. $\overline{RDDG}$	1.7 ± 0.5	2.1 ± 0.5
⟨DG⟩ vs. ($\overline{DG}$)m	1.9 ± 0.4	2.4 ± 0.4
⟨DGm⟩ vs. ($\overline{DGm}$)m	1.8 ± 0.5	2.2 ± 0.5
⟨RDDG⟩ vs. ($\overline{RDDG}$)m	2.0 ± 0.7	2.5 ± 0.7
(B) rms Shifts		
⟨DG⟩ vs. ⟨DGm⟩	1.1 ± 0.1	1.2 ± 0.1
⟨DG⟩ vs. ⟨RDDG⟩	2.1 ± 0.4	2.4 ± 0.5
⟨DGm⟩ vs. ⟨RDDG⟩	1.6 ± 0.4	1.9 ± 0.4
$\overline{DG}$ vs. $\overline{DGm}$	0.8	0.8
$\overline{DG}$ vs. $\overline{RDDG}$	1.2	1.4
$\overline{DGm}$ vs. $\overline{RDDG}$	0.7	0.8
$\overline{DG}$ vs. ($\overline{DG}$)m	1.2	1.4
$\overline{DGm}$ vs. ($\overline{DGm}$)m	0.9	1.2
$\overline{RDDG}$ vs. ($\overline{RDDG}$)m	1.1	1.2
($\overline{DG}$)m vs. ($\overline{DGm}$)m	0.9	1.2
($\overline{DG}$)m vs. ($\overline{RDDG}$)m	1.3	1.4
($\overline{DGm}$)m vs. ($\overline{RDDG}$)m	1.1	1.3

[a] The notation of the structures is as follows: ⟨DG⟩ comprise the eight converged distance geometry structures, ⟨DGm⟩ the structures derived from the DG structures by restrained energy minimization, and ⟨RDDG⟩ the structures derived from the DGm structures by restrained molecular dynamics (see text). $\overline{DG}$, $\overline{DGm}$, and $\overline{RDDG}$ are the mean structures obtained by averaging the coordinates of the DG, DGm, and RDDG structures, respectively. The estimated standard atomic rms error s_{mean} of these mean structures is given by $[\sum(\text{rmsd}_i)^2/n(n-1)]^{1/2}$, where rmsd_i is the atomic rms difference between the ith structure and the mean structure and n is the number of structures (see supplementary material). ($\overline{DG}$)m, ($\overline{DGm}$)m, and ($\overline{RDDG}$)m are the structures obtained by restrained energy minimization of the mean $\overline{DG}$, $\overline{DGm}$, and $\overline{RDDG}$ structures, respectively.

Table II: ϕ and ψ Angular rms Differences and Violations[a]

structure	rmsd_ϕ (deg)	viol_ϕ	rmsd_ψ (deg)	viol_ψ
(A) Distributions				
⟨DG⟩ vs. ⟨DG⟩	38 ± 4	9.5 ± 1.6	36 ± 4	9.9 ± 2.4
⟨DGm⟩ vs. ⟨DGm⟩	38 ± 4	8.8 ± 2.3	34 ± 4	8.2 ± 2.3
⟨RDDG⟩ vs. ⟨RDDG⟩	37 ± 5	8.8 ± 2.4	34 ± 3	8.1 ± 2.6
⟨DG⟩ vs. $\overline{DG}$	34 ± 3	5.5 ± 2.1	34 ± 4	6.2 ± 1.7
⟨DGm⟩ vs. $\overline{DGm}$	34 ± 2	4.2 ± 2.7	33 ± 4	3.7 ± 1.7
⟨RDDG⟩ vs. $\overline{RDDG}$	30 ± 4	5.1 ± 1.9	30 ± 5	4.1 ± 1.5
(B) Differences				
$\overline{DG}$ vs. ($\overline{DG}$)m	36	0	33	0
$\overline{DGm}$ vs. ($\overline{DGm}$)m	26	0	29	0
$\overline{RDDG}$ vs. ($\overline{RDDG}$)m	24	1	23	1
$\overline{DG}$ vs. $\overline{DGm}$	28	0	26	0
$\overline{DG}$ vs. $\overline{RDDG}$	34	2	30	1
$\overline{DGm}$ vs. $\overline{RDDG}$	23	1	19	1
($\overline{DG}$)m vs. ($\overline{DGm}$)m	23	3	30	2
($\overline{DG}$)m vs. ($\overline{RDDG}$)m	31	3	34	2
($\overline{DGm}$)m vs. ($\overline{RDDG}$)m	27	2	28	2

[a] The angular rms violations (viol_ϕ and viol_ψ) are defined as the number of angles for which the difference between the values for any given pair of structures is greater than 90°; these angles are not included in the calculation of the angular rms differences, rmsd_ϕ and rmsd_ψ. The notation of the structures is the same as that in Table I.

between the structures of the DG, DGm, and RDDG sets, however, are similar (Table II and supplementary material).

Looking at Figure 10, it is apparent that the positions of the backbone atoms in regions of regular secondary structure, namely, the two α-helices (residues 7–19 and 23–28) and the mini-antiparallel β-sheet (residues 2–4 and 32–34), are well defined with atomic rms deviations from the respective mean structures of <1 Å. Loops, turns, and irregular structures, on the other hand, tend to be relatively poorly defined. In particular, there are five regions in the structures where the atomic rms distribution of the backbone atoms about the mean structures is >2 Å, namely, the first N-terminal and last two C-terminal residues and residues 5–6, 19–21, and 35–40, which comprise three of the turns. The larger deviations in these regions arise from two contributory factors: first, a reduced network of interproton distance restraints, and second, the absence of stabilizing backbone hydrogen bonds. In this respect we note that none of the turns present in phoratoxin are classical in nature.

The positions of the side-chain atoms, as expected, are not as well defined as those of the backbone atoms (Figure 10), and this is most marked for external surface residues for which there are few or no distance restraints involving side-chain atoms (Figure 8). Side chains whose positions are restricted by stereochemistry, packing requirements within the protein interior, or a large network of interproton distance restraints are, however, well defined with atomic rms distributions about the mean structures of <2 Å (Figure 10). Thus, for example, the positions of the cystine side chains that are restricted by the requirements of disulfide bridge formation are well defined. Of those side chains that extend beyond the C^γ atom, the positions of Tyr-13, Arg-17, Leu-27, Ile-34, and Ile-35 are well defined due to a combination of packing requirements and interproton distance restraints. The position of Trp-44, on the other hand, is slightly less well defined than these other residues because, being a surface residue, its position is restricted only by the interproton distance restraints. Thus, the variability of the structures in certain areas reflects to a large extent the limitations imposed on the accuracy of the structure determination by the lack of suitable NOE restraints involving flexible parts of the molecule.

Refinement results in a reduction in the overall energy of the structures (Table IV). This principally involves the nonbonding (van der Waals, electrostatic, and hydrogen bonding) terms and the NOE restraints energies (Tables III and IV). The first-level improvement resulting from restrained energy minimization is achieved by only small rms shifts in the atomic positions, while the second-level improvement produced by restrained molecular dynamics is associated with much larger atomic rms shifts (Table I and supplementary material). Accompanying the successive reduction in the energies of the structure is a reduction in the radii of gyration (Table III), which arises from the improvement in the electrostatic energy and from the increased contribution of the attractive component of the van der Waals energy (Table IV).

The effects of refinement can be understood in terms of a two-level energy model. The distance geometry calculations generate a set of structures that are located in the region of the global minimum (tier 1). Each DG structure lies in the vicinity of a different local subminimum (tier 2), which is reached by restrained energy minimization (i.e., the DGm structures). Thus, only a very limited region of the tier 2 subspace is explored by restrained energy minimization due to its inability to overcome energy barriers. Restrained molecular dynamics, on the other hand, explores a much larger

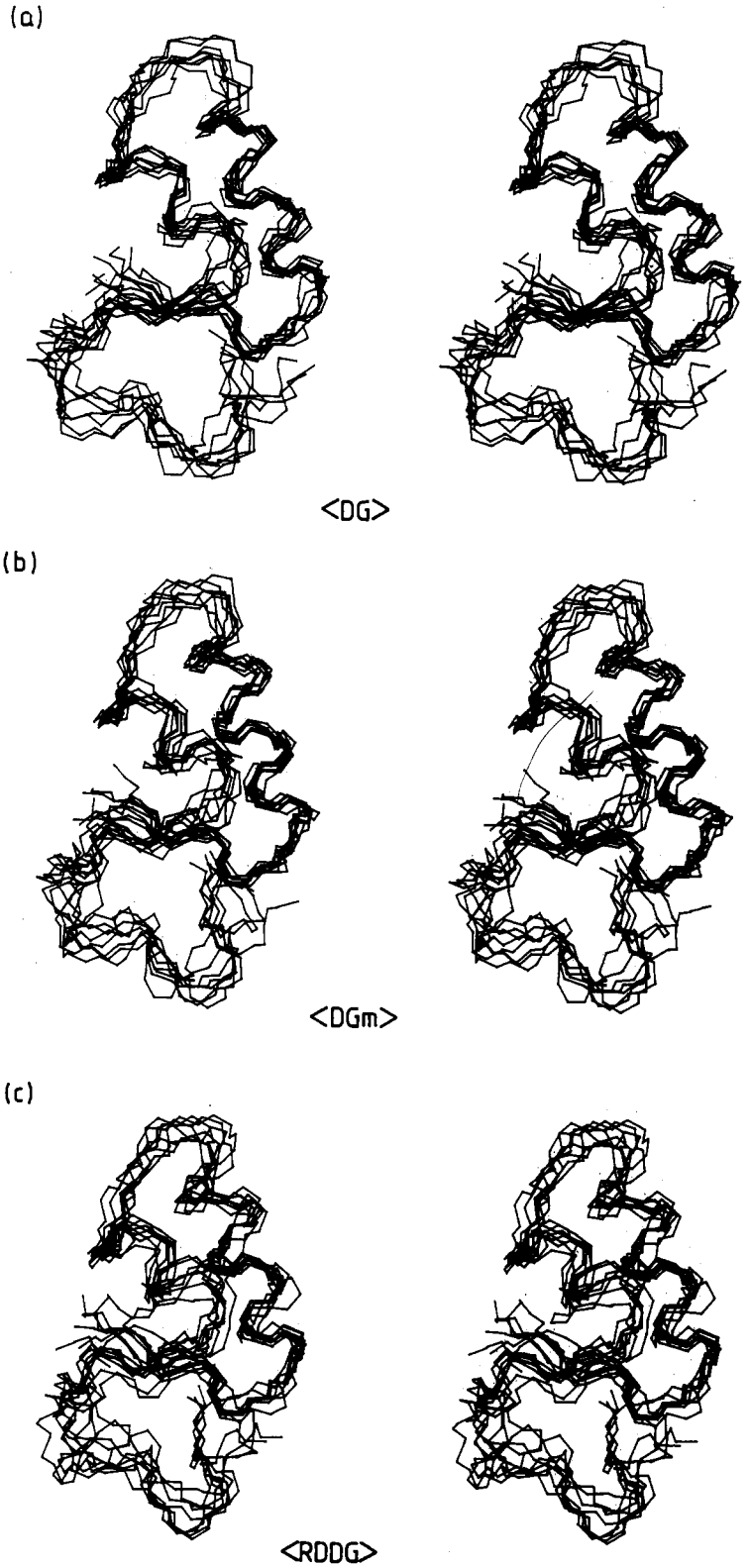

FIGURE 7: Best fit superpositions of the backbone (C, C^{α}, N) atoms of the eight converged (a) DG, (b) DGm, and (c) RDDG phoratoxin structures.

region of the tier 2 subspace, thereby locating lower energy local subminima. There are of course variations in the energies of the RDDG structures, but these are relatively small. Thus, the local subminima occupied by the RDDG structures are approximately equivalent energetically although they are located in different regions of subspace within the global minimum. In this respect, it is important to stress that restrained molecular dynamics per se does *not* result in convergence to an overall single global minimum structure defined by a single smooth potential well. Indeed, it is unlikely that such a minimum even exists. Rather, restrained molecular dynamics, starting from a number of DGm structures, permits the most efficient sampling of the lowest energy subminima (tier 2) of the global minimum region (tier 1), and the distribution of the RDDG structures provides a measure of the size of the tier 1 conformational space. This is probably slightly larger than that available to the actual protein in solution due to the increased structural variability afforded by the limited range

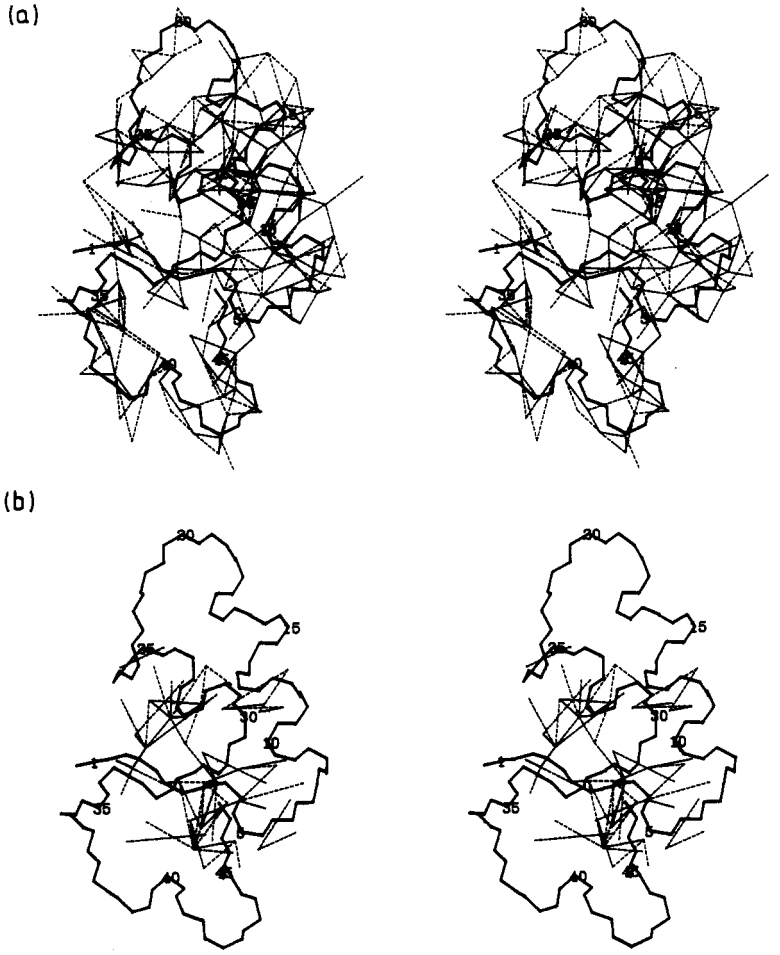

FIGURE 8: Stereoviews of the short- (a) and long- (b) range interproton distance restraints shown as dashed lines superimposed on a framework comprising the backbone (N, C^{α}, C) atoms of the phoratoxin structure (RDDG)m.

Table III: Interproton Distance Deviations and Radii of Gyration[a]

	rms difference between calculated and target interproton distance restraints (Å)				
		interresidue			
structure	all (331)	short range ($\|i-j\| \leq 5$) (135)	long range ($\|i-j\| > 5$) (60)	intraresidue (136)	radii of gyration (Å)
⟨DG⟩	0.56 ± 0.04	0.48 ± 0.05	0.96 ± 0.08	0.36 ± 0.01	10.09 ± 0.32
$\overline{DG}$	0.42	0.36	0.75	0.24	9.94
($\overline{DG}$)m	0.09	0.08	0.11	0.09	9.73
⟨DGm⟩	0.13 ± 0.02	0.10 ± 0.12	0.19 ± 0.06	0.12 ± 0.01	9.61 ± 0.21
$\overline{DGm}$	0.17	0.13	0.18	0.19	9.44
($\overline{DGm}$)m	0.10	0.08	0.12	0.11	9.53
⟨RDDG⟩	0.10 ± 0.01	0.08 ± 0.01	0.13 ± 0.02	0.10 ± 0.01	9.38 ± 0.11
$\overline{RDDG}$	0.18	0.16	0.18	0.19	9.13
($\overline{RDDG}$)m	0.10	0.09	0.12	0.09	9.48

[a] The notation of the structures is the same as that in Table I. The rms difference (rmsd) between the calculated (r_{ij}) and target restraints is calculated with respect to the upper (r_{ij}^{u}) and lower (r_{ij}^{l}) limits such that

$$\text{rmsd} = \begin{cases} [\Sigma(r_{ij} - r_{ij}^{u})^2/n]^{-1/2} & \text{if } r_{ij} > r_{ij}^{u} \\ 0 & \text{if } r_{ij}^{l} \leq r_{ij} \leq r_{ij}^{u} \\ [\Sigma(r_{ij} - r_{ij}^{l})^2/n]^{1/2} & \text{if } r_{ij} < r_{ij}^{l} \end{cases}$$

(<5 Å), accuracy, and number of the experimental interproton distance restraints. Nevertheless, the concept of tier 2 substates arising from the present analysis may have its counterpart in physical reality, as evidenced by the observation of power law kinetics for the rebinding of CO to myoglobin following flash photolysis at low temperatures (Austin et al., 1975; Ansari et al., 1985). [As indicated by Stein (1985), it also bears a natural analogy to spin glasses in condensed matter physics, in particular the properties of disorder and frustration leading to the presence of many inequivalent local energy minima separated by barriers within a global minimum region.]

An additional feature of the refinement concerns the mean structures, $\overline{DG}$, $\overline{DGm}$, and $\overline{RDDG}$, and the restrained energy minimized mean structures, ($\overline{DG}$)m, ($\overline{DGm}$)m, and ($\overline{RDDG}$)m. The atomic rms differences between the mean structures as well as between the restrained energy minimized

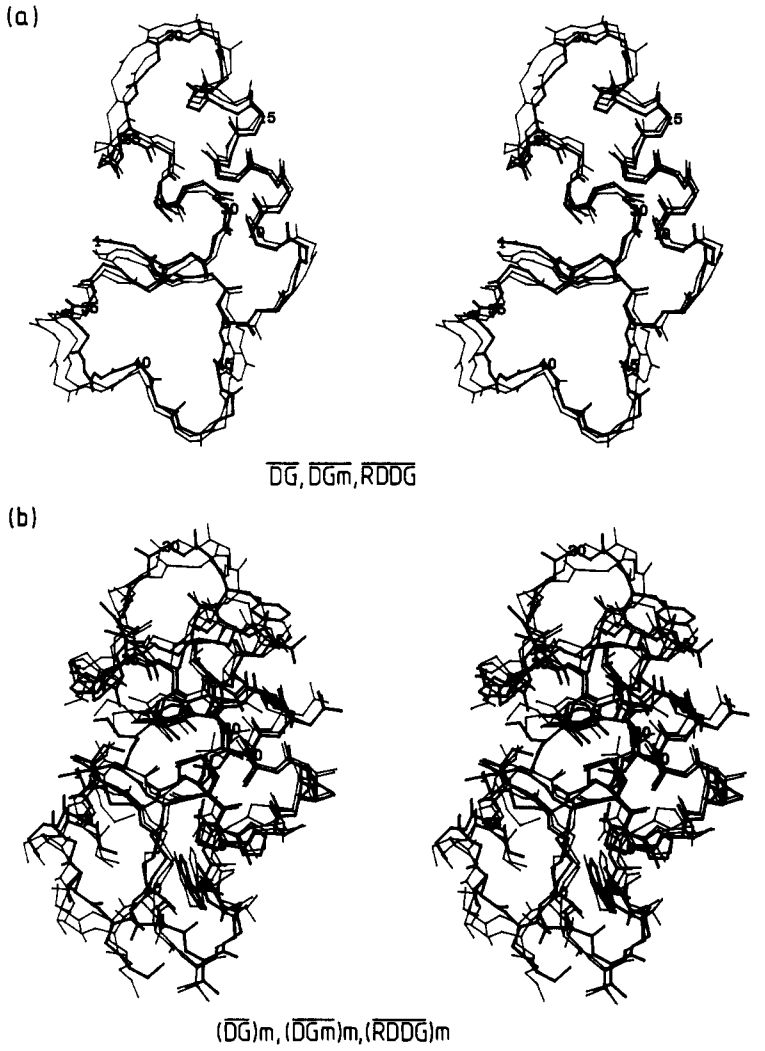

FIGURE 9: (a) Best fit superposition of the backbone (C, C^{α}, N) atoms of the mean phoratoxin structures $\overline{DG}$, $\overline{DGm}$, and $\overline{RDDG}$. (b) Best fit superposition of all atoms (except hydrogen atoms) of the restrained energy minimized mean structures ($\overline{DG}$)m, ($\overline{DGm}$)m, and ($\overline{RDDG}$)m. $\overline{RDDG}$ and ($\overline{RDDG}$)m are shown as thick lines whereas the other structures are shown as thin lines.

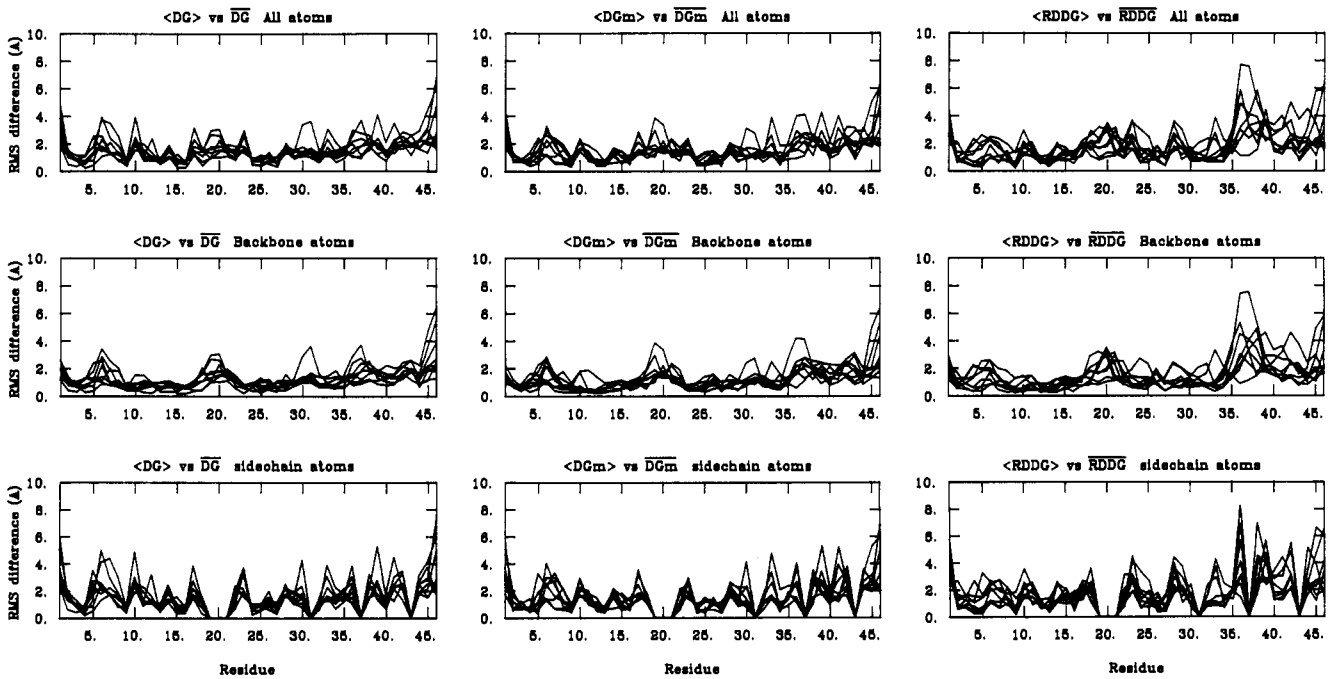

FIGURE 10: Atomic rms distribution of the eight DG, DGm, and RDDG structures about the mean $\overline{DG}$, $\overline{DGm}$, and $\overline{RDDG}$ structures, respectively, for all atoms, backbone atoms, and side-chain atoms as a function of residue number.

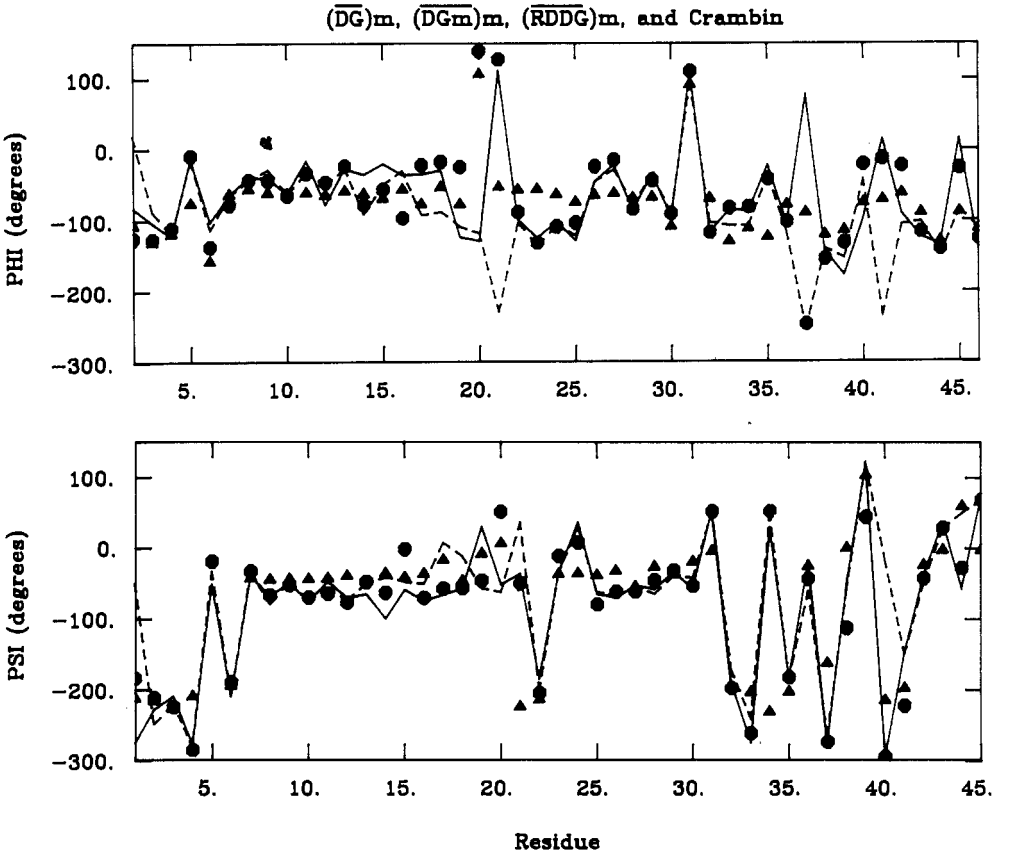

FIGURE 11: ϕ and ψ backbone torsion angles as a function of residue number for the (DG)m (--), (DGm)m (—), and (RDDG)m (●) phoratoxin structures and the X-ray structure of crambin (▲).

Table IV: Energies of the Structures[a]

structure	total	bond (678)[b]	angle (1219)[b]	dihedral (314)[b]	improper (162)[b]	van der Waals	electro-static	H bond	NOE restraints (331)[c]	ϕ torsion angle restraints (6)[c]
⟨DG⟩	5911 ± 2429[d]	88 ± 39	332 ± 87	317 ± 12	1.7 ± 3.4	1040 ± 2233[e]	−67 ± 44	−16 ± 4	4161 ± 551	54 ± 84[f]
DGm	390 ± 200	54 ± 14	418 ± 58	262 ± 23	34 ± 6	15 ± 48	−583 ± 40	−38 ± 6	226 ± 66	2 ± 4
RDDG	−116 ± 108	40 ± 6	345 ± 35	237 ± 18	32 ± 6	−51 ± 19	−784 ± 31	−67 ± 7	130 ± 28	0.2 ± 0.6
$\overline{DG}$	$>10^6$	3.6×10^4	6134	462	0.9	$>10^6$	−1641	−15	2348	4.2
$\overline{DGm}$	$>10^6$	1.9×10^4	3446	595	14	$>10^6$	−638	−26	479	0.0
$\overline{RDDG}$	$>10^6$	1.7×10^4	2949	551	60	$>10^6$	−1736	−35	383	0.0
$(\overline{DG})$m	116	49	343	252	39	−62	−575	−41	112	0.0
$(\overline{DGm})$m	−8	50	333	228	41	−46	−694	−55	135	0.0
$(\overline{RDDG})$m	−118	46	323	222	32	−56	−740	−69	124	0.0

[a] The notation of the structures is the same as that in Table I. The number of terms for the bond, angle, dihedral, and improper dihedral (planarity) potentials and for the effective NOE interproton distance and ϕ backbone torsion angle restraints potentials is given in parentheses. [b] The bond, angle, and dihedral potentials for the three disulfide bridges are included in these terms. [c] The restraint force constants (cf. eq 1) have values of 40 kcal mol^{-1} Å^{-2} for the NOE interproton distance restraints and 40 kcal mol^{-1} rad^{-2} for the ϕ backbone torsion angle restraints. [d] The range of the total energy of the DG structures extends from 4089 to 11652 kcal/mol. [e] The range of the van der Waals energy of the DG structures extends from 94 to 6560 kcal/mol. [f] The range of the ϕ backbone torsion angle restraint potential of the DG structures extends from 0.0 to 236 kcal/mol.

mean structures are much smaller than the atomic rms shifts between the individual sets of DG(i), DGm(i), and RDDG(i) structures (Table II and supplementary material). Indeed, the atomic rms differences between the restrained energy minimized mean structures are only slightly larger than those between the mean structures themselves, despite the fact that the restrained energy minimization procedure results in atomic rms shifts of ~1 Å for the backbone atoms. The energies of the restrained energy minimized mean structures display a hierarchy with $(\overline{DG})$m having the largest energy and $(\overline{RDDG})$m the lowest (Table IV). Interestingly, the van der Waals, hydrogen bonding, NOE, and ϕ restraint energies of all three restrained energy minimized mean structures are similar and are comparable to those of the individual RDDG structures (Table IV). The electrostatic energy of the $(\overline{DG})$m and $(\overline{DGm})$m structures, however, is significantly higher than that of the individual RDDG structures whereas that of the $(\overline{RDDG})$m structure is comparable to it. Thus, despite the relatively large distribution of the DG, DGm, and RDDG structures about their respective means, the mean structures are located in the vicinity of a local subminimum, and the lower the energy of the individual structures, the lower the energy of this local subminimum.

The coordinates of the mean structures are relatively well defined with atomic rms standard errors of <0.7 Å for all atoms (Table I). How, therefore, do the mean structures relate to the time-average structure in solution? We cannot provide an answer to this question as no independently determined

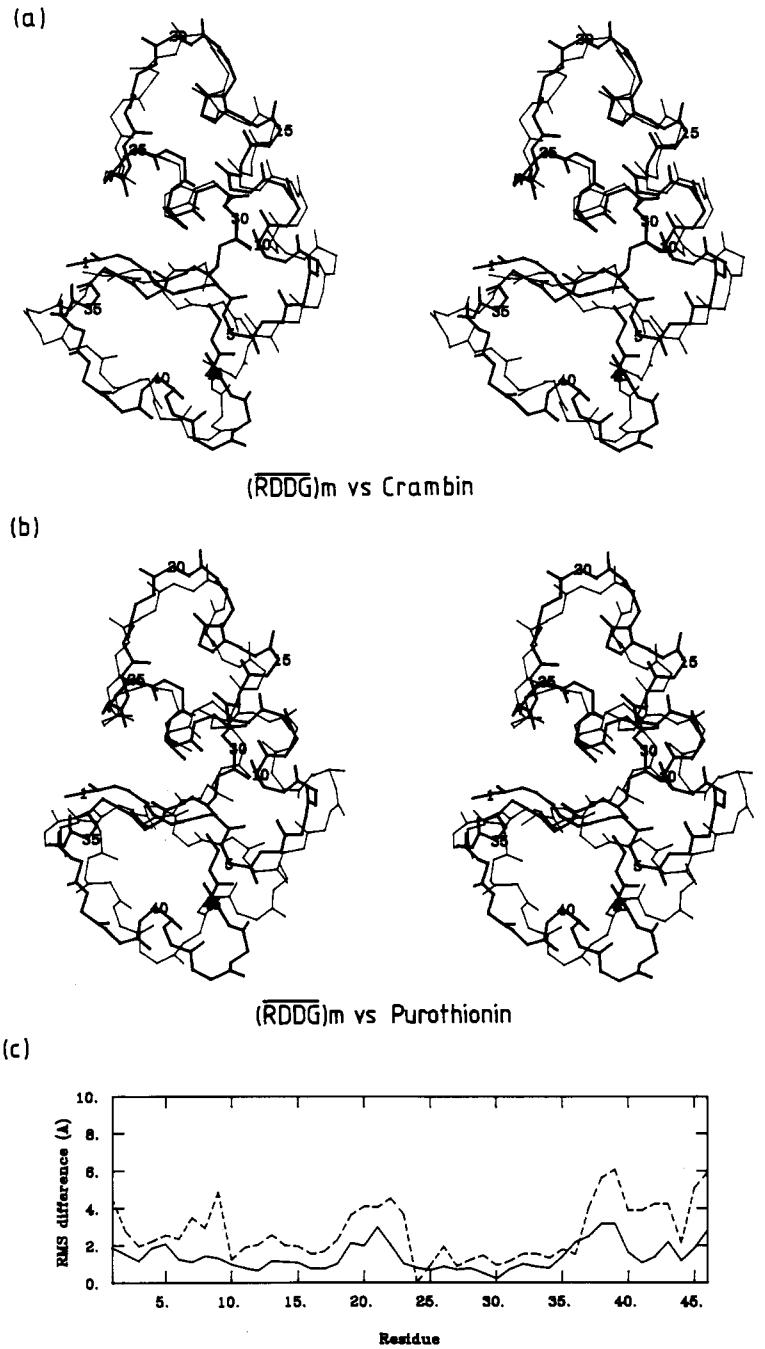

FIGURE 12: Best fit superposition of the backbone (C, C^{α}, N, O) atoms of the (RDDG)m structure of phoratoxin (thick lines) with (a) the X-ray structure of crambin (thin lines) and (b) the (RDDG)m structure of α_1-purothionin (thin lines). (c) Backbone (C, C^{α}, N, O) atomic rms difference between the (RDDG)m structure of phoratoxin on the one hand and the X-ray structure of crambin (—) and the (RDDG)m structure of α_1-purothionin (--) on the other. The 1.5-Å resolution X-ray structure of crambin is from Hendrickson and Teeter (1981), and the (RDDG)m solution structure of α_1-purothionin is from Clore et al. (1986a). [Note that the meaning of the notation (RDDG)m for α_1-purothionin is the same as that for phoratoxin.]

structure, such as an X-ray structure, is available in this case. Nevertheless, our model studies on crambin (Clore et al., 1986b) suggest the following interpretation. The individual DG, DGm, and RDDG structures are distributed about mean structures that are close to but not identical with the true time-average solution structure. They are not identical with the latter because of deficiencies in the empirical energy functions used which cannot be completely overcome by the experimental restraints available. These deficiencies are most severe for the DG structures as the only non-bonding interaction considered is a soft van der Waals repulsion term. Deficiencies in the empirical energy function used for restrained energy minimization and restrained molecular dynamics are clearly smaller but still present, for example, the relatively crude treatment of the electrostatic interactions, the assignment of partial charges, and the absence of solvent. These, however, are not as severe as one might imagine at first sight because they are almost completely compensated by the experimental restraint energies.

The mean DGm structure is thus intermediate in quality between the mean DG and RDDG structures because, although the empirical energy function used to generate the DGm structures is the same as that used to generate the RDDG structures, restrained energy minimization only enables one to locate the local submimima closest to the original DG

Table V: Backbone (C, C^{α}, N, O) Atom Backbone rms Differences between Solution Structures of Phoratoxin and α_1-Purothionin and X-ray Structure of Crambin[a]

	backbone (C, C^{α}, N, O) atomic rms difference (Å)		
		α_1-purothionin[c]	
structure	crambin[b]	$\overline{\text{RDDG}}$	$(\overline{\text{RDDG}})$m
phoratoxin			
$\overline{\text{DG}}$	1.7	3.0	3.1
$(\overline{\text{DG}})$m	1.9	3.3	3.5
$\overline{\text{DG}}$m	1.6	2.8	3.0
$(\overline{\text{DGm}})$m	1.8	3.1	3.2
$\overline{\text{RDDG}}$	1.7	2.7	2.9
$(\overline{\text{RDDG}})$m	1.6	2.9	3.1
α_1-purothionin			
$\overline{\text{RDDG}}$	2.3		
$(\overline{\text{RDDG}})$m	2.6		

[a] The notation of the structures is the same as that in Table I. [b] From the 1.5-Å resolution crystal structure of Hendrickson and Teeter (1981) deposited in the Brookhaven Protein Data Bank. [c] From Clore et al. (1986a). The meaning of $\overline{\text{RDDG}}$ and $(\overline{\text{RDDG}})$m for α_1-purothionin is the same as that for phoratoxin.

structures, whereas restrained molecular dynamics permits one to locate the lowest energy local subminima.

Comparison with Crambin and α_1-Purothionin. The secondary and tertiary structures of phoratoxin are similar to those of the X-ray structure of crambin (Hendrickson & Teeter, 1981) and the solution structure of α_1-purothionin (Clore et al., 1986a). All three proteins are L-shaped, with one arm of the L being formed by the two α-helices and the other arm by the mini-antiparallel β-sheet and the two turns and strand formed by the C-terminal residues 35–46. The angle between the long axes of the helices is $\sim$140°, and the angle between the plane formed by the two helices and the plane of the antiparallel β-sheet is $\sim$50°. The best fit superpositions of the backbone atoms of the $(\overline{\text{RDDG}})$m phoratoxin structure on the crambin X-ray structure and the $(\overline{\text{RDDG}})$m α_1-purothionin structure are shown in Figure 12. Also shown in Figure 12 is a plot of the respective backbone atomic rms differences as a function of residue number, which shows that the regions of closest agreement comprise the regular secondary structure elements. The mean and restrained energy minimized mean phoratoxin structures are closer to the crambin X-ray structure than the equivalent α_1-purothionin structures (Table V). This finding correlates with the observed amino acid sequence homologies (Figure 1). The phoratoxin and α_1-purothionin structures, however, are closer to crambin than they are to each other (Table V) although their amino acid sequences are more homologous to each other than to crambin (Figure 1). Given that the experimental restraints used to determine the phoratoxin and α_1-purothionin structures are similar both in quality and in quantity, this may be due to the deletion of residue 24 and the presence of an extra disulfide bridge between residues 12 and 29 in α_1-purothionin. Indeed, the latter may be responsible for the shorter helix A in α_1-purothionin, which extends from residues 10 to 19 while in phoratoxin and crambin this helix comprises residues 7–19.

The similarity between the phoratoxin structures and crambin extends to the ϕ and ψ backbone torsion angles (Figure 11 and supplementary material) and the backbone hydrogen bonds (see supplementary material). The ϕ and ψ angular rms differences between the mean and restrained energy minimized structures on the one hand and crambin on the other are $\lesssim$40° for angle pairs deviating by less than 90°. Refinement does not result in any significant reduction in this value but reduces the number of ϕ and ψ angles deviating by greater than 90° between crambin and the phoratoxin structures from 4 ± 0.8 for the $\overline{\text{DG}}$ and $(\overline{\text{DG}})$m structures to 2.5 ± 0.5 for the $\overline{\text{RDDG}}$ and $(\overline{\text{RDDG}})$m structures. The short-range ($|i-j| \leq 5$) backbone hydrogen bonds in the two helices are virtually identical in the restrained energy minimized mean phoratoxin structures and crambin, and four of the six long-range ($|i-j| > 5$) backbone hydrogen bonds present in crambin are also present in the phoratoxin structures (see the supplementary material for further details).

Conclusions

The results presented here as well as those in our previous paper on the solution structure of α_1-purothionin (Clore et al., 1986a) demonstrate that the similarity in the amino acid sequences of the three plant proteins phoratoxin, α_1-purothionin, and crambin is extended to their three-dimensional structures. All three structures are closely related with common secondary structure elements arranged almost identically in space. The main deviations that are observed are located in turns/loops connecting these elements and most notably at the C-terminal end of the polypeptide chain.

The observed microheterogeneity of phoratoxin at position 25 did not pose an obstacle with respect to the sequence-specific assignment. Additional cross-peaks arising from the minor species only extended over a short stretch of sequence and could be easily identified. Since no differences in relative cross-peak intensities were detected for this sequence, we conclude that the solution structure of the minor species, phoratoxin B, is indistinguishable from that of the major species, phoratoxin A, within the accuracy of the experimental data.

Acknowledgments

We thank Prof. Gunnar Samuellson for the gift of phoratoxin, Dr. A. T. Brünger and Prof. M. Karplus for useful discussions, and the Max-Planck-Institut für Plasma Physik (Garching) for computing facilities on the CRAY-XMP.

Supplementary Material Available

Five tables listing the complete resonance assignments, the complete list of NOEs used in the computation of the three-dimensional structure of phoratoxin, the atomic rms standard errors of the mean structures, angular rms deviations between the mean structures and crambin for the backbone torsion angles, and the list of backbone hydrogen bonds in the converged structures and four figures showing additional NMR data, atomic rms shifts between various converged structures as a function of residue number, and plots of backbone torsion angles of various converged structures as a function of residue number (16 pages). Ordering information is given on any current masthead page.

Registry No. Phoratoxin A, 65719-15-5; phoratoxin B, 97707-63-6; crambin, 78783-34-3; α_1-purothionin, 58239-09-1.

References

Ansari, A., Berendzen, J., Bowne, S. F., Frauenfelder, H., Iben, I. T. E., Sauke, T. B., Shyamsunder, E., & Young, R. D. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5000–5004.

Austin, R. H., Berson, K. W., Eisenstein, L., Frauenfelder, H., & Gunsalus, I. C. (1975) *Biochemistry 14*, 5355–5373.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 153*, 321–345.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93–106.
Braun, W., Wider, G., Lee, K. H., & Wüthrich, K. (1983) *J. Mol. Biol. 169*, 921–948.
Braun, W., Wagner, G., Wörgötter, E., Vasak, M., Kägi, J. H. R., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 125–129.
Braunschweiler, L., & Ernst, R. R. (1988) *J. Magn. Reson. 53*, 521–528.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.
Bruccoleri, R. E., & Karplus, M. (1986) *J. Comput. Chem. 7*, 165–175.
Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.
Carrasco, L., Vazquez, D., Fernandes-Lucus, C., Carbonero, P., & Garcie-Olmedo, F. (1971) *Eur. J. Biochem. 116*, 185–189.
Clore, G. M., & Gronenborn, A. M. (1985) *J. Magn. Reson. 61*, 158–164.
Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1985) *J. Mol. Biol. 186*, 435–455.
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., Martin, S. R., & Gronenborn, A. M. (1986c) *J. Mol. Biol. 191*, 553–561.
Clore, G. M., Sukuraman, D. K., Gronenborn, A. M., Teeter, M. M., Whitlow, M., & Jones, B. L. (1987) *J. Mol. Biol. 193*, 571–578.
Crippen, G. M., & Havel, T. F. (1978) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 282–284.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.
Dobson, C. H., Olejniczak, E. T., Poulsen, F. M., & Ratcliffe, R. G. (1982) *J. Magn. Reson. 48*, 87–110.
Havel, T. F. (1986) DISGEO, Quantum Chemistry Program Exchange Program No. 507, Indiana University, Bloomington, IN.
Havel, T. F., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.
Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 281–294.
Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.
Hendrickson, W. A., & Teeter, M. M. (1981) *Nature (London) 290*, 107–113.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.
Jones, B. L., Lockhart, G. L., Mak, A., & Cooper, D. B. (1982) *J. Hered. 73*, 143–144.
Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179–182.
Keepers, J., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.
Kline, A. D., & Wüthrich, K. (1985) *J. Mol. Biol. 185*, 503–507.
Kline, A. D., Braun, W., & Wüthrich, K. (1986) *J. Mol. Biol. 189*, 377–282.
Kumar, A., Wagner, G., Ernst, R. R., & Wüthrich, K. (1981) *J. Am. Chem. Soc. 103*, 3654–3658.
Macura, S., Huang, Y., Suter, D., & Ernst, R. O. (1981) *J. Magn. Reson. 43*, 259–281.
Mak, A. S., & Jones, B. L. (1976) *Can. J. Biochem. 54*, 834–842.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mellstrand, S. T., & Samuelsson, G. (1973) *Eur. J. Biochem. 32*, 143–147.
Mellstrand, S. T., & Samuelsson, G. (1974) *Acta Pharm. Suec. 11*, 347–360.
Nilsson, L., Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1986) *J. Mol. Biol. 188*, 455–475.
Okada, T., & Yoshisumi, H. (1973) *Agric. Biol. Chem. 37*, 2289–2294.
Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.
Rosell, S., & Samuelsson, G. (1966) *Toxicon 4*, 107–115.
Samuelsson, G., & Ekblad, M. (1967) *Acta Chem. Scand. 21*, 849–856.
Samuelsson, G., & Peterson, B. M. (1971) *Eur. J. Biochem. 21*, 86–89.
Samuelsson, G., Seger, L., & Olson, T. (1968) *Acta Chem. Scand. 22*, 2264–2642.
Sippl, M. J., & Scheraga, H. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2283–2287.
Stein, D. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3670–3674.
Strop, P., Wider, G., & Wüthrich, K. (1983) *J. Mol. Biol. 166*, 611–667.
Teeter, M. M., Mazer, J. A., & L'Italien, J. J. (1981) *Biochemistry 20*, 5437–5443.
Thunberg, E. (1983) *Acta Pharm. Suec. 20*, 115–122.
Van Etten, C. H., Nielsen, H. C., & Peters, J. E. (1965) *Phytochemistry 4*, 467–473.
Verlet, L. (1967) *Phys. Rev. 159*, 98–105.
Wagner, G., & Wüthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 160*, 343–361.
Wagner, G., Neuhaus, D., Wörgötter, E., Vasak, M., Kägi, J. H. R., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 131–135.
Weber, P. L., Wemmer, D. E., & Reid, B. R. (1985) *Biochemistry 24*, 4553–4562.
Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wüthrich, K. (1984) *J. Magn. Reson. 56*, 207–234.
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.
Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311–319.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 160*, 949–961.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.
Zarbock, J., Clore, G. M., & Gronenborn, A. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7628–7632.
Zuiderweg, E. R. P., Kaptein, R., & Wüthrich, K. (1983) *Eur. J. Biochem. 127*, 279–292.