

The three-dimensional structure of α 1-purothionin in solution: combined use of nuclear magnetic resonance, distance geometry and restrained molecular dynamics

G.Marius Clore, Michael Nilges, Dinesh K.Sukumaran, Axel T.Brünger¹, Martin Karplus¹ and Angela M.Gronenborn

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG and ¹Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

Communicated by R.Huber

The determination of the three-dimensional solution structure of α 1-purothionin using a combination of metric matrix distance geometry and restrained molecular dynamics calculations based on n.m.r. data is presented. The experimental data comprise complete sequence-specific proton resonance assignments, a set of 310 approximate interproton distance restraints derived from nuclear Overhauser effects, 27 Φ backbone torsion angle restraints derived from vicinal coupling constants, 4 distance restraints from hydrogen bonds and 12 distance restraints from disulphide bridges. The average atomic rms difference between the final nine converged structures and the mean structure obtained by averaging their coordinates is 1.5 ± 0.1 Å for the backbone atoms and 2.0 ± 0.1 Å for all atoms. The overall shape of α 1-purothionin is that of the capital letter L, similar to that of crambin, with the longer arm comprising two approximately parallel α -helices and the shorter arm a strand and a mini anti-parallel β sheet.

Key words: α 1-purothionin/3D structure/nuclear Overhauser effect/interproton distances/distance geometry/restrained molecular dynamics.

Introduction

α 1-Purothionin is a member of a group of low mol. wt water-soluble protein toxins which are ubiquitous throughout the plant kingdom (Mark and Jones 1976a, 1976b; Jones *et al.*, 1982). It possesses 33% sequence homology with respect to crambin (see Figure 1), a water-insoluble plant protein whose biological activity is not yet known (van Etten *et al.*, 1965) and whose crystal structure has been solved to very high resolution by X-ray and neutron diffraction (Hendrickson and Teeter, 1981; Teeter and Kossiakoff, 1982). In addition, the two proteins display the same pattern of disulphide linkages, with an additional one in α 1-purothionin. On this basis, it has been suggested that the three-dimensional (3D) structures of crambin and α 1-purothionin should be similar (Teeter *et al.*, 1981; Williams and Teeter, 1984; Whitlow and Teeter, 1985). To test this hypothesis we embarked on a ¹H-n.m.r. study of α 1-purothionin with the eventual aim of determining its 3D structure in solution. In a recent paper (Clore *et al.*, 1986a), we presented the complete assignment of the ¹H-n.m.r. spectrum of α 1-purothionin and the delineation of secondary structure elements on the basis of a qualitative interpretation of short range nuclear Overhauser effects (NOE). We showed that the secondary structure of α 1-purothionin is similar to that of crambin and consists of two

helices (residues 10–19 and 23–28), two short β strands (residues 3–5 and 31–34) which form a mini anti-parallel β -sheet and five turns. In this paper we extend our previous study to the determination of the 3D solution structure of α 1-purothionin based on interproton distance and dihedral angle restraints derived from the n.m.r. measurements using a combination of metric matrix distance geometry (Crippen and Havel, 1978; Havel and Wüthrich, 1984, 1985; Sippl and Scheraga, 1986) and restrained molecular dynamics (Clore *et al.*, 1985, 1986b; Brünger *et al.*, 1986a; Nilsson *et al.*, 1986; Kaptein *et al.*, 1985) calculations.

Results and Discussion

Interproton distance and dihedral angle restraints

A set of 310 approximate interproton distance restraints were derived from pure phase absorption two-dimensional NOE spectroscopy (NOESY) spectra recorded in D₂O and H₂O. Spectra recorded with mixing times of 200 ms and 300 ms were used for the assignment of NOESY cross peaks, whereas those recorded with mixing times of 100 ms and 200 ms were used for the classification of cross-peak intensities. Examples demonstrating the quality of the NOESY spectra are given in Clore *et al.* (1986a). The NOE interproton distance restraints, which comprised 116 intra-residue distances, and 135 short range ($|i-j| \leq 5$) and 59 long range ($|i-j| > 5$) inter-residue distances, were classified into three distance ranges, 1.8–2.5 Å, 1.8–3.5 Å and 3–5 Å, corresponding to strong, medium and weak NOEs, respectively. This classification procedure was carried out essentially as described previously (Williamson *et al.*, 1985; Clore *et al.*, 1985, 1986b; Kline *et al.*, 1986). The superposition of the inter-residue interproton distances on the final average structure is shown in Figure 2a. The number of inter-residue distances, both short and long range, is approximately the same as that used in our model studies on crambin (Clore *et al.*, 1986b, 1986c; Brünger *et al.*, 1986a), whereas the number of intra-residue distances is significantly larger.

The NOE restraints were supplemented by three other groups of restraints: (i) 27 Φ backbone torsion angle restraints derived from ³J_{HN α coupling constants (Pardi *et al.*, 1984) measured by double quantum filtered homonuclear correlated spectroscopy (DQF-COSY) in H₂O (viz. $\Phi = 0^\circ$ to -90° for ³J_{HN α < 6Hz and $\Phi = -80^\circ$ to -180° for ³J_{HN α > 9Hz; see Clore *et al.*, 1986a); (ii) four distance restraints for the two backbone hydrogen bonds between Cys 3 and Lys 32 in the mini anti-parallel β -sheet}}}



Fig. 1. Comparison of the amino acid sequences of α 1-purothionin and crambin. The numbering is that of α 1-purothionin which has one residue fewer than crambin. Unshaded boxes indicate conserved residues and shaded boxes conservative amino acid changes. The alignment is that which gives maximum homology (Teeter *et al.*, 1981). Fifteen residues are conserved giving a 33% sequence homology. In addition there are five residues for which the amino acid exchanges are of a conservative nature.

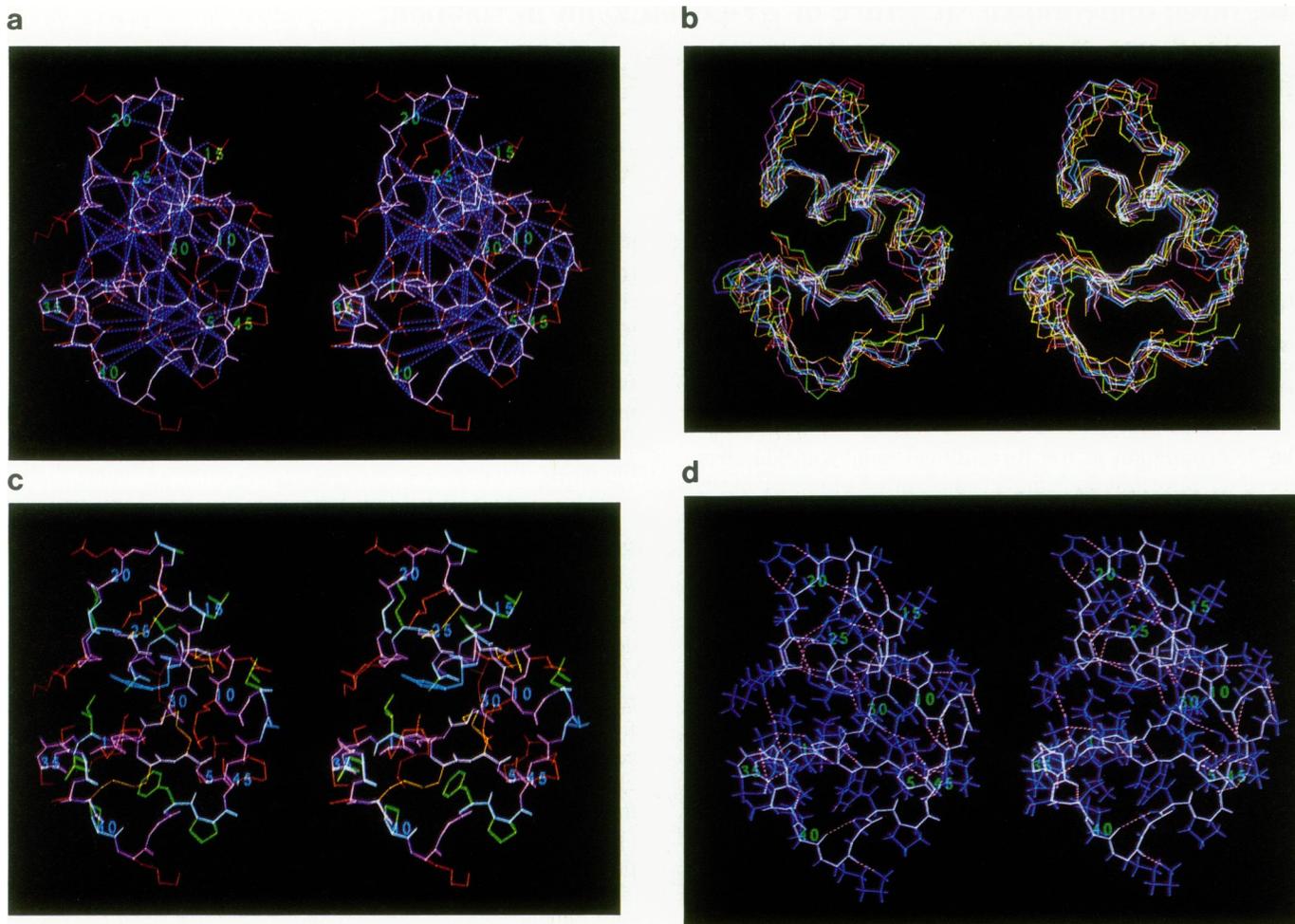


Fig. 2. The three-dimensional solution structure of $\alpha 1$ -purothionin. (a) Short and long range interresidue NOE interproton distance restraints superimposed on the structure (RDDG)m. For ease of viewing the distances have been directed to the directly bonded heavy atoms (colour code: interproton distances, dashed dark blue lines; backbone atoms, lilac; side chain atoms, red). (b) Best fit superposition of the backbone (C,N,C α) atoms of the nine converged RDDG structures. (c) Distribution of hydrophobic and hydrophilic side chains of the structure (RDDG)m: hydrophobic residues (Pro, Ala, Val, Leu, Ile, Phe) are shown in green, hydrophilic residues (Asn, Gln, Lys, Arg, Ser) in red, dual functionality residues (Thr, Tyr) in blue and Cys in orange. (d) Distribution of hydrogen bonds in the structure (RDDG)m (colour code: hydrogen bonds, lilac dashed lines; backbone atoms, light blue; side chain atoms, dark blue).

identified by Clore *et al.* (1986a) (for each hydrogen bond the N–O and NH–O distances are restrained to ranges of 2.3–3.3 Å and 1.3–2.3 Å, respectively); and (iii) 12 distance restraints for the four disulphide bonds between Cys 3 and Cys 39, Cys 4 and Cys 31, Cys 12 and Cys 29, and Cys 16 and Cys 25 (i.e. for each disulphide bond there are three distance restraints: S_i–S_j, S_i–C β _j and S_j–C β _i with values of 2.02 Å, 2.99 Å and 2.99 Å, respectively). The hydrogen bonds in α -helices cannot be assigned unequivocally from the NOE data alone and are therefore not included as restraints.

Calculational strategy

The generation of structures from the distance and dihedral angle restraints proceeded in two stages: (i) a structure determination stage using the DISGEO distance geometry algorithm (Havel and Wüthrich, 1984, 1985; Havel, 1986) based on the metric matrix (Crippen and Havel, 1978), and (ii) a refinement stage using a combination of restrained energy minimization and restrained molecular dynamics (Clore *et al.*, 1985, 1986b; Brünger *et al.*, 1986a; Nilsson *et al.*, 1986). This dual strategy was chosen as it represents the most efficient approach in terms of computational time requirements (Clore *et al.*, 1986c). It should be noted,

however, that the structure determination stage can also be carried out very effectively by restrained molecular dynamics as has been demonstrated using crambin as a model system (Clore *et al.*, 1986b; Brünger *et al.*, 1986a). Indeed a few such calculations have already been carried out on $\alpha 1$ -purothionin and will be presented in a future publication.

For the distance geometry calculations, the upper limits of the distance ranges involving protons for which stereospecific assignments could not be made (e.g. methylene, methyl and aromatic ring protons) were corrected for the pseudo-atom representation used by DISGEO as described by Wüthrich *et al.* (1983). The distance geometry calculations proceeded as follows. A complete set of upper and lower limits on the distances between all atoms of the molecule were determined by triangulation from the experimental distance and dihedral angle restraints and from the distance and planarity restraints obtained from the primary structure. A set of random substructures was then embedded which was consistent with the bounds corresponding to distances between a subset of all atoms comprising the C, C α , N and C α H backbone atoms and the non-terminal C β and C γ atoms. This was followed by the computation of a set of structures which approximately fitted all the distance data. This in-

volved a two step procedure. First, approximate distances between all pairs of atoms not in the substructure were chosen at random within the triangle limits, a procedure known as metrization in n dimensional distance space. Coordinates were then generated from all the distances by projection into 3D cartesian space. Finally, the structures generated in this manner, were subjected to 1500 cycles of restrained least squares refinement with respect to all the distances. The pseudo-atoms were then replaced by real atoms and all hydrogen atoms were built on using the HBUILD algorithm (Brünger, unpublished data) to generate a set of structures known as DG(i). This completed the structure determination stage.

The refinement stage comprised a combination of restrained energy minimization and restrained molecular dynamics calculations carried out using the program CHARMM (Brooks *et al.*, 1983), with an empirical energy function in which all hydrogen atoms were treated explicitly. The empirical energy function consisted of bond, angle, torsion, planarity and non-bonding (i.e. van der Waals, electrostatic and hydrogen bonding) potentials (Karplus and McCammon, 1983; Brooks *et al.*, 1983) supplemented by effective potential terms representing the NOE interproton distances, the assigned hydrogen bond distances, the disulphide bridge distances and Φ backbone torsion angles (Clare *et al.*, 1985, 1986b; Brünger *et al.*, 1986a). For the disulphide restraints a simple harmonic effective potential was used:

$$E_S = c(r_{ij} - r_{ij}^0)^2 \quad (1)$$

where r_{ij} and r_{ij}^0 are the calculated and target restraints respectively, and c is a force constant whose value was set to 200 kcal/mol/Å². In the case of the NOE interproton distance and the assigned hydrogen bond distance restraints we used a square well effective potential, rather than the skewed biharmonic potential that we described previously (Clare *et al.*, 1985), as the square well potential matches the form of the distance restraints penalty function used in DISGEO. This has the form:

$$E_{NOE} = \begin{cases} c(r_{ij} - r_{ij}^u)^2 & , \text{ if } r_{ij} > r_{ij}^u \\ 0 & , \text{ if } r_{ij}^l \leq r_{ij} \leq r_{ij}^u \\ c(r_{ij} - r_{ij}^l)^2 & , \text{ if } r_{ij} < r_{ij}^l \end{cases} \quad (2)$$

where r_{ij}^u and r_{ij}^l are the upper and lower limits of the target distance range, respectively. The force constant c was set to a value of 40 kcal/mol/Å². In the case of those distances involving protons that could not be stereospecifically assigned, a single $(\langle r^{-6} \rangle)^{-1/6}$ average distance was used (Clare *et al.*, 1985, 1986b). Similarly, the Φ backbone torsion angle restraints were represented by a square well effective potential:

$$E_\Phi = \begin{cases} c(\Phi_{ij} - \Phi_{ij}^u)^2 & , \text{ if } \Phi_{ij} > \Phi_{ij}^u \\ 0 & , \text{ if } \Phi_{ij}^l \leq \Phi_{ij} \leq \Phi_{ij}^u \\ c(\Phi_{ij} - \Phi_{ij}^l)^2 & , \text{ if } \Phi_{ij} < \Phi_{ij}^l \end{cases} \quad (3)$$

where Φ_{ij} is the calculated value of Φ , and Φ_{ij}^u and Φ_{ij}^l are the upper and lower limits of the target range of Φ , respectively. The force constant c was set to a value of 40 kcal/mol/rad².

The refinement proceeded in two phases: (i) 1000 cycles of restrained energy minimization to generate structures DGm(i); and (ii) 3 ps of equilibration and thermalization (Brooks *et al.*, 1983) followed by 12 ps of restrained molecular dynamics at 300 K. The coordinates of the last 8 ps of each trajectory were then averaged. To correct for minor distortions in the covalent structure produced by the averaging procedure, these structures were subjected to 500 cycles of restrained energy minimization, additionally constrained to their original structures by weak harmonic constraints (Brucoleri and Karplus, 1986), to generate the final structures RDDG (i). (Note that this last step is essen-

Table I. Atomic rms differences

	Atomic rms difference (Å)	
	Backbone atoms	All atoms
A. Distributions		
<DG> versus <DG>	2.1 ± 0.3	2.8 ± 0.3
<DGm> versus <DGm>	2.0 ± 0.3	2.7 ± 0.3
<RDDG> versus <RDDG>	2.3 ± 0.3	3.0 ± 0.3
<DG> versus \overline{DG}	1.4 ± 0.2	1.8 ± 0.2
<DGm> versus \overline{DGm}	1.3 ± 0.2	1.8 ± 0.1
<RDDG> versus \overline{RDDG}	1.5 ± 0.1	2.0 ± 0.1
B. Rms shifts		
<DG> versus <DGm>	1.0 ± 0.1	1.1 ± 0.1
<DG> versus <RDDG>	1.8 ± 0.1	2.2 ± 0.3
<DGm> versus <RDDG>	1.4 ± 0.2	1.8 ± 0.2
\overline{DG} versus \overline{DGm}	0.6	0.7
\overline{DG} versus \overline{RDDG}	1.2	1.4
\overline{DGm} versus \overline{RDDG}	0.9	1.0
(RDDG)m versus \overline{DG}	1.8	2.2
(RDDG)m versus \overline{DGm}	1.5	1.9
(RDDG)m versus \overline{RDDG}	1.0	1.2
C. Atomic rms standard errors		
\overline{DG}	0.5	0.6
\overline{DGm}	0.4	0.6
\overline{RDDG}	0.5	0.7
D. Rms difference with respect to crambin X-ray structure^a		
\overline{DG}	2.5	—
\overline{DGm}	2.6	—
\overline{RDDG}	2.3	—
(RDDG)m	2.6	—

The notation of the structures is as follows: <DG> comprise the nine 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 structures obtained by averaging the coordinates of the DG, DGm and RDDG structures, respectively. The standard atomic rms error of these average structures is given by rmsd/\sqrt{n} where rmsd is the average atomic rms difference between the n structures and the average structure. (RDDG)m is the structure obtained by restrained energy minimization of \overline{RDDG} .

^aTaken from the 1.5 Å resolution crystal structure of Hendrickson and Teeter (1981) deposited in the Brookhaven Protein Data Bank.

tially a regularization procedure and results in only very small atomic rms shifts of <0.2 Å for all atoms).

The converged structures

A total of nine converged DG(i) structures were generated and subjected to refinement. The course of the refinement is summarized in Tables I–III and the superposition of the backbone (C, C α , N) atoms of the final nine refined RDDG(i) structures is shown in Figure 2b.

The size of the conformational space sampled by the DG, DGm and RDDG structures is similar with the average atomic root mean square (rms) difference between the structures in each set ranging from 2.1 ± 0.3 Å to 2.3 ± 0.3 Å for the backbone atoms and 2.8 ± 0.3 Å to 3.0 ± 0.3 Å for all atoms. The first phase of the refinement, namely the restrained energy minimization, results in average atomic rms shifts of 1.0 ± 0.1 Å and 1.1 ± 0.1 Å for the backbone atoms and all atoms, respectively (Table I), and significant improvements in the interproton distance deviations (Table II), and in the energies of the distance and Φ backbone torsion angle restraints, the dihedral angles and the non-bonding terms (Table III). In addition, the radius of gyration which invariably tends to be slightly expanded in the DG(i)

Table II. Interproton distance deviations and radii of gyration

Structure	Rms difference between calculated and target distance restraints (\AA)					Radius of gyration (\AA)
	Interproton distances			Disulphide bridge restraints	Intraresidue	
	All	Inter-residue				
(310)	short range ($ i-j \leq 5$) (135)	long range ($ i-j > 5$) (59)	(116)	(12)		
<DG>	0.55 \pm 0.03	0.55 \pm 0.03	0.83 \pm 0.07	0.32 \pm 0.04	0.14 \pm 0.03	9.64 \pm 0.1
$\overline{\text{DG}}$	0.44	0.43	0.65	0.27	0.35	9.45
<DGm>	0.13 \pm 0.01	0.13 \pm 0.01	0.19 \pm 0.01	0.06 \pm 0.01	0.12 \pm 0.01	9.40 \pm 0.08
$\overline{\text{DGm}}$	0.21	0.23	0.15	0.22	0.39	9.21
<RDDG>	0.09 \pm 0.01	0.10 \pm 0.01	0.11 \pm 0.02	0.07 \pm 0.01	0.11 \pm 0.01	9.21 \pm 0.05
$\overline{\text{RDDG}}$	0.23	0.22	0.07	0.29	0.42	9.00
(RDDG)m	0.08	0.09	0.09	0.05	0.12	9.36

The notation of the structures is the same as that in Table I. In the case of the interproton distances, 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

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

In the case of the disulphide bridge restraints, the rmsd is calculated with respect to single target values.

Table III. Energies of the structures

Structures	Energy (kcal/mol)									
	Total	Bond	Angle	Dihedral	Improper	Van der Waals	Electrostatic	H-bond	Distance restraints	Φ torsion angle restraints
		(675)	(1222)	(312)	(139)				(326)	(27)
<DG>	8040 \pm 1200	79 \pm 9	455 \pm 83	326 \pm 22	0.2 \pm 0.3	1286 \pm 1055	-27 \pm 44	-10 \pm 3	4953 \pm 645	975 \pm 513
<DGm>	412 \pm 74	50 \pm 7	431 \pm 33	273 \pm 22	31 \pm 2	60 \pm 29	-593 \pm 52	-31 \pm 7	184 \pm 21	8 \pm 3
<RDDG>	47 \pm 68	40 \pm 6	366 \pm 45	245 \pm 23	32 \pm 4	21 \pm 24	-775 \pm 45	-49 \pm 6	161 \pm 22	5 \pm 2
(RDDG)m	106	45	425	262	51	22	-776	-52	128	0.7

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 potentials and for the effective distance and Φ torsion angle restraints potentials is given in parentheses. In addition to the NOE interproton distance effective potential, the distance restraints energy includes effective potentials for the four hydrogen bonding distance restraints and the 12 disulphide bridge distance restraints. The disulphide bridge effective potential is a simple harmonic potential (see Equation 1), whereas all the other restraints effective potentials are square well potentials (see Equations 2 and 3). The restraints force constants had values of 40 kcal/mol/ \AA^2 for the NOE interproton distance and hydrogen bonding restraints, 40 kcal/mol/ rad^2 for the Φ backbone torsion angle restraints and 200 kcal/mol/ \AA^2 for the disulphide bridge restraints.

structures (Havel and Wüthrich, 1985; Clore *et al.*, 1986c), is reduced in the DGm(i) structures (Table II). The mean structures $\overline{\text{DG}}$ and $\overline{\text{DGm}}$ about which the DG(i) and DGm(i) structures are distributed, however, are very similar, the atomic rms differences between them having approximately the same value as the atomic rms standard errors in their coordinates.

The second phase of the refinement, comprising the restrained molecular dynamics calculations, results in yet further significant improvements in both interproton distance deviations (Table II) and energies, particularly in the non-bonding terms (Table III). As expected the average atomic rms shifts from the DGm(i) to the RDDG(i) structures are larger than those from the DG(i) to the DGm(i) structures, and the mean structure $\overline{\text{RDDG}}$ about which the RDDG(i) structures are distributed is significantly different from either $\overline{\text{DG}}$ or $\overline{\text{DGm}}$ (Table I). The average radius of gyration of the RDDG(i) structures is slightly smaller than that of the DGm(i) structures; this is probably a manifestation of the increased contribution of the attractive component of the van der Waals energy.

The atomic rms distribution of all atoms, backbone atoms and side chain atoms of the RDDG(i) structures about the mean structure $\overline{\text{RDDG}}$ is shown in Figure 3, as a function of residue number, and plots of the Φ and Ψ backbone torsion angles are shown in Figure 4. Considering the backbone atoms, only four

regions are relatively ill defined: the first residue of the N terminus, residues 8 and 9 in the first turn, residues 18–20 comprising the end of helix A and the beginning of the second turn, and residues 41–45 comprising the fifth turn and C terminus. This is manifested by the larger atomic rms and Φ, Ψ torsion angle distributions for these regions. The precision with which the side chain atoms are defined is slightly worse than that for the backbone atoms. This is particularly marked for residues at the surface for which there are either no or very few NOE restraints. The atomic rms distribution for the sidechains of residues within the protein interior, however, is relatively small ($< 2 \text{\AA}$).

In stereochemical terms, the mean structure RDDG, not surprisingly, is rather poor, both with respect to bond lengths and angles and to non-bonded contacts. We therefore subjected $\overline{\text{RDDG}}$ to 1500 cycles of restrained energy minimization in which the van der Waals radii were slowly increased from a quarter of their usual values to their full values (Clore *et al.*, 1986c). This resulted in the structure $(\overline{\text{RDDG}})_m$ which in energy terms is as good as the individual RDDG(i) structures (Table II). $(\overline{\text{RDDG}})_m$ is rms shifted by 1.0 and 1.2 \AA for the backbone and all atoms, respectively, from $\overline{\text{RDDG}}$ but is still much closer to RDDG than to the two other mean structures $\overline{\text{DG}}$ and $\overline{\text{DGm}}$ (Table II). A stereoview of $(\overline{\text{RDDG}})_m$ showing both backbone and side chain atoms is shown in Figure 2c and d.

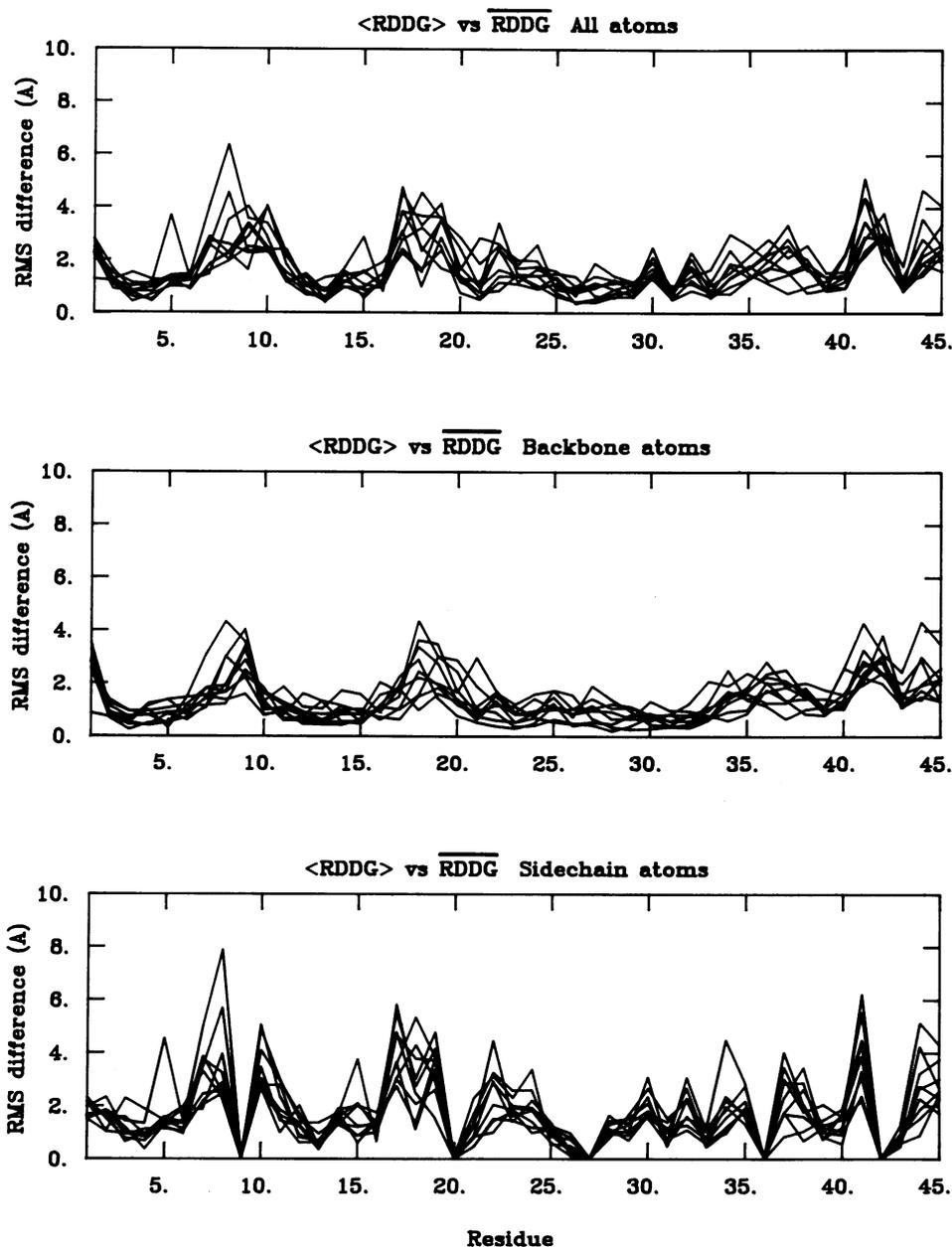


Fig. 3. Atomic rms distribution of the nine RDDG structures about the mean $\overline{\text{RDDG}}$ structure for all atoms, backbone atoms and side chain atoms as a function of residue number.

Structural features of $\alpha 1$ -purothionin

The principal features of the 3D solution structure of $\alpha 1$ -purothionin are illustrated in Figure 1b–d. The overall shape of the molecule (Figure 1b) is that of the capital letter L (shown in the \perp orientation), similar to that of crambin, with the longer arm comprising the two helices and the shorter one the mini anti-parallel β sheet and the C-terminal residues (35–45). The angle between the long axes of the two helices is $\sim 150^\circ$, and the angle between the plane formed by the two helices and the plane formed by the anti-parallel β -sheet is $\sim 50^\circ$.

The distribution of hydrophobic and hydrophilic residues is shown in Figure 2c. The hydrophobic residues are principally concentrated on the outer surface of the two helices which may represent the site of interaction of $\alpha 1$ -purothionin with lipid membranes, given that it exhibits haemolytic activity and lyses a wide

variety of mammalian cells (Anderson and Johansson, 1973). The outer surface of the corner of the L as well as the under surface of the shorter arm are hydrophilic with the exception of two hydrophobic residues, Pro 40 and Phe 43. The inner surface of the L is also mainly hydrophilic with a concentration of positively charged residues.

The distribution of hydrogen bonds in the structure $(\overline{\text{RDDG}})_m$ is shown in Figure 1d. In addition to the backbone hydrogen bonds stabilizing the two helices (residues 10–19 and 23–28) and the mini anti-parallel β sheet (residues 3–5 and 31–34), there are two tight turns stabilized by CO(i)-NH(i+3) backbone hydrogen bonds: the first is a type I turn comprising residues 6–9, and the second a type II turn comprising residues 40–43. The other three turns (residues 20–22, 29–30 and 35–38) are not classical in nature. There are also two long range backbone

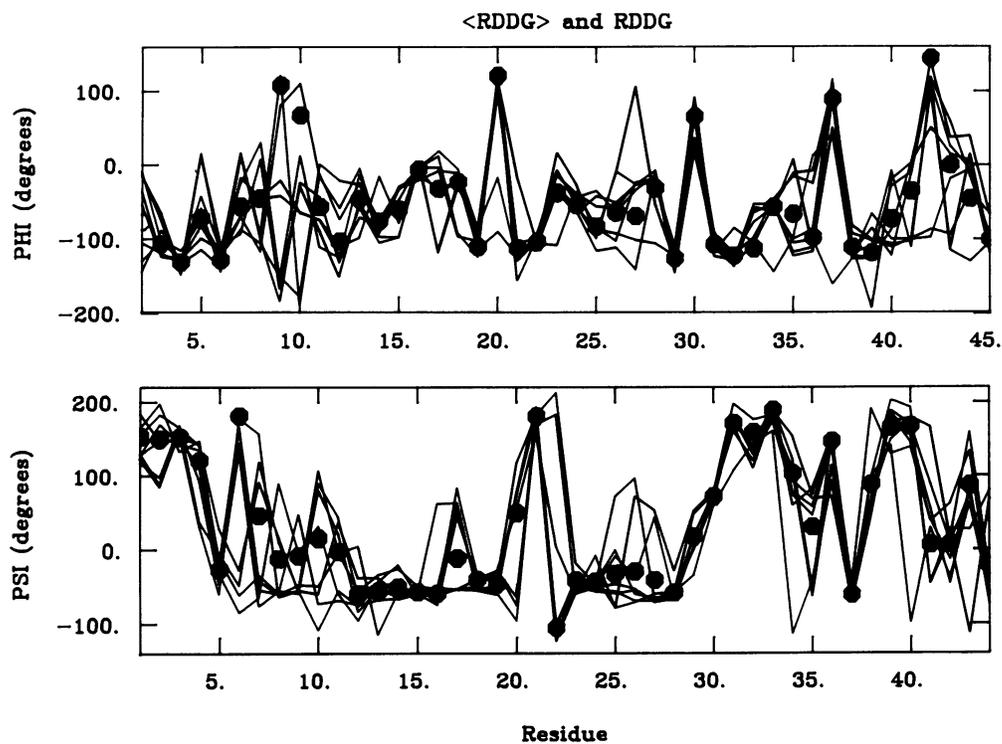


Fig. 4. Φ and Ψ backbone torsion angles of the nine RDDG structures (—) and the mean $\overline{\text{RDDG}}$ (●) structure as a function of residue number.

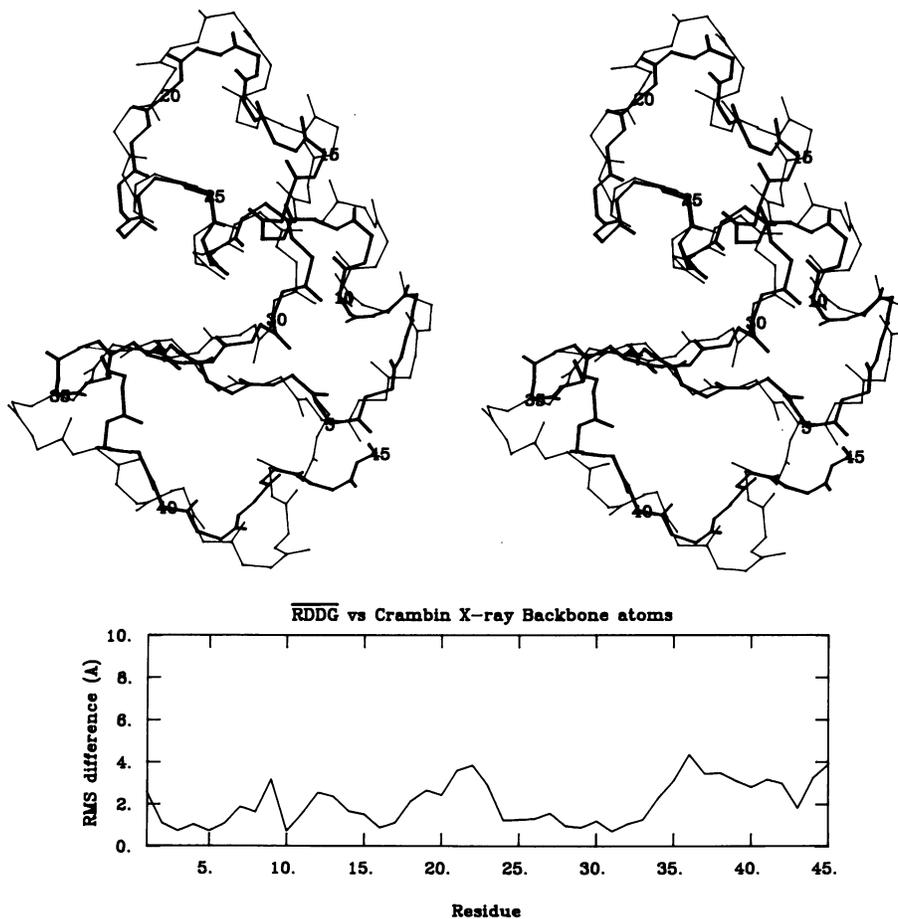


Fig. 5. Comparison of the solution structure of $\alpha 1$ -purothionin and the 1.5 Å resolution crystal structure of crambin (Hendrickson and Teeter, 1981). (a) Best fit superposition of the backbone (C,N,C α ,O) atoms of $\overline{\text{RDDG}}$ (thick line) and crambin (thin line). (b) Atomic rms difference of the backbone (C,C α ,N,O) atoms between RDDG and crambin as a function of residue number.

hydrogen bonds between the NH and CO groups of residues 4 and 43 and residues 45 and 5. The guanidinium groups of the five Arg residues are all involved in local ($|i-j| \leq 4$) hydrogen bonds to backbone carbonyl oxygen atoms. The guanidinium group of Arg 5 is also involved in a long range hydrogen bond to the backbone carbonyl oxygen atom of Arg 30. Finally, there is a hydrogen bond between the O^γH of Ser 34 and the C^δO of Gln 22 which bridges the two arms of the L.

Comparison with the X-ray structure of crambin

One of the initial aims of this study was to determine how the amino acid sequence homology between α 1-purothionin and crambin was reflected in their 3D structures. The best fit superposition of the backbone (C, C^α, N, O) atoms of RDDG and the 1.5 Å resolution crystal structure of crambin (Hendrickson and Teeter, 1981) is shown in Figure 5 together with a plot of the atomic rms difference between them as a function of residue number (with the alignment given in Figure 1). It is clear from this figure that the two structures are similar, the best fit overall atomic rms difference between the backbone atoms being 2.3 Å. This value is slightly smaller than that between crambin and the \overline{DG} , \overline{DGm} and $\overline{(RDDG)m}$ structures (Table I), but significantly larger than the values (< 1.2 Å) between crambin and the mean structures derived from the collection of restrained molecular dynamics (Clare *et al.*, 1986b; Brünger *et al.*, 1986a) and distance geometry (Clare *et al.*, 1986c) structures generated in model calculations using interproton distances derived from the crambin X-ray structure, similar in quality and quantity to those used here.

Concluding remarks

We have presented the 3D structure of α 1-purothionin in solution as determined from NOE interproton distance and ϕ backbone torsion angle restraints using a combination of distance geometry and restrained molecular dynamics calculations. A test of the quality of the structures obtained should soon be available as 2.5 Å resolution X-ray diffraction data on purothionin have been collected (M.M. Teeter and M. Whitlow, unpublished data). As no heavy atom derivatives have been obtained to date, an attempt is being made to solve the crystal structure directly by Patterson search techniques (Lattman, 1985; Machin, 1985) using the structures obtained in this paper as starting models to first obtain the orientation and position of the molecule in the unit cell and then to determine the initial X-ray phases for the calculation of an electron density map. A molecular replacement calculation based on NMR model structures for crambin has shown that a refined X-ray structure can potentially be obtained in this manner (Brünger *et al.*, 1986b).

Materials and methods

Samples for n.m.r. spectroscopy contained 6.8 mM α 1-purothionin (purified from Durum wheat as described by Mak and Jones 1976a, 1976b) in 500 mM sodium phosphate buffer pH 4.0. NOESY spectra (Jeener *et al.*, 1979; Macura *et al.*, 1981) were recorded in the pure phase absorption mode (Marion and Wüthrich, 1983) using the experimental conditions reported previously (Clare *et al.*, 1986a).

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

Acknowledgements

We thank the Max-Planck-Institut für Plasma Physik (Garching) for CRAY-XMP computing facilities. This work was supported by the Max-Planck-Gesellschaft and Grant No C1 86/1-1 from the Deutsche Forschungsgemeinschaft (GMC and AMG). M.N. acknowledges a pre-doctoral fellowship of the Max-Planck Gesellschaft.

References

- Anderson, K. and Johannson, M. (1983) *Eur. J. Biochem.*, **32**, 223–231.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.
- Brucoleri, R.E. and Karplus, M. (1986) *J. Comput. Chem.*, **7**, 165–175.
- Brünger, A.R., Clare, G.M., Gronenborn, A.M. and Karplus, M. (1986a) *Proc. Natl. Acad. Sci. USA*, **83**, 3801–3805.
- Brünger, A.T., Campbell, R.L., Clare, G.M., Gronenborn, A.M., Karplus, M., Petzko, G. and Teeter, M.M. (1986) *Science*, in press.
- Clare, G.M., Gronenborn, A.M., Brünger, A.T. and Karplus, M. (1985) *J. Mol. Biol.*, **185**, 435–455.
- Clare, G.M., Sukumaran, D.K., Gronenborn, A.M., Teeter, M.M. and Whitlow, M. (1986a) *J. Mol. Biol.*, in press.
- Clare, G.M., Brünger, A.T., Karplus, M. and Gronenborn, A.M. (1986b) *J. Mol. Biol.*, **191**, in press.
- Clare, G.M., Nilges, M., Brünger, A.T., Karplus, M. and Gronenborn, A.M. (1986c) *J. Mol. Biol.*, in press.
- Crippen, G.M. and Havel, T.F. (1978) *Acta Crystallogr.*, **A34**, 282–284.
- Havel, T.F. (1986) DISGEO, Quantum Chemistry Program Exchange Program. No. 507, Indiana University.
- Havel, T.F. and Wüthrich, K. (1984) *Bull. Math. Biol.*, **46**, 673–698.
- Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 281–294.
- Hendrickson, W.A. and Teeter, M.M. (1981) *Nature*, **290**, 107–113.
- Jeener, J., Meier, B.H., Bachmann, P. and 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. and Cooper, D.B. (1982) *J. Hered.*, **73**, 143–144.
- Kaptein, R., Zuiderweg, C.R.P., Scheek, R.M., Boelens, R. and van Gunsteren, W.F. (1985) *J. Mol. Biol.*, **182**, 179–182.
- Karplus, M. and McCammon, J. (1983) *Annu. Rev. Biochem.*, **53**, 263–300.
- Kline, A.D., Braun, W. and Wüthrich, K. (1986) *J. Mol. Biol.*, **189**, 377–382.
- Lattman, E. (1985) *Methods Enzymol.*, **115**, 55–80.
- Machin, P.A., ed. (1985) *Molecular Replacement – Proceedings of the Daresbury Study Weekend*. Science and Engineering Research Council.
- Macura, S., Huang, Y., Suter, D. and Ernst, R.R. (1981) *J. Magn. Resonance*, **43**, 259–281.
- Mak, A.S. and Jones, B.L. (1976a) *Can. J. Biochem.*, **54**, 835–842.
- Mak, A.S. and Jones, B.L. (1976b) *J. Sci. Food Agric.*, **27**, 205–213.
- Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
- Nilsson, L., Clare, G.M., Gronenborn, A.M., Brünger, A.T. and Karplus, M. (1986) *J. Mol. Biol.*, **188**, 455–475.
- Pardi, A., Billeter, M. and Wüthrich, K. (1984) *J. Mol. Biol.*, **180**, 741–751.
- Sippl, M.J. and Scheraga, H.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2283–2287.
- Teeter, M.M. and Kossiakoff, A.A. (1982) In Schoenborn, V. (ed.), *Neutrons in Biology*. Plenum Press, NY, pp. 335–348.
- Teeter, M.M., Mazer, J.A. and L'Italien, J.J. (1981) *Biochemistry*, **20**, 5437–5443.
- Van Elten, C.H., Nielsen, H.C. and Peters, J.E. (1965) *Phytochemistry*, **4**, 219–223.
- Whitlow, M. and Teeter, M.M. (1985) *J. Biomol. Struct. Dynam.*, **2**, 831–848.
- Williams, R.W. and Teeter, M.M. (1984) *Biochemistry*, **23**, 6796–6802.
- Williamson, M.P., Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 295–315.
- Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **160**, 949–961.

Received on 25 July 1986