

## Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry

Tadeusz A. HOLAK<sup>1</sup>, Michael NILGES<sup>1</sup>, James H. PRESTEGARD<sup>2</sup>, Angela M. GRONENBORN<sup>1</sup> and G. Marius CLORE<sup>1</sup>

<sup>1</sup> Max-Planck-Institut für Biochemie, Martinsried bei München

<sup>2</sup> Department of Chemistry, Yale University, New Haven, CT

(Received February 4, 1988) – EJB 88 0148

The solution conformation of acyl carrier protein from *Escherichia coli* (77 residues) has been determined on the basis of 423 interproton-distance restraints and 32 hydrogen-bonding restraints derived from NMR measurements. A total of nine structures were computed using a hybrid approach combining metric matrix distance geometry and dynamic simulated annealing. The polypeptide fold is well defined with an average backbone atomic root-mean-square difference of  $0.20 \pm 0.03$  nm between the final nine converged structures and the mean structure obtained by averaging their coordinates. The principal structural motif is composed of three helices: 1 (residues 3–12), 2 (residues 37–47) and 4 (residues 65–75) which line a hydrophobic cavity. Helices 2 and 4 are approximately parallel to each other and anti-parallel at an angle of  $\approx 150^\circ$  to helix 1. The smaller helix 3 (residues 56–63) is at an angle of  $\approx 100^\circ$  to helix 4.

Acyl carrier proteins (ACP) play an essential role in fatty acid biosynthesis [1–3]. There is also recent evidence suggesting that ACP has a function in the biosynthesis of membrane-derived oligosaccharides in *Escherichia coli* [4]. The discovery of two distinct roles for such a small protein enhances the interest in its structural characterization. Among all known ACPs, the best characterized is that from *E. coli* [1, 5]. It is a soluble protein of  $M_r$  8847 containing 77 amino acids, a large proportion of which are acidic ( $\approx 27\%$ ) and a small proportion ( $\approx 8\%$ ), predominantly clustered at the NH terminus, being positively charged. The amino acid sequences of ACP from *E. coli* strains E26 and K12 are known [5] (and S. Jackowski, J. E. Cronan and C. O. Rock, personal communication); the 4'-phosphopantetheine prosthetic group is known to be linked to Ser-36. Although crystals yielding suitable diffraction data for ACP have been obtained [6], no crystal structure has yet been published.

Recently, the <sup>1</sup>H-NMR spectrum of the free sulfhydryl species (ACPSH) has been assigned using two-dimensional NMR spectroscopy and its approximate secondary structure deduced on the basis of a qualitative analysis of short-range nuclear Overhauser enhancement (NOE) effects [7]. ACPSH was suggested to have three helices extending from residues 3–15, 37–51, and 65–75, with the possibility of the existence of a short  $\alpha$ -helix between residues 56–63. In addition, two short segments comprising residues 3–15 and 26–36 were examined individually using an energy minimization approach incorporating NMR restraints as pseudo-energies [8].

Correspondence to G. M. Clore, Laboratory of Chemical Physics, Building 2, Room 123, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

Abbreviations. ACP, acyl carrier protein; ACPSH, free sulfhydryl species of ACP; NOESY, two-dimensional NOE spectroscopy; rms, root mean square; SA, simulated annealing.

In the present paper we extend these studies to the determination of the full three-dimensional structure of ACP based on 455 approximate distance restraints derived from NMR data and calculations using a hybrid metric matrix distance geometry and dynamic simulated annealing approach [9].

### MATERIALS AND METHODS

ACP was isolated and purified as described previously [10, 11]. NOESY spectra were recorded on a Bruker WM-500 spectrometer using the experimental procedures and conditions reported elsewhere [8]. Intensities of the NOE cross peaks were determined from volume integrals [12].

Metric matrix distance geometry calculations were carried out using the program DISGEO [13] on a VAX 8550 computer. All minimization and dynamic simulated annealing calculations were carried out with the program XPLOR (A. T. Brünger, unpublished data) [14, 15] either on a VAX 8550 or on a CONVEX C1-XP. XPLOR is a version of the program CHARMM [16] which has been especially adapted for restrained molecular dynamics and simulated annealing. Displaying of the structures was carried out using a modified version of the function network of FRODO [17] interfaced with XPLOR on an Evans & Sutherland PS390 colour graphics system. The smooth backbone atom representations shown in Fig. 1 were obtained as described in [18].

### RESULTS AND DISCUSSION

#### Distance restraints

The basis of the calculations consisted of 423 approximate interproton distance restraints derived from pure phase absorption two-dimensional NOE (NOESY) spectra recorded with a mixing time of 80 ms. These comprised 279 short-range ( $|i-j| \leq 5$ ) and 48 long-range ( $|i-j| > 5$ ) interresidue restraints

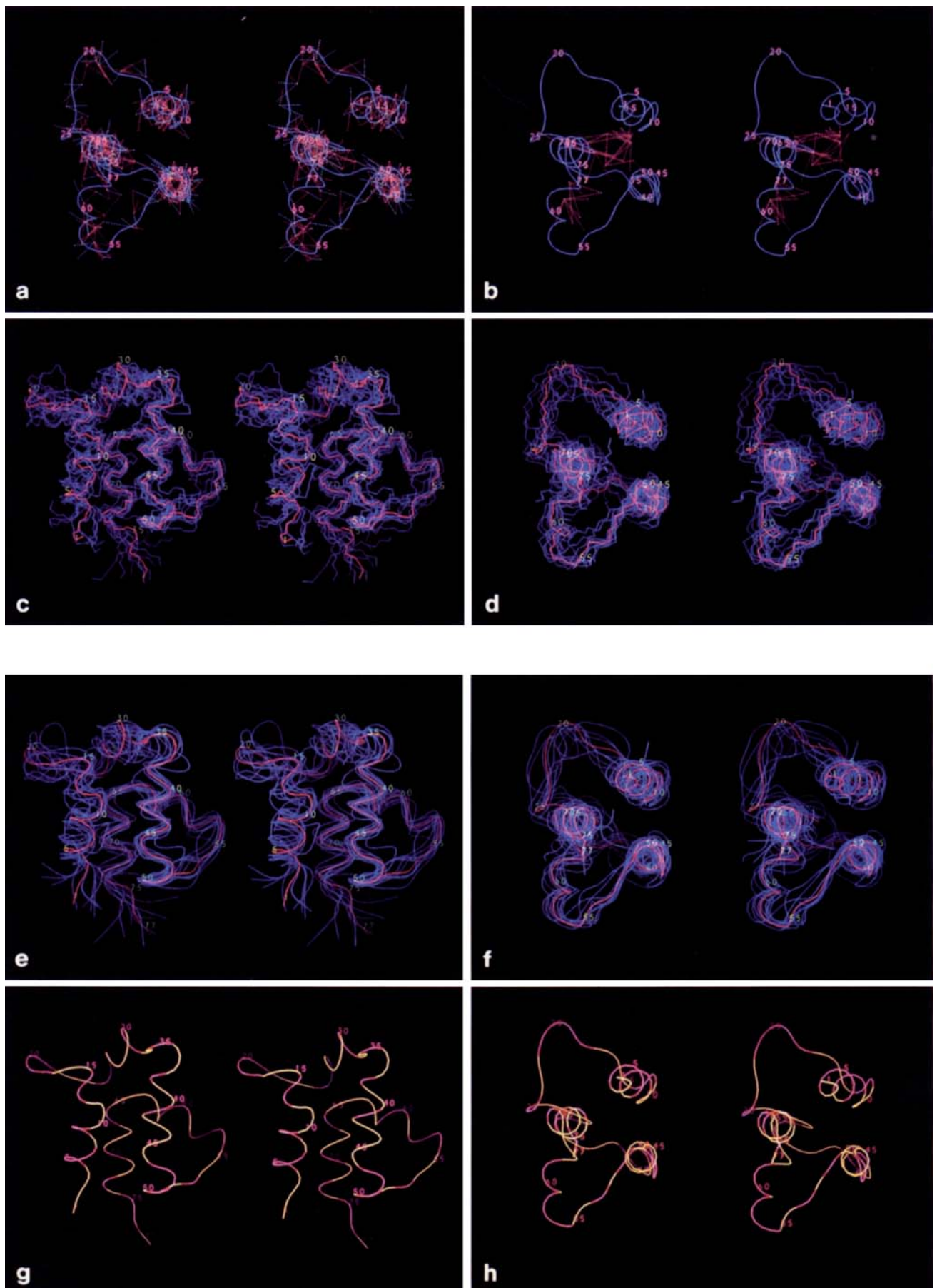


Fig. 1

and 96 intraresidue restraints, which were classified into four distance ranges, 0.18–0.26 nm, 0.18–0.31 nm, 0.18–0.37 nm and 0.18–0.50 nm. In the case of distances involving methylene, methyl or aromatic protons of Tyr and Phe, corrections were added to the upper limits of the distance restraints, as described in [19], to account for the pseudo-atom representation and centre averaging used in the structure calculations. The distance classification was carried out by converting the intensities of the NOESY cross peaks into interproton distances using a proportionality constant calculated from (a) the average intensities of the  $C^\alpha H(i)$ -NH( $i$ ) and NH( $i$ )-NH( $i+1$ ) NOEs in  $\alpha$ -helices which correspond to distances of about 0.27 and 0.28 nm, respectively, and from (b) the NOEs between the  $C^\alpha H_2$  geminal protons ( $r_{ij} = 0.18$  nm) of the two glycines in ACP whose NOE cross peaks showed first-order behavior. If only the  $C^\alpha H(i)$ -NH( $i$ ) NOEs are used as a single calibration distance of 0.27 nm, the NH( $i$ )-NH( $i+1$ ) and  $C^\alpha H_2$  glycine distances are calculated to be 0.28 and 0.187 nm, respectively. The interproton distance restraints were supplemented by 32 restraints for the 16 intrahelical NH( $i+4$ )-O( $i$ ) hydrogen bonds identified on the basis of slowly exchanging amide protons and the existence of  $C^\alpha H(i)$ -NH( $i+3,4$ ) and  $C^\alpha H(i)$ - $C^\beta H(i+3)$  NOEs. For each hydrogen bond the N( $i+4$ )-O( $i$ ) and NH( $i+4$ )-O( $i$ ) distances were restrained to ranges of 0.23–0.33 nm and 0.13–0.23 nm, respectively.

### Tertiary structure computation

The tertiary structure computation involved the use of a hybrid method combining distance geometry and dynamic simulated annealing [9].

The first two stages of the calculation are based on a metric matrix distance geometry algorithm operating in  $n$ -dimensional distance space [20–24] and make use of the program DISGEO [13]. [Note that the distance space dimensionality  $n$  is equal to the number of atom pairs given by  $N(N-1)/2$  where  $N$  is the number of atoms.] In the first stage a complete set of bounds on the distances between all atoms of the molecule is determined by triangulation from the experimental restraints and from the distance and planarity restraints obtained from the primary structure. The latter consist of assumed exact distances between all covalently bonded and geminal pairs of atoms, as well as lower limits on the distances between all pairs of atoms more than three bonds apart which are assumed to be no smaller than the sum of the hard-sphere Van der Waals radii. In the second stage, a set of random substructures is embedded, consistent with the bounds corresponding to distances between a subset of atoms comprising the main-chain C, N and  $C^\alpha H$  atoms together with all non-terminal  $C^\beta$  and  $C^\gamma$  atoms. This consists of randomly generating a set of distance values which lie within the distance bounds, followed by projection from  $n$ -dimensional distance space into cartesian coordinate space and 300 cycles of conjugate gradient minimization of the resulting coordinates. Because the randomly chosen distances are not checked with

respect to the triangle inequalities, this procedure is extremely fast requiring only about 15 min per substructure on a VAX 8550 computer. The coordinates of the substructures are a good approximation to those of the complete structures consistent with all the data and are used as the starting point for the subsequent calculations.

The third stage involves the application of a protocol of dynamic simulated annealing [9] in real cartesian coordinate space. This is similar in spirit to restrained molecular dynamics [14, 15, 25–29], but the form of the target function is much simpler than the full empirical energy function employed in conventional molecular dynamics. The total target function, which in effect represents the potential energy of the system (whose units in the present calculations are  $\text{kJ} \cdot \text{mol}^{-1}$ , although this choice is purely arbitrary), is made up of the following terms:

$$F_{\text{tot}} = F_{\text{covalent}} + F_{\text{repel}} + F_{\text{NOE}} \quad (1)$$

where  $F_{\text{covalent}}$  is the target function for maintaining correct bond lengths, angles, planes and chirality, and is given by

$$F_{\text{covalent}} = \sum_{\text{bonds}} k_b(r-r_0)^2 + \sum_{\text{angles}} k_\theta(\theta-\theta_0)^2 + \sum_{\text{impropers}} k_\omega(\omega-\omega_0)^2. \quad (2)$$

The values chosen for the force constants for the bond, angle and improper terms are  $2.508 \times 10^5 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ ,  $2.09 \times 10^5 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{rad}^{-2}$  and  $2.09 \times 10^5 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{rad}^{-2}$ , respectively. (Note that the improper torsion term is responsible for maintaining planarity, correct chirality, and planar and *trans* peptide bonds.)  $F_{\text{repel}}$  is a simple Van der Waals repulsion term used to prevent unduly close non-bonded contacts and is given by:

$$F_{\text{repel}} = \begin{cases} 0 & , \text{ if } r \geq s \cdot r_{\text{min}} \\ k_{\text{rep}}(s \cdot r_{\text{min}}^2 - r^2)^2 & , \text{ if } r < s \cdot r_{\text{min}} \end{cases} \quad (3)$$

The values of  $r_{\text{min}}$  are the standard values of the Van der Waals radii. Finally, the restraints potential  $F_{\text{NOE}}$  is a square-well potential [28, 30].

The protocol employed for the dynamic simulated annealing is given in Table 1 and takes about 2.5 h/structure on a VAX 8550 computer. The starting coordinates are obtained by best-fitting one residue at a time to the subset of atoms present in the substructures and subjecting the resulting coordinates to 200 cycles of Powell minimization to regularize the peptide bond. This is followed by 3.75 ps of dynamics at 1000 K during which time the force constants for the NOE ( $k_{\text{NOE}}$ ) and Van der Waals repulsion ( $k_{\text{rep}}$ ) terms are increased from their initial values (4.18 and  $83.6 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ , respectively) to their final maximum values (1672 and  $2.09 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ , respectively), 1.5 ps of dynamics at 300 K, and finally 200 cycles of restrained Powell minimization to yield the final SA structures.

As described previously [27], the coordinates of the individual SA structures were also averaged to yield the mean

Fig. 1. Stereoviews of the backbone (N,  $C^\alpha$ , C) atoms of the final SA structures of ACP. (a) Superposition of the intra-residue (light blue) and short-range ( $|i-j| \leq 5$ ) inter-residue (red) interproton distance restraints on the restrained minimized average structure (SA)r (blue). (b) Superposition of the long-range ( $|i-j| > 5$ ) inter-residue (red) interproton distance restraints on (SA)r (blue). (c–f) Superpositions of the individual SA structures (blue) on (SA)r (red). (g, h) Distribution of the charged (red), polar (lilac) and hydrophobic (yellow) residues of (SA)r. In (c) and (e) the actual backbone (N,  $C^\alpha$ , C) atom positions are shown whereas in the other panels a smooth backbone atom representation is displayed

structure  $\overline{SA}$ . The very poor stereochemistry and non-bonded contacts of the mean structures were easily corrected by 1000 cycles of Powell restrained minimization with only minor atomic rms shifts to generate the structure  $(\overline{SA})_r$ . This structure is closer to the corresponding mean structure than any of the individual structures, while satisfying the experimental distance restraints to the same extent as the individual structures from which it is derived.

Table 1. Protocol used for the dynamic simulated annealing calculations

Phase	Protocol
1.	200-cycles Powell minimization $k_{\text{NOE}} = 0 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ $k_{\text{repel}} = 4.18 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ <sup>a</sup>
2.	3.75-ps dynamics at 1000 K <sup>b</sup> $k_{\text{NOE}} = 83.6 \rightarrow 2.09 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ <sup>c</sup> $k_{\text{repel}} = 4.18 \rightarrow 1672 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ <sup>a,d</sup>
3.	1.5-ps dynamics at 300 K <sup>e</sup> $k_{\text{NOE}} = 2.09 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ $k_{\text{repel}} = 1672 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ <sup>d</sup>
4.	200-cycles Powell minimization $k_{\text{NOE}} = 2.09 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ $k_{\text{repel}} = 1672 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ <sup>d</sup>

<sup>a</sup> The scale factor  $s$  for the Van der Waals radii used to calculate the Van der Waals repulsion term,  $F_{\text{repel}}$ , given by Eqn (3) is 0.8.

<sup>b</sup> The initial velocities in phase 2 are assigned to a Maxwellian distribution at 1000 K and the velocities are then rescaled every 75 fs to 1000 K.

<sup>c</sup> The value of the force constant  $k_{\text{NOE}}$  for the NOE square-well potential  $F_{\text{NOE}}$  is increased every 75 fs by doubling its value until it reaches a maximum value of  $2.09 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ .

<sup>d</sup> The value of the force constant  $k_{\text{repel}}$  for the Van der Waals repulsion potential  $F_{\text{repel}}$  is increased every 0.75 ps by multiplying its value by  $e^{0.12}$ .

<sup>e</sup> The velocities in phase 3 are rescaled every 150 fs to 300 K.

### The converged structures

A total of nine SA structures was calculated. The course of the calculations is summarized in Tables 2–4. Best-fit superpositions of the backbone (N, C $\alpha$ , C) atoms of the SA structures as well as superpositions of the interproton distance restraints on the restrained minimized average structure  $(\overline{SA})_r$  are shown in Fig. 1.

The overall polypeptide fold is already approximately correct in the nine substructures and the average atomic rms difference between them is quite small ( $\approx 0.2$  nm for the backbone atoms; Table 2). These structures, however, have very poor non-bonded contacts, and fail to satisfy a large number of the experimental distance restraints with an average of more than 100 violations greater than 0.05 nm and very high values for  $F_{\text{NOE}}$  (Tables 3 and 4). Subjecting the substructures to dynamic simulated annealing leads to considerable improvement. The experimental restraints are satisfied (Table 3), the deviations from idealized geometry are very small (Table 4), and the non-bonded contacts are good as judged both by the small values of  $F_{\text{repel}}$  as well as the negative values of the Lennard-Jones Van der Waals energy (Table 4).

Examination of the data in Table 2 and Figs 1 and 2 indicates that the overall conformation is well defined with an average backbone atomic rms difference (residues 2–75) of  $0.20 \pm 0.03$  nm between the individual SA structures and the mean  $\overline{SA}$  structure. From the superpositions of the short- and long-range interproton distances on the restrained minimized average structure  $(\overline{SA})_r$  shown in Fig. 1a and 1b, respectively, it is clear that the concentration of short-range NOEs is much larger in the helices than in the loops, and that the long-range NOEs, although few in number, are strategically placed to enable the relative orientation between the various structural elements to be determined. In particular there are long-range NOEs between helix 1 (residues 3–12) and helix 2 (residues 37–47), helix 1 and helix 4 (residues 65–75), helix 2 and helix 4, and helix 3 (residues 56–63) and helix 4, as well as between the first loop (residues 13–36) and helices 2, 3 and

Table 2. Atomic rms distributions and shifts

The notation of the structures is as follows:  $\langle \text{Sub} \rangle$  comprises the nine substructures to which all atoms have been added by best fitting one residue at a time to the subset of atoms present in the original substructures generated by the program DISGEO [13];  $\langle \text{SA} \rangle$  comprises the nine dynamic simulated annealing structures derived using the nine  $\langle \text{Sub} \rangle$  structures as starting structures;  $\overline{SA}$  is the mean structure obtained by averaging the coordinates of the individual SA structures, best fitted to residues 2–75;  $(\overline{SA})_r$  is the restrained minimized average structure derived from the mean  $\overline{SA}$  structure. The average rms difference and standard deviation between all pairs of individual structures is related to the average rms difference and standard deviation between the individual structures and the mean structure derived from them by a factor approximately equal to  $(2N/N-1)^{1/2}$  where  $N$  is the number of individual structures

Structures	Atomic rms differences of			
	residues 2–75		residues 3–13, 38–50, 56–75	
	backbone atoms	all atoms	backbone atoms	all atoms
	nm			
rms distributions				
$\langle \text{Sub} \rangle$ vs $\langle \text{Sub} \rangle$	$0.20 \pm 0.01$	$0.33 \pm 0.02$	$0.20 \pm 0.02$	$0.32 \pm 0.03$
$\langle \text{SA} \rangle$ vs $\langle \text{SA} \rangle$	$0.30 \pm 0.05$	$0.42 \pm 0.05$	$0.21 \pm 0.04$	$0.33 \pm 0.05$
$\langle \text{SA} \rangle$ vs $\overline{SA}$	$0.20 \pm 0.03$	$0.28 \pm 0.04$	$0.14 \pm 0.02$	$0.22 \pm 0.04$
$\langle \text{SA} \rangle$ vs $(\overline{SA})_r$	$0.22 \pm 0.05$	$0.31 \pm 0.05$	$0.15 \pm 0.03$	$0.24 \pm 0.05$
rms shifts				
$\langle \text{Sub} \rangle$ vs $\langle \text{SA} \rangle$	$0.41 \pm 0.04$	$0.51 \pm 0.03$	$0.39 \pm 0.03$	$0.49 \pm 0.02$
$\overline{SA}$ vs $(\overline{SA})_r$	0.09	0.13	0.06	0.10

4, and the second loop (residues 48–55) and helix 4. However, because the local structure of the loops is not as well defined as that of the helices, which essentially form rigid rods, a much larger number of long-range NOEs would be required to define their atomic positions accurately. It is therefore not surprising to find that the backbone atomic rms distributions are smallest for the helices and largest for the loops (Figs 1 c–f and 2; Table 2).

As expected the side-chain positions are not as well defined as the backbone positions (Table 2 and Fig. 2). This is not too surprising as there are very few buried residues. Indeed, excluding glycines, only 22 of the 77 residues have surface-accessible areas  $\leq 0.50 \text{ nm}^2$ , of which 19 have side chains whose positions are distributed within  $\leq 0.25 \text{ nm}$  of the mean SA structure (Fig. 2). In addition, there are 10 other side chains, principally located in helices 2 and 4, whose positions are distributed within  $\leq 0.25 \text{ nm}$  of the mean structure.

### Structural features of the computed structure

It is interesting to compare the secondary structure elements obtained from the calculations with those derived on the basis of a qualitative interpretation of the short-range NOE data [7]. In general there is very good agreement and the differences are only minor in nature, reflecting the deficiencies involved in the exact delineation of regular secondary structure elements, particularly  $\alpha$ -helices, based on a qualitative interpretation of short-range NOE data [31, 32]. The first  $\alpha$ -helix extends from residue 3 to residue 12 in the computed structures, slightly shorter than predicted previously (residues 3–15). Helix 1 is followed by a turn comprising residues 12–16 in the computed structures. The predicted tight loop between residues 26–33 with a short contact between Ala<sup>34</sup>NH-Val<sup>29</sup>C<sup>2</sup>H of 0.23 nm is present in all the computed structures. This left-handed loop leads into a right-handed

Table 3. The rms differences between target and calculated values for the experimental interproton distance and hydrogen bonding restraints. The notation of the structures is the same as that in Table 2. 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

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

For distances involving methyl, methylene or aromatic protons the distances are calculated as arithmetic centre averages. The number of distance violations  $> 0.05 \text{ nm}$  (calculated with respect to the upper and lower limits) are given in parentheses next to the rms distance deviations

Structure	rms differences between calculated and target experimental distance restraints (violations $> 0.05 \text{ nm}$ )				
	all (455)	interresidue		intraresidue (96)	H-bond (32)
		short-range ( $ i-j  \leq 5$ ) (279)	long-range ( $ i-j  > 5$ ) (48)		
	nm				
<Sub>	0.1011 $\pm$ 0.005 (139.5 $\pm$ 9.4)	0.095 $\pm$ 0.009 (85.8 $\pm$ 5.7)	0.185 $\pm$ 0.053 (19.7 $\pm$ 3.7)	0.042 $\pm$ 0.002 (10.7 $\pm$ 1.6)	0.208 $\pm$ 0.030 (23.2 $\pm$ 2.9)
<SA>	0.006 $\pm$ 0.001 (0.33 $\pm$ 0.5)	0.006 $\pm$ 0.001 (0.33 $\pm$ 0.5)	0.003 $\pm$ 0.001 (0)	0.004 $\pm$ 0.0004 (0)	0.0047 $\pm$ 0.003 (0)
SA	0.013 (10)	0.013 (7)	0.000 (0)	0.018 (3)	0.000 (0)
(SA)r	0.005 (0)	0.006 (0)	0.004 (0)	0.007 (0)	0.002 (0)

Table 4. Deviations from ideality, and values of  $F_{\text{NOE}}$ ,  $F_{\text{repeI}}$ , the Lennard-Jones Van der Waals energy ( $E_{L-J}$ ) and the radii of gyration. The notation of the structures is the same as that in Table 2. The force constants used to calculate  $F_{\text{NOE}}$  and  $F_{\text{repeI}}$  are  $2.09 \times 10^4 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$  and  $1672 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ , respectively. The Lennard-Jones Van der Waals energy,  $E_{L-J}$ , is calculated using the CHARMM empirical energy function [16]

Structure	Deviations from ideality			$F_{\text{NOE}}$	$F_{\text{repeI}}$	$E_{L-J}$	Radii of gyration
	bonds (1175)	angles (2133)	impropers (309)				
	nm	degrees					
<Sub>	0.0014 $\pm$ 0.0001	2.64 $\pm$ 0.23	1.064 $\pm$ 0.103	1.2 ( $\pm$ 0.2) $\times 10^5$	1.4 ( $\pm$ 0.2) $\times 10^5$	$> 10^7$	1.191 $\pm$ 0.008
<SA>	0.0013 $\pm$ 0.0002	3.31 $\pm$ 0.25	0.433 $\pm$ 0.075	296 $\pm$ 100	167 $\pm$ 96	-456 $\pm$ 188	1.241 $\pm$ 0.009
SA	0.0559	26.49	30.11	1685	3.1 $\times 10^5$	$> 10^7$	1.208
(SA)r	0.0015	5.37	0.384	255	117	-460	1.243

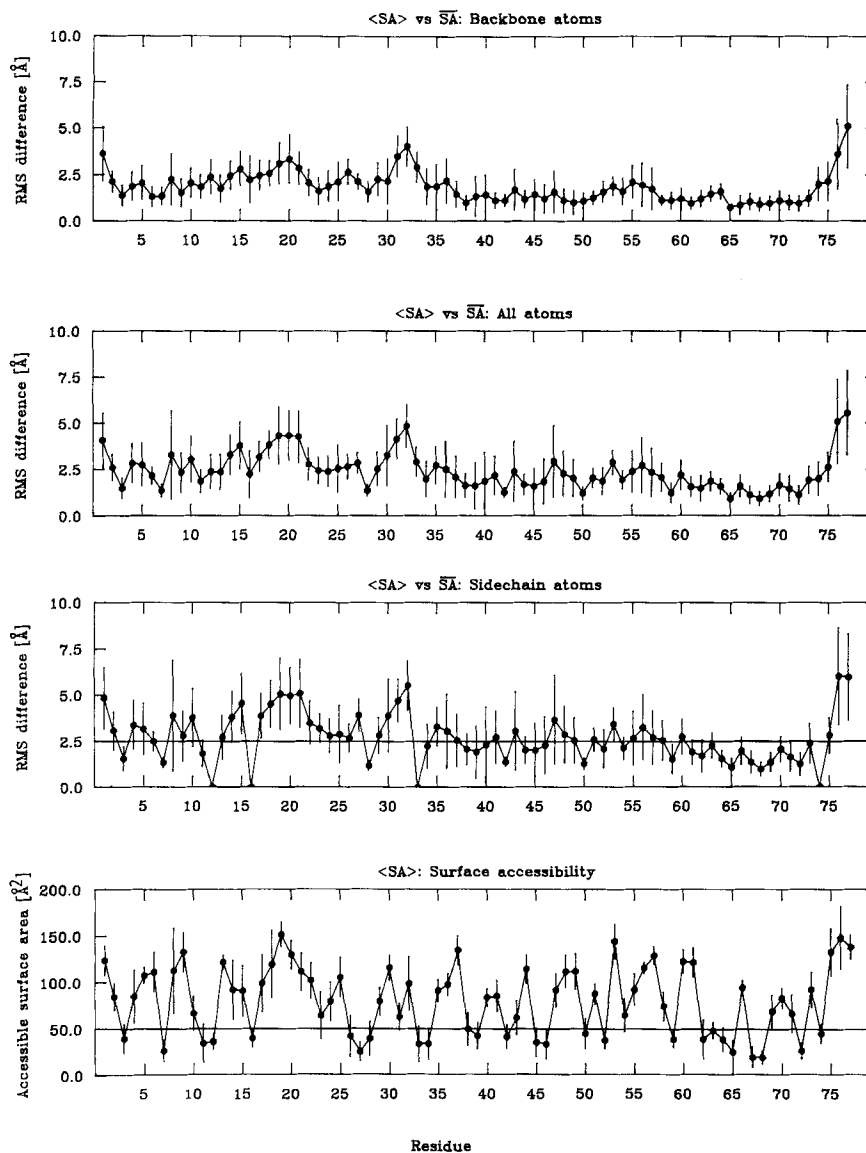


Fig. 2. Atomic rms distributions of the nine individual SA structures about the mean structure  $\overline{SA}$  best fitted to residues 2–75, and surface accessibility. The filled-in circles (●) represent the average rms difference at each residue between the individual structures and the mean structure, and the bars represent the standard deviations in these values; similarly for the surface accessibility.  $10 \text{ \AA}^2 = 0.1 \text{ nm}^2$

turn (residues 34–36) followed by the second  $\alpha$ -helix which extends from residue 37 to residue 47, also slightly shorter than predicted (residues 37–51). Helix 2 is followed by a turn (residues 48–52). The presence of a short  $\alpha$ -helix between residues 56 and 63 (helix 3) is now confirmed and the length of the last helix (helix 4) is the same as that predicted (residues 65–75).

The structure itself is dominated by helices 1, 2 and 4. Helices 2 and 4 are approximately parallel to each other and anti-parallel at an angle of about  $150^\circ$  to helix 1. When ACP is viewed down the axes of these three helices (Fig. 1 d, f, h) it can be seen that the three helices form an approximately equilateral triangle. The short helix 3 is at an angle of approximately  $100^\circ$  to helix 4, and together they form a helix-turn-helix motif not too dissimilar to that found in a number of DNA-binding proteins [33–36]. All four helices are amphipathic with hydrophobic residues lining their interior surface (Fig. 1 g, h). The three long helices form a hydrophobic

cavity with the back of the cavity lined by helix 4 and the sides by helices 1 and 2. This may provide the structural basis for understanding how ACP transports fatty acids. One should also bear in mind that the appearance of such a hydrophobic cavity is the result of the NOE data alone as there are no other factors in our methodology that could have skewed the distribution of residues in this way.

In the case of ACPSH the prosthetic group does not appear to make any contact with the protein other than the covalent linkage with Ser-36 as no NOEs could be detected between the prosthetic group and protein residues. The resonances for all of the protons of the prosthetic group are very sharp. These narrow line widths are due to the greater flexibility of the group resulting in a shorter effective correlation time and consequently narrower resonances for the residues. A substantial free motion for the entire group, together with the lack of the NOE contacts to the protein residues, suggest that in ACPSH the prosthetic group is extended away from the

protein into the bulk solution. When the SH of the prosthetic group is replaced by a fatty acid acyl chain, however, it seems probable that the fatty acyl moiety binds in the hydrophobic cavity formed by the three long helices. Preliminary data on the acylated species [37] support this hypothesis.

This work was supported by the *Max-Planck-Gesellschaft* and grant 321/4003/0318909A from the *Bundesministerium für Forschung und Technologie* (G. M. C. and A. M. G.) and by grant GM32243 from the National Institutes of Health (J. P.).

## REFERENCES

1. Prescott, D. J. & Vagelos, P. R. (1972) *Adv. Enzymol. Relat. Areas Mol. Biol.* **36**, 269–311.
2. Thompson, G. A. (1981) *The regulation of membrane lipid metabolism*, pp. 20–23, 33, CRC Press, Boca Raton, FL.
3. Wakil, S. J., Stoops, J. K. & Joshi, V. C. (1983) *Annu. Rev. Biochem.* **52**, 537–579.33.
4. Therisod, H., Weissborn, A. C. & Kennedy, E. P. (1986) *Proc. Natl Acad. Sci. USA* **83**, 7236–7240.
5. Vanaman, T. C., Wakil, S. J. & Hill, R. L. (1968) *J. Biol. Chem.* **243**, 6409–6431.
6. McRee, D. E., Richardson, J. S. & Richardson, D. C. (1985) *J. Mol. Biol.* **182**, 467–468.
7. Holak, T. A. & Prestegard, J. H. (1986) *Biochemistry* **25**, 5766–5774.
8. Holak, T. A., Prestegard, J. H. & Forman, J. D. (1987) *Biochemistry* **26**, 4652–4660.
9. Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988) *FEBS Lett.* **229**, 317–324.
10. Holak, T. A., Frederick, A. F. & Prestegard, J. H. (1987) *J. Biol. Chem.* **262**, 3685–3689.
11. Rock, C. O. & Cronan, J. E. Jr (1981) *Methods Enzymol.* **71**, 341–351.
12. Holak, T. A., Scarsdale, N. J. & Prestegard, J. H. (1987) *J. Magn. Reson.* **74**, 546–549.
13. Havel, T. F. (1986) *DISGEO, Quantum chemistry program exchange*, program no. 507, Indiana University.
14. Brünger, A. T., Kuryan, J. & Karplus, M. (1987) *Science (Wash. DC)* **235**, 458–460.
15. Brünger, A. T., Clore, G. M., Gronenborn, A. M. & Karplus, M. (1987) *Protein Eng.* **1**, 399–406.
16. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem.* **4**, 187–217.
17. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11**, 268–272.
18. Feldmann, R. J., Brooks, B. R. & Lee, B. (1986) *Tools for each age: understanding protein architecture through simulated unfolding*, Division of Computer Research and Technology, National Institutes of Health, Bethesda, MD.
19. Wüthrich, K., Billeter, M. & Braun, W. (1983) *J. Mol. Biol.* **169**, 949–961.
20. Crippen, G. M. & Havel, T. F. (1978) *Acta Crystallogr.* **A34**, 282–284.
21. Havel, T. F., Crippen, G. M. & Kuntz, I. D. (1979) *Biopolymers* **18**, 73–81.
22. Havel, T. F., Kuntz, I. D. & Crippen, G. M. (1983) *Bull. Math. Biol.* **45**, 665–720.
23. Havel, T. F. & Wüthrich, K. (1984) *Bull. Math. Biol.* **45**, 673–698.
24. Havel, T. F. & Wüthrich, K. (1985) *J. Mol. Biol.* **182**, 281–294.
25. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R. & Van Gunsteren, W. F. (1985) *J. Mol. Biol.* **182**, 179–182.
26. Clore, G. M., Brünger, A. T., Gronenborn, A. M. & Karplus, M. (1985) *J. Mol. Biol.* **186**, 435, 455.
27. Clore, G. M., Brünger, A. T., Karplus, M. & Gronenborn, A. M. (1986) *J. Mol. Biol.* **191**, 523–551.
28. Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M. & Gronenborn, A. M. (1986) *EMBO J.* **5**, 2729–2735.
29. Brünger, A. T., Clore, G. M., Gronenborn, A. M. & Karplus, M. (1986) *Proc. Natl Acad. Sci. USA* **83**, 3801–3805.
30. Clore, G. M., Sukumaran, D. K., Nilges, M. & Gronenborn, A. M. (1987) *Biochemistry* **26**, 1732–1745.
31. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, J. Wiley, New York.
32. Clore, G. M. & Gronenborn, A. M. (1987) *Protein Eng.* **1**, 275–288.
33. Anderson, W. F., Ohlendorf, D. H., Takeda, Y. & Matthews, B. W. (1981) *Nature (Lond.)* **290**, 754–758.
34. McKay, D. B. & Steitz, T. A. (1981) *Nature (Lond.)* **290**, 744–749.
35. Pabo, C. O. & Lewis, M. (1982) *Nature (Lond.)* **298**, 443–447.
36. Schevitz, R. W., Olwonowski, Z., Joachimiak, A., Lawson, C. L. & Sigler, P. (1985) *Nature (Lond.)* **317**, 782–786.
37. Jones, P.-J., Cioffi, E. H. & Prestegard, J. H. (1987) *J. Biol. Chem.* **262**, 8963–8965.