Mackinnon, C. M., Carter, P. E., Smyth, S. J., Dunbar, B., & Fothergill, J. E. (1987) *Eur. J. Biochem. 169*, 547–553.
March, S. C., Parikh, I., & Cuatrecasas, P. (1974) *Anal. Biochem. 60*, 149–152.
McKay, E. J., Laurell, A., Martensson, U., & Sjoholm, A. G. (1981) *Mol. Immunol. 18*, 349–357.
McRae, B. J., Lin, T. Y., & Powers, J. C. (1981) *J. Biol. Chem. 256*, 12362–12366.
Ohlin, A.-K., & Stenflo, J. (1987) *J. Biol. Chem. 262*, 13798–13804.
Poon, P. H., Schumaker, V. N., Phillips, M. L., & Strang, C. J. (1983) *J. Mol. Biol. 168*, 563–577.
Rapp, H. J., & Borsos, T. (1970) *Molecular Basis of Complement Action*, Appleton-Century-Crofts, New York.
Reid, K. B. M. (1983) *Biochem. Soc. Trans. 11*, 1–12.
Reid, K. B. M., Lowe, D. M., & Porter, R. R. (1972) *Biochem. J. 130*, 749–763.
Reid, K. B. M., Sim, R. B., & Faiers, A. P. (1977) *Biochem. J. 161*, 239–245.
Schumaker, V. N., Zavodszky, P., & Poon, P. H. (1987) *Annu. Rev. Immunol. 5*, 21–42.
Siegel, R. C., & Schumaker, V. N. (1983) *Mol. Immunol. 20*, 53–66.
Sim, R. B., Porter, R. R., Reid, K. B. M., & Gigli, I. (1977) *Biochem. J. 163*, 219–227.
Sim, R. B., Arlaud, G. J., & Colomb, M. G. (1980) *Biochim. Biophys. Acta 612*, 433–449.
Stenflo, J., Lundwall, A., & Dahlback, B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 368–372.
Tosi, M., Duponchel, C., Meo, T., & Julier, C. (1987) *Biochemistry 26*, 8516–8524.
Tschopp, J., Villiger, W., Fuchs, H., Kilchherr, E., & Engel, J. (1980a) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7014–7018.
Tschopp, J., Villiger, W., Lustig, A., Jaton, J. C., & Engel, J. (1980b) *Eur. J. Immunol. 10*, 529–535.
Valet, G., & Cooper, N. R. (1974a) *J. Immunol. 112*, 339–350.
Valet, G., & Cooper, N. R. (1974b) *J. Immunol. 112*, 1667–1673.
Villiers, C. L., Arlaud, G. J., & Colomb, M. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4477–4481.
Weiss, V., Fauser, C., & Engel, J. (1986) *J. Mol. Biol. 189*, 573–581.
Ziccardi, R. J. (1984) *Mol. Immunol. 22*, 489–494.
Ziccardi, R. J., & Tschopp, J. (1982) *Biochem. Biophys. Res. Commun. 107*, 618–623.

Three-Dimensional Structure of Acyl Carrier Protein Determined by NMR Pseudoenergy and Distance Geometry Calculations†

T. A. Holak,*,‡ S. K. Kearsley, Y. Kim, and J. H. Prestegard*
Department of Chemistry, Yale University, New Haven, Connecticut 06511
Received December 8, 1987; Revised Manuscript Received March 11, 1988

ABSTRACT: Distance constraints from two-dimensional NMR cross-relaxation data are used to derive a three-dimensional structure for acyl carrier protein from *Escherichia coli*. Several approaches to structure determination are explored. The most successful proves to be an approach that combines the early stages of a distance geometry program with energy minimization in the presence of NMR constraints represented as pseudopotentials. Approximately 450 proton to proton distance constraints including 50 long-range constraints were included in these programs. Starting structures were generated at random by the distance geometry program and energies minimized by a molecular mechanics module to give final structures. Seven of the structures were deemed acceptable on the basis of agreement with experimentally determined distances. Root-mean-square deviations from the mean of these structures for backbone atoms range from 2 to 3 Å. All structures show three roughly parallel helices with hydrophobic residues facing inward and hydrophilic residues facing outward. A hydrophobic cleft is recognizable and is identified as a likely site for acyl chain binding.

Acyl carrier protein (ACP)[1] is a small protein of interest because of its role in fatty acid synthesis. In *Escherichia coli* the protein is a cytosolic component of 8847 daltons that carries fatty acid chains via a thioester linkage to a phosphopantetheine prosthetic group as the chains are elongated by the fatty acid synthetase system (Ohlrogge, 1987; Thompson, 1980). Carriers are normally viewed as rather passive participants in biochemical events. However, in the case of ACP there is evidence to suggest that the presence of the protein can influence activities of enzymes associated with fatty acid synthesis. Consider, for example, thioesterases that cleave the fatty acid from the prosthetic group near the end of the elongation process. The phosphopantetheine prosthetic group is common to both ACP and coenzyme A, making acyl-CoAs homologous with acyl-ACPs for regions well displaced in both directions from the site of chain attachment. Yet, there appear to be significant differences in abilities of

† This work was supported by a research grant from the National Institutes of Health (GM32243), an instrumentation grant through the shared instrumentation program of the Division of Research Resources of the NIH (GM32243S1), and a grant of computer time at the Pittsburgh Supercomputer Center from the NSF (DMB-8612825).
‡ Present address: Max-Planck-Institut für Biochemie, 8033 Martinsried, München, West Germany.

[1] Abbreviations: ACP, acyl carrier protein; 2D-NMR, two-dimensional nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; AMBER, assisted model building with energy refinement; DISGEO, distance geometry; CoA, coenzyme A; RMS, root mean square.

0006-2960/88/0427-6135$01.50/0

these enzymes to distinguish acyl chains when they are bound to ACP as opposed to CoA (Barnes & Wakil, 1968; Spencer et al., 1978). Also, substituting ACP from one organism for ACP from another organism in an in vitro fatty acid synthetase system leads to significant variations in fatty acid distribution (Simoni et al., 1967). An interest in the structure of ACP and its possible modes of interaction with various acyl chains is, therefore, well justified.

Although there are both crystallographic and NMR studies in progress (McRee et al., 1985; Holak et al., 1987a), there is, as yet, no complete three-dimensional structure for the protein. We have reported over the last few years several stages in the development of a structure from one- and two-dimensional NMR work. This has included assignment of resonances in the proton NMR spectrum (Mayo et al., 1983; Holak & Prestegard, 1986), determination of secondary structure (Holak & Prestegard, 1986), and determination of three-dimensional characteristics of several short segments (Holak et al., 1987a). We present here a complete structure for the protein and a preliminary discussion of chemical aspects of potential importance in acyl chain binding. The determination of the structure has also provided an opportunity to explore various approaches to structure determination from cross-relaxation effects in two-dimensional NMR spectra. The structure presented is the result of a combined distance geometry (Williamson et al., 1985; Braun, 1987; Wagner et al., 1987) and pseudoenergy minimization scheme (Scarsdale et al., 1986; Clore et al., 1985; Kaptein et al., 1985; Billeter et al., 1987). This scheme seems to provide an approach to structure determination that may be useful in other applications.

Experimental Procedures

Distance Determination. Structural information in NMR data is extracted from the intensity of cross peaks in NOESY data sets (Wuthrich, 1986). These cross peaks arise because of the transfer of magnetization from one proton to another via through-space dipolar spin relaxation interactions. In the limiting case of a rigid spherical molecule, and sufficiently short periods of time for transfer, the intensities of these peaks are directly proportional to the inverse sixth power of the distances separating protons (Macura & Ernst, 1980). If a cross peak between protons at a known distance can be identified, ratios of the intensity for this cross peak to intensities for other cross peaks can be used to derive distances for other proton pairs. Real proteins and experimental conditions do not rigorously fulfill these limiting conditions, but the extreme distance dependence of cross relaxation allows rather severe violations without significant effects on distances derived (Olejniczak et al., 1984). In fact, much of the NMR-derived structural data on proteins has been obtained with rather crude classification of peak intensities in recognition of the fact that even approximate interproton distances may be adequate to define a structure.

In anticipation of more thorough theoretical descriptions of relaxation in the future, we have decided to treat cross-peak intensities as accurately as possible and accept any error that results from use of an approximate spin relaxation theory. We use a procedure that calculates volume integrals by combining column and row line integrals that have base lines chosen interactively (Holak et al., 1987b). Ratios of these volume integrals are equivalent to ratios of intensities. Intensities for known distances were taken from (a) the average intensities of the CαH(i)–NH(i) and NH(i)–NH(i + 1) NOEs in α helices, which correspond to distances of 2.7 and 2.8 Å, respectively, and from (b) the NOEs between the CαH$_2$ geminal protons (r_{ij} = 1.8 Å) of the two glycines in ACP whose NOE cross peaks showed first-order behavior. These were used to calculate a calibration constant for determination of unknown distances. If only the CαH(i)–NH(i) NOEs are used as a single calibration distance of 2.7 Å, the NH(i)–NH(i + 1) and CαH$_2$ glycine distances are calculated to be 2.8 and 1.9 Å, respectively.

Data were acquired on a 10 mM sample of ACP in the reduced, free sulfhydryl, form at pH 5.3 in 55 mM KH_2PO_4. The phase-sensitive NOESY experiment of States et al. was employed for data acquisition (States et al., 1982) using two different mixing times, 80 and 150 ms. Acquisition required approximately 72 h for the 150-ms set and 96 h for the 80-ms set. With the exception of a few long-range constraints, all distances were derived from the 80-ms data set, using the longer time data set largely to confirm the presence of very weak peaks and exclude peaks that clearly arose from secondary transfers. Partial analysis of this data has been presented previously (Holak et al., 1987a). Processing parameters are essentially the same as those used in our previous publication (Holak et al., 1987a), with the exception that the first data point of each column in the 2D data sets was multipled by a factor of 0.7 before Fourier transformation in the t_1 dimension in order to reduce t_1 ridges. This factor, which is somewhat larger than the theoretical value of 0.5, was found to provide better suppression (Otting et al., 1986).

Distances derived as described above can be used directly in many cases. However, for groups with high degrees of internal motion, such as a methyl group, and groups having protons that cannot be assigned stereospecifically, such as β-methylenes, it is more convenient to convert proton–proton distances to distances between corresponding pseudo atom or united atom representations of those protons (Wuthrich et al., 1983). The procedures we used for correcting distances are very similar to those in our previous publication except that pseudoatoms representing average proton positions in methyl groups and aromatic rings were introduced.

A total of 448 interproton distances were derived and assigned a sufficient confidence level to be used in the calculations. These can be classified as intraresidue, sequential, intermediate, and long-range distances. Intraresidue constraints constitute 101 of the 448 and are primarily amide to β-proton constraints and amide to γ-proton constraints. They were used only in the latter stages of our calculations. Sequential constraints number 205 of the 448 and are primarily amide to amide of i + 1 or i − 1 residue constraints, amide to α of the i − 1 residue constraints, and amide to β of the i − 1 residue constraints. Intermediate constraints number 87 of the 448 and include connectivities between residues removed by 2–4 positions in the sequence [$\leq(i + 5)$]. Both sequential and intermediate constraints are important in determining secondary structure. Long-range constraints number 55 and correspond to connectivities between protons on residues removed by more than 4 in the sequence. These prove to be very important in determining overall folding. Fifteen constraints were added for amide proton to amide i + 4 oxygen distances in cases where amide protons exchange slowly and belong to amino acids identified on a qualitative basis as being in an α-helical stretch. An additional 15 such NH(i)–O(i + 4) constraints, with lower weights, were added for residues with more rapid amide exchange rates that could still be assigned to the middle of regions with identifiable α-helical secondary structure. An additional 75 amide proton to amide i − 1 oxygen constraints were added to insure trans amide bond configurations in all amino acids except proline. A table

containing all the distances employed in our calculations is included as supplementary material.

Structure Solution. NMR distance constraints, no matter what the accuracy, seldom occur in sufficient number or sufficiently uniform distributions over the protein to independently define a structure. All approaches currently in the literature, therefore, supplement NMR data with sequence-derived information about which atoms are bonded to one another, what preferred bond angles exist, and what limits are imposed by van der Waals radii. In our initial attempts at structure solution, we used the very concise summary of chemical-bonding information contained in a readily available molecular modeling package AMBER[2] (Weiner & Kollman, 1981; Singh & Kollman, 1984). We will use this approach here both as an independent route to structure solution and as an adjunct to a distance geometry approach.

Utilization of a program like AMBER requires integration of NMR distance constraints with the normal energy terms describing molecular interactions. We have chosen to do this by representing distance constraints in terms of a pseudoenergy function written to mimic an error function for NOE-derived distances. This function is described more fully in previous publications (Holak et al., 1987a; Scarsdale et al., 1988). The united atom amino acid data base in AMBER also has to be modified slightly to allow specification of distances to discrete protons or pseudoprotons detected in NMR experiments but not normally included in representation of an amino acid. In our previous work we added discrete protons, giving them force field parameters derived from all atom descriptions of amino acids. We have since decided that it is better to leave the united atom force field unperturbed and add pseudoprotons, which are fixed in space relative to united atoms of the molecule but have no charge and no van der Waals interactions. This was done for α protons, β protons of valine and isoleucine, and γ protons of leucine.

We have utilized only the molecular mechanics module of AMBER in our calculations, but pseudoenergy approaches based on molecular dynamics are also possible (Clore et al., 1987). Molecular mechanics provides a rather efficient approach to finding a suitable structure by simply varying coordinates along a multidimensional path that leads to the nearest energy minimum. A molecular mechanics approach is susceptible to being trapped in local high-energy minima as opposed to a global minimum that fits all constraints. To minimize this possibility, minimization was effected in two stages. First, NMR energies were weighted heavily ($\sim$200 kcal/mol) and all molecular energy terms except those defining bonds and bond angles were neglected. This provides a simplified potential energy surface with fewer local minima. The highly asymmetric NMR pseudoenergy function we use also allows constraints involving residues close to one another to be satisfied early in the convergence with longer range constraints becoming more important later. In a second stage, minimization is effected after adding all energy terms. Pseudoenergy weightings are reduced in the midst of this step to a value we feel gives an appropriate balance of experimental and theoretical terms (for most contacts, $\sim$30 kcal/mol).

The calculations do require suitable starting structures. Our most successful attempts at applying a direct pseudoenergy approach began with incorporation of recognizable secondary structure elements (α helices) and an interactive folding to bring all long-range constraints within 15 Å of their ideal values. The first stages of the calculation were run on a VAX 11/750 or VAX 8600 with times for the 8600 being of the order of 1 h. The final minimizations with all energy terms present were done on a Cray X-MP supercomputer and required $\sim$1 h. While relatively fast, one must be concerned about possible prejudice in choosing starting structures and about the low probability of choosing one close enough to the global minimum when very large molecules are of interest.

Distance geometry algorithms provide a means of generating structures without the necessity of interactively choosing a starting structure. They do this by selecting atom–atom distances randomly within bounds for interatomic distances determined from a combination of experimental and bonding constraints and then projecting this high dimensionality distance space into Cartesian space, thus, arriving at atom positions. We therefore explore this approach with a program written by T. Havel.[3] Input followed literature procedures (Havel & Wuthrich, 1984, 1985). For these calculations the distances determined from the volume integrals were classified into one of five ranges: $\leq$2.5, 2.5–3.0, 3.0–3.5, 3.5–4.0, 4.0–5.0 Å. Distances were entered as upper bounds, and the program was allowed to set lower bounds based on van der Waals contact and other bonding constraints. For the long-range constraints a uniform upper bound of 5.0 Å was assumed (Wagner et al., 1987). To minimize memory requirements only the most heavily constrained portion of the molecule, residues 3–73, was included in the calculation. The program was run on a VAX 8600 and required between 5 and 6 h to complete each structure. Possible disadvantages of the distance geometry procedure include the fact that the program does not incorporate as complete a description of molecular energies and does not provide for a detailed balance of experimental and theoretical data.

A hybrid procedure was also followed in order to capitalize on the strong points of each approach. The distance geometry program used above is conveniently divided into several phases. The first three phases, which involve reading in experimental and sequential data, contracting bounds by application of triangle inequalities, and embedding one or more substructures, run rather quickly. The remaining two phases, which embed the complete structure and apply additional cycles to check inequalities and minimize departures from bounds, are rather time consuming. We found that substructures produced at the end of the first three phases are sufficiently close to global minima to allow efficient minimization using a pseudoenergy approach. In most of the examples presented below we actually embedded atoms corresponding to the complete structure in these first phases and did so using upper and lower bounds based on estimates in NOE errors rather than the more qualitative analysis used in the complete calculation. In setting bounds we used the following set of rules: (1) −0.1 and +0.2 were added to distances used in AMBER to produce lower and upper bounds when distances were less than 3.5 Å, (2) −0.2 and +0.4 were added to distances less than 4.5 Å, (3) −0.4 and +0.8 were added to distances over 4.5 Å, (4) an extra 0.2 Å was added and subtracted from bounds for each methyl proton pseudoatom involved, and (5) an extra 0.4 Å was added and subtracted from bounds for each non-methyl pseudoatom or united atom involved in representing the constraint. The asymmetry of bounds and increase in the width of bounds with increasing distance reflect the functional dependence and precision of NMR distance determination. The additions in

[2] AMBER 3.0 (U. C. Singh, P. K. Weiner, D. A. Case, J. Caldwell, and P. A. Kollman, 1986) is a program obtained through a licensing agreement with the Regents of the University of California at San Francisco.

[3] DISGEO (T. Havel, 1985) is a program obtained through the Quantum Chemistry Program Exchange of Indiana University, Bloomington, IN.

cases where pseudoatoms are used reflect imprecisions associated with our inability to precisely locate protons on groups that move rapidly or groups that require stereospecific assignments. It is not clear, however, that this degree of definition is required, and efficiency may be improved by widening bounds or embedding fewer atoms (Clore et al., 1987). We did, in fact, explore the option of embedding a main-chain subset of atoms in one of the structures (G) and found some improvement in speed.

The three phases of distance geometry on a VAX 8600 required ~1 h per substructure except for embedding of the main-chain subset which required 20 min. Embedding was followed by a stage of AMBER minimization using a simplified potential with high pseudoenergy weights requiring ~1 h. The final structures were produced by minimization with a complete potential on the Cray requiring ~1 h per structure. Indications are that most structural alterations during the AMBER minimizations take place in the first third of the time periods so it may be possible to reduce these times substantially.

Evaluation of Structures. Structures produced by the above calculations can be evaluated on the basis of several criteria. First, they can be evaluated by direct comparison of experimental distance constraints with interatomic distances in the final structures. We report these as deviations from our best estimate of experimental distances, rather than deviations from upper bounds, as is often done in distance geometry calculations. We regard violation of constraints by more than 2 Å as worrisome and report only these in the discussion that follows. We also report an RMS deviation calculated from all experimental constraints. Second, distance constraint violations are well represented by a rise in the total pseudoenergy contribution. The increase in penalties for violations as experimental distances increase is properly weighted by the pseudoenergy function, and one can equally well evaluate distance violations in structures by comparing a single total pseudoenergy term. Third, the real molecular energy of the final structure rises if constraints are met by placing atoms at positions that violate normal bonding geometries. Thus, the best structures have both low pseudoenergies and low real energies. Fourth, the structures can be evaluated by comparing RMS deviations of corresponding atoms in any structure to an average structure. This latter evaluation is important because it is difficult to ensure an adequate number of constraints to define a unique configuration for any one region of the protein. Regions that are underdefined will show large variations among structures, all of which may have low energies and distance violations within acceptable bounds. RMS deviations of atoms in a given structure from their positions in the average structure were calculated by first moving centroids of the structures to the origin and then formulating a rotation matrix that superimposes the structures for a minimum RMS deviation by using the least-squares method of Mackay (Mackay, 1983). Results are given as both deviations of backbone atoms, which are heavily constrained by NMR data, and deviations of all atoms, which include variations of the less well-constrained side chains.

RESULTS

The most successful of the above procedures in terms of efficiency of convergence and quality of results is the hybrid procedure, which utilizes the pseudo random selection of starting structures in the distance geometry approach and the relatively rapid convergence to an energetically stable form inherent in a molecular mechanics calculation. We will therefore emphasize structures generated by this procedure

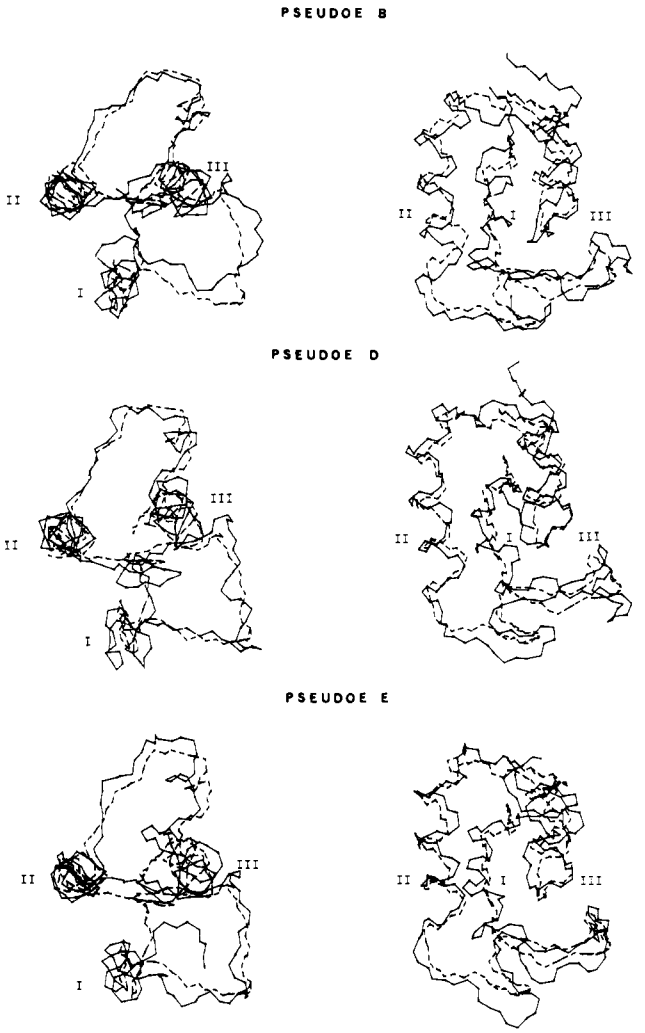

FIGURE 1: Comparison of individual ACP backbone structures to an average structure. In all cases, the average structure is indicated by dotted lines and the right member of each pair is produced by a rotation of ~90° about the X axis.

in our discussion. Seven structures were generated by this approach. They are labeled pseudoenergy A–G in the tables and figures. Figure 1 presents backbone tracings of three of these structures in comparison to the average of all seven. The structures are characterized by three roughly parallel α-helical segments. A helix from 3 to 14 is designated as helix I, a helix from 37 to 51 is designated helix II, and a helix from 65 to 75 is designated helix III. These helices are oriented with their axes coming out of the page in the left-hand view of each structure of Figure 1 and with their axes vertical in the page in the right-hand view. The portions of the backbone tracings associated with the helices follow one another closely with largest deviations in loops and near the C and N termini.

The direct use of a pseudoenergy minimization approach, despite our previous success with shorter peptide segments, proved cumbersome when dealing with the 77 amino acid sequence of ACP. Several attempts at convergence from manually folded starting structures resulted in local minimum structures with relatively high molecular energies and more than 20 distance constraint violations above 2 Å. Backbone foldings for the first two-thirds of the molecule did follow those shown in Figure 1 quite closely, with most distance violations localized to the C terminus. This suggests improper orientation of the C-terminus helix in the starting structure. However, since the effort in interactively generating new starting structures is relatively high, the approach was abandoned in

favor of distance geometry generated starting structures.

The direct use of distance geometry was carried to completion in two cases. Structures generated are labeled disgeo A and B in the following discussion. The structures show the same three parallel helices as those from pseudoenergy minimzation, but on closer inspection show an interesting divergence. They have a chain folding that is nearly the mirror image of the other structures. This is not a complete mirror image in that carbon centers have the proper local chirality. The structure does, however, have a global tracing that, when viewed as in Figure 1, proceeds in clockwise as opposed to counterclockwise direction. They also have segments of left-handed helices. The occurrence of a near mirror image might be the result of the rather generous bonds used in the direct distance geometry approach as opposed to the more restricted bounds used to generate starting structures for the hybrid approach. Generation of additional structures very likely would give a distribution of right- and left-handed tracings, but this rather time-consuming process was abandoned in favor of the hybrid approach (Havel & Wuthrich, 1984; Wagner et al., 1987).

Evaluation of the seven pseudoenergy and two distance geometry structures is presented in Tables I and II. Table I summarizes distance deviations and energetic terms for various structures. Specific deviations from experimental distances are listed only in cases where violations are greater than 2 Å and confidence in experimental data was sufficiently high to warrant weights for pseudoenergies greater than 10 kcal/mol. The pseudoenergy column gives a measure of how well distance constraints as a whole are satisfied. The total energy column gives a measure of efficiency in satisfying both distance and energetic constraints. The lower the energy in this column, the better the structure. The RMS distance deviation column gives a more conventional measure of the total deviation of interproton distances in various structures from experimental constraints.

Pseudoenergy structures B, D, and E are the best structures both in terms of having the smallest number of distance violations and having the smallest RMS distance deviations. All have very similar constraint energies, but B has a significantly lower molecular energy (obtained as the difference of total and constraint energies). These three best structures are depicted in Figure 1. Coordinates for the B structure are included as supplementary material. The distance geometry structures have a number of distance violations exceeding 30 in each case. They also have relatively high total constraint energies (−1.03 E04, −1.04 E04) and very high molecular energies. Energy minimization of structures did reduce numbers of violations and energies but did not invert the direction of folding.

Table II gives RMS deviations for backbone atoms, and all atoms, relative to an average structure computed from the coordinates of the seven pseudoenergy structures given above. Structures are very similar, as is evident for the three structures in Figure 1. In some cases, the A structure, for example, deviations near the C terminus are visually apparent and contribute to larger RMS atomic deviations. The trend toward divergence of structures near the C and N termini also occurs in the best structures as illustrated in Figure 2, which presents average atomic deviations from the average structure for each amino acid. There are in fact relatively few constraints beyond Ile-73. This may account for the variability of the C terminus. RMS deviations for the distance geometry structures after inversion in a mirror plane are also included. Agreement is not as close as among pseudoenergy structures because the

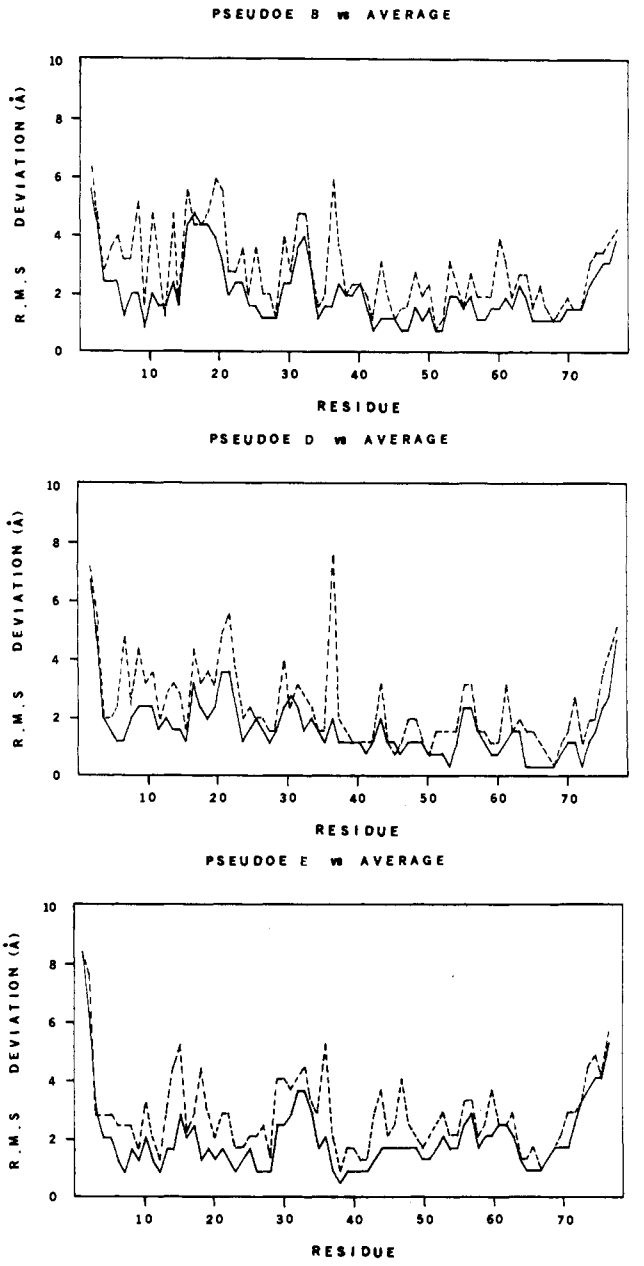

FIGURE 2: RMS deviations of atoms in pseudoe B, pseudoe D, and pseudoe E from the corresponding atoms in the average structure. Deviations from backbone atoms, N, CA, and C, were averaged and plotted versus residue number to produce the solid line. All atoms in a residue were averaged to produce the dotted line.

original structures are not complete optical isomers of pseudoenergy structures, but the values of 3.7 and 3.3 Å for the backbone atoms confirm the near mirror image chain tracings.

The convergence of seven of the nine structures to a similar minimum energy structure for all but the C and N termini suggests that molecular conformation is adequately defined by the combination of NMR and molecular energy constraints used in these calculations. None of the structures, however, meet all of the distance constraints as well as one might hope. Table I shows that even an energetically very good structure such as pseudoenergy structure B has six violations greater than 2 Å. There are several possible causes for these violations. One is that they could be the result of misassignment of resonances. Resonance assignments have, however, been carefully rechecked, and this seems an unlikely cause for all of the violations. Another is simply the existence of local minimum traps. In other words, further exploration of conformational space might yield a better structure. This pos-

Table I: Distance Deviations and Energies for ACP Structures[a]

structure	connectivity	dist deviation	total energy	const energy	RMS dist deviation
pseudoe A	F50 HZ–L46 HA	2.48	−1.56 E04	−1.80 E04	1.09
	F50 CZ–L46 HG	4.35			
	S27 HN–D31 CB	2.02			
	E30 HN–S27 CB	2.86			
	D31 HN–S27 HA	2.04			
	V40 HN–L37 HA	2.39			
	L46 HN–V43 HA	2.65			
	E48 HN–A45 HA	2.20			
	D51 HN–E48 HA	2.65			
	E47 HN–A45 HB	2.01			
	I11 HA–L42 HD1	2.79			
	F28 CZ–T63 HA	2.11			
	Y71 CG–T52 HG	2.12			
	Y71 CZ–E58 CB	6.16			
	Y71 CZ–E58 CG	6.03			
pseudoe B	E21 HN–Q19 HA	2.56	−1.82 E04	−1.85 E04	0.92
	Q66 HN–T64 HB	2.24			
	A67 HN–T64 HB	2.26			
	F50 CZ–I03 HD	2.33			
	F50 HZ–I03 HD	2.39			
	F28 CZ–T63 HA	2.83			
pseudoe C	Q66 HN–T64 HB	2.46	−1.60 E04	−1.73 E04	1.12
	A59 HN–D56 HA	2.40			
	H75 HN–I72 HA	2.06			
	I72 HG–I03 HG	2.32			
	I11 HA–L42 HD2	2.31			
	F28 CZ–L42 HD1	2.05			
	F28 HZ–L42 HD1	2.92			
	F28 CG–T63 HA	4.04			
	F28 CZ–T63 HA	3.93			
	F50 HZ–I72 HG	2.45			
	Y71 CG–I62 HD	5.92			
	Y71 CZ–I62 HD	4.65			
pseudoe D	G12 HN–K09 HA	2.11	−1.74 E04	−1.81 E04	0.99
	L46 HN–V43 HA	2.36			
	I11 HA–L42 HD2	2.99			
	F50 HZ–I03 HG	2.19			
	Y71 CZ–I54 HG	2.70			
	Y71 CZ–E58 CB	2.82			
	Y71 CZ–E58 CG	2.98			
pseudoe E	Q14 HN–I11 HA	2.17	−1.67 E04	−1.80 E04	1.01
	H75 HN–N73 HA	2.70			
	Y71 CG–T52 HG	2.52			
pseudoe F	A45 HN–L42 HA	2.34	−1.65 E04	−1.79 E04	1.15
	E49 HN–L46 HA	2.49			
	Q66 HN–T64 HB	2.20			
	A67 HN–T64 HB	2.13			
	F50 CZ–I03 HG	3.25			
	F50 HZ–I03 HG	3.60			
	F28 HZ–L42 HD2	2.21			
	F28 CG–T63 HA	4.11			
	F28 CZ–T63 HA	4.43			
	F50 CG–I72 HG	2.93			
	F50 CZ–I72 HG	2.81			
	F50 HZ–I72 HG	3.48			
	F50 HZ–I72 HD	4.70			
	F50 CG–I72 HD	6.10			
	F50 CZ–I72 HD	5.08			
pseudoe G	F50 HZ–L46 HG	2.75	−1.62 E04	−1.83 E04	1.06
	S27 HN–D31 CB	4.08			
	S27 HN–D31 HN	2.47			
	D31 HN–S27 HA	3.17			
	F50 HZ–I03 HG	2.36			
	F50 HZ–V07 HA	2.11			
	F28 CZ–L42 HD1	2.30			
	F28 HZ–L42 HD1	3.50			
	F28 HZ–L42 HD2	4.06			
	F28 CZ–L42 HD2	2.92			
	F28 CG–T63 HA	2.51			
	F28 CZ–T63 HA	3.61			

[a] Connectivity information is given in terms of a one-letter code for the amino acid type followed by a sequence number and atom type. Energies are given in kilocalories. Distance deviations are given in angstroms.

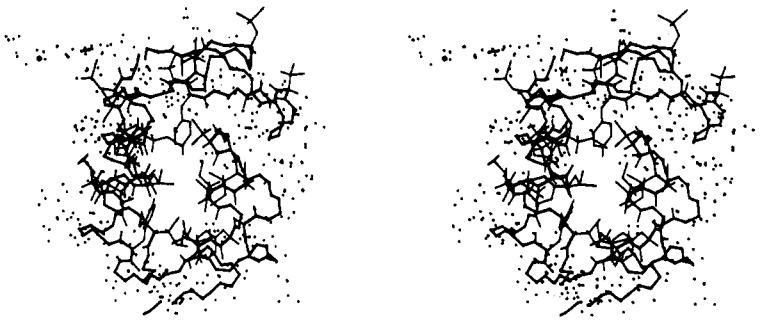

FIGURE 3: Stereopair for the pseudoe B structure. The view is produced by an ~180° rotation of the right-hand structure in Figure 1 about the X axis. Hydrophobic residues are shown in light solid lines, other residues are shown with one dot per atom, and the peptide backbone is traced in a heavy solid line.

Table II: RMS Deviations between Average Structure and Each Structure

	RMS deviation (Å)	
structure	backbone atoms	all atoms
pseudoe A	2.68	3.69
pseudoe B	2.32	3.37
pseudoe C	2.96	3.76
pseudoe D	2.02	3.10
pseudoe E	2.29	3.18
pseudoe F	2.83	3.71
pseudoe G	3.02	3.89
inverse A	3.66	4.65
inverse B	3.32	4.44

sibility is being pursued by using molecular dynamics methods. The ability to move a molecule out of a local minimum by increasing kinetic energy in these simulations provides an additional means of exploring conformational space (Clore et al., 1987; Holak et al., 1988). A more likely cause, however, is that distances determined are the result of averaging between two or more structures rather than the result of a single rigid conformation. Unlike many proteins previously studied by NMR, ACP is not cross-linked by sulfhydryl bonds and may be rather free to adopt alternate conformations. Evidence for averaging might be obtained by looking for geometrically incompatible constraints to a single residue. There are a rather large number of long-range constraints to Phe-50, among them, those involving Ile-3, Leu-46, and Ile-72. Examination of structures suggests that one cannot meet all of these constraints simultaneously. In pseudoenergy structures B and D, for example, contacts with Ile-3 are long, while others are short. In pseudoenergy structure A, contacts with Leu-46 are long, while others are met. In pseudoenergy structure C, contacts with Ile-72 are long, while others are met. A dynamic equilibrium involving contact of Phe-50 with these residues would lead to abnormally short distance predictions for each of these contacts and an inability to fit all three with a single reasonable structure.

Discussion

The chemical and structural properties of our best structures bear some further discussion. First, we have attempted structural solutions for segments of ACP previously (Holak et al., 1987a). It is of interest to see whether these structural elements persist when the entire structure is solved. The two elements examined previously were an α-helix from 3 to 13 and a right-handed turn from 26 to 36. Both of these elements are recognizable in the structures produced here. As a more quantitative measure, the RMS deviation between backbone atoms in the right-handed loop determined earlier and the corresponding atoms in the average of the seven pseudoenergy structures presented here is 2.2 Å. There are distortions that may well arise from inclusion of longer range constraints in the total structure calculation, but the overall similarity suggests that the short- and intermediate-range constraints used earlier are in many cases adequate to define local structures without the aid of longer range interactions.

The chemical properties of the best structures are also interesting. Figure 3 shows hydrophobic residues in solid lines and more hydrophilic residues in dotted lines for the B structure. Note the tendency for the hydrophilic residues to concentrate on the outer surface of the molecule with hydrophobic residues on the interior. The hydrophilic exterior is certainly not unexpected for a soluble protein, but its appearance in the structures produced here supports the validity of the methods in that there are no factors in our methodology that would have skewed the distribution in this way other than the raw data. The structure also shows a hydrophobic cleft running along the surfaces of helices II and III (left and right side of the figure) and bridged by the 52–65 unstructured segment on the bottom. This is an ideal candidate for an acyl chain binding site. The prosthetic group is attached to Ser-36 at the upper end of helix II in Figure 3. The prosthetic group is not in the cleft in the diagram because no NMR constraints were observed allowing placement of this group in the reduced, free sulfhydryl species. However, data on acylated species have indicated contact of the acyl chain with Ala-59 and Ile-54 near the 5-position of the chain (Jones et al., 1987). These amino acids are at the bottom of the cleft as viewed in Figure 3 and can only be reached by stringing the chain and prosthetic group down along helices II and III. An analysis of cross-relaxation data on an acylated species is underway and should lead to a more complete description of the interaction of the acyl chain with this region.

Supplementary Material Available

Tables containing distance constraints used in the calculations and coordinates for the pseudoenergy B structure (28 pages). Ordering information is given on any current masthead page.

References

Barnes, E. M., & Wakil, S. J. (1968) *J. Biol. Chem. 243*, 2955–2962.

Billeter, M., Havel, T. F., & Wuthrich, K. (1987) *J. Comput. Chem. 8*, 132–141.

Braun, W. (1987) *Q. Rev. Biophys. 19*, 1115–1157.

Clore, G. M., Gronenborn, A. M., Brunger, A. T., & Karplus, M. (1985) *J. Mol. Biol. 186*, 435–455.

Clore, G. M., Gronenborn, A. M., Kjaer, M., & Poulsen, F. M. (1987) *Protein Eng. 7*, 275–288.

Havel, T. F., & Wuthrich, K. (1984) *Bull. Math. Biol. 45*, 673–698.

Havel, T. F., & Wuthrich, K. (1985) *J. Mol. Biol. 182*, 281-294.
Holak, T. A., & Prestegard, J. H. (1986) *Biochemistry 26*, 5766-5774.
Holak, T. A., Prestegard, J. H., & Forman, J. D. (1987a) *Biochemistry 26*, 4652-4660.
Holak, T. A., Scarsdale, J. N., & Prestegard, J. H. (1987b) *J. Magn. Reson. 74*, 546-549.
Holak, T. A., Nilges, M., Prestegard, J. H., Gronenborn, A. M., & Clore, G. M. (1988) *Eur. J. Biochem.* (in press).
Jones, P. J., Cioffi, E. A., & Prestegard, J. H. (1987) *J. Biol. Chem. 262*, 8963-8965.
Kaptein, R., Zuiderweg, E. R. P., Scheck, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 196*, 611-639.
Mackay, A. L. (1983) *Acta Crystallogr., Sect. A: Found. Crystallogr. A40*, 165.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95-117.
Mayo, K. H., Tyrell, P. M., & Prestegard, J. H. (1983) *Biochemistry 22*, 4485-4493.
McRee, D. E., Richardson, J. S., & Richardson, D. C. (1985) *J. Mol. Biol. 182*, 467-468.
Ohlrogge, J. B. (1987) in *The Biochemistry of Plants*, Vol. 9, pp 137-157, Academic, New York.
Olejniczak, E. T., Dobson, C. M., Karplus, M., & Levy, R. M. (1984) *J. Am. Chem. Soc. 106*, 1923-1930.
Otting, G., Widmer, H., Wagner, G., & Wuthrich, K. (1986) *J. Magn. Reson. 66*, 187-193.
Scarsdale, J. N., Yu, R. K., & Prestegard, J. H. (1986) *J. Am. Chem. Soc. 108*, 6778-6784.
Scarsdale, J. N., Ram, P., Yu, R. K., & Prestegard, J. H. (1988) *J. Comput. Chem. 9*, 133-147.
Simoni, R. D., Criddle, R. S., & Stumpf, P. K. (1967) *J. Biol. Chem. 242*, 573-581.
Singh, V. C., & Kollman, P. A. (1984) *J. Comput. Chem. 5*, 129-145.
Spencer, A. K., Greenspan, A. D., & Cronan, J. E., Jr. (1978) *J. Biol. Chem. 253*, 5922-5926.
States, D. J., Habercorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.
Thompson, G. A., Jr. (1980) *The Regulation of Membrane Lipid Metabolism*, pp 19-43, CRC, Boca Raton, FL.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wuthrich, K. (1987) *J. Mol. Biol. 196*, 611-639.
Weiner, P. K., & Kollman, P. A. (1981) *J. Comput. Chem. 2*, 287-303.
Williamson, M. P., Havel, T. F., & Wuthrich, K. (1985) *J. Mol. Biol. 182*, 295-315.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wuthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.

Circular Dichroism and Potentiometry of FAD, Heme and Mo-Pterin Prosthetic Groups of Assimilatory Nitrate Reductase†

Christopher J. Kay,* Michael J. Barber, and Larry P. Solomonson
Department of Biochemistry, College of Medicine, University of South Florida, Tampa, Florida 33612
Received December 3, 1987; Revised Manuscript Received March 14, 1988

ABSTRACT: Oxidation-reduction midpoint potentials for flavin, heme, and molybdenum-pterin prosthetic groups of assimilatory nitrate reductase (NR) from *Chlorella vulgaris* were measured at room temperature by using CD and EPR potentiometry. The CD changes accompanying reduction of each prosthetic group were determined by using enzyme fragments containing either FAD or heme and molybdenum prosthetic groups, obtained by limited proteolysis, and by poising the enzyme at various redox potentials in the presence of dye mediators. Limited proteolysis did not appear to alter the environment of the prosthetic groups, as judged by their CD spectra. Also, CD potentiometric titration of FAD in intact NR (E_m' = −272 mV, n = 2) gave a similar value (E_m' = −286 mV) to the FAD of the flavin-containing proteolytic domain, determined by visible spectroscopy. Less than 1% of the flavin semiquinone was detected by EPR spectroscopy, indicating that E_m' (FAD/FAD$^{\bullet-}$) may be more than 200 mV lower than E_m' (FAD$^{\bullet-}$/FADH$^-$). Reduction of heme resulted in splitting of both Soret and α CD bands into couplets. The heme E_m' was −162 mV (n = 1) determined by both CD and visible spectroscopy. Reduction of Mo-pterin was followed by CD at 333 nm, and Mo(V) was monitored by room temperature EPR spectroscopy. Most of the change in the Mo-pterin CD spectrum was due to the Mo(VI)/Mo(V) transition. The E_m' values determined for Mo(VI)/Mo(V) were +26 mV by CD and +16 mV by EPR, whereas Mo(V)/Mo(IV) values were −40 mV by CD and −26 mV by EPR. In contrast, freezing samples to 100 K resulted in significant redistribution of electrons, particularly affecting the measured Mo(VI)/Mo(V) couple (−34 mV).

Assimilatory nitrate reductase (NR)[1] from *Chlorella vulgaris* catalyzes the initial and rate-limiting step, the NADH-dependent reduction of nitrate to nitrite, in the process of nitrate assimilation (Guerro et al., 1981). The enzyme exists as a homotetramer, with each subunit (M_r 96 000) containing FAD, a b-type cytochrome, and Mo-pterin prosthetic groups in a ratio of 1:1:1 (Howard & Solomonson, 1982). The strong visible absorbance of the b-type cytochrome of NR has enabled the redox properties of this center to be studied in detail (Kay

† This work was supported by Grants 1 RO1 GM32696 from the National Institutes of Health and 84-CRCR-1-1404 from the United States Department of Agriculture.

[1] Abbreviations: NR, nitrate reductase; EPR, electron paramagnetic resonance; EDTA, ethylenediaminetetraacetic acid; CD, circular dichroism; CIP, corn inactivating protease; MV$^{\bullet+}$, methylviologen radical; FAD$^{\bullet-}$, flavin adenine dinucleotide semiquinone anion; FADH$^-$, flavin adenine dinucleotide hydroquinone.

0006-2960/88/0427-6142$01.50/0