Determination of Local Protein Structure by Spin Label Difference 2D NMR: The Region Neighboring Asp61 of Subunit *c* of the F_1F_o ATP Synthase†

Mark E. Girvin and Robert H. Fillingame*

Department of Biomolecular Chemistry, University of Wisconsin Medical School, Madison, Wisconsin 53706

Received October 5, 1994; Revised Manuscript Received November 28, 1994®

ABSTRACT: Purified subunit *c* from the H^+-transporting F_1F_o ATP synthase of *Escherichia coli* folds as an antiparallel pair of extended helices in a solution of chloroform−methanol−water. A similar hairpin-like folding is predicted for the native protein in the multisubunit transmembrane F_o sector of the ATP synthase. A single Cys variant (A67C) of subunit *c* was created and modified with a maleimido-PROXYL [[3-(maleimidomethyl)-2,2,5,5-tetramethyl-1-pyrrolidinyl]oxy] spin label. Pairs of ^1H 2D correlation and NOE spectra were collected with the nitroxide oxidized (paramagnetic) and reduced (diamagnetic). The pairs of spectra were subtracted, yielding difference spectra containing only cross-peaks from ^1H within 15 Å of the spin label. These greatly simplified spectra were easily analyzed to provide complete assignments for residues 10−25 and 52−79 of the protein, 150 NOE distance restraints, and 27 hydrogen-bonding restraints. The chemical shifts and NOE patterns observed in the derivatized mutant were virtually identical to those which were resolved in the unmodified wild-type protein, strongly suggesting that the spin label was not perturbing the protein structure. The restaints enabled us to calculate a detailed structure for this region of subunit *c*. The structure consisted of two gently curved helices, crossing at a slight (30°) angle. The C-terminal helix was disrupted from Val60 to Ala62 near the essential Pro64. Asp61, the residue thought to undergo protonation−deprotonation with each H^+ transported across the membrane, was in ver der Waals contact with Ala24. The proximity of these residues had been predicted from mutant analyses, where H^+ translocation was retained on moving the Asp from position 61 to 24.

The F_1F_o ATP synthase reversibly catalyzes the synthesis of ATP from ADP and P_i (Senior, 1988). The driving force for ATP synthesis is an electrochemical proton gradient, typically generated by electron transport. The catalytic sites reside on the F_1 sector of the enzyme, while proton binding and translocation occur in the transmembrane F_o sector. The F_1 sector is composed of five types of subunits in an $\alpha_3\beta_3\gamma\delta\epsilon$ ratio. A structure of the beef heart F_1 domain, which resolves the α, β, and much of the γ, has recently been solved at 2.8 Å resolution by Abrahams et al. (1994). The model presents an asymmetric structure for the catalytic sites and most reasonably supports an alternating site model (Cross, 1981). The structure, in combination with an enormous amount of previous chemical and genetic data, should provide answers to questions regarding the mechanisms of catalysis and site−site cooperativity during ATP hydrolysis in F_1. Proton movement through the transmembrane F_o sector must in some way lead to conformational changes in F_1 which alter substrate binding and product release. Structural information for the F_o domain will be required if the coupling of these events is to be understood. The F_o sector of *Escherichia coli* is composed of three subunits, *a*, *b*, and *c*, in a 1:2:10 stoichiometry (Foster & Fillingame, 1982). Subunit *c*, the smallest subunit of F_o, is thought to play direct roles in proton translocation and conformational coupling to F_1 (Fillingame, 1990; Fillingame et al., 1992; Zhang et al., 1994). Subunit *c* is a hydrophobic protein of 79 residues which is predicted to fold through the lipid bilayer as an antiparallel pair of transmembrane helices. Asp61, lying in the center of the second predicted transmembrane helix, provides an essential carboxyl group that is thought to undergo protonation−deprotonation with each H^+ translocated across the membrane.

Structural information on membrane proteins has been difficult to obtain, largely because of difficulties encountered in crystallization. As an alternative approach to crystallization and X-ray diffraction, the structures of the small subunits, or fragments of larger subunits, could be solved by NMR. The arrangement of the subunits, or subunit fragments, would then be pieced together by chemical, genetic, and biophysical techniques. We have initiated such an approach with subunit *c*. Our previous ^1H NMR analysis of purified subunit *c* proved that the protein was folded as a pair of interacting helices in a chloroform−methanol−water solvent (Girvin & Fillingame, 1993). The purified protein was also shown to retain key biochemical properties seen in the native F_o complex (Girvin & Fillingame, 1994). We used a nitroxide analog of DCCD[1] to measure distances between the site of modification at Asp61 and the assigned ^1H in both transmembrane helices, and this data provided a preliminary

† This study was supported by U.S. Public Health Service Grant GM23105 from the National Institutes of Health. The National Magnetic Resonance Facility at Madison, supported by NIH Grant R02301, was used in these studies. Equipment in the facility was purchased with funds from the University of Wisconsin, the NSF Biological Instrumentation Program (Grant DMB8415048), the NIH Biomedical Research Technology Program (Grant RR02301), the NIH Shared Instrumentation Program (Grant RR02781), and the U.S. Department of Agriculture.

® Abstract published in *Advance ACS Abstracts,* January 15, 1995.

[1] Abbreviations: 2D, two dimensional; COSY, 2D correlated spectroscopy; CT, constant time; DCCD, *N,N'*-dicyclohexylcarbodiimide; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; PROXYL, (2,2,5,5-tetramethyl-1-pyrrolidinyl)oxy; rms, root mean square; TMS, tetramethylsilane; TOCSY, 2D total correlation spectroscopy.

0006-2960/95/0434-1635$09.00/0

model for the interaction of the two helices (Girvin & Fillingame, 1994).

During the analysis of NCCD-modified protein, it became obvious that reduced (diamagnetic) minus oxidized (paramagnetic) difference spectra of the nitroxide-modified protein could provide simplified subspectra of a 20−30 Å diameter region surrounding an attached spin label. Paramagnetic difference methods have been used frequently in 1D NMR spectroscopy, but their use in multidimensional studies has been rare [see review by Kosen (1989)]. Moonen et al. (1984) used reduced minus partially oxidized difference 2D ^{1}H NMR spectra to identify the resonances of several residues near the flavin binding site of flavodoxin. De Jong et al. (1988) and Folkers et al. (1993) have used difference spectra to identify the conformation of the residues of DNA binding proteins involved in the binding of spin-labeled oligonucleotides. In the present work we apply a method involving reduced minus oxidized spin label difference 2D NMR to subunit *c* and show that a detailed 3D structure of a 30 Å region can be derived readily from the difference NMR data. The difference method should be applicable to proteins much larger than subunit *c*, including the other subunits of the F_o domain of the ATP synthase.

EXPERIMENTAL PROCEDURES

Subunit c Purification and Modification. The Cys67 subunit *c* mutant was purified from overproducing strain MEG171 which carries a plasmid equivalent to pCP35 (Fillingame et al., 1991) in an A67C subunit *c*, AN346 chromosomal background (Miller et al., 1989). Cells were grown on M63 minimal medium, supplemented with 1.4% glucose (Miller et al., 1989). From a 4 L culture, 30 g of cells was harvested. The cells were suspended to a final volume of 50 mL in 50 mM aqueous ammonium acetate, pH 7.0. The suspension was blended with 300 mL of $CHCl_3$:CH_3OH (1:1) for 1 min. The extract was filtered, and the solvent ratio was adjusted to $CHCl_3$:CH_3OH:H_2O (8:4:3) to initiate a separation of aqueous and organic phases. The lower organic phase was concentrated, and subunit *c* was purified as described previously (Girvin & Fillingame, 1993). The final yield was 40 mg of subunit *c*.

Purified subunit *c* (10 mg) was quantitatively modified by reaction with 3 mM (maleimidomethyl)-PROXYL in 4.5 mL of $CHCl_3$:CH_3OH:H_2O (3:3:1), containing 25 mM ammonium acetate, pH 7.0, for 90 min at 25 °C. The extent of the modification was followed as loss of reactivity with Ellman's reagent, 5,5′-dithiobis(2-nitrobenzoic acid). The presence of H_2O in the solvent mixture was found to increase the reaction rate by more than 10-fold. The modified protein was separated from excess reagent on a Sephadex LH20 column in $CHCl_3$:CH_3OH (2:1).

NMR Spectroscopy and Data Analysis. The protein sample used for NMR spectroscopy was 1.6 mM subunit *c* in 0.5 mL of C^2HCl_3:C^2H_3OH:H_2O (4:4:1), containing 50 mM NaCl, at pD 5.5, prepared as described by Girvin and Fillingame (1993). Reduction of the PROXYL spin label was accomplished by the addition of 50 μL of 40 mM phenylhydrazine dissolved in C^2HCl_3:C^2H_3OH:H_2O (4:4:1) with the pD adjusted to 5.5. NMR spectra were acquired at 37 °C on a Bruker AM-500 spectrometer. Pairs of 2D CT-COSY, TOCSY, and NOESY spectra were recorded, the first with the PROXYL group oxidized and the second with the group reduced by phenylhydrazine. Spectra were obtained as follows: CT-COSY (Girvin, 1994) with a constant time period of 40 ms as 400 t_1 increments of 16 scans each; TOCSY (Bax & Davis, 1985) with a 75 ms mixing time as 480 t_1 increments of 48 scans each; and NOESY (Kumar et al., 1980; Macura et al., 1982) with a 150 ms mixing time as 448 t_1 increments of 80 scans. Quadrature detection in t_1 was achieved using the time-proportional phase incrementation scheme of Marion and Wüthrich (1983). All experiments included solvent suppression by presaturation, along with the SCUBA sequence for restoring magnetization to any H^α which were saturated by the suppression (Brown et al., 1988). The data size in the acquisition dimension was 1024 complex points, and the sweep widths in both dimensions were 5494.5 Hz. All chemical shifts are reported relative to tetramethylsilane.

The data were processed using Felix 2.0 (Hare Research, Bothell, WA). A 90° shifted sine bell squared window function was applied to the 1024 t_2 points, and the t_2 dimension was then zero-filled to 2048 points. After Fourier transformation in t_2, linear prediction was used to extend the t_1 data by 30%, a 90° shifted sine bell squared function was applied, and the data were zero-filled to 1024 real points.

After processing the data, difference spectra were obtained by subtracting the spectra of the oxidized (paramagnetic) sample from the reduced diamagnetic controls. The individual spectra were first smoothed in both dimensions using a seven-point binomial weighted smoothing function. While not absolutely required, smoothing did reduce artifacts from subtracting coarsely digitized 2D data. The base plane of each spectrum was corrected using the procedure of Chylla and Markley (1993), t_1 and low-level noise were removed using the procedure of Zolnai et al. (1986), and the intensities of the reduced spectra were scaled by a factor of 1.1 to compensate for dilution from the addition of phenylhydrazine. The phase sensitive CT-COSY data were converted to an absolute value mode. After shifting one data matrix of each pair by one or two data points in each dimension to optimize superposition of cross-peaks, the individual pairs of spectra were subtracted.

Resonance assignments were made as described by Wüthrich (1986). $H^NH^\alpha H^\beta$ correlations were identified using the difference CT-COSY spectrum. The spin system assignments were then checked and often extended further out on the side chains, using the difference TOCSY spectrum. Inter-residue H^NH^N, H^NH^α, $H^\beta H^N$, and $H^\alpha H^\beta$ NOEs were collated and used to make the sequence specific assignments.

Structural Analysis. The substructure of subunit *c* within the relaxation range of the nitroxide was calculated from the NMR data by the hybrid distance geometry-dynamical simulated annealing procedure of Nilges et al. (1988) using the program XPLOR 3.1 (Brünger, 1993). Data for the residues Leu9-Ile26 and Gln52-Ala79 were used. Initial distance geometry structures (200) were generated using the 107 short range inter-residue NOE restraints (intraresidue NOE cross-peaks were ignored throughout the analysis). Four of the strongest NOE restraints between the two segments were also included to provide a rough alignment of the two pieces of secondary structure. Each of the 200 distance geometry structures was coarsely refined by high-temperature simulated annealing, with a maximum temperature of 1500 K, using an additional 39 long range restraints between helices and 27 hydrogen-bonding restraints consistent with amide hydrogen exchange rates and the secondary

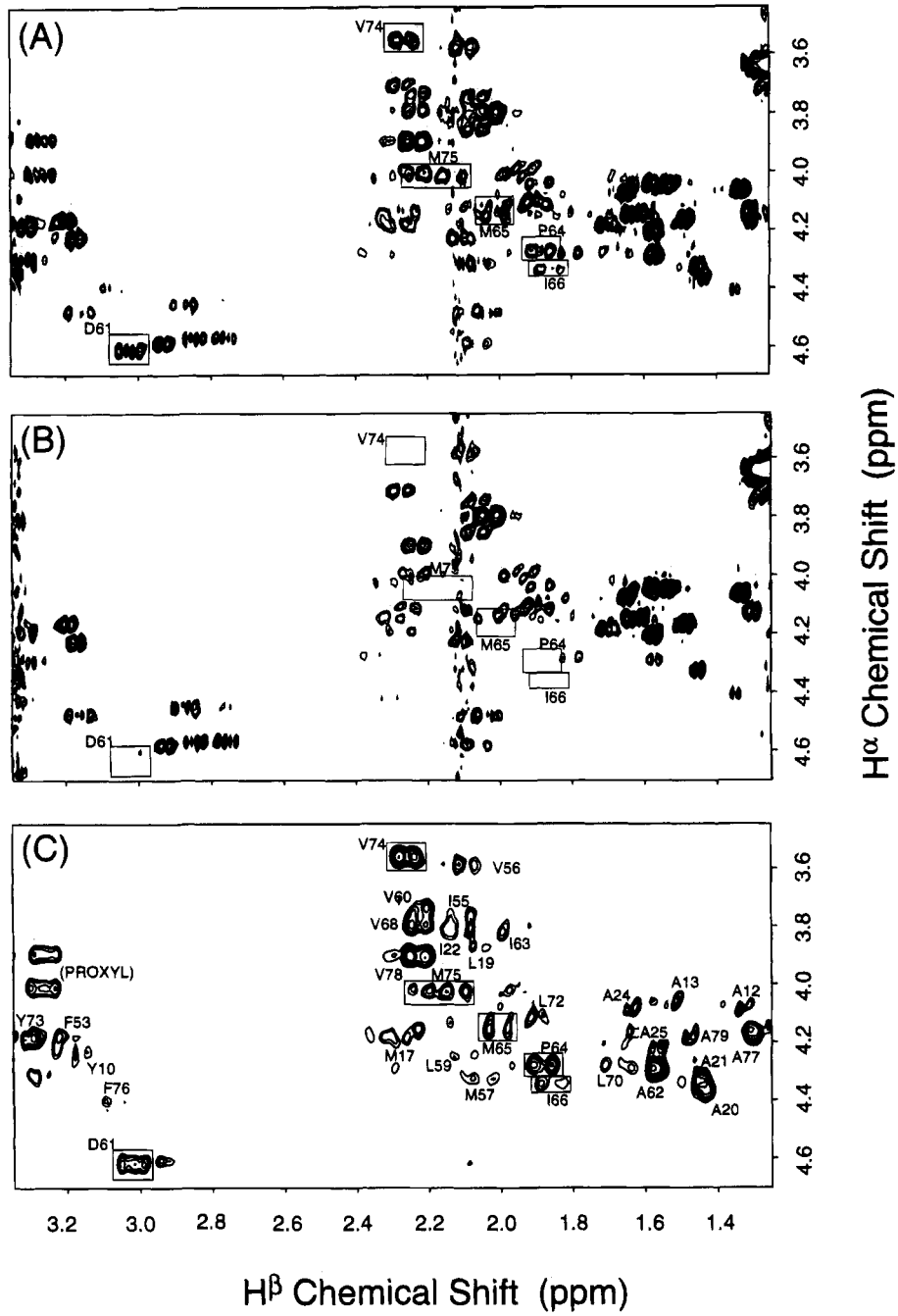

FIGURE 1: $H^{\alpha}-H^{\beta}$ region expansions of the 500 MHz ^1H CT-COSY spectra and difference spectrum of Cys67−(maleimidomethyl)-PROXYL subunit *c*. Representative cross-peaks in the spectrum of the reduced protein (A) which are greatly attenuated in the spectrum of the oxidized protein (B) are boxed and labeled. The reduced minus oxidized difference spectrum (C) contains only cross-peaks from ^1H within approximately 15 Å of the nitroxide attached to residue 67.

structures derived from the initial distance geometry calculations. All hydrogen bond restraints were modeled by a square well potential extending from 2.5 to 3.3 Å for the N−O distance. NOE restraints were also implemented as a square well potential extending from 1.8 to 5.5 Å. No attempt was made to quantitate the intensities of the NOE cross-peaks, since the intensities in the difference experiments were dependent on distance from the nitroxide, as well as interproton distance. Two more cycles of simulated annealing were performed on the 50 lowest energy structures, with maximal temperatures of 800 and finally 300 K. The nine structures of lowest energy were further refined by restrained molecular mechanics. The final nine structures have been deposited at the Brookhaven Protein Data Bank under the filename 1ATY.

RESULTS

Modification of Subunit c. We were interested in defining the structure of subunit *c* near Asp61. Earlier NMR studies provided an overall view of the association of the two transmembrane helices of subunit *c* but not a detailed picture of the region surrounding Asp61. A structural analysis of this region of subunit *c* appeared to be an ideal test of the feasibility of using spin label difference NMR spectroscopy to solve local protein structure. We wanted to introduce a single Cys for modification at a position near Asp61, using a modification that would not interfere with the expected packing (wild-type subunit *c* lacks Cys). We also wanted the mutant to be functional. An A67C mutation appeared to meet these criteria. The F_1F_o ATPase had normal function (Y. Zhang and R. Fillingame, unpublished), and previous

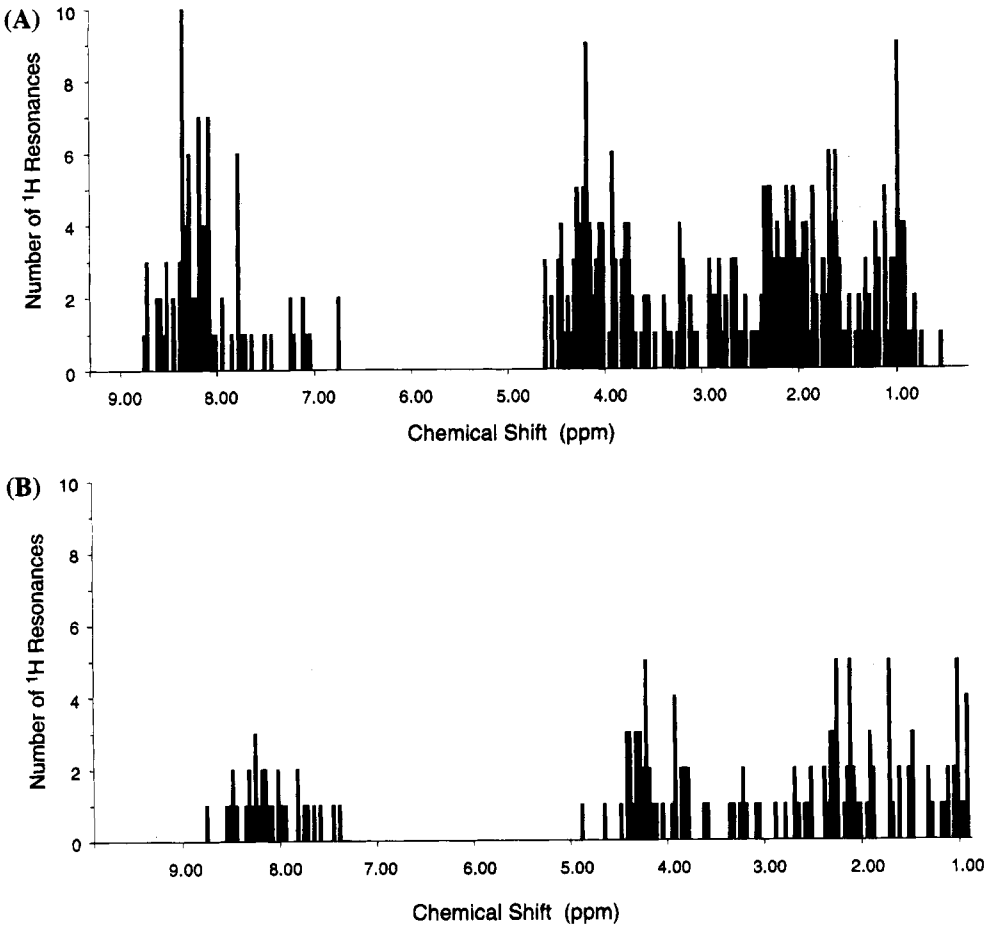

FIGURE 2: Distributions of chemical shifts in the complete spectrum of subunit *c* (A) and in the spin label difference spectrum (B). Histograms of the number of resonances within 0.01 ppm of a given chemical shift are plotted.

modeling, based on NMR data (Girvin & Fillingame, 1994), suggested that residue 67 should not lie at the interface between the two helices. (Maleimidomethyl)-PROXYL was selected as the modification reagent because of its reactivity with Cys67, the stability of the covalent linkage to reductant, and the relatively small size of the reagent.

NMR Difference Spectroscopy and Resonance Assignments. The presence of a paramagnetic nitroxide increases the relaxation rate (and hence broadens and reduces the intensities) of proton resonances within approximately 15 Å of the unpaired electron. Chemical reduction of the nitroxide restores resonances to their original line widths and intensities. In a reduced minus oxidized 2D NMR difference spectrum, one observes only the cross-peaks from groups within approximately 15 Å of the spin label probe.

The spectral simplification achieved with spin label difference spectroscopy is illustrated in Figures 1−3. The control CT-COSY spectrum, with the nitroxide reduced, is shown in Figure 1A. The cross-peaks from several residues near the labeling site at position 67 are marked for reference. The cross-peaks from residues within 12−15 Å of the nitroxide are missing in the CT-COSY spectrum of the oxidized sample (Figure 1B). Subtraction of the oxidized from the reduced spectrum leads to a greatly simplified spectrum of subunit *c* (Figure 1C). The usefulness of this simplification is illustrated in Figure 2, where the incidence of degenerate chemical shifts (within 0.01 ppm) for the entire protein (Figure 2A) is compared with the subset of resonances observed in the difference experiment (Figure 2B). The reduction in the number of overlapped resonances in the H^{α}, H^{β}, and H^{N} regions of the spectrum was crucial for making quick and reliable sequential assignments using the standard 2D NMR methods. Spin label difference spectra also helped resolve cross-peaks in the overlapped regions of the NOESY spectrum (Figure 3), in addition to resolving the ambiguity associated with the identification of the spin pairs responsible for any given NOE cross-peak. This increased resolution was particularly helpful in analysis of the $H^{N}-H^{\beta}$, $H^{N}-H^{\alpha}$, and $H^{N}-H^{N}$ regions of the NOESY spectrum.

The region defined by the difference experiments of subunit *c* labeled at position 67 included a total of 44 residues: Tyr10−Ala25 of the N-terminal region and Gln52 through the C-terminal Ala79. Within this region, the most strongly affected resonances were from Tyr10 and residues 17−25 and 56−74. The assignments for Cys67 subunit *c* are presented in Table 1. Although we have previously reported many of the ^{1}H assignments for this protein, the chemical shifts, as well as the type and sequential identities, were determined by the standard methods (Wüthrich, 1986) without reference to previous results. This second determination provided a check of previous assignments, as well as testing of the ease of using the difference procedure to make assignments in a new protein. With the exceptions of residues adjacent to the modification site and slight shifts of the H^{N} resonances at pD 5.5 relative to pD 6.8, the assignments coincided with those previously reported (Girvin & Fillingame, 1993). The assignments for residues 20, 21, 57, 59, 63, 65, 66, and 79 had not been possible previously but were easily accomplished in the present study.

Difference NOE Analysis and Distance Restraints. Long range NOEs, those NOE cross-peaks between ^{1}H on

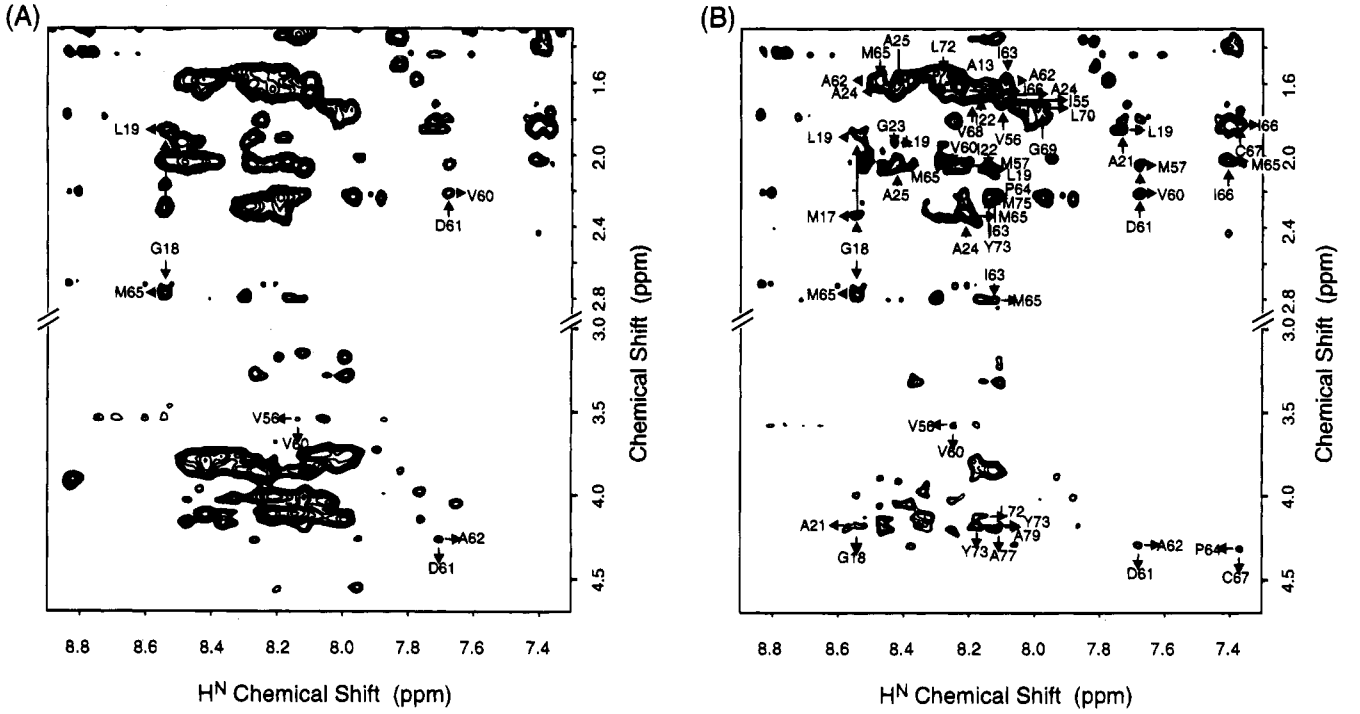

FIGURE 3: H^N-H^α and H^N-H^β regions of the 500 MHz ^1H NOESY spectrum and difference spectrum of Cys67−(maleimidomethyl)-PROXYL subunit *c*. The spectrum of the sample in which the spin label has been reduced (A) is compared with that of the reduced minus oxidized difference spectrum (B). Inter-residue cross-peaks, but not intraresidue cross-peaks, are labeled in the difference spectrum. For reference, several cross-peaks are labeled in the spectrum of the reduced protein.

residues distant from each other in the protein sequence, are essential for determining the three-dimensional structure of a protein from NMR data. The analysis of long range NOEs in the difference spectrum was complicated by two factors not ordinarily present in NOESY data. First, some ^1H on residues within the sphere of strong relaxation were within NOE distance of ^1H on residues which were not strongly affected by the spin label. The result was that a number of NOE cross-peaks were observed between ^1H assigned in this study and ^1H not assigned. These NOEs were ignored.

The other complicating factor involved quantitation of NOE cross-peak volumes. The NOE is a distance dependent phenomenon which is strongest at close distance and falls off with a reciprocal sixth power dependence on distance until it is not observable by the time the distance between ^1H reaches 5−6 Å. One typically integrates the volume of an NOE cross-peak and calculates or predicts an approximate distance range between the two ^1H from that value. In the difference NOE experiment, the intensity of a cross-peak is also a function of the distance of both nuclear spins from the unpaired electron. Hence the NOE cross-peaks in this study were used in a plus−minus fashion. The observation of a cross-peak was taken to mean the two ^1H were between 1.8 and 5.5 Å from each other. In principle, distance ranges could be approximated later, after a rough structure had been calculated. Distances from the nitroxide to a pair of interacting spins could be estimated, and a correction to the observed cross-peak volume could be calculated. This was not done in the present work, in part because the PROXYL group did not appear to be fixed in a single orientation (see below).

Analysis of the NOESY data provided a total of 150 inter-residue restraints for the structure calculation. Of these, 32 between ^1H in residues 10−25, 75 were in residues 52−78, and 43 were between residues in these two groups. An additional 27 backbone hydrogen-bonding restraints were based on slow H^N exchange when the protein was dissolved in ^2H solvent and were included only in later stages of the calculations. Cross peaks in the NOESY spectrum were observed for ^1H on the (maleimidomethyl)-PROXYL attached to the protein, and restraints from these NOEs were included in initial structural calculations. As no single conformation of the (maleimidomethyl)-PROXYL was found which would simultaneously satisfy all restraints, it was concluded that the probe was present in multiple conformations. The residues giving rise to NOE cross peaks with the (maleimidomethyl)-PROXYL (Ile63, Pro64, Ile66, Leu72, and Val74) were later found to lie on an external face of the second helix, where there would be little restriction on probe rotations.

Structural Results and Analysis. The best structures resulting from the distance geometry-simulated annealing protocol are presented in Figures 4 and 5. These were the nine structures which best fit the restraints (no restraint violations greater than 0.5 Å) and had the lowest calculated energies. The backbone was well-defined for residues 10−25 and 60−77 and more poorly defined for residues 52−59 (Figure 4). The sequence from Gln52 to Val60 is probably not disordered in solution. The segment appears to be helical from the short range NOEs and slow H^N exchange rates (Girvin & Fillingame, 1993). The different orientations calculated for this segment almost certainly result from the lack of long range NOE restraints to the first helix, which are beyond the detection range of the difference experiments. Such restraints would fix the orientation of this segment. The well-defined region of the protein backbone is shown in Figure 5, along with selected side chains of interest. The structural statistics are summarized in Table 2. In particular, the root mean square deviations from the restraints were less than 0.05 Å, and the pairwise rms deviations between corresponding atoms in the individual structures varied from 0.4 Å for backbone atoms in the well-defined region to 0.8

Table 1: Proton Chemical Shifts for PROXYL−Cys67 Subunit *c*[a]

residue	H^N	H^α	H^β	H^γ	H^δ and other
Tyr10		4.24	3.16		2,6H: 7.12; 3,5H: 6.73
Ala12		4.08	1.33		
Ala13		4.06	1.52		
Ala14		4.15	1.63		
Met16		4.46	2.07, 2.22	2.29, 2.48	
Met17		4.24	2.31	2.56, 2.66	
Gly18	8.53	4.21, 4.35			
Leu19	8.33	3.88	2.06, 2.25	1.86	0.97
Ala20	7.80	4.37	1.44		
Ala21	7.73	4.34	1.45		
Ile22	8.15	3.82	2.12	1.44, 0.93	0.88
Gly23	8.45	3.78			
Ala24	8.21	4.09	1.64		
Ala25	8.43	4.23	1.57		
Phe53	8.22	4.18	3.22		
Phe54	8.07	4.19	3.32		
Ile55		3.76	2.09	1.70, 1.47, 1.11	0.98
Val56	8.12	3.59	2.09	1.09, 0.94	
Met57	8.45	4.32	2.05, 2.16	2.48, 2.64	
Gly58	7.55	4.37			
Leu59	7.93	4.25	2.10, 2.28	1.91	
Val60	8.22	3.74	2.22	1.08, 1.01	
Asp61	7.67	4.62	3.03, 2.85		
Ala62	7.78	4.29	1.57		
Ile63	8.11	3.82	1.98	1.69, 1.27, 0.98	0.88
Pro64		4.29	1.88	2.36, 2.25	3.89
Met65	8.49	4.16	2.01, 2.30	2.62, 2.75	
Ile66	7.41	4.35	1.86	1.68, 1.26, 0.88	
Cys67	7.35	4.85	3.20		
Val68	8.17	3.80	2.23	1.15, 1.01	
Gly69	7.97	3.90			
Leu70	8.73	4.28	1.70, 1.89	0.89	
Gly71	8.30	3.87	4.14		
Leu72	8.29	4.11	1.90, 2.10	1.69	
Tyr73	8.14	4.19	3.22, 3.31		2,6H: 7.04; 3,5H: 6.74
Val74	8.24	3.56	2.26	1.00	
Met75	7.95	4.03	2.12, 2.22	2.35, 2.53	
Phe76	8.24	4.40	3.06, 3.17		3,4,5H: 7.18
Ala77	7.61	4.17	1.30		
Val78	7.98	3.91	2.23	0.97	
Ala79	8.06	4.19	1.48		

[a] All chemical shifts for subunit *c* at pD 5.5 and 37 °C are reported in ppm relative to tetramethylsilane (TMS).

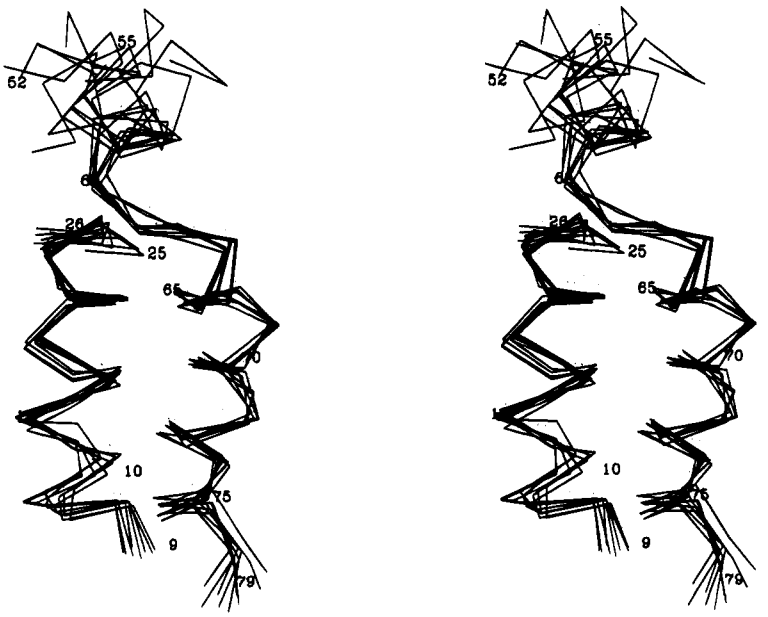

FIGURE 4: Stereoviews of the C^α traces of the nine lowest energy XPLOR substructures of subunit *c*. The substructure includes residues 9−26 and 52−79. The C^α traces were superimposed to minimize the rms deviations between backbone atoms for residues 10−25 and 60−74.

FIGURE 5: Stereoview of the backbone from the well-defined region of the nine best subunit *c* structures, along with selected side chains. The protein backbone is shown from residue 10 to 25 and 60 to 78. Side chains are shown for Tyr10, Ala20, Ala21, Ala24, Asp61, Pro64, Tyr73, and Phe76.

Table 2: Structural Statistics for PROXYL–Cys67 Subunit *c*

NOE rms deviations from restraints (Å)[a]	0.045 ± 0.008
rms deviations from ideal bond lengths (Å)	0.008 ± 0.000
rms deviations from ideal bond angles (deg)	1.154 ± 0.060
all observed residues (9–26 and 52–79)	
backbone heavy atom rms deviations (Å)	0.636 ± 0.171
all heavy atom rms deviations (Å)	0.817 ± 0.181
well-defined residues (10–25 and 60–77)	
backbone heavy atom rms deviations (Å)	0.376 ± 0.091
all heavy atom rms deviations (Å)	0.494 ± 0.097

[a] No NOE restraint violations were greater than 0.5 Å.

Å for all atoms in the entire region examined. The sequential atomic deviations and restraints are plotted in Figure 6. For the well-defined portion of subunit *c*, the backbone atomic rms deviations were 0.2–0.4 Å, but at the extreme edge of detection in the difference spectra (e.g., Gln52), they were as large as 5 Å. The number of restraints per residue (Figure 6) generally varied inversely with the deviations. The helical nature of this protein is reflected in the periodicity in the number of restraints per residue, with significantly more long range restraints observed for residues at the interface between the two helices. The backbone dihedral angles (Figure 7) are also consistent with a highly helical protein.

General Description of the Structure. The substructure of the well-defined region of subunit *c* consists of two gently curved α helices. The two helices cross at an angle of 30° at the point where Met17 and Ala21 are opposite Gly69 and Met65, respectively (Figure 8). The C-terminal helix is unraveled from Val60 to Ala62, changes direction by 25° at Pro64, and then continues through Phe76. The two helices are packed together throughout their length. The aromatic residues (Tyr10, Tyr73, and Phe 76) are clustered at the helix–helix interface, with distances from the edge of one ring to the centroid of the adjacent ring of 3.5–4.5 Å. Extensive interhelical van der Waals packing is observed throughout the contact region. The Tyr10 side chain extends between the side chains of Leu72 and Phe76. The Leu72 side chain packs against the shorter Ala13 side chain and is sandwiched between Tyr10 and Met17. Met17, in turn, packs against Gly69 and makes lengthy contacts with Met65 and Leu72 of the C-terminal helix. Met65 is opposite Ala20 and Ala21, again making extensive contact with Met17.

The overall folding of the protein backbone is illustrated in Figure 9, along with the locations of several residues of biological interest. The ribbon diagram highlights the helical nature of the structure, as well as the marked disruption of the C-terminal helix between Val60 and Ala62. The Asp61 side chain, which is thought to be involved in proton translocation, faces outward from the two helices and is at van der Waals contact distance from the Ala24 side chain. Ala24 is the site of a mutation which confers resistance to DCCD modification of Asp61 in native F_o (Fillingame et al., 1991). The Pro64 and Ala20 side chains are angled outwards, toward each other, with the nitrogen of each at a comparable distance from the Asp61 carboxyl group (6.8 Å for Ala20 and 5.8 Å for Pro64). A Pro residue is required at one of these two positions for a functional enzyme (Fimmel et al., 1983; Fraga, 1990). The Ile63–Pro64 peptide bond is *trans*. With the local unfolding over residues 60–62, the Ile63 and Asp61 side chains are on the same side of the backbone. The Val60 side chain is also near that of Asp61. The Cβ–Cβ distances from Asp61 are 4.8 Å for Val60, 5.4 Å for Ile63, 4.9 Å for Ala62, and 3.2 Å for Ala24.

DISCUSSION

Subunit *c* of the F_1F_o ATP synthase has been studied extensively by genetic and biochemical methods [for recent reviews, see Fillingame (1990) and Fillingame et al. (1992)]. On the basis of these studies, the protein has been predicted to fold as a pair of interacting transmembrane helices in the native F_o complex. In the present work, the structure of the region of subunit *c* within 15 Å of a spin label attached to residue 67 has been solved by spin label difference 2D NMR spectroscopy in a chloroform–methanol–water solvent. This region was indeed folded as a pair of antiparallel helices in solution. The C-terminal transmembrane helix exhibited local unfolding near Asp61, as had been observed previously in the secondary structure determined by Norwood et al. (1992) for subunit *c* in trifluoroethanol. The structure determined here confirms many predictions made from genetic studies and should help explain the effects observed in some mutant forms of the protein.

An alignment for the two transmembrane helices has been predicted from genetic studies which places residue 24 opposite residue 61. Miller et al. (1990) showed that a functional ATP synthase is obtained when the required

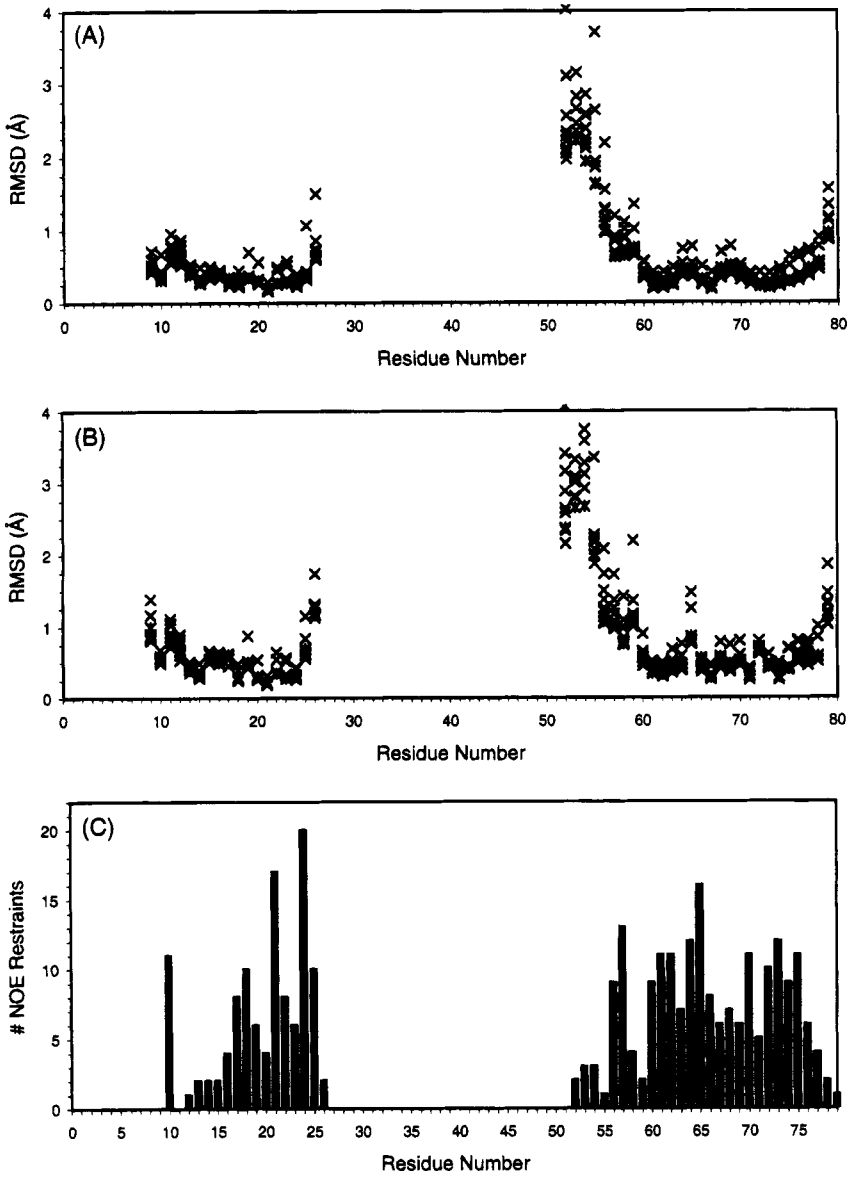

FIGURE 6: Atomic deviations and NOE restraints as a function of residue number. Pairwise rms deviations for the backbone atoms (C^{α}, C, N, and O) of each residue (A) and all non-hydrogen atoms (B) were calculated for each atom in each of the nine structures relative to the same atom in the other eight structures. The numbers of inter-residue NOE restraints for each residue in the sequence are also shown (C).

aspartic acid side chain was moved from position 61 to position 24 in an A24D, D61G double mutant. In the structure determined in the present work, the Ala24 side chain makes physical contact with the side chain of Asp61. If an Asp is substituted for Ala24 in this molecular model, the new carboxyl group is easily superimposed on the carboxyl group from Asp61.

An example of the usefulness of the substructure presented here for the interpretation of results from genetic experiments comes from an extension of the work of Miller et al. A selection for third-site mutations which optimize the function of the D24 G61 double mutant was performed, and two groups of optimizing mutations in subunit *c* were found (Fraga et al., 1994). The first group included residues near position 61. The second group included M65V, G71V, and M75I. The effects of substitutions at positions 65, 71, and 75 were originally interpreted in terms of altered flexibility of the C-terminal helix. In the present structure these residues are seen to be in the interface between the two transmembrane helices (Figure 8), and the substitutions would be expected to affect the packing of the two helices and possibly the orientation or exposure of the Asp at position 24.

The structure of a single form of subunit *c* is not sufficient to explain the effects observed in other mutants, however. The importance of a Pro at position 64 or 20 in subunit *c* from *E. coli*, demonstrated by Fimmel et al. (1983) and confirmed by Fraga (1990), provides an example. Replacement of Pro64 by Leu or Ala resulted in an inactive ATP synthase. In a second-site partial revertant, the compensating substitution was Pro for Ala20. Examination of the structural model suggests that Pro64 could be important for proper folding around Asp61, e.g., the unraveling and change in direction of the helix, the increased flexibility in this region, and/or a more polar environment along the protein backbone exposed by the extended conformation, conceivably including water molecules hydrogen bonding to the backbone. Partial restoration of function in the P64L/A20P or P64A/A20P double mutants might be rationalized as altering packing of the two helices to compensate for the effects of Pro replacement or as increasing the exposure or flexibility of a portion of the backbone.

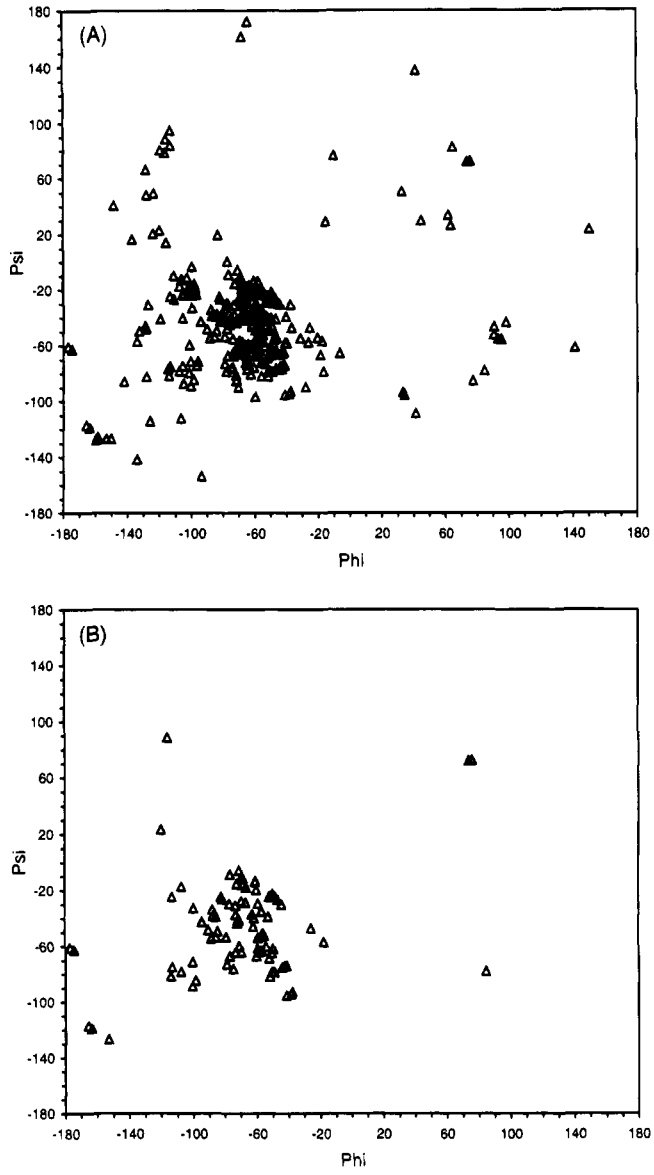

FIGURE 7: Ramachandran plot for substructures of subunit *c*. The backbone dihedral angles are shown for all residues in all nine structures (A) and for the well-ordered residues (10–25 and 60–74) in the lowest energy structures (B).

The structure calculated for subunit *c* demonstrates that one can fold a polytopic membrane protein in a detergent free medium with retention of tertiary structure. We have shown previously that subunit *c* dissolved in a chloroform–methanol–water mixture does fold in a way which retains biochemical properties observed in the native F_o complex. We have also calculated some long range interhelix distances from paramagnetic relaxation of resonances from ^1H in the N-terminal helix of subunit *c* by a nitroxide DCCD analog attached to Asp61 in the C-terminal helix (Girvin & Fillingame, 1994), as well as a small number of tertiary contacts [long range NOE cross-peak(s), reported in Girvin and Fillingame (1993)]. In the present study we have observed a complete local subset of such close contacts. The ability to study the structure of membrane proteins in organic solvent systems offers significant advantages over the use of detergent solutions for NMR, if native folding can be achieved. Detergents increase the effective size of a protein, which broadens resonances and could potentially increase background signals if the detergent is not completely deuterated.

The demonstration of tertiary structure for membrane proteins dissolved in organic mixtures has been remarkably difficult. In other NMR studies of subunit *c* in chloroform–methanol (Moody et al., 1987) and trifluoroethanol (Norwood et al., 1992), no long range NOESY cross peaks between residues in adjacent helices could be identified. Because of extensive spectral overlap, these authors could not distinguish whether long range interactions between helices were present. Similarly, in NMR studies of an N-terminal proteolytic fragment (residues 1–71; Pervushin et al., 1994) and a C-terminal fragment (residues 163–231; Barsukov et al., 1992) of bacterioopsin carried out in chloroform–methanol, many short range NOE contacts were identified and secondary structures could be modeled, but no long range contacts were detected. The reasons for the difficulty in observing tertiary contacts are not yet resolved. Spectral overlap can be severe with these membrane proteins (e.g., Figure 2). The amino acid composition of transmembrane segments are highly enriched in a small subset of residue types (Leu, Ile, Val, Ala, Gly, and Met). The dispersion of chemical shifts would be expected to be small in a protein of uniform amino acid composition and secondary structure. Orekhov et al. (1994) have speculated that, at least in the case of the 1-71 fragment of bacterioopsin, conformational exchange detected by dynamics measurements could be responsible for the lack of observed long range NOE cross peaks. The spin label difference experiments described here can at least alleviate the difficulties arising from spectral overlap. It remains to be determined whether other solvent conditions are conducive to tertiary folding of transmembrane proteins and whether other membrane proteins fold properly in aqueous–organic solvent mixtures.

The use of 2D spin label difference NMR spectroscopy enabled us to quickly and accurately interpret the ^1H COSY, TOCSY, and NOESY spectra of the region of subunit *c* near the modification site and use this information to calculate a detailed structure for this region of the protein. Our earlier attempts to determine covalent connectivities and through space distances resulted in only limited success (Girvin & Fillingame, 1993). Many correlation cross-peaks, and most NOE cross-peaks, could not be assigned to a unique pair of spins in the full spectra of subunit *c*. The ease and confidence with which assignments could be made from the difference spectra were striking. De Jong et al. (1988) and Folkers et al. (1993) have already demonstrated the feasibility of observing clean difference cross-peaks from spin label difference spectra and obtaining useful structural information about the DNA binding site on M13 gene V protein from such data. Their approach involved subtracting spectra of a protein complexed with spin-labeled DNA from spectra of the protein alone. One difficulty with their method was the need to detect and compensate for altered chemical shifts resulting from conformational changes of the protein on ligand binding. In reduced minus oxidized difference spectra, the same group is present in both the paramagnetic and diamagnetic spectra, so conformational changes are no longer a concern. A valid concern, however, is that a modifying reagent perturbs native folding of the protein. As stated earlier, the modification site on subunit *c* was chosen to avoid the expected packing interface. Experimental results indicate that the modification by (maleimidomethyl)-PROXYL at Cys67 did not alter the folding of subunit *c*. Identical NOE cross peaks involving the aromatic residues observed in the unmodified protein (Girvin & Fillingame,

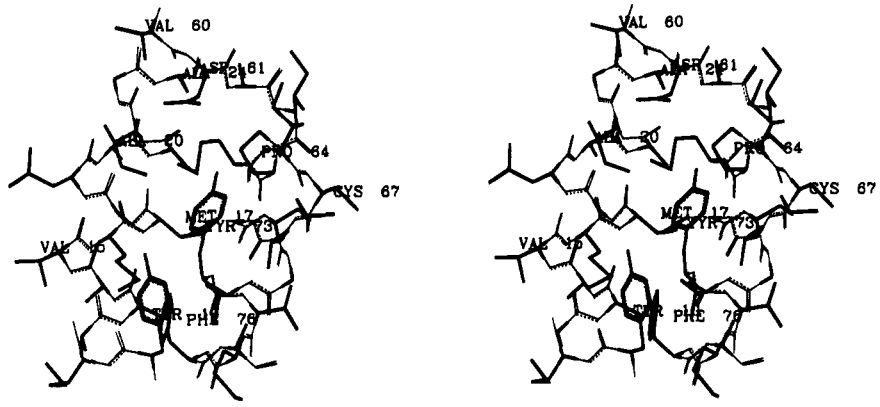

FIGURE 8: Stereoview of the well-defined region of the best structure of subunit *c* shown with all side chains. This structure had the lowest calculated energy and best fit to the NOE constraints. Residues from Tyr10 to Ala25 and Val60 to Ala77 are shown. Hydrogens are omitted for clarity.

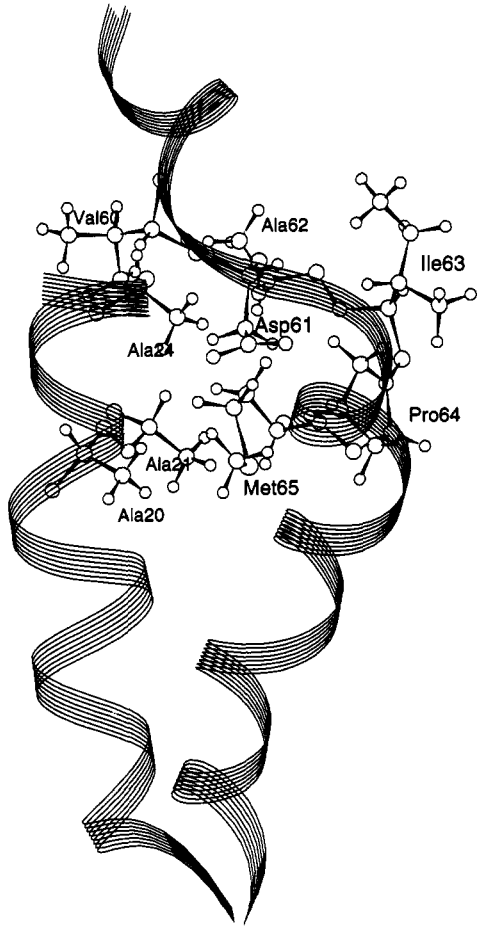

FIGURE 9: Ribbon representation of the backbone of subunit *c*. The two helices and the disruption of the C-terminal α helix between Val60 and Ala62 can best be seen using a ribbon representation. Also shown are side chains with hydrogens for residues neighboring Asp61.

1993) were also observed in the modified protein. Furthermore, with the exception of residues adjacent to the site of mutation and modification (residue 67), no significant changes in ^{1}H chemical shifts were detected.

Spin label difference 2D NMR spectroscopy should provide a way to use NMR for detailed structural studies of moderately large proteins. Resonance overlap is currently the major limitation in NMR studies of large proteins. Subspectra of a single 15 Å radius spherical region of a much larger molecule would eliminate most problems of overlap. The other factor limiting NMR study of large proteins is the increase in line widths resulting from longer correlation times. As described by Campbell et al. (1973), subtracting a paramagnetic spectrum from a spectrum without the paramagnetic species results in a resolution enhancement (narrowing of line widths in the spectrum) analogous to calculation of a convolution difference spectrum. Mathematically broadening the spectra of subunit *c* to simulate a 50 kDa protein still gave assignable 2D spectra, and in fact, we routinely broadened our spectra by approximately 10 Hz to improve the subtraction. Furthermore, techniques are available to reduce line widths, such as fractional deuteration (LeMaster & Richards, 1988). The combination of site-directed mutagenesis to create a single labeling site with spin label difference spectroscopy holds great promise for determining structural and dynamic features of domains of interest in larger proteins.

ACKNOWLEDGMENT

We thank Ying Zhang (a Ph.D. student) for creating the original A67C mutation and Ms. Mary Gillis for constructing the overproducing strain. The NMR spectra were obtained at the National Magnetic Resonance Facility at Madison, without whose existence and support this work would not have been possible. We thank Facility Director Dr. John Markley for his support. We also thank Dr. Axel Brünger (and Yale University) for making his molecular simulation and distance geometry software available at modest cost and in portable source code form.

REFERENCES

Abrahams, J. P., Leslie, A. G. W., Lutter, R., & Walker, J. E. (1994) *Nature 370*, 621.

Barsukov, I. L., Nolde, D. E., Lomize, A. L., & Arseniev, A. S. (1992) *Eur. J. Biochem. 206*, 665.

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

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166.

Brünger, A. T. (1993) *XPLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Campbell, I. D., Dobson, C. M., Williams, R. J. P., & Xavier, A. V. (1973) *J. Magn. Reson. 11*, 172.

Chylla, R. A., & Markey, J. L. (1993) *J. Magn. Reson., Ser. B 102*, 148.

Cross, R. L. (1981) *Annu. Rev. Biochem. 50*, 681.

De Jong, E. A. M., Claesen, C. A. A., Daemen, C. J. M., Harmsen, B. J. M., Konigs, R. M. H., Tesser, G. I., & Hilbers, C. W. (1988) *J. Magn. Reson. 88*, 197.

Fillingame, R. H. (1990) in *The Bacteria* (Krulwich, T. A., Ed.) Vol XII, pp 345−391, Academic Press, New York.
Fillingame, R. H., Oldenburg, M., & Fraga, D. (1991) *J. Biol. Chem. 266*, 20934.
Fillingame, R. H., Girvin, M. E., Fraga, D., & Zhang, Y. (1992) *Ann. N. Y. Acad. Sci. 671*, 323.
Fimmel, A. L., Jans, D. A., Langman, L., James, L. B., Ash, G. R., Downie, J. A., Senior, A. E., Gibson, F., & Cox, G. B. (1983) *Biochem. J. 213*, 451.
Folkers, P. J. M., van Duynhoven, J. P. M., Lieshout, H. T. M., Harmsen, B. J. M., van Boom, J. H., Tesser, G. I., Konings, R. N. H., & Hilbers, C. W. (1993) *Biochemistry 32*, 9407.
Foster, D. L., & Fillingame, R. H. (1982) *J. Biol. Chem. 257*, 2009.
Fraga, D. (1990) Ph.D. Thesis, University of Wisconsin, Madison, WI.
Fraga, D., Hermolin, J., & Fillingame, R. H. (1994) *J. Biol. Chem. 269*, 2562.
Girvin, M. E. (1994) *J. Magn. Reson., Ser. A 108*, 99.
Girvin, M. E., & Fillingame, R. H. (1993) *Biochemistry 32*, 12167.
Girvin, M. E., & Fillingame, R. H. (1994) *Biochemistry 33*, 665.
Kosen, P. A. (1989) *Methods Enzymol. 177*, 86.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1.
LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142.
Macura, S., Wüthrich, K., & Ernst, R. R. (1982) *J. Magn. Reson. 46*, 269.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967.
Miller, M. J., Fraga, D., Paule, C. R., & Fillingame, R. H. (1989) *J. Biol. Chem. 264*, 305.
Miller, M. J., Oldenburg, M., & Fillingame, R. H. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4900.
Moody, M. F., Jones, P. T., Carver, J. A., Boyd, J., & Campbell, I. D. (1987) *J. Mol. Biol. 193*, 759.
Moonen, C. T. W., Scheek, R. M., Boelens, R., & Müller, F. (1984) *Eur. J. Biochem. 141*, 323.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317.
Norwood, T. J., Crawford, D. A., Steventon, M. E., Driscoll, P. C., & Campbell, I. D. (1992) *Biochemistry 31*, 6285.
Orekhov, V. Y., Pervushin, K. V., & Arseniev, A. S. (1994) *Eur. J. Biochem. 219*, 887.
Pervushin, K. V., Orekhov, V. Y., Popov, A. I., Musina, L. Y., & Arseniev, A. S. (1994) *Eur. J. Biochem. 219*, 571.
Senior, A. E. (1988) *Physiol. Rev. 68*, 177.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zhang, Y., Oldenburg, M., & Fillingame, R. H. (1994) *J. Biol. Chem. 269*, 10221.
Zolnai, Zs., Macura, S., & Markley, J. L. (1986) *Comput. Enhanced Spectrosc. 3*, 141.

BI942335X