Heteronuclear ^1H–^{15}N Nuclear Magnetic Resonance Studies of the c Subunit of the *Escherichia coli* F_1F_0 ATP Synthase: Assignment and Secondary Structure[†]

Timothy J. Norwood,[‡] D. Arthur Crawford, Mary E. Steventon, Paul C. Driscoll,[§] and Iain D. Campbell*

Department of Biochemistry, University of Oxford, South Parks Road, Oxford, U.K. OX1 3QU

Received December 9, 1991; Revised Manuscript Received March 26, 1992

ABSTRACT: Nuclear magnetic resonance (NMR) studies of the c subunit of F_1F_0 ATP synthase from *Escherichia coli* are presented. A combination of homonuclear (^1H–^1H) and heteronuclear (^1H–^{15}N) 2D and 3D methods was applied to the 79-residue protein, dissolved in trifluoroethanol. Resonance assignment for all the backbone amide groups and many $C^{\alpha}H$ side-chain protons was achieved. Analysis of inter- and intraresidue ^1H–^1H nuclear Overhauser effect (NOE) data and scalar coupling constant information indicates that this protein contains two extended regions of predominant α-helical character (residues 10–40 and 48–77) separated by an eight-residue segment which displays little evidence of ordered secondary structure. This model is consistent with information about the molecular motion of the protein deduced from ^{15}N–^1H heteronuclear NOE data and observed pK_a values of carboxylic acid groups.

The synthesis of ATP, the final step in the electron-transport chain, is achieved by the F_1F_0 ATP synthase. This enzyme allows the reentry of protons down a potential gradient into the cell or mitochondrion and uses the energy released to drive ATP synthesis from ADP and P_i. The ATP synthase is a large complex consisting of two components: the water-soluble F_1 which is the site of ATP synthesis (Kagawa & Racker, 1966), and the membrane-bound F_0, which is thought to form the proton channel (Yoshida et al., 1977). Although this enzyme has been the subject of many investigations, little is known of the structure of the complex at atomic resolution, especially the membrane-spanning portion. In *Escherichia coli*, the F_0 section of the F_1F_0 ATP synthase is made up of three subunits named a, b, and c (in order of decreasing molecular weight), all three being required to reconstitute a functional protein (Schneider & Altendorf, 1985). It has been reported (Foster & Fillingame, 1982) that the stoichiometry in *E. coli* is a_1, b_2, c_{10}, but other ratios have been reported, particularly in other organisms [e.g., see Senior (1988)].

The *E. coli* c subunit [also called the dicyclohexylcarbodiimide (DCCD)[1] binding peptide and the proteolipid] is 79 amino acid residues in length, molecular weight 8290 (Senior, 1988). This very hydrophobic protein is insoluble in water and was first purified from *E. coli* in chloroform/methanol mixtures (Fillingame, 1976). It has also been extracted using butanol (Sigrist et al., 1977) and detergents (Schneider & Altendorf, 1985). The sequences of a large number of proteolipid subunits from a wide range of sources have been determined [e.g., see Sebald and Hoppe (1987)], and they display a number of interesting common characteristics. The c subunit is short (70–82 residues); 2 stretches of about 25 amino acids in length have been suggested to form transmembrane α-helices (Hoppe & Sebald, 1984; Mao et al., 1982; Senior, 1983). Within these proposed transmembrane helices are the helix-destabilizing proline residue at position 64 and the charged residue Asp-61 (this position is occupied by glutamate in many organisms) which is thought to be the site of DCCD binding. The section lying between the proposed helices, which may function as a turn or loop (Mao et al., 1982; Senior, 1988; Hoppe & Sebald, 1984), contains an invariant proline residue at position 43 and, in *E. coli*, an additional proline at position 47.

The structure of the c subunit has been studied using circular dichroism (Mao et al., 1982) both in solution in TFE and in DMPC vesicles. In both cases, the protein appeared to be largely α-helical. Structure prediction (Mao et al., 1982; Senior, 1983; Hoppe & Sebald, 1984) and chemical modification experiments (Hoppe & Sebald, 1984) also suggest that the protein forms two α-helices which cross the membrane separated by a hairpin bend or turn region. A proton NMR study of the c subunit in chloroform/methanol (Moody et al., 1987) also presented evidence for helical character.

Here we report an assignment of the homonuclear (^1H–^1H) and heteronuclear (^1H–^{15}N) NMR spectra of the ^{15}N-labeled protein in TFE with information about the protein's secondary structure and dynamics.

Materials and Methods

Sample Preparation. Protein was prepared essentially by the method of Fillingame (1976). Cell lysis was achieved by incubation at 37 °C with lysozyme (1 mg/g of cells) and DNase 1 (20 μg/g of cells) for 3 h (Hoppe et al., 1980). Alternatively, cells were lysed by beating for 5 min using 150–212-μm glass beads, the chamber being cooled by ice. On elution from the Sepharose Cl-6B column, the protein was collected in 0.5-mL fractions. Fractions containing protein and exhibiting an absorbance at 280 nm were pooled and air-dried. The protein was judged to be greater than 95% pure on the basis of gel electrophoresis and NMR experiments.

[†] This is a contribution from the Oxford Centre for Molecular Sciences which is supported by the SERC and MRC.

* Address correspondence to this author.

[‡] Present address: Department of Chemistry, Leicester University, Leicester, U.K. LE1 7RH.

[§] P.C.D. is a Royal Society Research Fellow.

[1] Abbreviations: CD, circular dichroism; COSY, correlation spectroscopy; DCCD, dicyclohexylcarbodiimide; HMQC, heteronuclear multiple-quantum coherence; HOHAHA, homonuclear Hartman–Hahn spectroscopy; HSQC, heteronuclear single-quantum coherence; *N*-Cbz, *N*-carbobenzoxy; *N-t*-Boc, *N-tert*-butoxycarbonyl; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; SQC, single-quantum coherence; SQMBC, single-quantum multiple-bond correlation; TFE, trifluoroethanol; TFE-CD_2OH, partially deuterated trifluoroethanol.

0006-2960/92/0431-6285$03.00/0

Unlabeled protein was prepared from the strain of *E. coli* CM2786, an overproducer donated by Kasper Von Meyenberg at the Technical University of Denmark (Von Meyenberg et al., 1984). An auxotrophic strain of *E. coli* DL39G (auxotrophic for Asp, Gly, Ile, Leu, Phe, Tyr, and Val) was a gift from David Lemaster (Yale University). This was transformed with the plasmid from CM2786, prepared by the alkaline lysis method (Maniatis et al., 1982) with ampicillin and tetracycline for selection. Auxotrophic overproducers were assessed for protein activity by preparing membrane vesicles according to Hertzberg and Hinkle (1974) and assaying the vesicles for ATP hydrolytic activity by coupling ADP production to NADH oxidation using pyruvate kinase and lactate dehydrogenase.

Selective labeling at glycine residues was achieved by providing the auxotrophic overproducer with [^{15}N]glycine. Uniformly ^{15}N-labeled protein was purified from CM2786 cells fed with ^{15}N-labeled ammonium chloride.

NMR Spectroscopy. All NMR experiments were recorded on a home-built NMR spectrometer operated by a GE 1280 computer and using an 11.8-T vertical-bore Oxford Instruments magnet operating at 500.1 MHz and ^{1}H and at 50.7 MHz for ^{15}N at 25 °C. For 2D experiments, quadrature detection in F_1 was achieved according to the States method (States et al., 1982). For the initial assignment work, both homonuclear ^{1}H and heteronuclear ^{1}H-detected ^{1}H–^{15}N correlation spectra were recorded. The former included COSY, HOHAHA, and NOESY, while the latter included ^{15}N SQC–COSY, SQC–NOESY, and HOHAHA (Norwood et al., 1989, 1990a), the SQMBC long-range ^{1}H–^{15}N correlation experiment (Norwood et al., 1990b), HMQC–COSY (Gronenborn et al., 1989), and the heteronuclear ^{1}H–^{15}N NOE experiment (Kay et al., 1989b). For all ^{15}N experiments, $\Delta = 5.25$ ms, and for SQMBC, $\tau = 50$ ms; for NOESY experiments, a mixing time $t_m = 200$–400 ms was used. For all experiments, 2K t_2 and either 256 or 512 t_1 data points were acquired. Data were processed using a double-exponential (Lorentzian To Gaussian) function in F_2 and a shifted sine-bell in F_1.

The $^3J_{HN\alpha}$ coupling constants were extracted from the F_1 dimension of the HMQC–COSY spectrum by fitting line shapes to the multiple-quantum doublets (Kay et al., 1989a). Typical examples are shown in Figure 3.

To confirm some assignments and to resolve some of the overlap in the C$^\alpha$H region, three-dimensional NOESY–HMQC and HOHAHA–HMQC (Marion et al., 1989) experiments were recorded. A version of the three-dimensional HMQC–NOESY–HMQC experiment (Frenkiel et al., 1990) using HSQC sequences (i.e., HSQC–NOESY–HSQC) was also carried out.

pH Titration. The ionization states of the acidic residues in the c subunit were analyzed by adjusting the pH* (uncorrected glass electrode reading) of the sample with DCl and NaOD dissolved in TFE-CD$_2$OH and monitoring the chemical shift of the amide nitrogens. Care was taken to maintain a constant salt concentration and water content as there were found to affect the chemical shift. A two-dimensional autocorrelation HSQC spectrum was acquired at each pH*, and the ^{15}N shifts were plotted against pH* to give titration curves for the residues of interest. The pK_a values of some of the residues were estimated from those of neighboring residues due to difficulties in assignment of these peaks at certain pH* values. Titrations of the blocked amino acids *N*-Cbz-aspartate and *N*-*t*-Boc-L-glutamic acid α-benzyl ester were also carried out under similar conditions to allow comparison between the ionization states of the titrating groups in the protein and the corresponding free carboxyl side chains.

Results

Assignment Strategy. In general, the ^{1}H spectra of the c subunit were characterized by relatively broad lines and regions of limited chemical shift dispersion leading to considerable overlap. The spectral assignment thus required a combination of homonuclear and heteronuclear techniques even though the protein is relatively small.

Glycine residues were identified from ^{15}N HSQC–COSY spectra of a selectively enriched sample. Other residues were identified from a uniformly enriched sample using a combination of ^{15}N HSQC–COSY and HSQC–HOHAHA together with ^{1}H COSY and HOHAHA spectra. For side-chain resonances beyond C$^\beta$H, low signal intensities in the ^{15}N SQC–HOHAHA spectrum made greater reliance on the ^{1}H COSY and HOHAHA spectra necessary. Side-chain assignment beyond this point is incomplete due to the poor ^{1}H spectral dispersion.

A sequence-specific assignment of spin systems was achieved using interresidue NOEs and long-range ^{1}H–^{15}N interresidue scalar couplings. The NOEs of greatest importance in this respect were $d_{NN}(i, i+1)$ and $d_{\alpha N}(i, i+1)$. These were obtained from ^{15}N SQC–NOESY and ^{1}H NOESY spectra. The ^{15}N SQC–NOESY spectrum is shown in Figure 1. Interresidue ^{1}H–^{15}N scalar coupling correlations were observed in the SQMBC spectrum. A list of the resulting sequential assignments is given in Table I.

NOE Restraints and Coupling Constants. Due to limited assignment of side-chain protons beyond C$^\beta$H, most NOEs were extracted from ^{15}N SQC–NOESY spectra and consequently are restricted to correlations involving amide protons. Short-range NOEs are given in Figure 2, and $^3J_{HN\alpha}$ ^{1}H–^{15}N coupling constants, measured by line-shape analysis, are listed in Table I.

These results permit the definition of some regions of secondary structure (Wüthrich, 1986). The best defined of these, in terms of both NOEs and scalar coupling constants, is a relatively long α-helix between residues 10 and 40. This is characterized by coupling constants of 5 Hz or less and numerous $d_{\alpha N}(i, i+3)$ NOEs. There is also evidence for α-helical regions between residues 48 and 76. It should be noted that the relative paucity of longer range NOEs is in some places due to severe cross-peak overlap rather than the absence of NOEs. The proportion of the protein, thus defined as α-helical, is in approximate agreement with circular dichroism studies (Mao et al., 1982) (72%). The region between helices one and two, around residue 43, might correspond to the turn suggested by structure prediction studies. The helical region following the turn is disrupted by Pro-64, as might be expected.

Dynamic Information. A number of studies have illustrated the utility of ^{15}N labeling of proteins to probe the dynamics of the polypeptide backbone (Kay et al., 1989b; Clore et al., 1990). The ^{15}N–^{1}H heteronuclear NOE experiment proves particularly suitable for probing parts of the molecule which are relatively dynamically disordered with respect to the main core. Figure 4 illustrates the ^{15}N–^{1}H NOEs measured for the c subunit at 500.1 MHz (for ^{1}H). The ^{1}H–^{15}N NOE varies from +0.8 to −3.5 for correlation times (τ_c) in the range from 1.6×10^{-8} to 1×10^{-10} s. The NOE is therefore sensitive to the effective correlation time in this range with a small or negative NOE corresponding to relatively mobile residues.

Several distinct features are apparent in Figure 4. The ^{1}H–^{15}NOEs are lowest at the two termini of the protein, presumably reflecting the relatively high mobility of residues

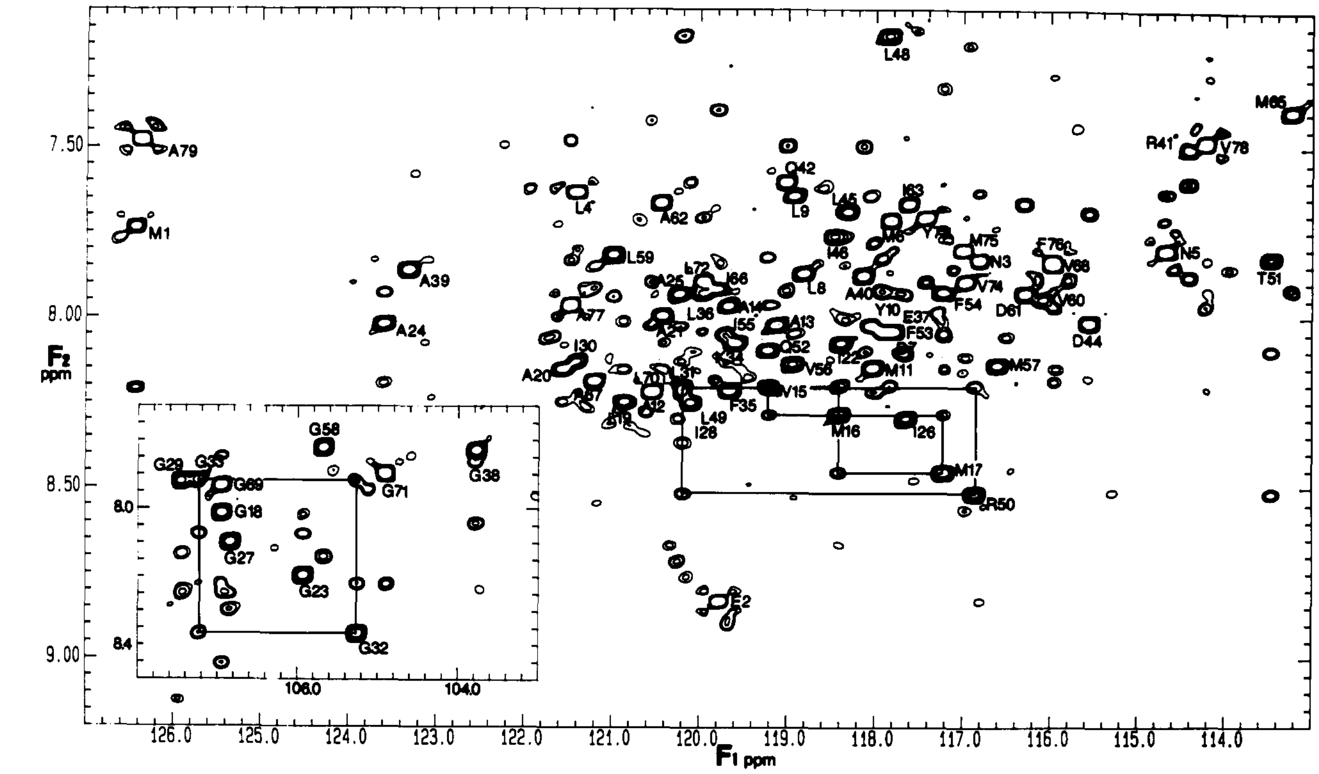

FIGURE 1: ^{15}N SQC–NOESY spectrum of the c subunit in TFE-CD_2OH at pH* 4.8 and 47 °C. F_1 and F_2 axes represent ^{15}N and ^1H chemical shifts, respectively. Assignments corresponding to Table I are indicated. (The glycine region of the spectrum is shown in the inset.) Representative NOE connectivities are indicated between residues 15, 16, and 17; between 32 and 33; and between 49 and 50.

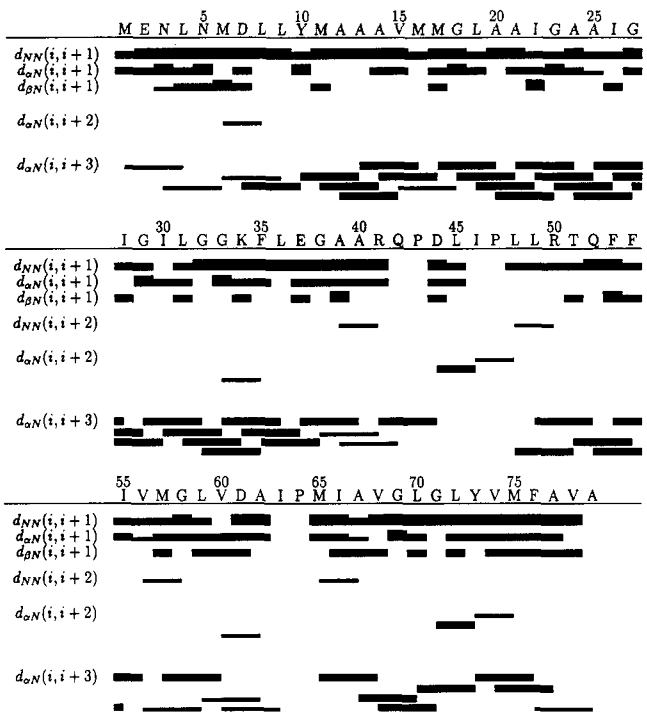

FIGURE 2: Sequence of the c subunit from *E. coli* with a summary of short-range NOE connectivities. All (i, i+1) NOEs are represented by black bars directly beneath the residue number. Connectivities (i, i+2), (i, i+3), and i, i+4) are represented by tie-bars between connected residues. The thickness of the bars indicates the presence of a strong, medium, or weak NOE.

in these regions. The NOEs observed in residues away from the ends rapidly increase, reaching the edge of a plateau at residues 8 and 75. The greater number of residues preceding the plateau at the N-terminus (8 instead of 4) suggests that this section might, in situ, protrude from the nonpolar membrane, a hypothesis that is consistent with the presence of hydrophilic residues such as Glu-2 and Asp-8. The most prominent feature of the plateau is a minimum at residues 41 and 42. This minimum was also observed in ^{15}N T_1 experiments collected at a ^1H frequency of 500.1 MHz (data not shown) and in HSQC line-width measurements at 360 and 500.1 MHz (^1H). The smaller NOE values in this region reflect increased mobility, and coincide with the region where a turn is predicted in the hairpin model of the protein. There is also some evidence for a decrease in the NOE values immediately prior to Asp-61, the residue thought to be involved in proton translocation. This apparent decrease is sustained through to the C-terminus, implying that this region also has increased flexibility possibly due to the presence of Pro-64.

Ionization States. pK_a values were measured for the five titrating aspartate and glutamate carboxylate side chains. Where possible, this was done by observing the amide chemical shift directly (see, for example, the results for Glu-37 shown in Figure 5). In cases of severe overlap, the assignment of pK_a values was assisted by observing the titration behavior of neighbouring groups. The observed values for the aspartates at positions 44, 61, and 7 were 4.0, 5.3, and 5.5, respectively. These should be compared with a value of 4.3 obtained for the *N*-Cbz-aspartate side chain in TFE. The pK_a values for the glutamate residues at positions 2 and 37 were measured to be 5.5 and 4.0, respectively, to be compared with the value of 5.29 observed for *N*-*t*-Boc-L-glutamic acid α-benzyl ester. The pK_a of the C-terminal alanine was found to be 5.0.

Discussion

Previous evidence for a helical "hairpin" conformation of the c subunit relied chiefly on chemical labeling, structural prediction, and circular dichroism experiments. By using ^{15}N-labeled protein and two- and three-dimensional heteronuclear ^1H–^{15}N experiments, we have obtained evidence for many of the predicted secondary structural elements for this protein. There is a long relatively well-defined helical stretch

Table I: ^{15}N and ^{1}H Chemical Shifts for the c Subunit (ppm, ±0.01) in TFE-CD_2OH at pH 4.8, 47 °C

residue	amide NH ^{15}N	amide NH ^{1}H	$C^{\alpha}H$	$C^{\beta}H$	other	$^{3}J_{HN\alpha}$ (Hz)	error
Met-1	126.45	7.74	4.53	2.15, 2.025	$C^{\gamma}H$ 2.57, 2.62		
Glu-2		8.83	4.22	2.06	$C^{\gamma}H$, 2.425, 2.37	5.03	±0.04
Asn-3	117.10	7.83	4.66	2.91, 2.77			
Leu-4	121.50	7.63	4.27	1.725		4.64	±0.01
Asn-5	114.95	7.80	4.53	2.83		5.69	±0.10
Met-6	117.95	7.72	4.33	2.14	$C^{\gamma}H$ 2.58, 2.65		
Asp-7	117.70	8.04	4.49	3.00, 2.83		4.11	±0.08
Leu-8	118.90	7.87	4.20	1.89		4.30	±0.01
Leu-9	118.90	7.64	4.16	1.85		4.85	±0.01
Tyr-10	117.90	8.02	4.32	3.20			
Met-11	118.00	8.15	4.24	2.225, 2.30	$C^{\gamma}H$ 265, 276	4.65	
Ala-12	120.60	8.22	4.08	1.55			
Ala-13	119.05	8.02	4.08	1.52		4.04	±0.01
Ala-14	119.70	7.97	4.075	1.525		4.10	
Val-15	119.20	8.21	3.75	2.29	$C^{\gamma}H_3$	5.10	±0.01
Met-16	118.30	8.28	4.24	2.29, 2.22	$C^{\gamma}H$ 2.65, 2.76	3.92	
Met-17	117.20	8.455	4.275/4.29	2.15, 2.275	$C^{\gamma}H$ 2.66, 2.78	4.52	±0.01
Gly-18	107.10	8.02	≈3.95				
Leu-19	120.85	8.25	4.23	1.99			
Ala-20	121.55	8.15	4.155	1.575		4.42	±0.01
Ala-21	120.40	7.99	4.18	1.61		4.44	±0.02
Ile-22	118.35	8.075	3.925	2.015		4.97	±0.02
Gly-23	105.95	8.195	3.84				
Ala-24	123.50	8.025	4.23	1.55		4.73	
Ala-25	120.20	7.93	4.12	1.61			
Ile-26	117.60	8.30	3.90	2.00		5.12	±0.01
Gly-27	106.80	8.10	≈3.86				
Ile-28	120.05	8.25	3.94	1.985		4.78	±0.05
Gly-29	107.55	7.92	3.84				
Ile-30	121.30	8.13	4.00	2.08		4.76	±0.01
Leu-31	120.15	8.23	4.15	1.80, 1.915			
Gly-32	105.35	8.37	3.96				
Gly-33	107.20	7.92	3.94				
Lys-34	119.60	8.075	4.245			4.73	±0.04
Phe-35	119.50	8.22	4.405	3.28		4.60	
Leu-36	119.80	7.93	4.13	1.835, 1.925			
Glu-37	117.65	8.04	3.98	2.16, 2.28	$C^{\gamma}H$ 2.48	3.92	±0.05
Gly-38	104.10	7.825	≈3.88				
Ala-39	123.25	7.87	4.145	1.40		4.74	±0.01
Ala-40	118.15	7.875	4.21	1.48		4.94	±0.01
Arg-41	114.85	7.50	4.40	1.88, 2.05	$C^{\gamma}H$ 1.725		
Gln-42	119.10	7.60	4.57	2.24, 2.125	$C^{\gamma}H$ 2.48	5.74	
Pro-43							
Asp-44	115.90	8.02	4.54	2.775		6.12	
Leu-45	118.70	7.69	4.39	1.77		6.13	±0.03
Ile-46	118.40	7.765	4.00	2.22		3.61	±0.01
Pro-47							
Leu-48	117.80	7.175	4.20			6.04	±0.06
Leu-49	120.15	8.19	4.14				
Arg-50	117.00	8.52	4.075	1.95	$C^{\gamma}H$ 1.715	3.71	±0.02
Thr-51	113.60	7.825	≈4.10	4.46	$C^{\gamma}H_3$ 1.38	4.69	±0.01
Gln-52	119.10	8.09	4.08	≈2.26	$C^{\gamma}H$ 2.50, 2.38	3.77	
Phe-53	117.80	8.04	4.26	3.18		4.11	±0.08
Phe-54	117.20	7.925	4.27/4.29	3.31			
Ile-55	119.85	8.05	3.81	2.10			
Val-56	119.05	8.135	3.725	2.145	$C^{\gamma}H_3$ 1.09	4.44	±0.01
Met-57	116.60	8.14	4.27			4.23	±0.04
Gly-58	105.55	7.825	3.86				
Leu-59	121.10	7.825	4.31	1.92		5.85	±0.02
Val-60	116.10	7.94	3.90	2.225	$C^{\gamma}H_3$ 1.06	5.04	±0.01
Asp-61	116.20	7.93	4.66	2.90		5.98	±0.03
Ala-62	120.50	7.665	4.38	1.53		6.40	
Ile-63	117.60	7.665	4.09	2.13		3.96	±0.01
Pro-64							
Met-65	113.40	7.39	4.30	2.22	$C^{\gamma}H$ 2.68, 2.62	5.97	±0.01
Ile-66	119.65	7.92	3.88	2.05			
Ala-67	121.20	8.18	4.09	1.51		3.73	
Val-68	115.95	7.84	3.90	2.24	$C^{\gamma}H_3$ 1.12		
Gly-69	106.95	7.93	≈3.95				
Leu-70	120.45	8.22	4.36	1.84, 1.70		5.05	±0.04
Gly-71	104.90	7.89	3.87				
Leu-72	119.90	7.89	4.20				
Tyr-73	117.30	7.71	4.30	3.26, 3.175		4.65	±0.01
Val-74	116.95	7.89	3.73	2.21	$C^{\gamma}H_3$ 1.11, 0.97	4.89	±0.05
Met-75	117.00	7.80	4.15	2.06, 1.92	$C^{\gamma}H$ 2.51, 2.39	4.80	±0.03

Table I (Continued)

residue	amide NH ^{15}N	amide NH ^{1}H	$C^{\alpha}H$	$C^{\beta}H$	other	$^3J_{HN\alpha}$ (Hz)	error
Phe-76	115.80	7.83	4.55	3.225, 3.065			
Ala-77	121.40	7.965	4.23	1.34		6.07	
Val-78	114.15	7.48	4.125	2.22	$C^{\gamma}H$ 0.97		
Ala-79	125.70	7.48	4.26	1.41			

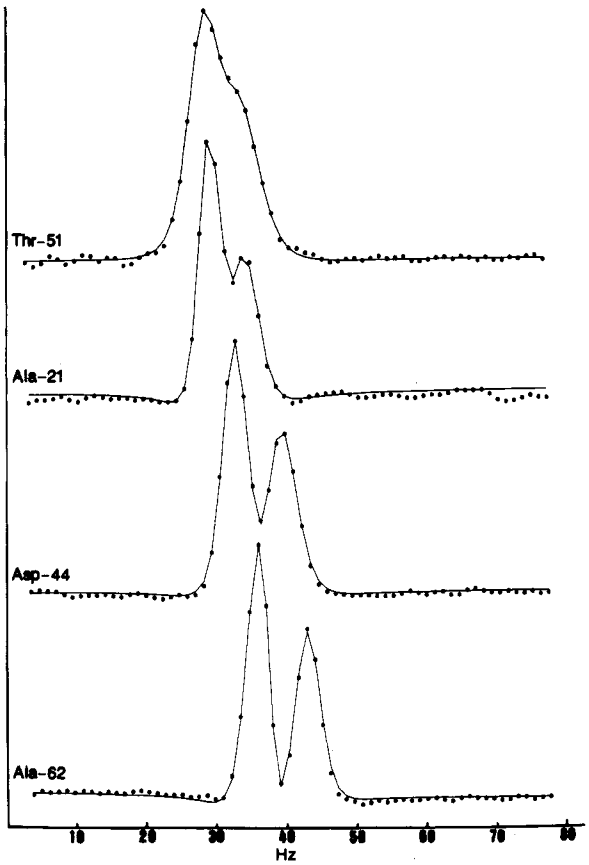

FIGURE 3: Representative data showing calculated fits to F_1 (^{15}N) line shapes of a HMQC-COSY experiment. Solid circles indicate experimental data and the solid line traces the calculted line shape. $^1H-^{15}N$ $^3J_{NH\alpha}$ coupling constants calculated in this way are listed in Table I. The experimental conditions used were the same as in Figure 1.

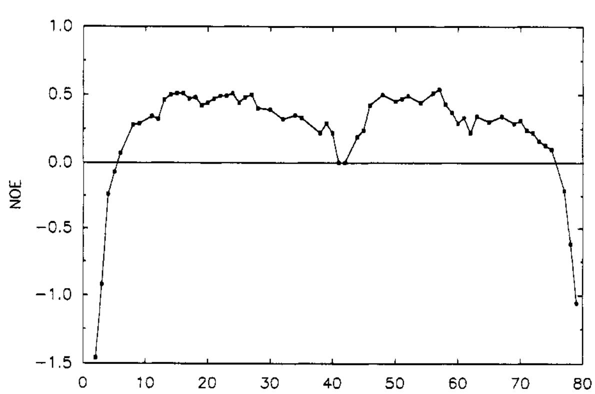

FIGURE 4: Heteronuclear $^1H-^{15}N$ NOEs obtained using the experimental procedures described in Kay et al. (1989b), recorded using the same sample conditions as in Figure 1.

between residues 10 and 40, a turn around residues 40–48, and a less well-defined helical segment between residues 48 and 76, disrupted by Pro-64. NMR evidence for long-range interaction between helices is still inconclusive because of

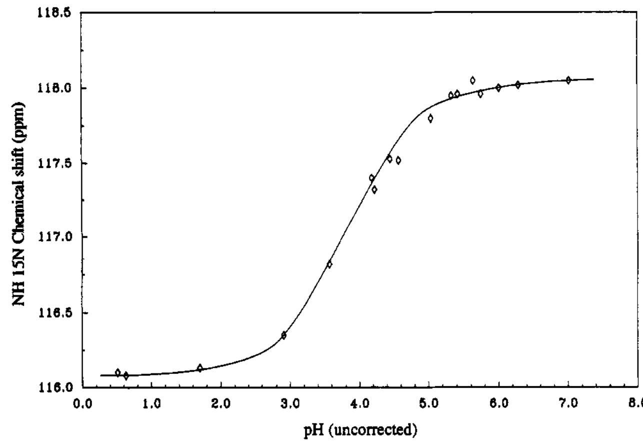

FIGURE 5: Representative titration curve showing the ^{15}N chemical shift of the resonance of Glu-37 of the c subunit in TFE-Cd_2OH. The pK_a values of other titrating residues were estimated from similar data. The pH* values quoted correspond to uncorrected glass electrode readings.

spectral overlap problems, and rigorous structural calculation requires a more comprehensive set of restraints. In these studies, it must be remembered that the TFE solvent used is not ideal for mimicking a membrane bilayer but there is good evidence that this solvent does not induce helical structures indiscriminately [e.g., see Breeze et al. (1990)].

The structural information derived from NOEs and coupling constants is supplemented by independent information on ionization states and on dynamics from $^1H-^{15}N$ NOEs.

The dynamic data suggest a relatively mobile region around residue 42, joining two less mobile, helical regions. There is also increased flexibility near the N- and C-termini. The second helix could be considered as two helical regions disrupted by Pro-64. Relative movement of two such helical regions, facilitated by the missing H-bond, arising from the presence of a proline, has been noted previously in putative membrane-spanning helices and been studied both by NMR and by molecular dynamics (Bazzo et al., 1988; Pastore et al., 1988; Dempsey et al., 1991).

Some of the pK_a values observed in the protein are significantly different from those observed in the free amino acid, although larger deviations can be observed for carboxylic acids in proteins. Asp-44 and Glu-37 have relatively low pK_a's which could be consistent with ionic stabilization of the carboxylate side chain in the proposed central turn region of the protein. Some of the pK_a values of the carboxylic acid side chains are elevated (residues 2, 7, and 61). This could be due to a localized hydrophobic environment, possibly via inter- or intrahelical interactions. Asp-61 has been implicated in proton translocation (Cox et al., 1986). The elevation of its pK_a by 1 unit and the observed flexibility in the region between Asp-61 and the C-terminus could be of relevance to such mechanisms.

ACKNOWLEDGMENTS

We thank Jane Heritage for help with sample preparation and Jonathan Boyd and Nick Soffe for help with the NMR experiments.

Registry No. ATP synthase, 37205-63-3.

REFERENCES

Bazzo, R., Tappin, M. J., Pastore, A., Harvey, T. S., Carver, J. A., & Campbell, I. D. (1988) *Eur. J. Biochem. 173*, 139–146.

Breeze, A. L., Harvey, T. S., Bazzo, R., & Campbell, I. D. (1991) *Biochemistry 30*, 575–582.

Clore, G. M., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 7387–7401.

Cox, G. B., Fimmel, A. L., Gibson, F., & Hatch, L. (1986) *Biochim. Biophys. Acta 849*, 62–69.

Dempsey, C. E., Bazzo, R., Harvey, T. S., Syperek, I., Boheim, G., & Campbell, I.D. (1991) *FEBS Lett. 281*, 240–244.

Fillingame, R. H. (1976) *J. Biol. Chem. 251*, 6630–6637.

Foster, D. L., & Fillingame, R. H. (1982) *J. Biol. Chem. 257*, 2009–2015.

Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B., & Fenney, J. (1990) *J. Magn. Reson. 90*, 420–425.

Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.

Hertzeberg, E. L., & Hinkle, P. C. (1974) *Biochem. Biophys. Res. Commun. 58*, 178–184.

Hoppe, J., & Sebald, W. (1984) *Biochim. Biophys. Acta 768*, 1–27.

Hoppe, J., Schairer, H. U., & Sebald, W. (1980) *Eur. J. Biochem. 112*, 17–24.

Kagawa, Y., & Racker, E. (1966) *J. Biol. Chem. 241*, 2475–2482.

Kay, L. E., Brooks, B., Sparks, S. W., Torchia, D. A., & Bax, A. (1989a) *J. Am. Chem. Soc. 111*, 5488–5490.

Kay, L. E., Torchia, D. A., Bax, A., & Lewis, E. K. (1989b) *Biochemistry 28*, 8972–8979.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mao, D., Wachter, E., & Wallace, B. A. (1982) *Biochemistry 21*, 4960–4968.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 6150–6156.

Moody, M. F., Jones, P. T., Carver, J. A., Boyd, J., & Campbell, I. D. (1987) *J. Mol. Biol. 193*, 759–774.

Norwood, T. J., Boyd, J., & Campbell, I. D. (1989) *FEBS Lett. 255*, 369–371.

Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., & Campbell, I. D. (1990a) *J. Magn. Reson. 87*, 488–501.

Norwood, T. J., Boyd, J., Soffe, N., & Campbell, I. D. (1990b) *J. Am. Chem. Soc. 112*, 9638–9640.

Pastore, A., Harvey, T. S., Dempsey, C. E., & Campbell, I. D. (1989) *Eur. Biophys. J. 16*, 363–367.

Schneider, E., & Altendorf, K. (1985) *EMBO J. 5*, 515–518.

Sebald, W., & Hoppe, J. (1981) *Curr. Top. Bioenerg. 12*, 1–64.

Senior, A. E. (1983) *Biochem. Biophys. Acta 726*, 81–95.

Senior, A. E. (19885) *Physiol. Rev. 68*, 177–231.

Sigrist, H., Sigrist-Nelson, K., & Gitler, C. (1977) *Biochem. Biophys. Res. Commun. 74*, 178–184.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–298.

Von Meyenberg, K., Jorgenson, B. B., & van Deurs, B. (1984) *EMBO J. 3*, 1791–1797.

Wüthrich, K. (1986) in *NMR of Proteins and Nucleic acids*, Wiley-Interscience, New York.

Yoshida, M., Okamoto, H., Sone, N., Hirata, H., & Kagawa, Y. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 936–940.

Use of a Potential of Mean Force To Analyze Free Energy Contributions in Protein Folding[†]

F. Avbelj[‡]

Institute of Chemistry, Hajdrihova 19, 61001 Ljubljana, Slovenia

Received October 21, 1991; Revised Manuscript Received March 19, 1992

ABSTRACT: A method for calculation of the free energy of residues as a function of residue burial is proposed. The method is based on the potential of mean force, with a reaction coordinate expressed by residue burial. Residue burials are calculated from high-resolution protein structures. The largest individual contributions to the free energy of a residue are found to be due to the hydrophobic interactions of the nonpolar atoms, interactions of the main chain polar atoms, and interactions of the charged groups of residues Arg and Lys. The contribution to the free energy of folding due to the uncharged side chain polar atoms is small. The contribution to the free energy of folding due to the main chain polar atoms is favorable for partially buried residues and less favorable or unfavorable for fully buried residues. Comparison of the accessible surface areas of proteins and model spheres shows that proteins deviate considerably from a spherical shape and that the deviations increase with the size of a protein. The implications of these results for protein folding are also discussed.

The evaluation of the free energies of a protein in the denatured and folded conformations is crucial for the understanding the protein folding process. The free energy of a protein can be approximated by the contributions arising from hydrophobic interactions, configurational entropy, hydrogen bonding, electrostatic interactions of charged residues, and van der Waals' interactions. Unfortunately, the quantitative contributions of the configurational entropy and of the individual interactions to the free energy of a protein are not well understood.

[†] Supported by Science Foundation of Slovenia.

[‡] Present address: Center for Advanced Research in Biotechnology, 9600 Gudelsky Drive, Rockville, MD 20850.

0006-2960/92/0431-6290$03.00/0