Journal of Biomolecular NMR (2006) 36:279–293
DOI 10.1007/s10858-006-9094-x


*Article*

# Signal assignment and secondary structure analysis of a uniformly [$^{13}$C, $^{15}$N]-labeled membrane protein, $H^+$-ATP synthase subunit *c*, by magic-angle spinning solid-state NMR

Masatoshi Kobayashi[a,b], Yoh Matsuki[a,c], Ikuko Yumen[a,d], Toshimichi Fujiwara[a] & Hideo Akutsu[a,b,*]

[a]*Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita 565-0871, Japan;* [b]*CREST, Japan Science and Technology Agency, Kawaguchi-shi, Saitama, Japan;* [c]*BIRD, Japan Science and Technology Agency, Kawaguchi-shi, Saitama, Japan;* [d]*Japan Biological Information Consortium, Tokyo, Japan*

Received 25 May 2006; Accepted 19 September 2006

*Key words:* transmembrane protein, $F_O$ subunit *c*, MAS solid-state NMR, non-crystal solid, sequential assignment, spectral simulation

**Abstract**

Signal assignment and secondary structural analysis of uniformly [$^{13}$C, $^{15}$N] labeled $H^+$-ATP synthase subunit *c* from *E. coli* (79 residues) in the solid state were carried out by two- and three-dimensional solid-state NMR under magic-angle spinning. The protein took on a unique structure even in the solid state from the $^{13}$C linewidths of about 1.7 ppm. On the basis of several inter- and intra-residue $^{13}$C–$^{13}$C and $^{13}$C–$^{15}$N chemical shift correlations, 78% of $C^{\alpha}$, 72% of $C^{\beta}$, 62% of C′ and 61% of $N^{H}$ signals were assigned, which provided the secondary structure information for 84% of the 79 residues. Here, inter-residue correlations involving Gly, Ala, Pro and side-chains and a higher resolution in the 3D spectrum were significantly useful for the sequence specific assignment. On top of this, the $^{13}$C–$^{13}$C correlation spectra of subunit *c* was analyzed by reproducing experimental cross peaks quantitatively with chemical shift prediction and signal-intensity calculation based on the structure. It revealed that the subunit *c* in the solid state could be specified by α-helices with a loop structure in the middle (at sequence 41–45) as in the case of the solution structure in spite of additional extended conformations at 76–79 at the C-terminus.

## Introduction

$H^+$-ATP synthase is an important enzyme in energy transduction systems in organisms. It comprises membrane-embedded ($F_O$) and peripheral ($F_1$) parts. $F_OF_1$-ATP synthase uses a transmembrane $H^+$ gradient to synthesize ATP from ADP and phosphate. Subunit *c* of $F_O$ from *E. coli* ($EF_Oc$) is a hydrophobic protein composed of 79 amino acid residues. Subunit *c* forms a ring structure in $F_O$ and plays a key role by transporting $H^+$ across the membrane through $F_O$ (Seelert et al., 2000; Jiang et al., 2001; Stahlberg et al., 2001; Mitome et al., 2004). The structures of $EF_Oc$ and $F_Oc$ from thermophilic *Bacillus* PS3 in organic solvents were determined by solution NMR to be hairpin folds with two helical segments connected by a short loop (Girvin et al., 1998; Nakano et al., 2006). Recently, the crystal structure of the F-type $Na^+$-ATP synthase subunit *c* ring was reported (Meier et al., 2005), revealing a cylindrical ring-structure composed of eleven

*To whom correspondence should be addressed.
E-mail: akutsu@protein.osaka-u.ac.jp

subunit *c* in a hairpin fold with two membrane-spanning α-helices. To elucidate the mechanisms of $H^+$-translocation and energy conversion, solid-state NMR (SSNMR) is one of the most powerful methods because structural analysis can be performed in a membrane environment.

Solid-state NMR for structural determination of biomolecules has been rapidly developed in the last decade. One of the recent strategies for structural determination of peptides and proteins by SSNMR is the analysis of uniformly [$^{13}C$, $^{15}N$] labeled samples under magic-angle spinning (MAS). A number of pulse sequences have been developed for signal assignment and structural analysis of uniformly labeled molecules (Straus et al., 1998; Hong, 1999a; Castellani et al., 2002; Fujiwara et al., 2004). Proteins in microcrystalline states provided high resolution spectra, leading to successful signal assignment for small proteins (Pauli et al., 2001; Böckmann, et al., 2003; Igumenova et al., 2004). The crystallization of membrane proteins is usually performed in the presence of a surfactant. However, the structure and function of a membrane protein in detergent micelles may be different from those in membranes. It is important to apply the SSNMR method to a membrane protein in non-crystalline membranes.

In SSNMR spectral analysis of transmembrane proteins, small dispersion of the NMR signals could be a serious problem. Generally, the amino acid composition of an integral membrane protein includes more hydrophobic residues with low variety. For example, the primary sequence of $EF_Oc$, which is the subject of this work, is

M E N L N M D L L Y M A A A V M M G L A AIGAAIGIGI LGGKFLEGAA

RQPDLIPLLR TQFFIVMGLV DAIPMIAVGL GLYVMFAVA.

This contains 13 Ala, 12 Leu, 10 Gly, 8 Met and 8 Ile, accounting for 65% of the 79 residues. Furthermore, small transmembrane proteins prefer α-helical conformations, which leads to serious signal overlapping because of low chemical shift dispersion in contrast to in the case of water soluble proteins. To overcome this problem, a series of amino acid- and site-specific labelings has been utilized (Hong, 1999b; Castellani et al., 2002; van Gammeren et al., 2005). Although they are effective, a certain number of samples with different isotope-labeling are required for complete signal assignment. In the case of a uniformly labeled sample, we can obtain an all inter-residual aliphatic carbon correlation, which is useful for specific signal assignment. Furthermore, spectral simulation based on the prediction of structure-based chemical shifts and the calculation of cross peak intensities can be applied to assign the signals and to determine the secondary structure of a uniformly labeled sample.

Our final objective is the signal assignments and structural analysis of $EF_Oc$ in membranes by SSNMR. As a first step towards working on membrane samples, we have investigated the pure subunit *c* in the solid state in this work. Since this sample gave intense signals, it was suitable for magnetization transfer experiments with low sensitivity such as the inter-residue $(C^{\alpha}C^{\beta})_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ correlation ones. We have assigned the SSNMR signals of subunit *c* by means of several kinds of chemical shift correlation experiments. Quantitative spectrum simulation was performed to confirm the signal assignments and to estimate the secondary structures in subunit *c*.

## Materials and methods

### *Preparation of uniformly $^{13}C$, $^{15}N$-labeled subunit c*

*E. coli* (MEG119 strain) cells transformed by plasmid pCP35 harboring the $EF_O$ subunit *c* gene were cultured in a $^{13}C$ and $^{15}N$ labeled CHL medium (CHLORELLA Industry Co., Japan) with a [$^{13}C$, $^{15}N$] amino acid mixture, 0.5 g/l, and [$^{13}C$] glucose, 4.0 g/l, for 24–26 h. $EF_Oc$ was purified according to the reported method (Girvin and Fillingame, 1993). Subunit *c* was extracted from the homogenized cells (wet 27.2 g) with 12 wet cell volumes of a chloroform/methanol (1:1) mixture for 2 h at 4 °C. The supernatant was collected after centrifugation. Chloroform and water were added to the supernatant. The resultant chloroform/methanol/water (8:4:3) mixture was left standing still overnight at 4 °C. The chloroform fraction was concentrated with a rotary evaporator and then precipitated with five volumes of diethyl ether at −30 °C for 2 days. The crude subunit *c* was applied to a carboxymethyl cellulose

column and purified subunit *c* was eluted with a chloroform/methanol/water (5:5:1) solution. The yield was about 10 mg/4 l culture. The purity of the subunit *c* was confirmed by Tricine-SDS-PAGE and MALDI-TOF-MASS. For SSNMR experiments, purified subunit *c* was precipitated gradually from a chloroform/methanol (2:1) solution with five volumes of diethyl ether. The sample used in this study amounted to 9 mg (1.0 μmol).

*NMR experiments*

NMR measurements were performed with Varian Infinity-plus 500, 600, and 700 spectrometers operating at 11.74, 14.09 and 16.44 T static magnetic fields, respectively. Broadband double and triple resonance MAS probes for a 3.2 mm rotor were used. The MAS frequency was 12.5–15.0 kHz. The sample temperature was set at −40 to −20 °C. The $^{1}$H RF amplitude was 70–75 kHz for TPPM decoupling (Bennett et al., 1995) during the evolution, and data acquisition periods and 90–100 kHz for CW decoupling during the dipolar mixing period. The repetition time was 3.5 s.

We performed two-dimensional $^{13}$C homonuclear correlation experiments with RFDR (Bennett et al., 1998) and SPC$^{Z}$5 (Hohwy et al., 1999; Matsuki et al., 2003), and a heteronuclear correlation experiment, an N–C$^{\alpha}$C$^{\beta}$ one, to obtain the intra-residue correlation. In order to obtain the inter-residue correlation, $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ and $(C^{\alpha}C^{\beta})_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ experiments were performed (Figure 1). Most spectra were recorded at 11.74 T. RFDR and three-dimensional $N_{i+1}$–$C'_i$–$C^{\alpha}_i$ correlation spectra were recorded at 14.09 and 16.44 T, respectively. Dipolar interactions for $^{13}$C and $^{15}$N transferred the magnetization during the mixing period. The initial magnetization was prepared by cross-polarization from the proton magnetization with a contact time of 1.0–2.0 ms and the RF field amplitude ramped near the first sideband for the Hartmann–Hahn condition, $\gamma B_1^{X} = \gamma B_1^{H} - \omega_R$. Here $\gamma B_1^{X}$ and $\gamma B_1^{H}$ are the RF field amplitudes of $^{13}$C/$^{15}$N and $^{1}$H, respectively, and $\omega_R$ is the sample spinning frequency.

Broadband $^{13}$C–$^{13}$C correlation spectra were obtained with zero-quantum $^{13}$C dipolar mixing under RFDR applied at offset frequency 105 ppm with an RF amplitude of 50 kHz. A band-selective correlation experiment for the aliphatic carbons was performed by the SPC$^{z}$5 sequence with a double-quantum $^{13}$C–$^{13}$C dipolar mixing. We adopted $a = 240°$ in the C$^{z}$-element pulses $(a/2 - \pi/2)_{\phi}(a)_{\phi+\pi}(a/2 + \pi/2)_{\phi}$. The $B_1$ field amplitude for SPC$^{z}$5 mixing was 41.7 kHz at $\omega_R = 12.5$ kHz with the frequency at the center of the aliphatic carbon region, 40 ppm. The mixing time was 0.6–2 ms for RFDR and 1–2 ms for SPC$^{z}$5. The data matrices, $600(t_1) \times 600(t_2)$ for RFDR and $430(t_1) \times 430(t_2)$ for SPC$^{z}$5 experiments were zero-filled to $1024 \times 1024$. The number of acquisitions was 16 for each FID. The states method was used for quadrature detection in the indirect dimension.

Frequency selective cross-polarization (CP) from $^{15}$N to $^{13}$C was used in $^{13}$C–$^{15}$N correlation experiments (Baldus et al., 1998; Petkova et al., 2003). The magnetization was transferred under the condition of $\gamma B_1^{13C} \pm \gamma B_1^{15N} = \omega_R$, where $\gamma B_1/\omega_R$ was about half-integer. The $^{13}$C frequency was set in the aliphatic carbon region

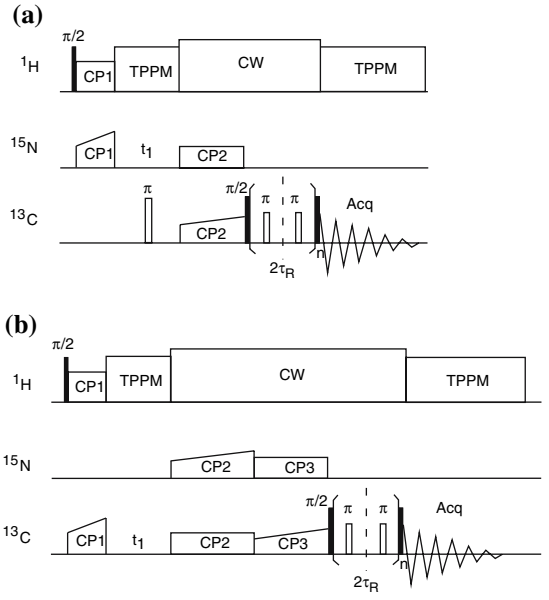


*Figure 1.* Pulse sequence for the inter-residue heteronuclear $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ experiment (a), and that for the inter-residue $(C^{\alpha}C^{\beta})_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ experiment (b) under MAS. Subscripts *i* and *i* + 1 stand for the residue numbers. CP, CW, TPPM and $\tau_R$ indicate cross polarization, continuous wave decoupling, two pulse phase modulation decoupling and the rotor spinning period, respectively. The sequences consist of the evolution period for $^{13}$C or $^{15}$N isotropic chemical shifts and the mixing period under $^{13}$C–$^{13}$C or $^{13}$C–$^{15}$N dipolar coupling.

for the N–$C^{\alpha}C^{\beta}$ correlation and in the carbonyl carbon region for the $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ correlation. $^{15}N$ RF was set at 120 ppm. In the N–$C^{\alpha}C^{\beta}$ experiments, the contact time was 6 ms with $(\gamma B_1^{13C}/2\pi, \gamma B_1^{15N}/2\pi)$ = (30.5 kHz, 18 kHz). The data matrix was $320(t_1) \times 512(t_2)$ with spectral widths of 50 kHz for both the $^{13}C$ and $^{15}N$ axes. In the $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ pulse sequence (Figure 1A), the contact time was 4 ms with $(\gamma B_1^{13C}/2\pi, \gamma B_1^{15N}/2\pi)$ = (19 kHz, 31.5 kHz). The RFDR mixing time from C′ to $C^{\alpha}C^{\beta}$ was 2.24 ms with an amplitude of 50 kHz. The data matrix was 240 $(t_1) \times 512(t_2)$, and the spectral widths were 50 and 40 kHz for the $^{13}C$ and $^{15}N$ axes, respectively. In both N–$C^{\alpha}C^{\beta}$ and $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ spectra, the data matrices were zero-filled to $1024 \times 1024$. The numbers of acquisitions for N–$C^{\alpha}C^{\beta}$ $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ experiments were 88 and 128 for each FID, respectively.

The 3D $N_{i+1}$–$C'_i$–$(C^{\alpha}_i C^{\beta}_i)_i$ experiment was performed with the pulse sequence which were made by inserting the second evolution time into the 2D $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ pulse sequence. The $^{13}C$–$^{15}N$ mixing time was 2 ms with $(\gamma B_1^{13C}/2\pi, \gamma B_1^{15N}/2\pi)$ = (8 kHz, 23 kHz). RFDR mixing time from C′ to $C^{\alpha}C^{\beta}$ was 2.4 ms with an amplitude of 50 kHz. The spectral widths were 10, 16, and 70 kHz for the $^{15}N$, $^{13}C'$, and $^{13}C^{\alpha}C^{\beta}$ axes, respectively. The experimental data matrix, $17(t_1) \times 20(t_2) \times 512(t_3)$, was zero-filled to $64 \times 64 \times 1024$. The number of acquisitions was 64 for each FID.

The pulse sequence for 2D $(C^{\alpha}C^{\beta})_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ correlation is given in Figure 1B (Fujiwara et al., 2004). $^{13}C$ RF was set at 40, 175 and 105 ppm for the two CP and the RFDR mixing, respectively. $^{15}N$ RF was set at 120 ppm. The RF amplitudes were $(\gamma B_1^{13C}/2\pi, B_1^{15N}/2\pi)$ = (20.5 kHz, 8 kHz). The contact times from $C^{\alpha}_{i+1}$ to $N_{i+1}$ and $N_{i+1}$ to $C'_i$ were 5 and 4 ms, respectively. The RFDR mixing time was 1.6 ms. The number of acquisitions was 128 for each FID. The data matrix, $216(t_1) \times 512$ $(t_2)$ was zero-filled to $1024 \times 1024$.

The obtained FID signals were processed with Felix2000 (Accelrys Inc., USA). Fourier-transformations were performed with cosine-bell broadening functions. The $^{13}C$ chemical shift was referenced to DSS by using the methine carbon signal of adamantane under MAS at 31.5 ppm relative to DSS in $D_2O$ (Morcombe and Zilm, 2003). The $^{15}N$ chemical shift was referenced to that of liquid $NH_3$ deduced from the $^{13}C$ chemical shift using the $\gamma^{15N}/\gamma^{13C}$ ratio following the IUPAC recommendation (Markley et al., 1998).

### *NMR spectral analysis and structure prediction based on chemical shifts*

We utilized structure-based chemical shift prediction programs, SHIFTS (Xu and Case, 2001), SHIFTX (Neal et al., 2003), and PROSHIFT (Meiler, 2003) to analyze structure-dependent shifts. In order to predict backbone dihedral angles $\phi$ and $\Psi$ from the chemical shifts of $^{13}C^{\alpha}$, $^{13}C'$, $^{15}N^{H}$ and $^{13}C^{\beta}$, we utilized the TALOS program (Cornilescu et al., 1999).

Two-dimensional RFDR spectra for $C'-C^{\alpha}$, $C'-C^{\alpha}C^{\beta}$ and $C^{\alpha}C^{\beta}$ regions were simulated as a function of atomic coordinates of the protein. To obtain the chemical shift of each residue for $\alpha$-helix and extended conformations, respectively, spectral simulations were carried out for artificial all-helix and all-extended structures of $EF_oc$. The signal intensities were determined from the magnetization transfer calculated by a time-dependent Hamiltonian including terms for dipolar interaction, $^{13}C$ chemical shifts and the effect of RF pulses (Fujiwara et al., 1998). In the calculation, 20 amino acid residues were categorized into 12 groups, G/A/C/S/T/QEM/PLKR/DN/FY/WH/V/I (single letter abbreviations for amino acids) in terms of spin systems and chemical shift patterns. The chemical shifts for $C^{\alpha}$, $C^{\beta}$ and carbonyl C′ were predicted with SHIFTX and those for side chain carbons were predicted with PROSHIFT on the basis of the assumed tertiary structure. Signals were convoluted by means of Gaussian or Lorentzian function with a linewidth of about 1.7 ppm, which was estimated from experimental spectra. The $\alpha$-helix $(\phi, \Psi)$ = (−60°, −40°) and extended $(\phi, \psi)$ = (−120°, 140°) conformations were assumed for the secondary structure. The solution structure (PDB 1a91, Girvin et al., 1998) was used for side chain conformations. RMSD between experimental and simulated spectra was defined as

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i,j}(kS_{i,j}^{\mathrm{exp}} - S_{i,j}^{\mathrm{sim}})^2}$$

where $S_{i,j}^{exp}$ and $S_{i,j}^{sim}$ are the signal intensity at point ($i$, $j$) in experimental and simulated 2D spectra, respectively, and $N$ is the number of data matrix points for the spectrum. Intensity factor, $k$, was calculated so as to minimize RMSD.

## Results

### *Intra-residue spin connectivities for the signal assignments to amino acid types*

Figure 2a and b presents 2D $^{13}C$–$^{13}C$ RFDR spectra of subunit *c* with a mixing time of 1.85 ms. RFDR dipolar mixing gave the broadband cross peaks between carbonyl and aliphatic carbons, and among aliphatic carbons (Figure 2a), and among aromatic carbons, and between aromatic and aliphatic carbons (Figure 2b). The latter led to the direct assignment of aromatic and $C^{\beta}$ carbons of Phe and Tyr. Otherwise, we started with the BMRB database that gives the typical chemical shifts of amino acid residues. Then, intra-residue spin connectivities of specific amino acids were traced on a SPC$^{z}$5 spectrum, as shown in Figure 2c, where the aliphatic carbon spin systems of Leu and Ile are connected as examples. In the DQ dipolar mixing spectrum, one-step magnetization transfer gave rise to negative intensity, colored red, and two-step transfer to positive intensity, colored black. Other spin connectivities were also obtained considering the intensity of cross peaks and the composition of amino acid residues in subunit *c*, as shown in Figure 2c. Although the resolution of cross peaks in the carbonyl region (Figure 2a) was insufficient due to small chemical shift dispersion, the isolated cross peaks, such as Gly $C'$–$C^{\alpha}$ and Ala $C'$–$C^{\beta}$, were assigned. The assignments are given in Figure 2. The short RFDR mixing times (0.64 and 0.96 ms) gave one bond correlations, allowing us to distinguish $C'-C^{\alpha}$ from $C'-C^{\beta}$ (data not shown).

Figure 3a presents a 2D $^{13}C$–$^{15}N$ spectrum for the intra-residue $N-C^{\alpha}C^{\beta}$ correlation. The $^{15}N$ chemical shifts can be specified from the $^{13}C$ chemical shifts of isolated cross peaks. The cross peak at ($^{13}C$, $^{15}N$) = (66, 117) ppm was assigned to Val and Ile residues because these $C^{\alpha}$ spins resonated at about 65 ppm, as shown in Figure 2a and c. The numbers of Val and Ile residues were six and eight, respectively. The linewidth in the $^{15}N$ dimension was 6.5 ppm, covering 14 residues. Because of the characteristic $^{15}N$ chemical shift, the cross peak at (65, 133) ppm was assigned to Pro residues taking on α-helix. The linewidth for $^{15}N$ was 4.0 ppm, representing two residues. The broad shoulder peak at around (63, 137) ppm was assigned to Pro taking on an extended conformation. The most intense cross peak at (55, 120) ppm was mainly due to $C^{\alpha}-N$ of 13 Ala because most $C^{\alpha}$ of Ala resonated at about 55 ppm (Figure 2a). This assignment could be extended to Ala $C^{\beta}-N$ at (18, 120) ppm. The linewidths for $^{15}N$ of Ala were 5.2 and 4.9 ppm, respectively. The cross peak at (58, 119) ppm could be assigned to Leu and Met residues because their $C^{\alpha}$ resonated at about 58 ppm. The linewidth for $^{15}N$ was 6.4 ppm and 12–20 residues are included in this region. The intense cross peak at (47, 107) ppm was assigned to Gly residues because of characteristic $^{15}N$ chemical shifts. The linewidth for $^{15}N$ was 5.4 ppm. The side chain signals of, Lys $C^{\varepsilon}-N^{\zeta}$, Arg $C^{\delta}-N^{\varepsilon}$ and Arg $C^{\gamma}-N^{\varepsilon}$ (not shown) were observed in isolation. The $^{15}N$ chemical shift of Lys $N^{\zeta}$ at 32.8 ppm indicated that the side chain of Lys residue took on the $-NH_3^+$ form (Levy and Lichter, 1979).

### *Inter-residue spin connectivities for sequence specific assignments*

Although the resolutions for RFDR and SPC$^{Z}$5 spectra were relatively poor, it can be improved using inter-residual correlations in two and three dimensional spectra. Especially, correlations with specific amino acid residues such as Gly and Ala, and those with side chain signals will contribute to the improvement. Figure 3b shows a 2D $^{13}C$–$^{15}N$ spectrum for the inter-residue $N_{i+1}-(C'C^{\alpha}C^{\beta})_i$ correlation. The $^{15}N$ chemical shifts of the $(i+1)$th residue could be specified from the $^{13}C$ chemical shifts of the *i*th residue. Characteristic signals such as $X_iG_{i+1}$ could be easily identified. There are ten $X_iG_{i+1}$ pairs, namely, three IG, two LG, two MG, one EG, one VG and one GG. Using the $^{15}N$ chemical shift of Gly (107 ppm) determined in Figure 3a, the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts of the residues linked to Gly could be determined. The isolated cross peak of G32G33 was easily identified. The cross peak at ($^{13}C$, $^{15}N$) = (66, 107) ppm could be ascribed to $X_i^{\alpha}G_{i+1}^{N}$, where X was I or V, with $i$ = 22, 26, 28 and 68. The cross peak at (38, 107) ppm was assigned to $I_i^{\beta}G_{i+1}^{N}$ with $i$ = 22, 26, 28.

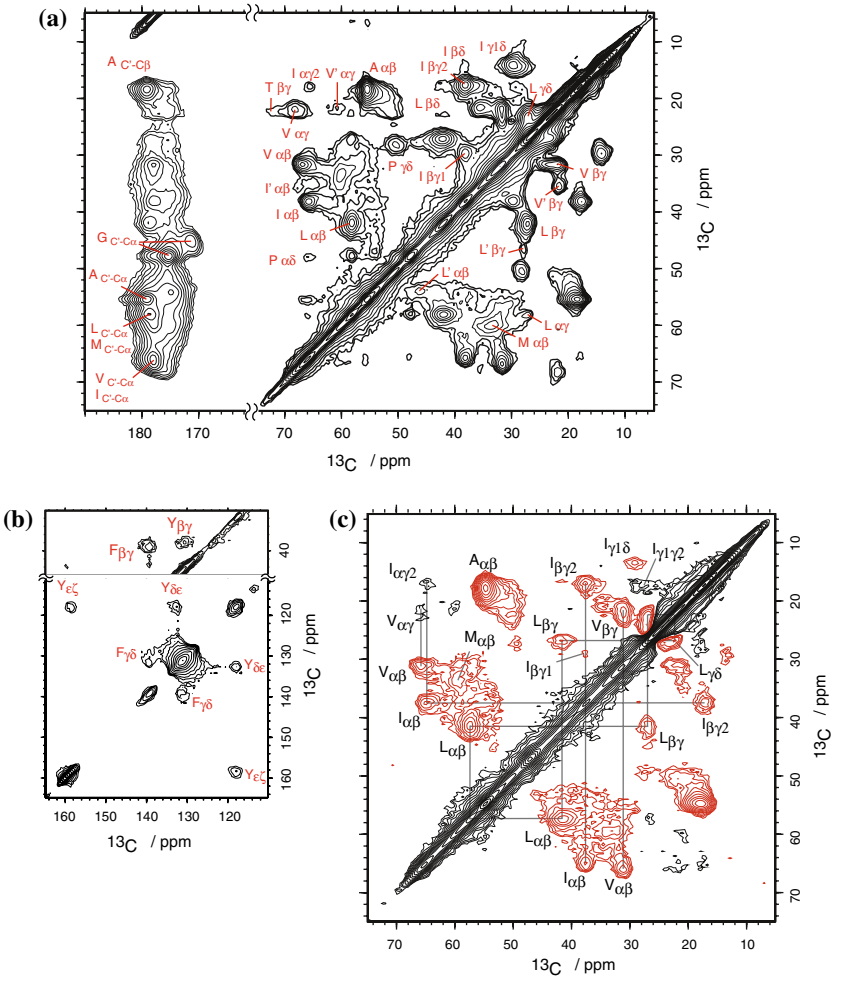


*Figure 2.* Intra-residue 2D $^{13}C$–$^{13}C$ correlation spectra of [U-$^{13}C$, $^{15}N$] subunit*c* by RFDR and SPC$^{z}$5 dipolar mixing. (a) Correlations between carbonyl and aliphatic carbons, and among aliphatic carbons, and (b) those among aromatic carbons, and between aromatic and aliphatic carbons obtained by RFDR with a mixing time of 1.85 ms. (c) Correlations among aliphatic carbons obtained by SPC$^{z}$5 with a mixing time of 1.92 ms. Black and red lines indicate the positive and negative contour levels, respectively. The spin connectivity was presented for Leu and Ile.

The cross peaks of $G_i^{\alpha}X_{i+1}^{N}$ and $A_i^{\beta}X_{i+1}^{N}$ were observed at about (47, 121) and (18, 119) ppm, respectively. The $^{15}N$ linewidth of $G^{\alpha}X^{N}$ was 5.0 ppm and the $^{15}N$ signals of nine residues, i.e., four Leu, two Ala, two Ile, and one Lys, overlapped in this region. The $^{15}N$ linewidth of $A^{\beta}X^{N}$ was 5.0 ppm and the $^{15}N$ signals of 12 residues, i.e., five Ala, three Ile, three Val, and one Arg, overlapped in this region. The cross peaks of $I_i^{\beta}X_{i+1}^{N}$ were observed at about (38, 118 ppm) for $X_{i+1}$ = L31, V56, A67. Similarly, the $^{15}N$ shifts of $X_{i+1}$ for $L_i^{\beta}X_{i+1}^{N}$ and $V_i^{\beta}X_{i+1}^{N}$ were 118 and 119 ppm, respectively. The side chain signal of Arg $C^{\zeta}-N^{\eta}$ was clearly observed at (160, 71) ppm. In spite of the reduction of cross peak intensities due to the two magnetization transfer process, the signal of single $G^{\alpha}32G^{N}33$ and the side chain signal of two Arg $C^{\zeta}-N^{\eta}$ were observed. The 3D $N_{i+1}$–$C'_i$–$C_i^{\alpha}C_i^{\beta}$ experiment gave a better resolution as shown in Figure 3c and d. The correlation information supported the assignment obtained in the 2D $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ and RFDR experiments. Furthermore, inter-residue correlations of $Q^{\alpha}42P^{N}43$ and $M_i^{\beta}G_{i+1}^{N}$ were observed, as indicated in Figure 3c. $G_i^{\alpha}X_{i+1}^{N}$ at ($F_3$, $F_2$, $F_1$) = (45, 171, 123) ppm corresponding to the weak $G_{C'-C\alpha}$ cross peak in Figure 2a was also observed (Figure 3d).

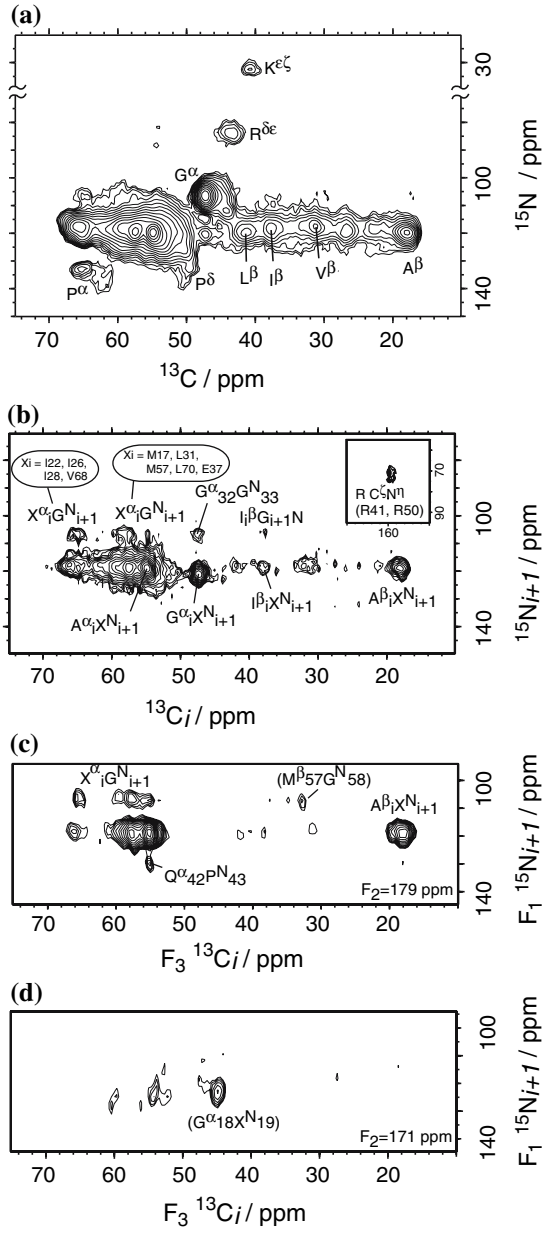


*Figure 3*. Intra-residue 2D N–C$^{\alpha}$C$^{\beta}$ (a), inter-residue 2D $N_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ (b), and inter-residue 3D $N_{i+1}$–$C'_i$–$C^{\alpha}_i C^{\beta}_i$ (c and d) correlation spectra of [U-$^{13}$C, $^{15}$N] subunit *c*. (a) The mixing time was 6 ms. (b) The mixing times were 4 and 2.24 ms for the first (CP2) and second ($2n\tau_R$) mixing periods. The spectral region for the cross peaks of the Arg side chain is shown in the inset. (c) and (d) 2D cross sections sliced at $F_2$ = 179 and 171 ppm, respectively. The mixing times were 2 and 2.4 ms for the first and second mixing periods. Parentheses show eventual assignments.

Figure 4a presents an inter-residue $(C^{\alpha}C^{\beta})_{i+1}$–$(C'C^{\alpha}C^{\beta})_i$ correlation spectrum of subunit *c*. C$^{\alpha}$ signals appearing from 52 to 63 ppm in both the $F_1$ and $F_2$ dimensions are due to A, L, M, F, Y, D, E, N and Q. About 33 cross peaks were located in this region. The C$^{\alpha}$ signals of Gly were well isolated. Ten Gly residues were resolved into four peaks as $X_iG_{i+1}$ in the $F_2$ axis and into five peaks as $G_iX_{i+1}$ in the $F_1$ axis. The typical chemical shift of Gly C$^{\alpha}$ was 47.7 ppm. A small peak at 44–45 ppm was assigned to Gly C$^{\alpha}$ as well. We could determine the C$^{\alpha}$ chemical shifts of Gly in $X_{i-1}G_iZ_{i+1}$, referring to the amino acid sequence. The two cross peaks *c* and *e* are connected by sequential assignments, as indicated by green lines in Figure 4a, because the intensities of both correspond to one pair of residues. Therefore, peaks *c* and *e* originate from a single $X_{i-1}G_iZ_{i+1}$ sequence. Candidate sequences were M17G18L19, E37G38A39, M57G58L59, and L70G71L72.

As the C$^{\beta}$ signals of Ala appeared clearly, the signals from $X^{\alpha}_i A^{\beta}_{i+1}$ or $A^{\beta}_i X^{\alpha}_{i+1}$ could be assigned. The chemical shifts of AA were determined from the $A^{\alpha}_i A^{\beta}_{i+1}$ or $A^{\beta}_i A^{\alpha}_{i+1}$ cross peaks. Two $G^{\alpha}_i A^{\beta}_{i+1}$ with $i$ = 23 and 38, resonated at (47, 18) ppm. Since the chemical shifts of the observed $G^{\alpha}_i A^{\beta}_{i+1}$ and $A^{\alpha}_i A^{\beta}_{i+1}$ cross peaks do not fit to those of peaks *c* and *e*, E37G38A39(A40) should be eliminated from the candidates for those peaks. Weak cross peaks at (55, 31) and (38, 47) ppm were tentatively assigned to $A^{\alpha}_i V^{\beta}_{i+1}$ with $i$ = 14 and 67, and $I^{\beta}_i G^{\alpha}_{i+1}$ with $i$ = 22, 26 and 28. $A^{\beta}_i I^{\alpha}_{i+1}$ and $A^{\beta}_i V^{\alpha}_{i+1}$ resonated at (19, 66) ppm. A77 was not included in $A^{\alpha}_i V^{\beta}_{i+1}$ at (55, 31) because two isolated cross peaks at (59.7, 50.8) and (50.8, 62.0) ppm could be assigned to F$^{\alpha}$76A$^{\alpha}$77 and A$^{\alpha}$77V$^{\alpha}$78, respectively.

In EF$_O$*c*, there are successions of the same residues, such as five AA, two LL, one FF, one GG, and one MM, giving rise to diagonal peaks. Major amino acid residues such as A, V, I and M provided cross peaks with strong intensity. An intense cross peak at ($F_2$, $F_1$) = (55, 66) ppm was assigned to three pairs of $A^{\alpha}_i I^{\alpha}_{i+1}$ and $A^{\alpha}_i V^{\alpha}_{i+1}$. A broad peak at (67, 60) ppm and a weak one at (31, 60) ppm were assigned to three pairs of $V^{\alpha}_i M^{\alpha}_{i+1}$ and $V^{\beta}_i M^{\alpha}_{i+1}$, respectively. Although other correlations such as $P^{\alpha}_i L^{\alpha}_{i+1}$, $I^{\alpha}_i L^{\alpha}_{i+1}$, $F^{\alpha}_i L^{\alpha}_{i+1}$, and $P^{\alpha}_i M^{\alpha}_{i+1}$ would be overlapping at around (67, 60) ppm, the assignment of $M^{\alpha}_{i+1}$ was supported by the cross peak of $V^{\beta}_i M^{\alpha}_{i+1}$. The chemical shifts of M$^{\alpha}$65, I$^{\alpha}$66, I$^{\beta}$66 and A$^{\beta}$67 could be uniquely determined from weak peaks at (58.5, 38.3) and (64.7, 18.1) ppm assigned to M$^{\alpha}$65I$^{\beta}$66 and I$^{\alpha}$66A$^{\beta}$67, respectively (Figure 4a).

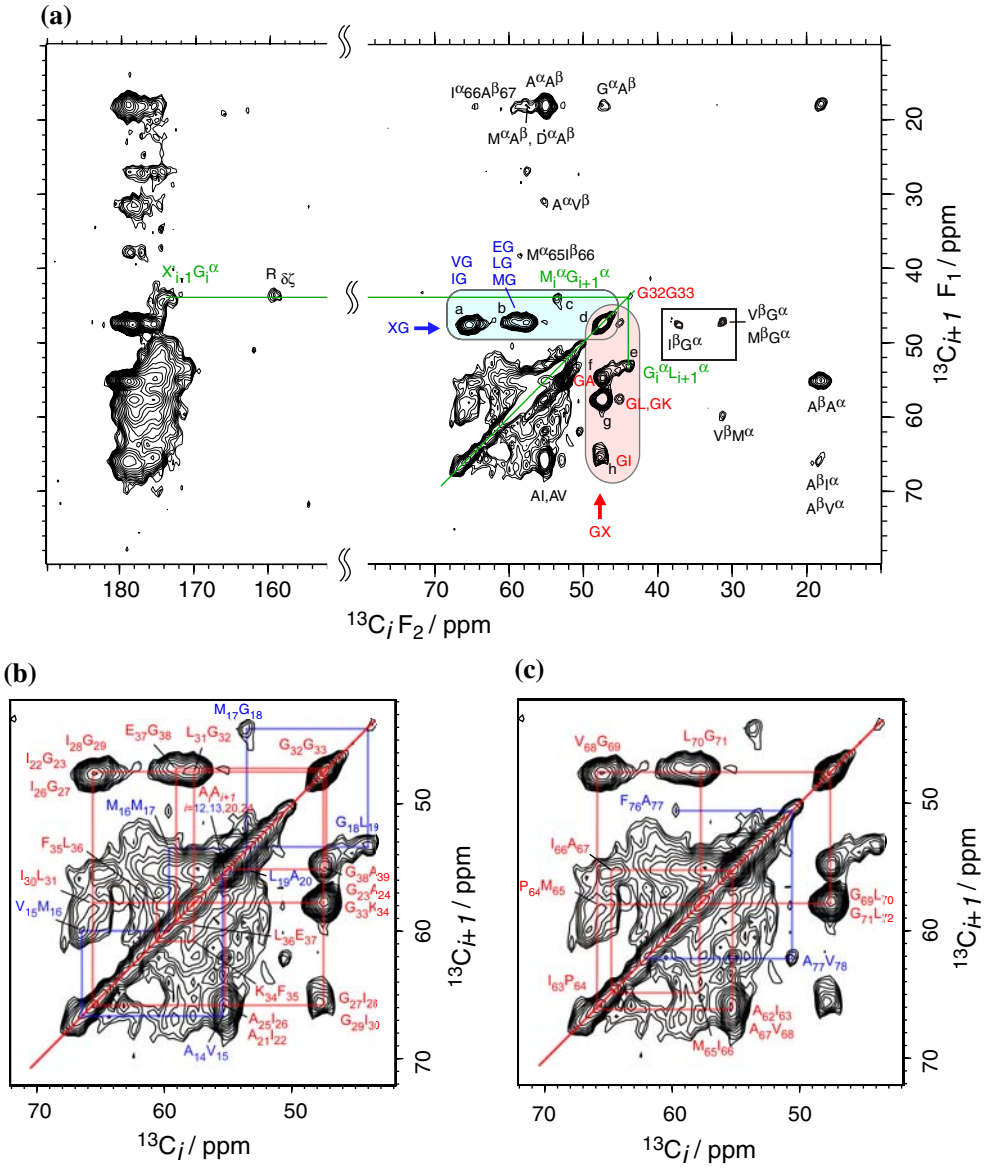


*Figure 4*. (a) 2D $(C^{\alpha}C^{\beta})_{i+1}-(C'C^{\alpha}C^{\beta})_i$ correlation spectrum of [U-$^{13}$C, $^{15}$N] subunit *c*. The mixing times for CP2, CP3 and $2n\tau_R$ were 5.0, 4.0 and 1.6 ms, respectively. The green line connects the inter-residue correlation of non-helical Gly. Correlations XG and GX are given by blue and red letters, respectively. The lowest contour level for the $I^{\beta}G^{\alpha}$ and $V^{\beta}G^{\alpha}/M^{\beta}G^{\alpha}$ cross peaks in a square is 63% of that in the other area. (b) The sequential assignment of $C^{\alpha}$ of A20 through A39 is indicated by red lines. The additional assignment of $C^{\alpha}$ of A12 through L19 is indicated by blue lines. (c) The sequential assignments of $C^{\alpha}$ of A62 through L72 and F76 through V78 are indicated by red and blue lines, respectively.

## Discussion

The structure distribution could be estimated from the linewidth. The apparent $C^{\alpha}$ linewidths of cross peaks are summarized in Table 1. The average linewidth of all Gly $C^{\alpha}$ was 1.6 $\pm$0.1 ppm. The average linewidth of all Ala $C^{\beta}$ was 2.0 $\pm$ 0.3 ppm. The $^{13}$C linewidth of 1.5–2.0 ppm at 11.74 T was typical of well-structured and rigid non-crystalline peptides and proteins, while a disordered structure gives a 4–5 ppm $^{13}$C linewidth (Petkova et al., 2002; Fujiwara et al., 2004). The correlation of chemical shifts and secondary structures in proteins has been studied extensively (Markley et al., 1967; Sepra and Bax, 1991; Wishart et al., 1991; de Dios et al., 1993;

*Table 1.* $^{13}C$ chemical shifts, signal intensities and assignments of Gly from inter-residue $(C^{\alpha}C^{\beta})_{i+1}-(C'C^{\alpha}C^{\beta})_i$ correlations

| Peak[a] | Chemical shift | | Linewidth | | Peak vol. | Assignment[b] |
|---|---|---|---|---|---|---|
| | F2 | F1 | F2 | F1 | | |
| *a* | 65.5 | 47.7 | 2.0 | 1.8 | 4.2 | 3IG, 1VG |
| *b* | 57.8 | 47.3 | 3.6 | 1.7 | 4.6 | 2LG, 1MG, 1EG |
| *c* | 53.5 | 44.2 | 1.7 | 1.8 | 0.88 | *1MG* |
| *e* | 44.2 | 53.5 | 1.6 | 1.6 | 0.89 | *1GL* |
| *f* | 47.7 | 55.0 | 1.5 | 1.6 | 2.2 | 2GA |
| *g* | 47.7 | 57.8 | 1.5 | 1.6 | 4.1 | 3GL, 1GK |
| h | 48.0 | 65.5 | 1.6 | 1.7 | 2.0 | 2GI |

[a]The peak *d* assigned to GG is not shown because diagonal components are overlapping. [b]The assignments in italics indicate pairs of candidate residues that take on non-helical forms.

Oldfield, 1995). Since Gly and Ala are widely scattered in subunit *c*, the structure of $EF_O$ subunit *c* in the solid state should be basically uniform.

When $EF_oc$ in solid takes a unique structure, the inter-residue $(C^{\alpha}C^{\beta})_{i+1}-(C'C^{\alpha}C^{\beta})_i$ correlations in Figure 4 can provide the main chain chemical shifts from A20 to A39 and from A62 to L72 almost sequentially on the basis of the assignment made in this work. Although most signals are not resolved, relatively well-identified signals involving Gly and Ala can work as check points within resolution limits. A red line connected cross peaks along the amino acid sequence from A20 to A39, as shown in Figure 4b, starting from the unique assignment G32G33. The signal of I30L31 was not resolved in the spectrum. However, the signal could be identified because of the determined chemical shifts for G29I30 and L31G32. The chemical shift of $F^{\alpha}35$ was assumed to be 61 ppm on the basis of database for helical conformations according to the conclusion discussed later.

Sequential assignment from A62 to L72 was performed, as shown in Figure 4c, starting from the unique assignment M65I66. Although the cross peaks of I63P64 and P64M65 were not resolved, the $C^{\alpha}$ chemical shift of P64 was determined from the intra-residue $^{15}N$–$^{13}C$ correlation experiment. In this connectivity, $V^{\alpha}68G^{\alpha}69$, $G^{\alpha}69L^{\alpha}70$ and $L^{\alpha}70G^{\alpha}71$ were attributed to peaks *a*, *g* and *b*, respectively. Since peaks *c* and *e* could not fit into this connectivity, L70G71L72 should be removed from the candidates for them. To assign peaks *c* and *e*, we connected them into the sequential assignments in Figure 4b and c as M17G18L19 and M57G58L59, respectively. Although the former fitted into the sequential walking, as indicated by blue lines, the latter did not. There was no cross peak expected for V56M57 at all. Thus, peaks *c* and *e* were assigned to M17G18L19. However, in view of the distorted signal shapes and suppressed signal intensities for peaks *c* and *e*, M17G18L19 may take heterogeneous conformations, namely, helical and non-helical ones. Since the structure of subunit *c* is basically uniform as discussed above, the heterogeneity should be local. The results provide unique assignments of the cross peaks in Figure 3c and d to $M^{\beta}57G^{N}58$ and $G^{\alpha}18L^{N}19$, respectively.

The secondary structural information was deduced from the chemical shifts of the assigned signals by TALOS. The chemical shifts of A20–A39 and A62–L72 indicated that these regions took on helices. The chemical shifts of F76–V78 indicated an extended structure in the C-terminal region. There are four Phe. The $C^{\beta}$ chemical shift of the major Phe peak in Figure 2b is typical for α-helix. Thus, the three Phe other than F76 (F35, F53, and F54) should be included in helical regions. Since the chemical shift of $Q^{\alpha}42$ determined in the 3D $N_{i+1}-C'_i-C^{\alpha}_i$ experiment was exactly the same as the value in solution, R41–D44 would take on a loop structure similar to the solution structure.

In Figure 5a, major $C'-C^{\alpha}$ cross peaks of Gly residues appeared at (175, 48) ppm, i.e., typical chemical shifts for α-helix. In addition to this, a weak cross peak was found at (172, 45) ppm, which is connected to cross peaks *c* and *e* in Figure 4a. The chemical shifts for different secondary structures were calculated by SHIFTS and

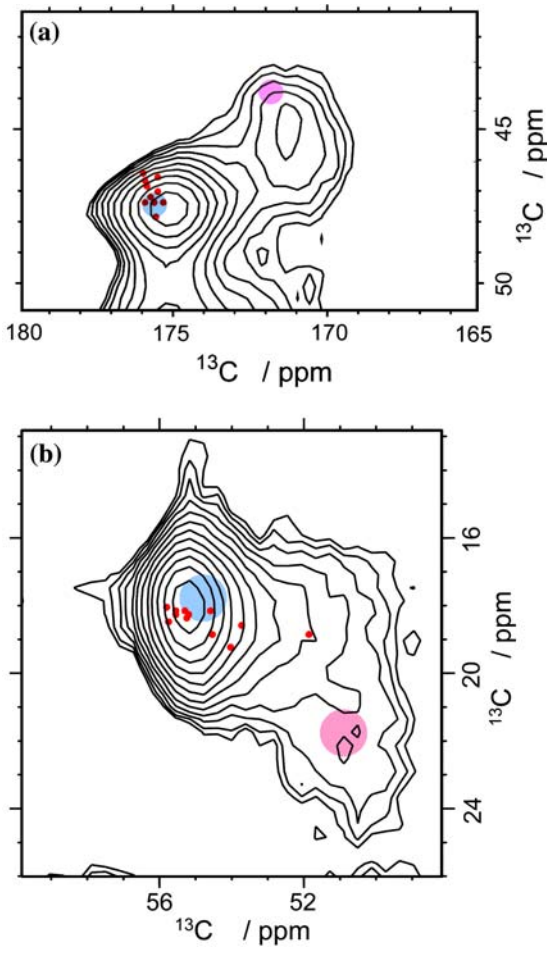


*Figure 5*. Secondary structural analysis of Gly and Ala residues based on RFDR spectra. The $C'-C^{\alpha}$ cross peak region of Gly (a) and the $C^{\alpha}-C^{\beta}$ cross peak region of Ala (b) were obtained with mixing times of 0.96 and 1.92 ms, respectively. The chemical shift values of solution structure available from BMRB are plotted as red spots. The chemical shifts predicted for helical and extended conformations by SHIFTX are shown as light blue and magenta circles, respectively.

SHIFTX. Those for helical and extended conformations are indicated by light-blue and magenta circles in Figure 5a, respectively. The results revealed that the weak cross peak could be assigned to a Gly residue taking on an extended-like non-helical form. The signal intensity ratio of the helical to extended forms was 20:3, indicating that one of the ten Gly residues is non-helical. In RFDR spectra, a signal intensity would depend on the chemical shift and the position of a chemical group in the residue. However, when we compare signals from the same chemical groups with similar chemical shifts, the signal intensity will be proportional to number of the chemical group within a certain error (Fujiwara et al., 2004). The result indicates that G18 is the sole non-helical Gly.

Secondary structural analysis of the cross peaks of Ala $C^{\alpha}-C^{\beta}$ was also performed in the same way, as shown in Figure 5b. The signal intensity ratio of the helical to extended forms was 10:1, suggesting that one or two out of the 13 Ala residues is extended. The two extended Ala should be A77 and A79 because F76–V78 takes on the extended form and A79 is the terminal residue. The isolated $C^{\beta}-C^{\gamma}$ cross peaks of Val at (22, 32) and (22, 36) ppm were ascribed to the helical and extended forms, respectively, as shown in Figure 2a. From the peak intensity, two of the six Val would be extended. Similarly, the relative peak intensities of $C^{\beta}-C^{\gamma}$ cross peaks of Leu at (25, 42) and (28, 46) ppm indicated that two of the 12 Leu take on extended conformations. From the assignments of M17G18L19, one extended Leu is L19 and the other would be L45 around the loop region. These isolated cross peaks gave us quantitative information on its secondary structure.

SHIFTX simulation gave the following further information. Judging from the $C^{\beta}$ chemical shifts, two Tyr, Y10 and Y73, took on helical conformations. The cross peaks of $A^{\alpha}A^{\beta}$ at (55, 18) ppm and $G^{\alpha}A^{\beta}$ at (48, 18) ppm in Figure 4a indicated that two GAA sequences, G23A24A25 and G38A39A40, took on helical conformations. Judging from the signal intensities and chemical shifts of the $P^{\alpha N}$ cross peaks in Figure 3a, two Pro, P47 and P64, should take on helical conformations. The cross peak at (31, 60) ppm assigned to $V_i^{\beta}M_{i+1}^{\alpha}$ observed in the $(C^{\alpha}C^{\beta})_{i+1}-(C'C^{\alpha}C^{\beta})_i$ correlation spectrum (Figure 4a) suggested that three VM sequences, V15M16, V56M57 and V74M75, took on helical conformations. The helical conformation for M57G58L59 is consistent with the assignment of peaks *c* and *e* to the M17G18L19 sequence. Furthermore, we could fit V15M16 into the sequential walking and could extend it to A12, as shown in Figure 4b. Then, the cross peak at (67, 60) ppm could be assigned to V56M57 and V74M75. The chemical shifts obtained in the experiments and the results of quantitative analyses are summarized in Table 2 (BMRB accession number, 10021). So far, the secondary structure was determined for 84% of the 79 residues.

*Table 2.* $^{13}C$ and $^{15}N$ chemical shifts of $EF_{O}c$

| a.a. | $N^{H}$ | $C'$ | $C^{\alpha}$ | $C^{\beta}$ | $C^{\gamma}$ | $C^{\delta}$ | Others | |
|---|---|---|---|---|---|---|---|---|
| M1 | | | | | | | | |
| E2 | | | | | | | | |
| N3 | | | | | | | | |
| L4 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| N5 | | | | | | | | |
| M6 | | | | | | | | |
| D7 | | | | | | | | |
| L8 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| L9 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| Y10 | | | | 38 | 131 | 133 | $C^{\varepsilon}$118 | |
| M11 | | 179 | 58 | 34 | | | | |
| A12 | 120 | 179 | 55 | 18 | | | | |
| A13 | 120 | 179 | 55 | 18 | | | | |
| A14 | 120 | 179 | 55 | 18 | | | | |
| V15 | 117 | 178 | 67 | 32 | 22 | | | |
| M16 | | 179 | 60 | 34 | | | | |
| M17 | | | (53.5) | | | | | |
| G18 | (112) | (172) | (44.2) | | | | | |
| L19 | (123) | | (53.5) | (46) | (28) | | | |
| A20 | 120 | 179 | 55 | 18 | | | | |
| A21 | 120 | 179 | 55 | 18 | | | | |
| I22 | 117 | 178 | 66 | 38 | 30/18 | 14 | | |
| G23 | 107 | 175 | 48 | | | | | |
| A24 | 120 | 179 | 55 | 18 | | | | |
| A25 | 120 | 179 | 55 | 18 | | | | |
| I26 | 117 | 178 | 66 | 38 | 30/18 | 14 | | |
| G27 | 107 | 175 | 48 | | | | | |
| I28 | 117 | 178 | 66 | 38 | 30/18 | 14 | | |
| G29 | 107 | 175 | 48 | | | | | |
| I30 | 117 | 178 | 66 | 38 | 30/18 | 14 | | |
| L31 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| G32 | 107 | 175 | 47.7 | | | | | |
| G33 | 107 | 175 | 47.3 | | | | | |
| K34 | | | 58 | | | | $N^{\zeta}$32.8 | $C^{\varepsilon}$42.2 |
| F35 | | | | 39 | 140 | 132 | | |
| L36 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| E37 | 119 | 178 | 59 | 27 | | | | |
| G38 | 107 | 175 | 48 | | | | | |
| A39 | 120 | 179 | 55 | 18 | | | | |
| A40 | 120 | 179 | 55 | 18 | | | | |
| R41 | | | | | 27 | 44 | $N^{\varepsilon}$84<br>$C^{\zeta}$160 | $N^{\eta}$71/74 |
| Q42 | | | 55.0 | | | | | |
| P43 | 137 | | 62.5 | | | 50.1 | | |
| D44 | | | | | | | | |
| L45 | | | 54 | 46 | 28 | | | |
| I46 | | 178 | 66 | 38 | 30/18 | 14 | | |
| P47 | 133 | | 65 | | | 49 | | |
| L48 | 120 | 179 | 58 | 42 | 27 | 24 | | |

*Table 2.* Continued

| a.a. | $N^H$ | C′ | $C^\alpha$ | $C^\beta$ | $C^\gamma$ | $C^\delta$ | Others | |
|---|---|---|---|---|---|---|---|---|
| L49 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| R50 | | | | | 27 | 44 | $N^\varepsilon$84 | $N^\eta$71/74 |
| | | | | | | | $C^\zeta$160 | |
| T51 | | | | 71.3 | 22.3 | | | |
| Q52 | | | | | | | | |
| F53 | | | | 39 | 140 | 132 | | |
| F54 | | | | 39 | 140 | 132 | | |
| I55 | 117 | 178 | 66 | 38 | 30/18 | 14 | | |
| V56 | 117 | 178 | 67 | 32 | 22 | | | |
| M57 | | 179 | 60 | 32.8 | | | | |
| G58 | 107 | 175 | 48 | | | | | |
| L59 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| V60 | 117 | 178 | 67 | 32 | 22 | | | |
| D61 | | | 58 | | | | | |
| A62 | 120 | 179 | 55 | 18 | | | | |
| I63 | 117 | 179 | 66 | 38 | 30/18 | 14 | | |
| P64 | 133 | | 65 | | | 49 | | |
| M65 | | 179 | 58.5 | | | | | |
| I66 | 117 | 179 | 64.7 | 38.3 | 30/18 | 14 | | |
| A67 | 120 | 179 | 55 | 18.1 | | | | |
| V68 | 117 | 178 | 66 | 31 | 22 | | | |
| G69 | 107 | 175 | 47 | | | | | |
| L70 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| G71 | 107 | 175 | 48 | | | | | |
| L72 | 120 | 179 | 58 | 42 | 27 | 24 | | |
| Y73 | | | | 38 | 131 | 133 | $C^\varepsilon$118 | |
| V74 | 117 | 178 | 67 | 32 | 22 | | | |
| M75 | | | 60 | 34 | | | | |
| F76 | | | 59.7 | 43.3 | 139 | 132 | | |
| A77 | | | 50.8 | 21.9 | | | | |
| V78 | | | 62.0 | 35.5 | 21.5 | | | |
| A79 | | | | | | | | |

Underlined: Uniquely determined
Parentheses: Maybe heterogeneous

To confirm the signal assignments in Table 2 and analyze the secondary structure of whole subunit *c*, three spectra, the $C'-C^\alpha C^\beta$ regions of RFDR spectra with mixing times of 0.64 and 1.60 ms, and the $C^\alpha-C^\beta$ region with a mixing time of 1.6 ms were simulated using the assigned chemical shifts. For the spectral simulation, non-assigned chemical shifts were predicted by SHIFTX. We could estimate the conformations at specific residues from one of the $C^\alpha$, $C^\beta$, and C′ chemical shifts. For instance, the experimentally obtained $C^\beta$ chemical shift suggested that Y10 took on a helical conformation. Then, the $C^\alpha$ and C′ chemical shifts for the helical conformation were obtained by SHIFTX.

Since the chemical shifts for M1, E2, N3, N5, M6, D7, R41, D44, R50, and Q52 could not be obtained, spectral simulation was carried out using the predicted chemical shifts of helical or extended conformations. For M1–N3 and N5–D7, signals predicted for extended structures were not found in the experimental spectrum. Therefore, the extended conformation was ruled out in these regions. R50 and Q52 were assumed to take on helical conformations, because the residues around them took on helical conformations. The

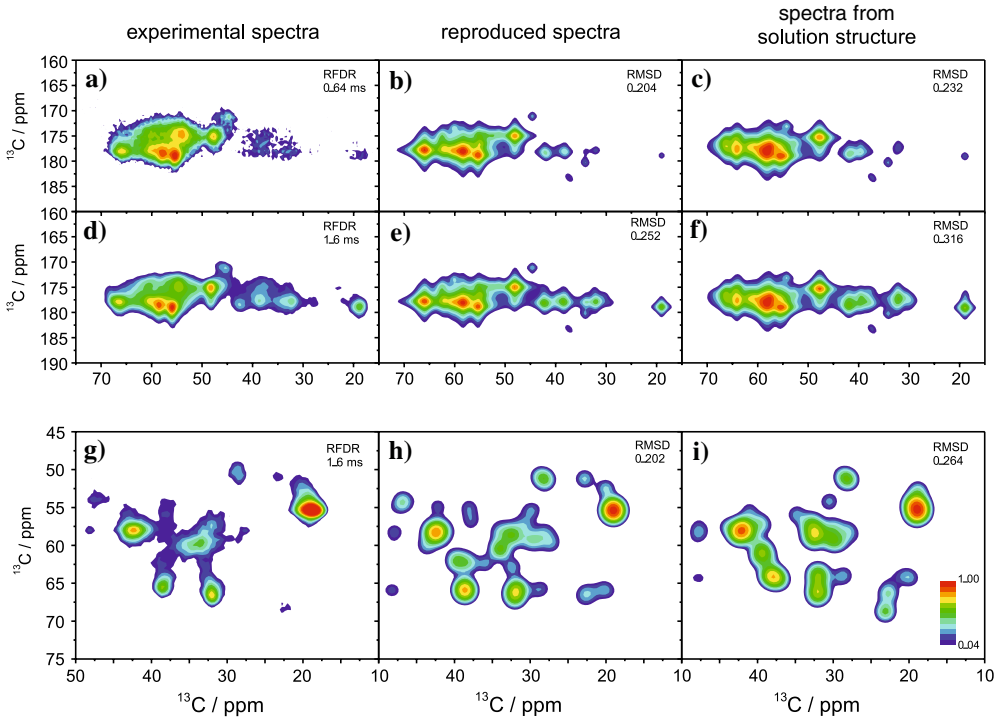


*Figure 6*. Comparison of the $C'-C^{\alpha}$, $C'-C^{\alpha}C^{\beta}$, and $C^{\alpha}-C^{\beta}$ regions in experimental and simulated RFDR spectra. (a), (d), and (g), Experimental spectra with mixing times of 0.64, 1.6 and 1.6 ms for $C'-C^{\alpha}$, $C'-C^{\alpha}C^{\beta}$, and $C^{\alpha}-C^{\beta}$, respectively; (b), (e), and (h), optimized simulation spectra; and (c), (f), and (i), spectra simulated for the solution structure. The intensity scale is logarithmic and divided into a twelve color scale, as illustrated in (i). Root mean square deviations (RMSD) were calculated for the experimental and simulated spectra.

experimentally obtained chemical shift indicated that Q42P43 took a loop structure like in the solution. Thus, for R41 and D44, we adopted chemical shifts of the solution structure, which reproduced the spectra well fitted to the experimental ones. In the next step, we modified the chemical shifts of non-assigned signals such as $C^{\alpha}$, and $C^{\beta}$ of Asn, and $C^{\alpha}$ of Tyr and Phe to improve the simulated spectrum without changing their α-helical conformations. Furthermore, $C^{\alpha}$, $C^{\beta}$ linewidths of M1 and L48–T51 were modified from 1.7 ppm to 4 ppm.

Comparison of the optimized simulation spectra with the observed one is presented in Figure 6. The observed spectra (a), (d), and (g), and the simulated ones (b), (e), and (h) correspond to the $C'-C^{\alpha}$, $C'-C^{\alpha}C^{\beta}$, and $C^{\alpha}C^{\beta}$ regions, respectively. For reference, corresponding spectra simulated from the solution structure are shown in Figure 6c, f, and i. In the experimental spectra (a) and (d), $C'$ signals appeared at around 175 ppm, suggesting the presence of extended structures. These spectral features were reproduced quantitatively in the simulated spectra based on assigned chemical shifts. In the spectral region of $C^{\alpha}C^{\beta}$, the signal separation was better than that in the carbonyl region. The peak positions of Phe and Tyr at (38, 62) ppm and extended Ala at (23, 51) ppm in the experimental spectrum (g) were reproduced in the simulated spectrum (h). In all spectral regions, the RMSD values of simulated spectra were smaller than those between the spectra of the solid and solution structures, showing a difference in the secondary structure.

The dihedral angles ϕ and Ψ were predicted by TALOS using the chemical shifts obtained on the spectral simulation and are plotted along the sequence in Figure 7. The estimated secondary structure of $EF_Oc$ in the solid state is illustrated along with the crystal structure of Yeast Mitochondrial subunit *c* (Stock et al., 1999) and the solution structure of $EF_Oc$ (Girvin et al., 1998). The major parts take on α-helix. The helix structures are not formed in the R41–L45 and F76–A79

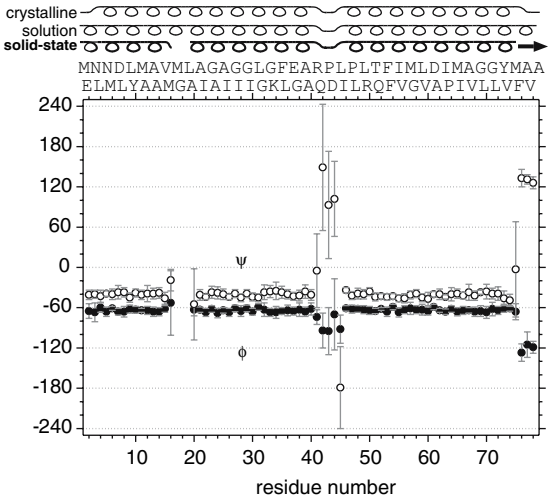


*Figure 7*. Predicted backbone dihedral angles and the deduced secondary structure of $EF_Oc$ based on the optimized simulation spectrum. The dihedral angles, ϕ and Ψ were predicted by TALOS. The error bars indicate RMSD for the ten predicted dihedral angles. The secondary structures of yeast mitochondrial subunit *c* in the crystal (Stock et al., 1999) and $EF_Oc$ in an organic solvent (Girvin et al., 1998) are also presented at the top.

regions. The R41–L45 region does not take on typical extended conformations. As discussed above, R41–L45 would take on a loop conformation as in solution, suggesting a hairpin fold for the whole peptide. Since M17–L19 seems to take heterogeneous conformations, they are not specified.

The results of this work have shown that membrane proteins of this size can be analyzed by use of high resolution SSNMR methodology. Subunits *c* from various biological species have been reported to take on helix-loop-helix type hairpin structures not only in the crystal but also in an organic solution. This work has revealed that the major parts of $EF_oc$ took on helices with a loop-like conformation in the middle, suggesting a hairpin-like structure even in the non-crystal solid. Therefore, the elemental structure in the subunit *c* ring seems quite stable. However, the appearance of a distorted region in the N-terminal helix for the same sample suggests that the stability of the helix structure is not uniform along the sequence.

## Conclusion

A strategy for signal assignment and secondary structural analysis of a uniformly [$^{13}C$, $^{15}N$] labeled helix-rich membrane protein, $EF_Oc$, in solid was shown feasible. Two- and three-dimensional SSNMR spectra under MAS were useful for the signal assignment. Especially, inter-residue correlations involving Gly, Ala, Pro and side-chains and a higher resolution in the 3D spectrum were useful for the sequence specific assignment. In addition to the NMR measurement, structure dependent chemical-shift calculation and spectral simulation were useful for the secondary structure analysis. As a result, 78, 72, 62, and 61% of signals were assigned for $C^\alpha$, $C^\beta$, $C'$, and $N^H$, respectively. The secondary structure could be obtained for 84% of the 79 residues based on ϕ and Ψ obtained from the assigned $C^\alpha$ and/or $C^\beta$ chemical shifts. Quantitative spectrum simulation provided us the information about the secondary structure of the whole subunit *c*. These results have shown that the major parts of $EF_oc$ took on helices with a loop-like conformation in the middle, suggesting a hairpin-like structure even in the non-crystal solid.

## Acknowledgements

We are grateful to Profs. R. H. Fillingame and M. Yoshida for providing us *E. coli* MEG119 strain transformed by plasmid pCP35 harboring the gene of subunit *c* from *E. coli* and for their encouragement. We would also like to thank Prof. D. S. Wishart for releasing the source program of SHIFTX to us. This work was partly supported by Grants-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Technology, Sport and Culture of Japan (HA), and grants from JST (Core Research for Evolutional Science and Technology) (HA and MK) and the Japan New Energy and Industrial Technology Development Organization (HA).

## References

Baldus, M., Petkova, A.T., Herzfeld, J. and Griffin, R.G. (1998) *Mol. Phys.*, **95**, 1197–1207.

Bennett, A.E., Rienstra, C.M., Auger, M., Lakshmi, K.V. and Griffin, R.G. (1995) *J. Chem. Phys.*, **103**, 6951–6958.

Bennett, A.E., Rienstra, C.M., Griffiths, J.M., Zhen, W.G., Lansbury, P.T. and Griffin, R.G. (1998) *J. Chem. Phys.*, **108**, 9463–9479.

Böckmann, A., Lange, A., Galinier, A., Luca, S., Giraud, N., Juy, M., Heise, H., Montserret, R., Penin, F. and Baldus, M. (2003) *J. Biomol. NMR*, **27**, 323–339.

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K. and Oschkinat, H. (2002) *Nature*, **420**, 98–102.
Cornilescu, G., Delaglio, F. and Bax, A. (1999) *J. Biomol. NMR*, **13**, 289–302.
de Dios, A.C., Pearson, J.G. and Oldfield, E. (1993) *Science*, **260**, 1491–1495.
Fujiwara, T., Shimomura, T., Ohigashi, Y. and Akutsu, H. (1998) *J. Chem. Phys.*, **109**, 2380–2393.
Fujiwara, T., Todokoro, Y., Yanagishita, H., Tawarayama, M., Kohno, T., Wakamatsu, K. and Akutsu, H. (2004) *J. Biomol. NMR*, **28**, 311–325.
Girvin, M. and Fillingame, R.H. (1993) *Biochemistry*, **32**, 12167–12177.
Girvin, M., Rastogi, V.K., Abildgaard, F., Markley, J.L. and Fillingame, R.H. (1998) *Biochemistry*, **37**, 8817–8824.
Hohwy, M., Rienstra, C.M., Jaroniec, C.P. and Griffin, R.G. (1999) *J. Chem. Phys.*, **110**, 7983–7991.
Hong, M. (1999a) *J. Biomol. NMR*, **15**, 1–14.
Hong, M. (1999b) *J. Magn. Reson.*, **139**, 389–401.
Igumenova, T.I., McDermott, A.E., Zilm, K.W., Martin, R.W., Paulson, E.K. and Wand, A.J. (2004) *J. Am. Chem. Soc.*, **126**, 6720–6727.
Jiang, W., Hermolin, J. and Fillingame, R.H. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 4966–4971.
Levy, G.C. and Lichter, R.L. (1979) *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, John Wiley & Sons, Inc, New York.
Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *Pure Appl. Chem.*, **70**, 117–142.
Markley, J.L., Meadows, D.H. and Jardetzky, O. (1967) *J. Mol. Biol.*, **27**, 25–40.
Martin, R.W. and Zilm, K.W. (2003) *J. Magn. Reson.*, **165**, 162–174.
Matsuki, Y., Akutsu, H. and Fujiwara, T. (2003) *J. Magn. Reson.*, **162**, 54–66.
Meier, T., Polzer, P., Diederichs, K., Welte, W. and Dimroth, P. (2005) *Science*, **308**, 659–662.
Meiler, J. (2003) *J. Biomol. NMR*, **26**, 25–37.
Mitome, N., Suzuki, T., Hayashi, S. and Yoshida, M. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 12159–12164.
Morcombe, C.R. and Zilm, K.W. (2003) *J. Magn. Reson.*, **162**, 479–486.
Nakano, T., Ikegami, T., Suzuki, T., Yoshida, M. and Akutsu, H. (2006) *J. Mol. Biol.*, **358**, 132–144.
Neal, S., Nip, A.M., Zhang, H. and Wishart, D.S. (2003) *J. Biomol. NMR*, **26**, 215–240.
Oldfield, E. (1995) *J. Biomol. NMR*, **5**, 217–225.
Pauli, J., Baldus, M., van Rossum, B., de Groot, H. and Oschkinat, H. (2001) *Chembiochem*, **2**, 272–281.
Petkova, A.T., Ishii, Y., Balbach, J.J., Antzutkin, O.N., Leapman, R.D., Delaglio, F. and Tycko, R. (2002) *Proc. Natl. Acad. Sci. USA*, **97**, 13045–13050.
Petkova, A.T., Baldus, M., Belenky, M., Hong, M., Griffin, R.G. and Herzfeld, J. (2003) *J. Magn. Reson.*, **160**, 1–12.
Ponder, J.W. and Richards, F.M. (1987) *J. Comput. Chem.*, **8**, 1016–1024.
Rastogi, V.K. and Girvin, M.E. (1999) *Nature*, **402**, 263–268.
Seavey, B.R., Farr, E.A., Westler, W.M. and Markley, J.L. (1991) *J. Biomol. NMR*, **1**, 217–236.
Seelert, H., Poetsch, A., Dencher, N.A., Engel, A., Stahlberg, H. and Muller, D.J. (2000) *Nature*, **405**, 418–419.
Sepra, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.
Stahlberg, H., Muller, D.J., Suda, K., Fotiadis, D., Engel, A., Meier, T., Matthey, U. and Dimroth, P. (2001) *EMBO Rep.*, **2**, 229–233.
Stock, D., Lesile, A.G. and Walker, J.E. (1999) *Science*, **286**, 1700–1705.
Straus, S.K., Bremi, T. and Ernst, R.R. (1998) *J. Biomol. NMR*, **12**, 39–50.
van Gammeren, A.J., Hulsbergen, F.B., Hollander, J.G. and de Groot, H.J.M. (2005) *J. Biomol. NMR*, **31**, 279–293.
Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) *J. Mol. Biol.*, **222**, 311–333.
Xu, X-P. and Case, D.A. (2001) *J. Biomol. NMR*, **21**, 321–333.