67

The use of $^1J_{C\alpha H\alpha}$ coupling constants as a probe for protein backbone conformation*

Geerten W. Vuister, Frank Delaglio and Ad Bax

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, U.S.A.

SUMMARY

Simple pseudo-3D modifications to the constant-time HSQC and HCACO experiments are described that allow accurate ($\pm$ 0.5 Hz) measurement of one bond $J_{C\alpha H\alpha}$ coupling constants in proteins that are uniformly enriched with ^{13}C. An empirical φ,ψ-surface is calculated which describes the deviation of $^1J_{C\alpha H\alpha}$ from its random coil value, using 203 $^1J_{C\alpha H\alpha}$ values measured for residues in the proteins calmodulin, staphylococcal nuclease, and basic pancreatic trypsin inhibitor, for which φ and ψ are known with good precision from previous X-ray crystallographic studies. Residues in α-helical conformation exhibit positive deviations of 4–5 Hz, whereas deviations in β-sheet are small and, on average, slightly negative. Data indicate that $^1J_{C\alpha H\alpha}$ depends primarily on ψ, and that $^1J_{C\alpha H\alpha}$ may be useful as a qualitative probe for secondary structure. Comparison of $^1J_{C\alpha H\alpha}$ coupling constants measured in free calmodulin and in its complex with a 26-amino-acid peptide fragment of myosin light-chain kinase confirm that the calmodulin secondary structure is retained upon complexation but that disruption of the middle part of the 'central helix' is even more extensive than in free calmodulin.

INTRODUCTION

During recent years a large array of 3D and 4D heteronuclear NMR experiments have been introduced which allow assignment and structure determination of $^{15}N/^{13}C$-labeled proteins with molecular weights of up to 30 kDa. The availability of labeled proteins together with their high-resolution structures has made it possible to investigate the correlation between conformation and a range of NMR parameters. For example, recently it was shown that the secondary ^{13}C

---

* *Supplementary material available from the authors:* One table listing 352 $^1J_{C\alpha H\alpha}$ and δ^1J-values, together with φ,ψ-values for 203 residues of known conformation. Two figures showing (a) a Ramachandran plot of the φ,ψ-values of 203 residues used in deriving $\Delta^1J(\varphi,\psi)$, and (b) the r.m.s.d. $\Delta^1J(\varphi,\psi)$ distribution.

chemical shift of C^α and C^β carbons correlates with the protein backbone conformation (Spera and Bax, 1991; Wishart et al., 1991). Coupling constants provide another rich source of structural information (Bystrov, 1976), and a large number of experiments have been proposed for measuring hetero- and homonuclear coupling constants in labeled proteins (Montelione et al., 1989; Wider et al., 1989; Chary et al., 1991; Delaglio et al., 1991; Edison et al., 1991; Griesinger and Eggenberger, 1992; Vuister and Bax, 1992a). Many of these experiments employ E.COSY (Griesinger et al., 1986)-type cross-peak patterns between two spins that are both coupled to a passive third spin. The present work focuses on the measurement of $^1J_{C\alpha H\alpha}$ and describes simple pseudo-3D modifications of the constant-time HSQC (Santoro and King, 1992; Van de Ven and Philippens, 1992; Vuister and Bax, 1992b) and HCACO (Powers et al., 1991; Palmer III et al., 1992) experiments that permit accurate measurement of the $^1J_{C\alpha H\alpha}$ coupling constants.

Although the one-bond $J_{C\alpha H\alpha}$ coupling constants frequently are thought to have uniform values, significant variations in their magnitude occur. Variations in $^1J_{C\alpha H\alpha}$ previously have been studied in conformationally constrained cyclic peptides (Egli and von Philipsborn, 1981). A recent report on the measurement of $^1J_{C\alpha H\alpha}$ in proteins shows a range of 131.6–150.3 Hz and indicates that the magnitude of $^1J_{C\alpha H\alpha}$ correlates with the backbone conformation (Vuister et al., 1992). These data also show that the variation in $^1J_{C\alpha H\alpha}$ is determined primarily by ψ, and to a lesser extent by φ. Here, the analysis is extended, using a larger database and taking into account the different 'random coil' values of the amino acids. An empirical φ,ψ-surface is calculated for the measured $^1J_{C\alpha H\alpha}$ data and the fit to this surface is compared with the fit to a previously reported trigonometric expression (Vuister et al., 1992).

MEASUREMENT OF $^1J_{C\alpha H\alpha}$

$^1J_{C\alpha H\alpha}$ couplings can be measured from heteronuclear ^1H–^{13}C correlation spectra in which ^{13}C spins are not decoupled during ^1H detection, either from in-phase or from antiphase doublet splitting, depending on the experimental scheme used. Although conceptually simple, such measurements are limited by the accuracy with which the individual peak positions can be established. Especially for large $^{15}N/^{13}C$-labeled proteins, resonance overlap, the presence of carbon–carbon couplings, small phase-errors, and broad natural line widths decrease the accuracy of J couplings measured in this manner. Here we describe a different approach which is less sensitive to the problems indicated above, and which provides very accurate values for $J_{C\alpha H\alpha}$. The new methods are simple extensions of the CT-HSQC (Santoro and King, 1992; Van de Ven and Philippens, 1992; Vuister and Bax, 1992a,b) and HCACO (Powers et al., 1991; Palmer III et al., 1992) experiments. In the present case, a set of such 2D spectra is recorded as a function of a third time variable, λ, and the intensities of correlations in the 2D spectra are modulated by $\sin(2\pi^1J\lambda)$.

The pulse schemes used in the present work are shown in Fig. 1. The mechanisms of the two experiments are very similar and only the simplest one (Fig. 1a) will be discussed in some detail. The scheme of Fig. 1b includes an additional magnetization relay step, from C^α to CO, yielding spectra that correlate H^α and CO, instead of H^α and C^α. The initial section of the pulse scheme of Fig. 1a, which differs from that of the regular CT-HSQC experiment, will be described in terms of the product operator formalism (Sørensen et al., 1983) for the (simplified) case of two J-coupled protons, I_1 and I_2, and a single ^{13}C, labeled S, which is coupled to I_1. Additional spins, as present in practice, alter the complexity of the description, but do not introduce any terms of a
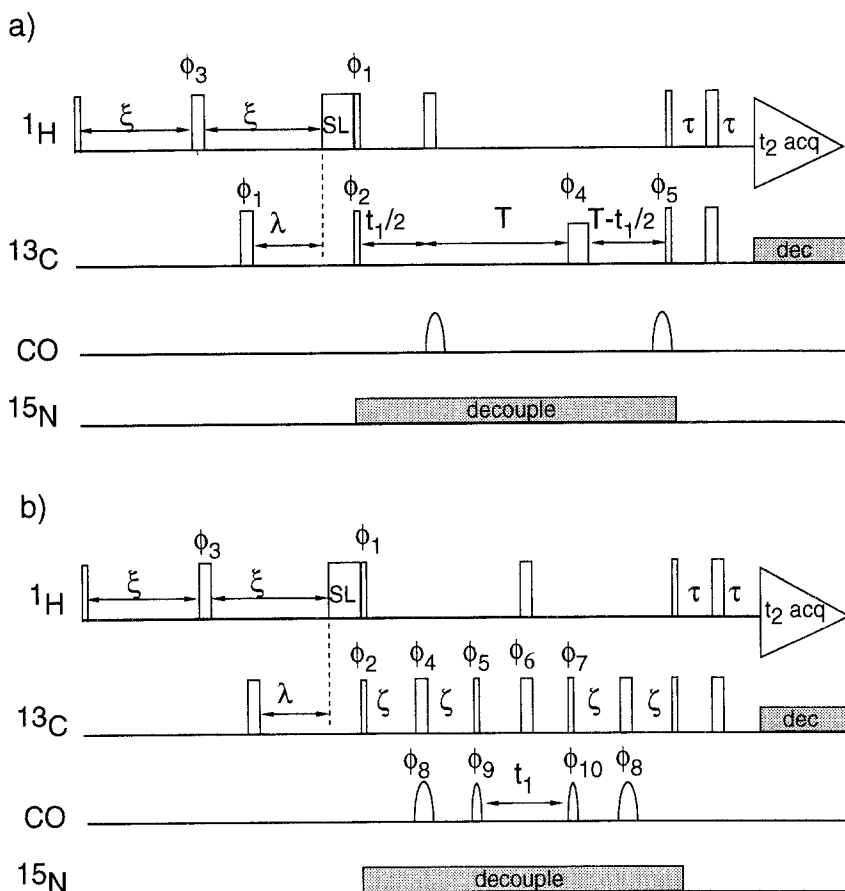
69

Fig. 1. Pulse sequence of (a) the J-modulated constant-time HSQC experiment and (b) the J-modulated HCACO experiment. Narrow and wide pulses denote 90° and 180° flip angles, respectively. Carbonyl 90° and 180° pulses were implemented as phase-modulated pulses with an amplitude profile corresponding to the center lobe of a sine function. For the J-modulated constant-time HSQC experiment, the carrier for the carbon 90° and 180° pulses was positioned at 43 ppm and the power of the 180° φ_4 pulse was adjusted in such a way that it did not excite the ^{13}CO nuclei. For the J-modulated HCACO experiment, the carbon carrier was positioned at 57 ppm and all 90° and 180° pulses were adjusted in such a way that they did not exite the ^{13}CO nuclei. SL denotes a 0.5-ms spin-lock pulse to suppress the residual HDO resonance (Messerle et al., 1989), dec denotes synchronous ^{13}C GARP decoupling (Shaka et al., 1983a), using a RF field strength of 4.5 kHz. ^{15}N decoupling was accomplished with an asynchronous WALTZ-16 decoupling scheme (Shaka et al., 1983b), using a RF field strength of 1.5 kHz. Unless indicated otherwise, all pulses are applied along the x-axis. The phase cycle for (a) is: $\varphi_1 = $ y, $-$y; $\varphi_2 = $ 4(x), 4($-$x); $\varphi_3 = $ 2(x), 2($-$x), 2(y), 2($-$y); $\varphi_4 = $ 2(y), 2($-$y), 2($-$x), 2(x); $\varphi_5 = $ 16(x), 16($-$x); receiver = 4(x, $-$x), 8($-$x,x), 4(x,$-$x). The phase cycle for (b) is: $\varphi_1 = $ y, $-$y; $\varphi_2 = $ 8(x), 8($-$x); $\varphi_3 = $ 2(x), 2($-$x), 2(y), 2($-$y); $\varphi_4 = $ 4(x), 4(y); $\varphi_5 = $ 4(y), 4($-$y); $\varphi_6 = $ x, $-$x, y, $-$y; $\varphi_7 = $ y; $\varphi_8 = $ 4(x), 4($-$x); $\varphi_9 = $ 2(x), 2($-$x); $\varphi_{10} = $ 16(x), 16($-$x); receiver = (x, $-$x, $-$x, x), 2($-$x, x, x, $-$x), (x, $-$x, $-$x, x), ($-$x, x, x, $-$x), 2(x, $-$x, $-$x, x), ($-$x, x, x, $-$x). In both experiments quadrature detection in t_1 was obtained by the States-TPPI technique (Marion et al., 1989) incrementing phases φ_2 and φ_9, respectively. The following delay durations were used: $\tau = $ 1.7 ms, T = 13.3 ms, $\zeta = $ 3.3 ms; $\xi = \lambda_{max} + $ 10 μs. The effective duration of λ^1 during which heteronuclear $J_{C\alpha H\alpha}$ evolves must be compensated for the effect of finite pulse widths and is given by $\lambda' = \lambda + (\tau_{180}(^{13}C) + (2/\pi)\tau_{90}(^1H))/2$, where $\tau_{180}(^{13}C)$ and $\tau_{90}(^1H)$ are the ^{13}C 180° and ^{1}H 90° pulse widths, respectively.

different nature than discussed below for the simple three-spin system. During the first ξ-180°(^{1}H)-ξ part of the pulse sequence, the heteronuclear coupling is active for a period 2λ and I_1 (H$^\alpha$) proton coherence dephases during this time due to J_{I1S}:

$$I_{1y} \rightarrow I_{1y} \cos(2\pi J_{I1S}\lambda) - 2I_{1x} S_z \sin(2\pi_{I1S}\lambda: \tag{1}$$

Both terms at the right hand side of Eq. 1 are also subject to homonuclear J coupling ($^3J_{I1I2}$), active during the ξ-180° (^{1}H)-ξ period, but only the last term in Eq. 1 needs to be considered as the other term cannot be transferred to ^{13}C by the subsequent part of the pulse scheme:

$$2I_{1x}S_z \sin(2\pi J_{I1S}\lambda) \rightarrow 2I_{1x}S_z \sin(2\pi J_{I1S}\lambda)\cos(2\pi J_{I1I2}\xi) + 4I_{1y}I_{2z}S_z \sin(2\pi J_{I1S}\lambda)\sin(2\pi J_{I1I2}\xi) \tag{2}$$

The $I_{1x}S_z$ term at the right-hand side in Eq. 2 represents the desired transfer into transverse ^{13}C magnetization. This term is not affected by the short spin lock (SL in Fig. 1) and is transformed to antiphase carbon magnetization by the 90°$_{\varphi1}$ (^{1}H), 90°$_{\varphi2}$ (^{13}C) pulse pair. Subsequently, it will evolve as described before for the CT-HSQC (Vuister and Bax, 1992b) and HCACO experiments (Powers et al., 1991) and give rise to a doubly absorptive F_1/F_2 line shape in the 2D spectrum. The last term of Eq. 2 is affected by the short spin-lock pulse, of total flip angle α:

$$4I_{1y}I_{2z}S_z \xrightarrow{\alpha°_x} 4\cos^2\alpha(I_{1y}I_{2z}S_z) - 4\sin^2\alpha(I_{1z}I_{2y}S_z) + 4\sin\alpha\cos\alpha(I_{1z}I_{2z}S_z - I_{1y}I_{2y}S_z) \tag{3a}$$

The flip angle, α, is many times 360° and varies over different parts of the sample due to RF inhomogeneity. Consequently, only terms with coefficients $\cos^2\alpha$ and $\sin^2\alpha$ 'survive' this type of trimpulse and Eq. 3a can be rewritten as:

$$4I_{1y}I_{2z}S_z \xrightarrow{\alpha°_x} 2I_{1y}I_{2z}S_z - 2I_{1z}I_{2y}S_z \tag{3b}$$

The subsequent 90°$_{\varphi1}$ pulse converts the ^{1}H (I spin) terms at the right-hand side of Eq. 3b into zero quantum coherence, and these terms give rise to weak artifacts in the final spectrum that are displaced in the F_1 dimension by $\pm$ ZQ from the S spin frequency, where ZQ is the zero quantum I_1I_2 frequency. The multiplet components of these artifacts are antiphase to each other and 90° out of phase with the signals of interest. These zero quantum artifacts cannot be eliminated easily by phase cycling or purge pulses, but in practice they are strongly attenuated by relaxation and further dephasing caused by homonuclear J coupling during the relatively long period 2T (Fig. 1a) or 4ζ+t$_1$ (Fig. 1b). Equation 2 indicates that the absorptive signal component is attenuated by a factor $\cos(2\pi J_{I1I2}\xi)$ as a result of the homonuclear couplings. For an 11-ms ξ duration, this loss in practice is less than $\sim$30%, except for glycines which suffer more because of their large geminal J coupling.

As is clear from Eq. 1, the intensity of the I_1-S correlation is modulated by $\sin(2\pi J_{I1S}\lambda)$. This allows $^1J_{C\alpha H\alpha}$ to be extracted from a series of spectra in which λ is varied. The simple dependence of ^{1}J on λ makes the data extremely suitable for a non-linear minimization procedure with only a limited set of adjustable parameters. Since the complete cross peak in each of the spectra of the time series can be used as input for the fitting procedure, a highly overdetermined system is obtained which yields very accurate data.

ANALYSIS OF $^1J_{C\alpha H\alpha}$ DATA

The quantitative description of the magnitude of coupling constants in terms of lone pair effects has historically received much attention in the field of organic chemistry. In such a description, lone pair effects appear as a correction to a classical localized molecular orbit (MO) or valence bond (VB) treatment. Molecular geometry can then be related to the orientation of the quasi-localized lone pairs (Gil and von Philipsborn, 1989). In such a description, $^1J_{C\alpha H\alpha}$ is influenced mainly by the interactions between the H_α-C_α bonding orbital and the carbonyl π orbitals and the lone pair N electrons, i.e. the $^1J_{C\alpha H\alpha}$ coupling constant is expected to depend both on φ and ψ.

The variation in the magnitude of the $^1J_{C\alpha H\alpha}$ coupling constants as a function of φ and ψ can be described by a simple trigonometric function (Vuister et al., 1992):

$$^1J_{C\alpha H\alpha} = A + B\sin(\psi+138°) + C\cos2(\psi+138°) + D\cos2(\varphi+30°) \tag{4}$$

In our previous study, we used A = 140.3 (independent of the type of amino acid), B = 1.4, C = −4.1 and D = 2.0. The larger $J_{C\alpha H\alpha}$ values observed for proline residues, as compared to all other residue types, prevented these from being included. Indeed, from studies on small peptides it has

TABLE 1

RANDOM COIL VALUES OF $^1J_{C\alpha H\alpha}$ FOR AMINO ACIDS IN RANDOM COIL PEPTIDES, FROM THE ROOT-MEAN-SQUARE DEVIATION OF MEASURED $^1J_{C\alpha H\alpha}$ VALUES AND THE $\Delta^1J(\varphi,\psi)$ AND $\Delta^1J_j(\varphi,\psi)$ DISTRIBUTIONS[a]

Residue	Entries	Random coil	r.m.s.d. using	
			$\Delta^1J(\varphi,\psi)$	$\Delta^1J_j(\varphi,\psi)$
Ala	22	143.7	1.5	1.8
Arg	10	141.5	2.3	2.7
Asn	7	141.5	1.3	2.3
Asp	14	142.5	1.7	2.9
Gln	7	141.1	1.5	1.6
Glu	21	141.9	1.3	1.4
His	3	143.8	1.2	1.2
Ile	14	141.3	1.5	1.8
Leu	13	141.1	1.6	2.0
Lys	18	141.5	1.3	1.6
Met	9	142.2	2.6	2.8
Phe	8	142.9	1.3	1.3
Pro	9	148.4	1.5	1.8
Ser	6	142.1	2.1	2.3
Thr	19	141.4	1.7	2.0
Trp	1	143.0	0.3	0.4
Tyr	6	143.0	0.9	1.1
Val	16	141.3	1.9	2.3

[a] Values for glycine residues are not included because of lack of stereospecific assignment. Cysteine residues are not included because no random coil value was available.

been suggested that the constant term of Eq. 4 is influenced by the C_α substituents (Hansen, 1981). Here, we accommodate for this effect by replacing the A value by the sum of a new constant, A', and the $^1J_{C\alpha H\alpha}$ values measured in random coil peptides (Table 1).

The difference between a measured $^1J_{C\alpha H\alpha}$ coupling and its random coil value is denoted δ^1J. From the couplings in our database, obtained for residues with well-defined backbone angles, an experimental φ,ψ-distribution ($\Delta^1J(\varphi,\psi)$) is calculated by summation over all $\delta^1J(\varphi_k,\psi_k)$-values convoluted with a 2D Gaussian profile $G(\varphi,\psi,\varphi_k,\psi_k)$:

$$\Delta^1J(\varphi,\psi) = \Sigma_k \, d^1J(\varphi_k,\psi_k) \times G(\varphi,\psi,\varphi_k,\psi_k) \, / \, \Sigma_k \, G(\varphi,\psi,\varphi_k,\psi_k) \tag{5}$$

where $\Sigma_k \, G(\varphi,\psi,\varphi_k,\psi_k)$ denotes a 'residue density function', with $G(\varphi,\psi,\varphi_k,\psi_k)$ defined as:

$$G(\varphi,\psi,\varphi_k,\psi_k) = \exp\{-[(\varphi - \varphi_k)^2 + (\psi - \psi_k)^2] \, / \, 200\} \tag{6}$$

$\Delta^1J(\varphi,\psi)$ describes the variation in the average δ^1J value as a function of φ and ψ.

The root-mean-square deviation distribution, r.m.s.d. $\Delta^1J(\varphi,\psi)$, is calculated as:

$$\text{r.m.s.d.} \, \Delta^1J(\varphi,\psi) = \{\Sigma \, [\delta^1J(\varphi_k,\psi_k) - \Delta^1J(\varphi,\psi)]^2 \times G(\varphi,\psi,\varphi_k,\psi_k) \, / \, \Sigma \, G(\varphi,\psi,\varphi_k,\psi_k)\}^{1/2} \tag{7}$$

The r.m.s.d. distribution shows the spread in the δ^1J values for a given set of φ,ψ values. This spread will depend strongly on the width of the Gaussian convolution function of Eq. 6. For example, if a very narrow Gaussian function were used (e.g., by replacing the factor 200 in Eq. 6 by 0.2), a very 'rough' surface would be obtained but all data would be much closer to this surface because the magnitude of $\Delta^1J(\varphi_k,\psi_k)$ would be determined almost exclusively by the residue with angles φ_k,ψ_k. Therefore, it is also useful to calculate the r.m.s.d. for the function $\Delta^1J_j(\varphi,\psi)$, which is composed in the manner of Eq. 5, but excluding amino acid type j (for which the r.m.s.d. is to be calculated) from the distribution. $\Delta^1J_j(\varphi,\psi)$ is the empirical function, composed of all residue types except j, which predicts δ^1J of residue type j, i.e., it indicates how well the δ^1J of residue type j is described by the data for all other residue types.

EXPERIMENTAL

NMR

All NMR spectra were recorded at 35 °C on a Bruker AMX-600 spectrometer operating at a ^1H resonance frequency of 600.13 MHz. The natural abundance ^1H–^{13}C correlation spectra of bovine pancreatic trypsin inhibitor (BPTI) (11 mM, pD 5.8), angiotensin II, and a random coil peptide comprising residues 74–83 of the 'central helix' of calmodulin were recorded with a HMQC (Bax et al., 1983) pulse sequence in which a 90° (^{13}C) purge pulse (Sørensen and Ernst, 1983) was inserted just prior to data acquisition and no ^{13}C decoupling was used during ^1H data acquisition. Spectra for all three compounds were recorded with the ^{13}C transmitter positioned at 56 ppm, and 128 scans were acquired for each complex increment. For angiotensin and the central helix peptide, the acquired data matrix comprised $256*(t_1) \times 512*(t_2)$ points (where n* indicates n complex points), with acquisition times of 25 (t_1) and 81 (t_2) ms. Squared sine-bell windows shifted by 60° were used for both domains, and the spectra were zero-filled to $1024*(t_1)$

$\times$ 2048*(t_2). For the BPTI spectrum, a 384* $\times$ 384* data matrix was acquired with acquisition times of 37 (t_1) and 61 (t_2) ms. Again, squared sine-bell windows shifted by 60° were used for both domains. The spectrum was zero-filled to 1024* $\times$ 1024*. The $^1J_{C\alpha H\alpha}$ coupling constants were obtained from the F_2 J splittings by surface fitting of the peak locations, using the program NMR2 (NMRi, Syracuse, NY).

Series of CT-HSQC and HCACO spectra with varying delays, λ, were recorded with the pulse sequences sketched in Fig. 1 for a 1.0 mM solution in D_2O calmodulin (CaM) (148 residues) complexed with 4 molar equivalents of Ca^{2+}, pD 6.8, a 1.5 mM solution in D_2O of staphylococcal nuclease (SNase) (149 residues) complexed with Ca^{2+} and thymidine 3',5'-diphosphate, pD 7.0, and a 1.5 mM solution in D_2O of a 1:1 complex of calmodulin with a 26-amino acid peptide fragment of skeletal muscle myosin light-chain kinase (CaM/M13), pD 6.8. Both CaM and SNase were uniformly enriched (>95%) in ^{15}N and ^{13}C.

Details regarding the CT-HSQC experiments have been described previously (Vuister and Bax, 1992b). In the present set of experiments, the constant-time duration 2T was set to 26.6 ms. For each λ duration, a 128*(t_1) $\times$ 384*(t_2) data matrix was recorded with 64 scans per complex t_1 increment. Squared sine-bell windows shifted by 70° were used in both domains, prior to zero-filling to 1024* $\times$ 1024* points, and Fourier transformation. Seven CT-HSQC spectra of CaM were recorded with λ' values of 5.269, 6.519, 7.019, 7.519, 8.769, 10.019 and 10.519 ms. Fourteen CT-HSQC spectra of SNase were recorded with λ' values of 6.269, 6.519, 6.769, 7.269, 7.519, 7.769, 8.769, 9.519, 10.019, 10.269, 10.519, 10.769, 11.019 and 11.519 ms. Finally, six CT-HSQC spectra of CaM/M13 were recorded with λ' values of 6.019, 7.019, 8.019, 8.769, 9.519 and 10.519 ms.

For the HCACO spectra, 200*(t_1) $\times$ 384*(t_2) data matrices were recorded with 64 scans per complex increment. Squared sine-bell windows, shifted by 66°, were used in both domains, prior to zero-filling to 1024* $\times$ 1024* points, and Fourier transformation. For CaM, eleven spectra were recorded with λ' delays of 5.765, 6.515, 7.265, 8.015, 8.765, 9.515, 9.765, 10.265, 11.015, 11.265 and 11.765 ms. For CaM/M13, fourteen spectra were recorded with λ' delays of 5.765, 6.015, 6.515, 7.265, 7.765, 8.015, 8.765, 9.515, 9.765, 10.265, 10.515, 11.015, 11.265 and 11.765 ms.

Cross peaks in the λ-dependent series of spectra were fitted to Gaussian profiles in the F_1 and F_2 dimensions and a sinusoidal λ'-dependent function in the third dimension, using a non-linear least-squares minimization routine (Delaglio, F., unpublished results). The fit with six adjustable parameters (F_2-position, F_2-width, F_1-position, F_1-width, amplitude and frequency) typically accounted for more than 95% of the observed intensity. In case of partially overlapping peaks, fitting for these peaks was done simultaneously. As discussed below, estimates of the errors in J were obtained by comparison of the results obtained from different experiments.

The assignments reported previously for BPTI (Wagner and Brühwiler, 1986; Hansen, 1981), CaM (Ikura et al., 1990), CaM/M13 (Ikura et al., 1991a), and SNase (Torchia, private communication) were used for identification of the ^1H–^{13}C correlations.

Statistics

Our database comprises 352 residues for which couplings could be determined with sufficient accuracy (see Table 2 in the Supplement). For BPTI, calmodulin, and SNase, the crystal structures are known at a resolution of 1.0 Å (Wlodawer et al., 1987), 2.2 Å (Babu et al., 1988) and 1.65

Å (Loll and Lattman, 1989), respectively. Amino- and carboxy-terminal residues (Ala[1]–Leu[4] and Ser[147]–Lys[148] for CaM; Arg[1] and Ala[58] for BPTI; Ala[1]–Leu[7] and Glu[142]–Gln[149] for SNase) were not used for calculation of the $\Delta^1 J(\varphi,\psi)$ distributions, nor were residues for which the X-ray and NMR studies reported different structures, and regions with substantial disorder (i.e. large B factors or low values of the generalized order parameter, S^2), including part of the ω-loop of SNase (Glu[43]–Tyr[54]) and residues Lys[75]–Ser[81] of calmodulin. Values for glycine residues were not included because measurement of the $^1 J_{C\alpha H\alpha}$ using the J-modulated experiments gives poor results when T is set to 13.3 ms, and because no stereospecific assignments of the H_α resonances have been made at this time. No data were obtained for cysteine residues using the J-modulated experiments because both calmodulin and SNase lack this residue. Reliable values for cysteine residues in the 2D spectrum of BPTI were only obtained for Cys[51] and Cys[55]. No reliable values could be determined for the remaining cysteines because of overlap with the residual HDO (Cys[38]), or because the multiplet components were relatively weak and did not meet the 'goodness-of-fit' criterion in the surface fitting procedure (3 residues).

Random coil $^1 J_{C\alpha H\alpha}$ values were obtained from the HMQC spectra of angiotensin II, the
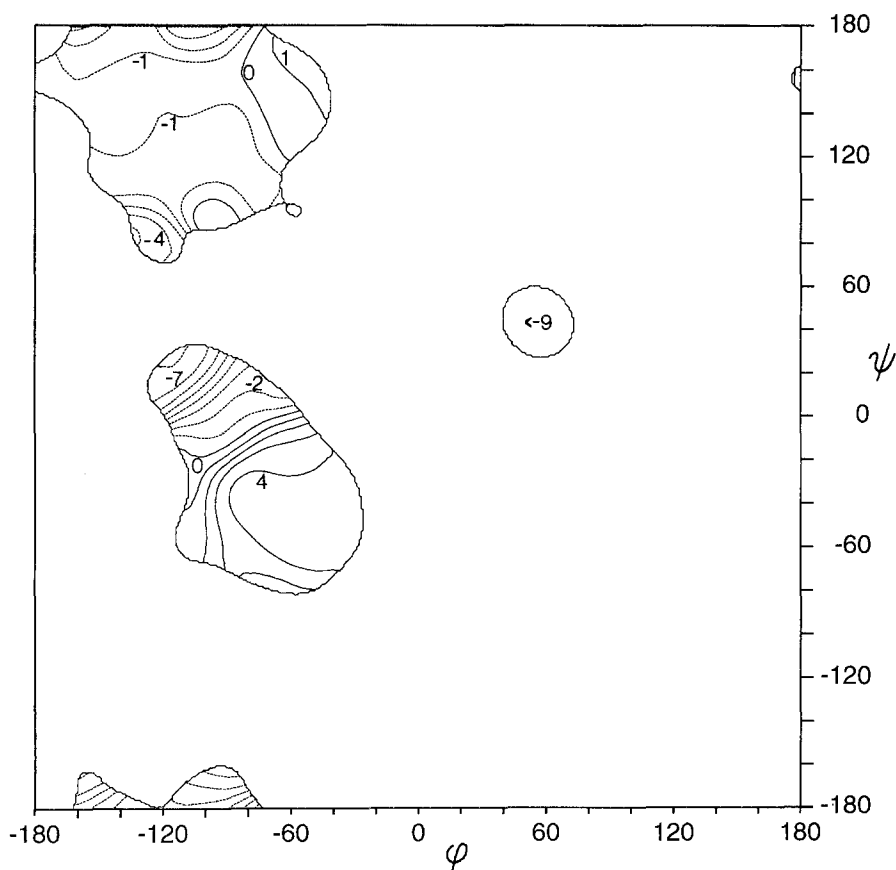

Fig. 2. $\Delta^1 J(\varphi,\psi)$ distribution calculated from Eq. 5 for 203 residues with accurately known φ and ψ angles. Equidistant contours from −9 to 4 Hz are shown for those areas where the residue density function exceeded 1.0. Positive and negative levels are shown as solid and dashed lines, respectively.

peptide comprising the 'central helix' of calmodulin, and from the unstructured N- and C-terminal tails of SNase. No random coil value for oxidized cysteine residues was available.

RESULTS AND DISCUSSION

Fitting the cross peaks in the λ-dependent spectra to the model function yields highly reproducible results. Evaluation of the fitting procedure on simulated data, added to regions of the spectra where no real signal is present, suggests that the accuracy of this method is about 0.3 Hz. An independent estimate of the error is obtained by comparing data from different experiments. For CaM, $^1J_{C\alpha H\alpha}$ coupling constants measured with the J-modulated CT-HSQC method show a r.m.s.d. of 1.05 Hz ($n = 57$) with the values obtained from the $^1J_{C\alpha H\alpha}$ multiplet splitting in a 3D CT- GDHCACO spectrum (Vuister et al., 1992). For SNase, a r.m.s.d. of 0.92 Hz ($n = 74$) was found. Comparison of the data obtained from the series of J-modulated CT-HSQC spectra with the series of J-modulated HCACO spectra yielded a r.m.s.d. of 0.57 Hz ($n = 39$) for CaM and 0.72 Hz ($n = 56$) for CaM/M13. Finally, by using two different set of only 7 (of the 14) SNase CT-HSQC spectra in two separate fittings, we obtained a r.m.s.d. of 0.3 Hz ($n = 105$) between those two sets. These comparisons of J values, measured from different sets of experimental data, indicated that fitting of non-overlapping peaks yields highly reproducible results, with an accuracy better than 0.5 Hz. For resonances with partial overlap, the accuracy is somewhat lower and only resonances that reproduce better than ±1 Hz are included in the database.

For 203 (out of 352) residues for which $^1J_{C\alpha H\alpha}$ was measured, φ and ψ are known with good precision from high-resolution X-ray work. A Ramachandran plot and a complete list of φ, ψ angles, $^1J_{C\alpha H\alpha}$, and δ^1J values is available as supplementary material. After subtraction of the random coil values (cf. Table 1), the set of 203 δ^1J values was used to calculate the experimental $\Delta^1J(\varphi,\psi)$ distribution according to Eq. 5. The result is shown in Fig. 2 for those regions where the
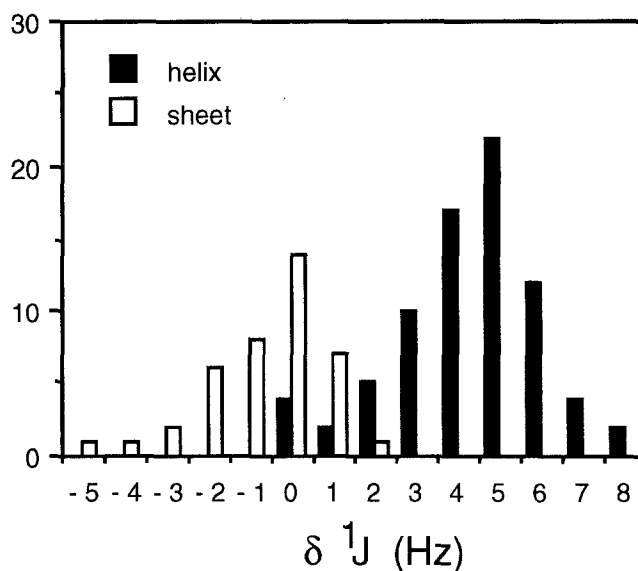

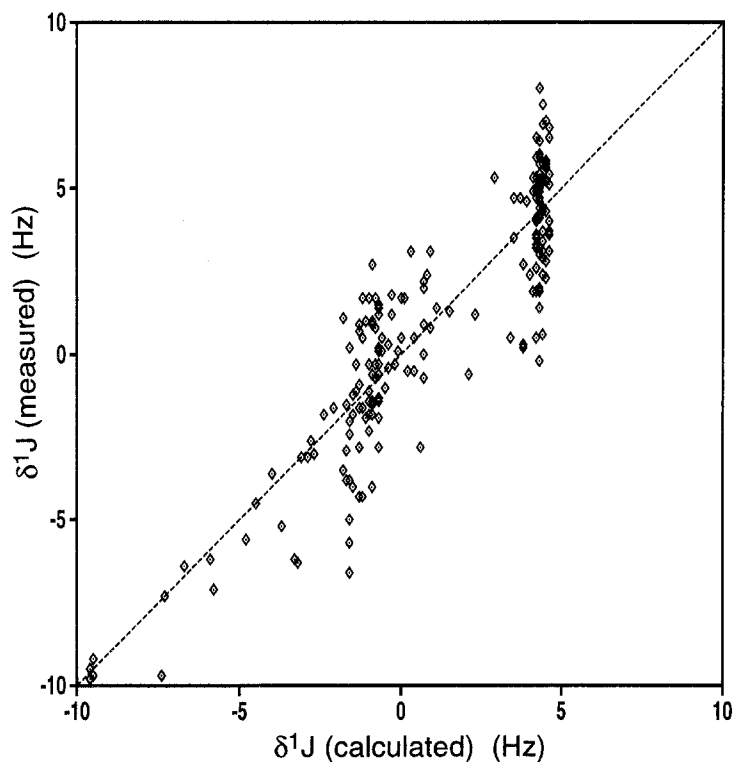Fig. 3. Histogram of the δ^1J values for residues in α-helical and β-sheet conformation.

Fig. 4. Measured $\delta^1 J$ versus $\delta^1 J$ expected on the basis of the $\Delta^1 J(\varphi,\psi)$ distribution of Fig. 2. The r.m.s.d. is 1.6 Hz.

residue density function exceeds 1.0. This figure shows that α-helical residues, i.e. φ,ψ about $-60°$, experience a positive deviation of $\sim$4–5 Hz from their random coil value. In contrast, $^1J_{C\alpha H\alpha}$ in a β-sheet or in residues with an extended conformation, on average, is very close to its random coil value. Large negative deviations from the random coil $^1J_{C\alpha H\alpha}$ coupling constants (< -8 Hz) are observed for residues with $\psi \sim 40°$, i.e. in practice often for residues with both positive ψ and φ angles.

$J_{C\alpha H\alpha}$ for residues in calmodulin for which ^{15}N relaxation studies previously found low values of the generalized order parameter, S^2 (Barbato et al., 1992), are all close to their random coil value. However, by itself, a $J_{C\alpha H\alpha}$ close to its random coil value cannot be used as evidence for a flexible structure because extended structures, such as those found in a β-sheet, also have $J_{C\alpha H\alpha}$ close to the random coil value. Residues near the amino- and carboxy-terminal ends of the polypeptide backbone of SNase, for which very small order parameters have been found (Kay et al., 1989), are indistinguishable from values measured in short unstructured peptides.

Qualitative inspection of Fig. 2 indicates that a large fraction of the variation in the $^1J_{C\alpha H\alpha}$ coupling constants can be attributed to the ψ dependence. This suggests that $\delta^1 J$ may be useful as a qualitative probe for secondary structure. In Fig. 3, the histograms of residues in helical domains ($-83° < \varphi < -38°$ and $-73° < \psi < -16°$) ($n = 78$) and β-sheet ($-151° < \varphi < -84°$ and $105° < \psi < 175°$) ($n = 40$) are shown. The two maxima are clearly distinct, although some overlap in the two distributions is observed for $\delta^1 J$ values in the 0–2 Hz range. Although it may be

tempting to describe the variation in $^1J_{C\alpha H\alpha}$ solely in terms of ψ dependence, the distribution of Fig. 2 also indicates at least a small φ dependence. Unfortunately, the relatively small range in φ-values accessible in proteins prohibits a better insight into the φ dependence of the $^1J_{C\alpha H\alpha}$ coupling constant.

The r.m.s.d. surface calculated using Eq. 7 is given in the supplementary material. It shows that the helical region, the β-sheet domain, and the region of positive φ,ψ have the lowest r.m.s.d. (<2 Hz) and are therefore determined most accurately. These also are the regions that show the smallest φ,ψ gradients in $\Delta^1J(\varphi,\psi)$. The correlation between values expected on the basis of the $\Delta^1J(\varphi,\psi)$ distribution versus the measured values is shown in Fig. 4, and shows a good correlation with a r.m.s.d. of 1.6 Hz. In contrast, a fitting of the 203 δ^1J values to Eq. 1 using singular value decomposition yields a r.m.s.d. of 1.9 Hz, slightly lower than the r.m.s.d. of 2.0 Hz reported previously for Eq. 4 (Vuister et al., 1992), using a smaller database and a single random coil value.

The 18 $\Delta^1J_j(\varphi,\psi)$ distributions have been used to calculate the r.m.s.d. between observed and expected δ^1J for each residue type (cf. Table 1). This gives an indication how well the δ^1J values for a certain residue type can be predicted, based on the data for all other residue types. The r.m.s.d. values calculated from the $\Delta^1J_j(\varphi,\psi)$ distributions follow the trend of the r.m.s.d. values calculated from the $\Delta^1J(\varphi,\psi)$ distribution, but are, on average, about 0.3 Hz higher. This indicates that there is no systematic bias towards specific residue types.

As a test case we probed the changes in the $^1J_{C\alpha H\alpha}$ coupling constants that occur for the protein calmodulin upon complexation with a 26-amino acid peptide, comprising the CaM-binding domain of skeletal muscle myosin light-chain kinase. Eight α-helical segments are present together with two short anti-parallel β-sheets (Ikura et al., 1991b). The location of these elements is indicated in Fig. 5, together with the measured δ^1J-values for both CaM (solid lines) and CaM/M13 (broken lines). It is evident from the figure that the δ^1J-values for CaM follow the pattern expected from its secondary structure. Helical regions clearly show positive δ^1J-values of about 4 Hz. The beginnings and ends of the helices, as determined previously from the NMR data, agree fairly well with the pattern observed from the δ^1J-values. The δ^1J-values provide further support for the disruption of the so-called 'central helix' in CaM for Asp[78]–Ser[81]. In fact, residues Met[76] and Lys[77] also have δ^1J values which are significantly lower than expected for an α-helix (vide infra). Comparing the δ^1J-values for residues 1–73 with those of 74–146 clearly reveals the homology in secondary structure of the two globular domains (residues 6–73 and 83–146) of calmodulin.

The recently reported low resolution solution structure of the CaM-peptide complex (Ikura et al., 1992) shows that the conformation of the two globular domains of calmodulin remains essentially unchanged upon complexation. The δ^1J-values support this notion, showing the same overall pattern as the values obtained for CaM. The solution structure indicates that the two domains of this globular approximately ellipsoidal complex are connected by a flexible loop consisting of residues 74–82. The δ^1J-values are in qualitative agreement with this finding, although low values for δ^1J are observed as early as Met[72], whereas Glu[82] has a large value, suggesting that this residue may already be part of helix E. The carboxy-terminal ends of both helices D and H are exposed to solvent both in free CaM, and in the CaM/M13 complex. For the last residues of these helices, the $^1J_{C\alpha H\alpha}$ values are smaller than expected, whereas NOE connectivities are indicative of helical structure. It is unclear at present whether the decrease in $^1J_{C\alpha H\alpha}$ is caused by solvent effects (Barfield, 1973), or by fraying of the last residues of these helices.
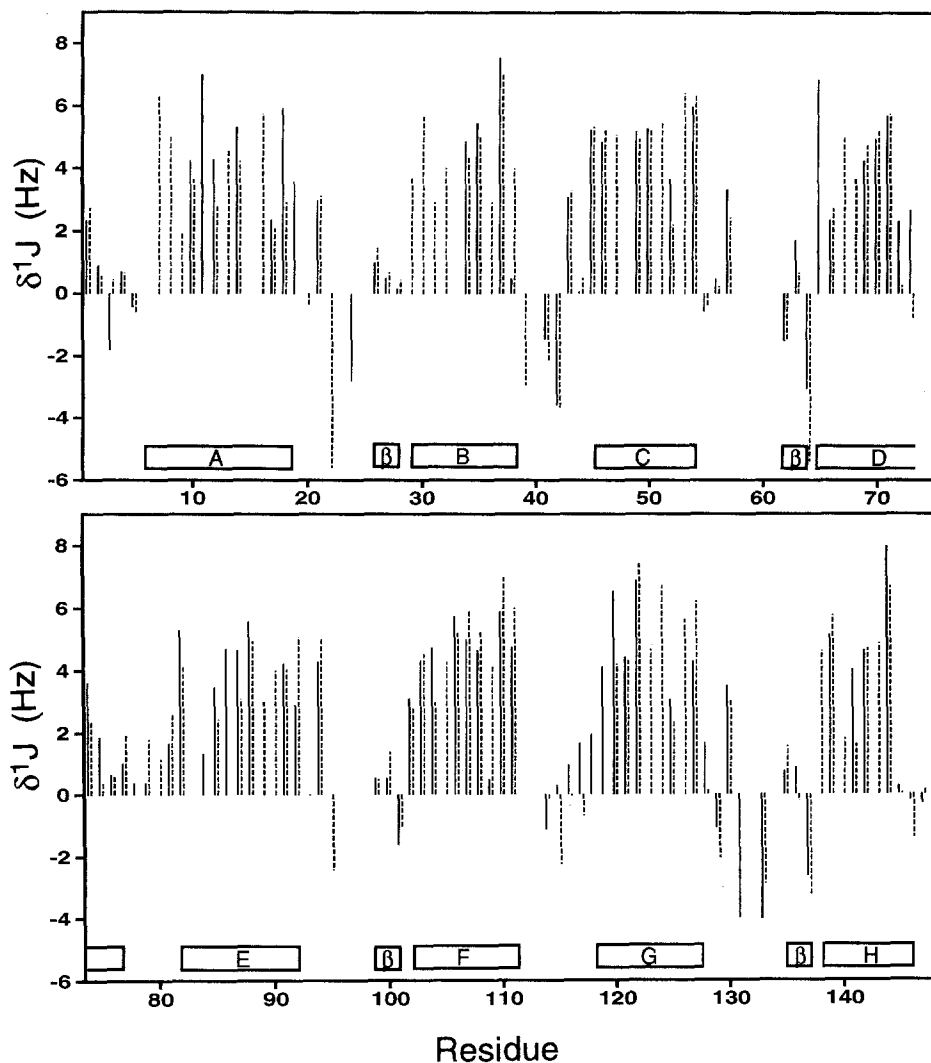
Fig. 5. Measured $\delta^1 J$ values for CaM (solid lines) and the CaM/M13 complex (broken lines). Indicated are the secondary structural elements for CaM as inferred from previous NMR studies (Ikura et al., 1991b); A–H denote α-helices, β denotes anti-parallel β-sheet.

CONCLUSIONS

For ^{13}C-labeled proteins, one-bond $J_{C\alpha H\alpha}$ couplings can be measured accurately from the above described pseudo-3D experiments. Non-linear least-squares fitting routines allow the $^1J_{C\alpha H\alpha}$ coupling to be extracted with an estimated error of less than 0.5 Hz for non-overlapping H^α-C^α and H^α-CO correlations, which is at least twofold better than can be obtained from direct measurements of the multiplet splitting in the 3D spectrum.

The experimental $\Delta^1 J(\varphi, \psi)$ distribution calculated from 203 experimental couplings shows the dependence of $\delta^1 J$ on the dihedral angles, φ and ψ. For the case where the data of amino acid type

j are not used for calculating the shape of the surface, the r.m.s.d. between the surface and the measured values has an average value of 1.98 Hz. This is almost identical to the r.m.s.d. of 1.93 Hz, obtained by fitting $\delta^1 J$ values to Eq. 4, yielding A = 1.7, B = 1.4, C = −4.1 and D = 1.7. This indicates that Eq. 4 provides an adequate description of $\delta^1 J$. This remaining r.m.s.d. is significantly larger than the uncertainty in the measurement, and factors other than φ and ψ clearly have a significant influence on the $^1J_{C\alpha H\alpha}$ value. These other effects may include the presence of solvent and electric fields on the C^α–H^α binding orbitals.

ACKNOWLEDGEMENTS

REFERENCES

Babu, Y.S., Bugg, C.E. and Cook, W.J. (1988) *J. Mol. Biol.*, **204**, 191–204.
Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W. and Bax, A. (1992) *Biochemistry*, **31**, 5269–5278.
Barfield, M. and Johnston, M.D. (1973) *Chem. Rev.*, **73**, 53–73.
Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) *J. Magn. Reson.*, **55**, 301–315.
Bystrov, V.F. (1976) *Prog. NMR Spectrosc.*, **10**, 44–81.
Chary, K.V., Otting, G. and Wüthrich, K. (1991) *J. Magn. Reson.*, **93**, 218–224.
Delaglio, F., Torchia, D. and Bax, A. (1991) *J. Biomol. NMR*, **1**, 439–446.
Edison, A.S., Westler, W.M. and Markley, J. (1991) *J. Magn. Reson.*, **92**, 434–438.
Egli, H. and von Philipsborn, W. (1981) *Helv. Chim. Acta*, **64**, 976–988.
Gil, V.M.S. and von Philipsborn, W. (1989) *Magn. Reson. Chem.*, **27**, 409–430.
Griesinger, C. and Eggenberger, U. (1992) *J. Magn. Reson.*, **97**, 426–434.
Griesinger, C., Sørensen, O.W. and Ernst, R.R. (1986) *J. Chem. Phys.*, **85**, 6837–6843.
Hansen, P.E. (1981) *Prog. NMR Spectrosc.*, **14**, 175–296.
Hansen, P.E. (1991) *Biochemistry*, **30**, 10457–10466.
Ikura, M., Kay, L.E. and Bax, A. (1990) *Biochemistry*, **29**, 4659–4667.
Ikura, M., Kay, L.E., Krinks, M. and Bax, A. (1991a) *Biochemistry*, **30**, 5498–5504.
Ikura, M., Spera, S., Barbato, G., Kay, L.E., Krinks, M. and Bax, A. (1991b) *Biochemistry*, **30**, 9216–9228.
Ikura, M., Clore, G.M., Gronenborn, A.M., Klee, C.B., Zhu, G. and Bax, A. (1992) *Science*, **256**, 632–638.
Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.
Loll, P.J. and Lattman, E.E. (1989) *Proteins: Struct. Funct. Genet.*, **5**, 183–201.
Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.
Messerle, B.A., Wider, G., Otting, G., Weber, C. and W#uthrich, K. (1989) *J. Magn. Reson.*, **85**, 608–613.
Montelione, G.T., Winkler, M.E., Rauenbühler, P. and Wagner, G. (1989) *J. Magn. Reson.*, **82**, 198–204.
Palmer III, A.G, Fairbrother, W., Cavanagh, J., Wright, P.E. and Rance, M. (1992) *J. Biomol. NMR*, **2**, 103–108.
Powers, R., Gronenborn, A.M., Clore, G.M. and Bax, A. (1991) *J. Magn. Reson.*, **94**, 209–213.
Santoro, J. and King, G.C. (1992) *J. Magn. Reson.*, **97**, 202–207.
Shaka, A.J., Keeler, J., Frenkiel, T. and Freeman, R. (1983a) *J. Magn. Reson.*, **52**, 335–338.
Shaka, A.J., Baker, P. and Freeman, R. (1983b) *J. Magn. Reson.*, **53**, 313–340.
Sørensen, O.W., Eich, G.W., Levitt, M.H., Bodenhausen, G. and Ernst, R.R. (1983) *Prog. NMR Spectrosc.*, **16**, 163–192.
Sørensen, O.W. and Ernst, R.R. (1983) *J. Magn. Reson.*, **51**, 477–489.
Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.
Van de Ven, F.J.M. and Philippens, M.E.P. (1992) *J. Magn. Reson.*, **97**, 637–644.

Vuister, G.W. and Bax, A. (1992a) *J. Biomol. NMR,* **2,** 401–405.

Vuister, G.W. and Bax, A. (1992b) *J. Magn. Reson.,* **98,** 428–435.

Vuister, G.W., Delaglio, F. and Bax, A. (1992) *J. Am. Chem. Soc.,* in press.

Wagner, G. and Brühwiler, D. (1986) *Biochem stry,* **25,** 5839–5843.

Wider, G., Neri, D., Otting, G. and Wüthrich, K. (1989) *J. Magn. Reson.,* **85,** 426–431.

Wishart, D.S., Sykes, B.D., Richards, F.M. (1991) *J. Mol. Biol.,* **222,** 311–333.

Wlodawer, A., Deisenhofer, J. and Huber, R. (1987) *J. Mol. Biol.,* **193,** 145–156.