

An Efficient Experiment for Sequential Backbone Assignment of Medium-Sized Isotopically Enriched Proteins

STEPHAN GRZESIEK* AND AD BAX

Laboratory of Chemical Physics, National Institutes of Diabetes and Digestive and Kidney Diseases,
National Institutes of Health, Bethesda, Maryland 20892

Received May 1, 1992

Recently proposed triple-resonance experiments (1-14) make it possible to obtain complete backbone assignments of isotopically enriched proteins without recourse to NOE information. A main practical problem with the triple-resonance assignment approach is the necessity to record a substantial number of 3D spectra, each requiring typically at least one or two days of measuring time. A 4D triple-resonance approach has been proposed to remove the need for recording more than just two 4D experiments for completing the backbone assignments (11). However, if there is overlap in the $^1\text{H}_\alpha$ - $^{13}\text{C}_\alpha$ region of the 2D spectrum, these and other (12) 4D experiments also do not offer unique connections between sequential amides. In cases of such overlap, an experiment that relays magnetization from the amide to the intraresidue carbonyl can resolve the ambiguity caused by the C_α - H_α degeneracy (13). Alternatively, an experiment named CBCA(CO)NH, which correlates amide signals with the C_α and C_β resonances of the preceding residue by using relay via the intervening ^{13}C nucleus, provides information on the type of amino acid preceding each amide and thus frequently resolves the ambiguity caused by C_α - H_α degeneracy (14). Here we describe a similar experiment, named CBCANH, which correlates the amide ^1H and ^{15}N signals with both the intraresidue C_α and C_β signals and with the C_α and C_β of the preceding residue. For small- and medium-sized proteins, up to $\sim 20\text{kDa}$, this 3D experiment alone can yield virtually complete assignments of the backbone and the C_β resonances.

The pulse scheme for the CBCANH experiment is sketched in Fig. 1. Before describing the experiment in more detail, we first outline its general mechanism. An INEPT transfer enhances polarization of both C_α and C_β spins. Chemical shifts of both C_α and C_β are encoded during the constant-time evolution period, $2T_{\text{AB}}$ (7, 8, 10, 14). When the 90° ^{13}C pulse is applied, at time b in Fig. 1, C_β magnetization is transferred to C_α , but at the same time a fraction of C_α magnetization remains on C_α . Subsequently, C_α magnetization is transferred to its intraresidue ^{15}N and to the ^{15}N of the next residue, via the $^1J_{\text{C}_\alpha\text{N}}$ and $^2J_{\text{C}_\alpha\text{N}}$ couplings, respectively. Following a constant-time ^{15}N evolution period, the ^{15}N magnetization is transferred to the amide protons with a reverse INEPT sequence, prior to detection of the H_N signals.

* On leave from F. Hoffmann-LaRoche Ltd., Basel, Switzerland.

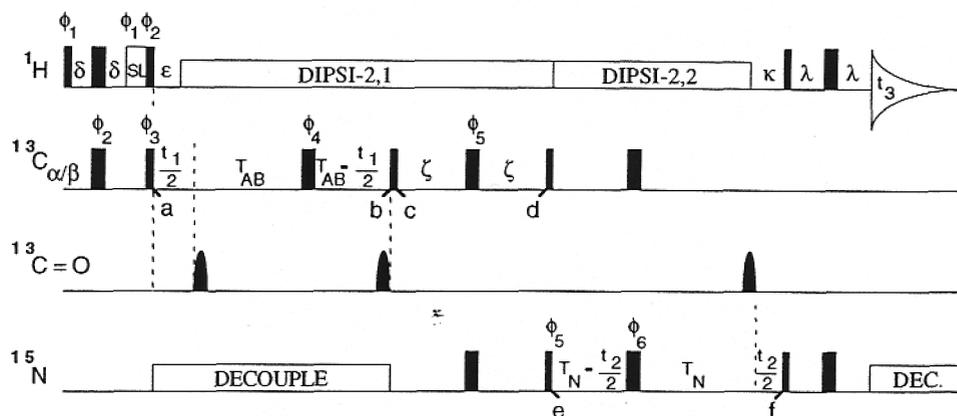


FIG. 1. Pulse scheme of the CBCANH experiment. Narrow and wide pulses correspond to 90° and 180° flip angles, respectively. Pulses for which the phase is not indicated are applied along the x axis. The ^1H carrier is set to the H_2O frequency for the first part of the pulse sequence, up to time d , and switched to the center of the amide region (8.4 ppm) thereafter. Thus, the frequency of the ^1H broadband decoupling (with a 5 kHz RF field) is switched at this point, and the two modes of decoupling, using a coherent DIPSI-2 scheme (22), are marked DIPSI-2,1 and DIPSI-2,2. ^{15}N decoupling is accomplished using WALTZ-16 modulation with a 1.5 kHz RF field. Carbonyl pulses have a shaped amplitude profile, corresponding to the center lobe of a $\sin x/x$ function and a duration of 202 μs . The carrier for the $^{180^\circ} \text{C}_{\alpha/\beta}$ pulses is positioned at 46 ppm, and the power of the $^{180^\circ} \text{C}_{\alpha/\beta}$ pulses is adjusted such that they do not excite the ^{13}CO nuclei (11.5 kHz RF field for 150.9 MHz ^{13}C frequency). Phase cycling is as follows: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = x$; $\phi_4 = 8(x), 8(y), 8(-x), 8(-y)$; $\phi_5 = 2(x), 2(-x)$; $\phi_6 = 4(x), 4(-x)$; Acq. = $x, 2(-x), 2(x), 2(-x), x, -x, 2(x), 2(-x), 2(x), -x$. Quadrature in the t_1 and t_2 domains is obtained by changing the phases ϕ_3 and ϕ_5 , respectively, in the usual States-TPPI manner (14). Delay durations are $\delta = 1.5$ ms; $\epsilon = 2.1$ ms; $T_{AB} = 3.3$ ms; $\zeta = 11$ ms; $T_N = 11.2$ ms; $\kappa = 5.4$ ms; $\lambda = 2.25$ ms.

The relevant terms of the magnetization-transfer pathway can be conveniently expressed in terms of the product-operator formalism (15). For clarity, relaxation terms are not included and constant multiplicative factors are omitted. Only terms that result in observable magnetization during the detection period, t_3 , are retained. The spin operators used are B for C_β , A for C_α , N_1 for the intrasidue ^{15}N spin, and N_2 for the ^{15}N spin of the succeeding residue. The two amide protons attached to these two ^{15}N nuclei are labeled H_1 and H_2 , respectively. The resonance offset of a nucleus, X , is denoted δ_X .

The first INEPT part of the pulse sequence, which includes a water-purge pulsed labeled SL (16), transfers ^1H magnetization into antiphase $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ magnetization (time a). Refocusing of this antiphase magnetization follows a different time dependence for methine, methylene, and methyl carbons. A compromise value of 2.1 ms (17) was used for this refocusing delay, ϵ , in the scheme in Fig. 1.

In the $\text{C}_\alpha/\text{C}_\beta$ constant-time evolution period (between points a and b), of total duration $2T_{AB} = 6.6$ ms, the effect of J coupling to carbonyl carbons is eliminated by applying a selective ^{13}CO pulse at time $t_1/2$ after the $^{13}\text{C}_{\phi_3}$ pulse and by adjusting the power of the $^{180^\circ}_{\phi_4}$ pulse in such a way that it does not excite the ^{13}CO resonances. During this period, evolution of the transverse C_α and C_β magnetization, refocused with respect to its attached proton(s), is then described by

$$A_x \rightarrow A_x \cos^m(2\pi J_{AB}T_{AB}) \cos(2\pi\delta_A t_1) \quad [1a]$$

$$B_x \rightarrow B_y A_z \sin(2\pi J_{AB}T_{AB}) \cos^n(2\pi J_{BC}T_{AB}) \cos(2\pi\delta_B t_1), \quad [1b]$$

where J_{AB} and J_{BC} are the C_α - C_β and C_β - C_γ one-bond J couplings, and the exponent m is 0 for glycine and 1 for all other residues. The exponent n is 2 for valine and isoleucine, 0 for alanine, aspartic acid, asparagine, and serine, and 1 for all other residues. The second shaped ^{13}CO pulse, applied just before time b, merely serves to eliminate the phase error induced by a Bloch-Siegert effect (18) of the first shaped 180° ^{13}CO pulse on the transverse C_α/C_β magnetization, ensuring a pure cosinusoidal modulation in the t_1 dimension. Temporarily omitting the trigonometric terms from Eq.[1], the 90_x° C_α/C_β pulse at point b converts these terms according to

$$A_x \rightarrow A_x \quad [2a]$$

$$B_y A_z \rightarrow B_z A_y. \quad [2b]$$

During the rephasing/dephasing interval, between time points c and d, C_α magnetization of [2b] rephases with respect to spin B, whereas [2a] partially dephases. At the same time, dephasing of C_α magnetization caused by coupling to the intrareidue ^{15}N spin and to the ^{15}N of the next residue takes place. These processes are described by

$$A_x \rightarrow A_y N_{1z} \sin(2\pi J_{N1A}\zeta) \cos(2\pi J_{N2A}\zeta) \cos^m(2\pi J_{AB}\zeta) \\ + A_y N_{2z} \sin(2\pi J_{N2A}\zeta) \cos(2\pi J_{N1A}\zeta) \cos^m(2\pi J_{AB}\zeta) \quad [3a]$$

$$B_z A_y \rightarrow A_y N_{1z} \sin(2\pi J_{N1A}\zeta) \cos(2\pi J_{N2A}\zeta) \sin(2\pi J_{AB}\zeta) \\ + A_y N_{2z} \sin(2\pi J_{N2A}\zeta) \cos(2\pi J_{N1A}\zeta) \sin(2\pi J_{AB}\zeta). \quad [3b]$$

Apart from the trigonometric coefficients, terms originating from C_α and C_β magnetization are indistinguishable at this time, and we need to consider only the fate of either $A_y N_{1z}$ or $A_y N_{2z}$ to describe the remainder of the pulse sequence.

At time d the pair of simultaneous 90° $^{13}\text{C}_\alpha/^{15}\text{N}$ pulses transforms the antiphase C_α magnetization according to

$$A_y N_{1z} \rightarrow A_z N_{1y} \quad [4a]$$

$$A_y N_{2z} \rightarrow A_z N_{2y}. \quad [4b]$$

During the subsequent ^{15}N constant-time evolution period, of total duration $2T_N$, the $A_z N_{1y}$ term rephases with respect to its C_α coupling partner, but dephases because of the coupling via $^2J_{N1C_\alpha}$ to its preceding C_α (note that this C_α is not the spin labeled A, but the C_α of the preceding residue). The effect of the ^{13}CO - ^{15}N J coupling is eliminated during the first $2T_N - t_2$ fraction of the ^{15}N evolution period by the 180° ^{15}N pulse and by the fact that the 180° C_α/C_β pulse is adjusted such that it does not excite the carbonyl spins. During the last fraction, t_2 , of the constant-time ^{15}N evolution period, the shaped 180° ^{13}CO pulse decouples the $^{15}\text{N}/^{13}\text{CO}$ interaction and evolution caused by the ^{15}N chemical shift takes place. During the delay, κ , ^1H decoupling is switched off and ^{15}N magnetization becomes antiphase with respect to its attached proton spin. These effects are all summarized by

$$N_{1y}A_z \rightarrow N_{1y}H_{1z}\sin(2\pi J_{N1A}T_N)\cos(2\pi J_{N1C\alpha}T_N)\sin(\pi J_{N1H1}\kappa)\cos(2\pi\delta_{N1}t_2) \quad [5a]$$

$$N_{2y}A_z \rightarrow N_{2y}H_{2z}\sin(2\pi J_{N2A}T_N)\cos(2\pi J_{N2C\alpha}T_N)\sin(\pi J_{N2H2}\kappa)\cos(2\pi\delta_{N2}t_2). \quad [5b]$$

$^1J_{NH}$ is quite uniform in proteins and typically falls in the 92–94 Hz range. With $\kappa = 5.4$ ms, the $\sin(\pi J_{N1H1}\kappa)$ and $\sin(\pi J_{N2H2}\kappa)$ terms in Eq.[5] are very close to 1 and therefore can be safely omitted. $J_{N1C\alpha}$ is the two-bond coupling between N_1 and the C_α of the preceding residue and $J_{N2C\alpha}$ is the one-bond coupling between N_2 and its intraresidue C_α .

The final reverse INEPT, applied following time f , transforms $N_{1y}H_{1z}$ into H_{1x} and $N_{2y}H_{2z}$ into H_{2x} , which are both observed during the detection period, t_3 . As is clear from the above, for a given amide proton, the CBCANH experiment yields correlations to its intraresidue C_α and C_β nuclei and to the C_α and C_β of the preceding residue. Neglecting losses during the first INEPT and the final reverse INEPT, the amplitudes of the four corresponding cross peaks are obtained by combining the trigonometric terms of Eqs. [1], [3], and [5]:

$$A \rightarrow H_{1x}\cos^m(2\pi J_{AB}T_{AB})\sin(2\pi J_{N1A}\zeta)\cos(2\pi J_{N2A}\zeta)\cos^m(2\pi J_{AB}\zeta)\sin(2\pi J_{N1A}T_N) \\ \times \cos(2\pi J_{N1C\alpha}T_N)\cos(2\pi\delta_{A1}t_1)\cos(2\pi\delta_{N1}t_2) \quad [6a]$$

$$A \rightarrow H_{2x}\cos^m(2\pi J_{AB}T_{AB})\sin(2\pi J_{N2A}\zeta)\cos(2\pi J_{N1A}\zeta)\cos^m(2\pi J_{AB}\zeta)\sin(2\pi J_{N2A}T_N) \\ \times \cos(2\pi J_{N2C\alpha}T_N)\cos(2\pi\delta_{A1}t_1)\cos(2\pi\delta_{N2}t_2) \quad [6b]$$

$$B \rightarrow H_{1x}\sin(2\pi J_{AB}T_{AB})\cos^n(2\pi J_{BC}T_{AB})\sin(2\pi J_{N1A}\zeta)\cos(2\pi J_{N2A}\zeta)\sin(2\pi J_{AB}\zeta) \\ \times \sin(2\pi J_{N1A}T_N)\cos(2\pi J_{N1C\alpha}T_N)\cos(2\pi\delta_{B1}t_1)\cos(2\pi\delta_{N1}t_2) \quad [6c]$$

$$B \rightarrow H_{2x}\sin(2\pi J_{AB}T_{AB})\cos^n(2\pi J_{BC}T_{AB})\sin(2\pi J_{N2A}\zeta)\cos(2\pi J_{N1A}\zeta)\sin(2\pi J_{AB}\zeta) \\ \times \sin(2\pi J_{N2A}T_N)\cos(2\pi J_{N2C\alpha}T_N)\cos(2\pi\delta_{B1}t_1)\cos(2\pi\delta_{N2}t_2). \quad [6d]$$

$J_{C\alpha C\beta}$ couplings fall in the 35–40 Hz range, $^1J_{NC\alpha}$ falls in the 9–12.5 Hz range, and $J_{NC\alpha}$ falls in the 6–9 Hz range (19). Based on these values, Eq.[6] indicates that significant intensities are expected for all four correlations if $T_{AB} = 3.3$ ms, $\zeta = 11$ ms, and $T_N = 11$ ms. If $T_N = 11$ ms does not allow for a sufficiently long ^{15}N constant-time evolution period, i.e., the ^{15}N resolution obtainable with $T_N = 11$ ms is lower than required, a value that is r ms longer can be used by applying the $^{13}\text{C}_{\alpha/\beta}$ pulse r ms after the $180^\circ_{\phi_6}$ ^{15}N pulse.

The CBCANH experiment is demonstrated for a sample containing 1.5 mM calmodulin, saturated with Ca^{2+} and complexed with a 26-residue peptide fragment of skeletal muscle myosin light chain kinase. Total molecular weight of the complex is 20 kDa. Experiments are recorded at 600 MHz ^1H frequency on a Bruker AMX-600, operating at 35°C. A 3D time-domain matrix consisting of $52^* (t_1, ^{13}\text{C}) \times 32^* (t_2, ^{15}\text{N}) \times 512^* (t_3, ^1\text{H})$ data points was recorded, where N^* refers to N complex data points. Acquisition times were 6.16 ms (t_1), 21.12 ms (t_2), and 55.3 ms (t_3). For each pair of t_1/t_2 values 192 scans were recorded, and the total acquisition time was ~93 hours. Mirror image linear prediction (20) was used to extend the time domain data to 128^* in the t_1 dimension and to 64^* in the t_2 dimension, prior to cosine-squared bell apodization in the t_1 and t_2 dimensions and Fourier transformation. The absorptive part of the final 3D matrix corresponds to $256 \times 128 \times 1024$ data points.

Figure 2 shows strips taken through the 3D CBCANH spectrum for residues Glu54–Phe65. Each strip represents a narrow F_3 region of an F_1/F_3 cross section of the 3D spectrum at the $^1\text{H}/^{15}\text{N}$ frequency of an amino acid. Thus, each strip displays the C_α and C_β correlations observed for the amide of a given residue. As can be seen in Fig. 2, except for the two glycine residues and the two residues following these glycines, all amides show the four expected correlations to both the intraresidue C_α and C_β and to the C_α and C_β of the preceding residue. These latter correlations are marked by solid dots and can be tentatively differentiated from the intraresidue ones by their weaker intensities (caused by the smaller value of $^2J_{\text{NC}\alpha}$ compared to that of $^1J_{\text{NC}\alpha}$), or can be unambiguously identified using the CBCA(CO)NH relay experiment, described elsewhere (14).

C_α resonances of glycines and C_β resonances of all other residues are opposite in phase relative to the other C_α correlations in Fig. 2. This is caused by the fact that $\cos(2\pi J_{\text{AB}}\zeta)$ in Eq.[6] is negative and opposite in sign to $\sin(2\pi J_{\text{AB}}\zeta)$. This feature is useful for discriminating between C_α and C_β resonances of serine residues, which can resonate in the same spectral region.

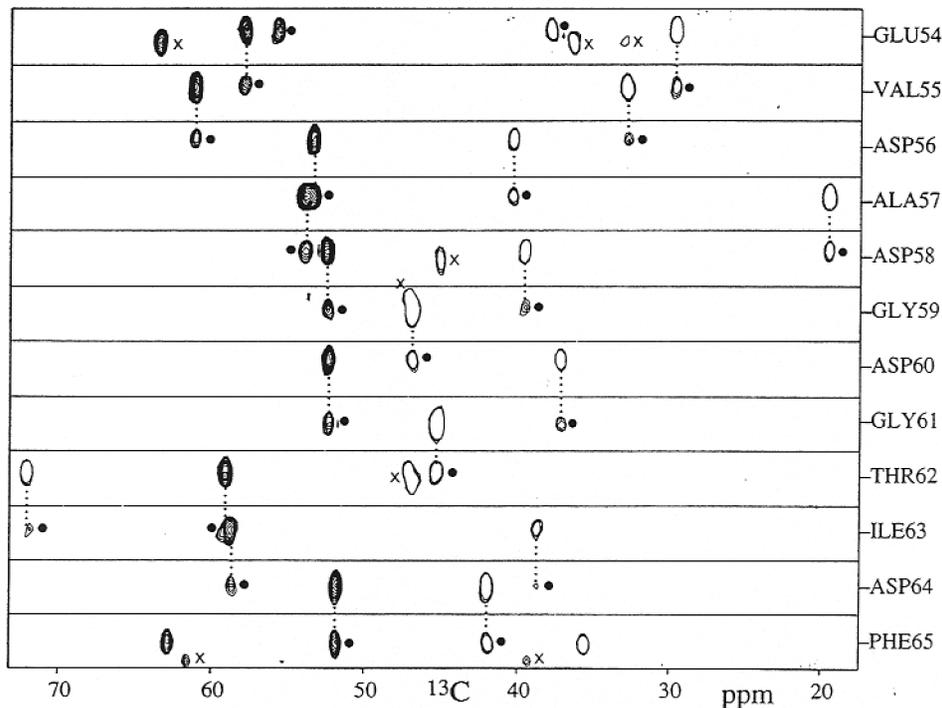


FIG. 2. Strip plot of the correlations observed for the amides of residues Glu54–Phe65 of a calmodulin-peptide complex. Each amide correlates with its intraresidue C_α and C_β nuclei and with the C_α and C_β of the preceding residue. Interresidue connectivities are marked by solid dots. Negative peaks, corresponding to C_β resonances and to C_α of glycine residues, are displayed with a maximum of three contours. Peaks marked “x” correspond to correlations to amide ^1H – ^{15}N pairs that are close in frequency to that for which the strip has been selected.

The CBCANH experiment is particularly powerful for making sequential backbone and C_β assignments in isotopically enriched proteins because two independent connectivity pathways, marked by dotted lines in Fig. 2, are available in the same spectrum. As chemical-shift dispersion, on average, is even better for C_β than for C_α , the CBCANH spectrum yields a large number of unambiguous connectivities between sequential amino acids. Moreover, the unique C_α and C_β shifts of glycine, threonine, serine, and alanine residues immediately identify the amides of these residues. For other amides, combined knowledge of both the intraresidue C_α and C_β shifts frequently narrows down the possible residue types to only a small number. This extensive knowledge of residue type generally will solve ambiguities that may occur for two residues with identical C_α chemical shifts and degenerate C_β shifts. Therefore, the CBCANH spectrum alone, or combined with the CBCA(CO)NH spectrum, provides sufficient information for making complete backbone assignments of isotopically enriched proteins. Analysis of the CBCANH spectrum indicated that residues Glu83 and Glu84 as well as Glu7 and Glu14 had been interchanged in the original assignment (21).

In contrast to the CBCA(CO)NH experiment, which has been applied successfully to a 31 kDa protein (14), a major limitation of the CBCANH experiment is its relatively low sensitivity, which is affected most strongly by the transverse relaxation of C_α magnetization during the relatively long dephasing period, 2ζ . For this reason, we expect that the CBCANH experiment will rapidly decrease in effectiveness for proteins larger than ~ 20 kDa. For Gln3 through Lys148 of the 20 kDa calmodulin-peptide complex, only two $H_N-C_{\alpha,i}$, six $H_N-C_{\alpha,i-1}$, no $C_{\beta,i}$, and eight $C_{\beta,i-1}$ correlations were below the noise threshold. For isotopically enriched proteins of up to ~ 20 kDa, the CBCANH experiment is therefore probably the most effective single method for obtaining sequential resonance assignments published to date.

ACKNOWLEDGMENTS

This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health. S.G. acknowledges funding by the ROCHE Research Foundation.

REFERENCES

1. M. IKURA, L. E. KAY, AND A. BAX, *Biochemistry* **29**, 4659 (1990).
2. W. P. NIEMCZURA, G. L. HELMS, A. S. CHESNICK, R. E. MOORE, AND V. BORNEMANN, *J. Magn. Reson.* **81**, 635 (1989).
3. L. E. KAY, M. IKURA, R. TSCHUDIN, AND A. BAX, *J. Magn. Reson.* **89**, 596 (1990).
4. B. J. STOCKMAN, M. D. REILLY, W. M. WESTLER, E. L. ULRICH, AND J. L. MARKLEY, *Biochemistry* **28**, 230 (1989).
5. L. E. KAY, M. IKURA, AND A. BAX, *J. Magn. Reson.* **91**, 84 (1991).
6. A. BAX AND M. IKURA, *J. Biomol. NMR* **1**, 99 (1991).
7. R. POWERS, A. M. GRONENBORN, G. M. CLORE, AND A. BAX, *J. Magn. Reson.* **94**, 209 (1991).
8. A. G. PALMER III, W. J. FAIRBROTHER, J. CAVANAGH, P. E. WRIGHT, AND M. RANCE, *J. Biomol. NMR* **2**, 103 (1992).
9. B. T. FARMER II, R. A. VENTERS, L. D. SPICER, M. G. WITTEKIND, AND L. MÜLLER, *J. Biomol. NMR* **2**, 195 (1992).
10. S. GRZESIEK AND A. BAX, *J. Magn. Reson.* **96**, 432 (1992).
11. W. BOUCHER, E. D. LAUE, S. CAMPBELL-BURK, AND P. J. DOMAILLE, *J. Am. Chem. Soc.* **114**, 2262 (1992).
12. L. E. KAY, M. IKURA, G. ZHU, AND A. BAX, *J. Magn. Reson.* **91**, 422 (1991).
13. R. T. CLUBB, V. THANABAL, AND G. WAGNER, *J. Magn. Reson.* **97**, 213 (1992).

14. S. GRZESIEK AND A. BAX, *J. Am. Chem. Soc.*, submitted.
15. R. R. ERNST, G. BODENHAUSEN, AND A. WOKAUN, "Principles of Nuclear Magnetic Resonance in One and Two Dimensions," pp. 25-29, Clarendon Press, Oxford, 1987.
16. B. A. MESSERLE, G. WIDER, G. OTTING, C. WEBER, AND K. WÜTHRICH, *J. Magn. Reson.* **85**, 608 (1989).
17. D. P. BURUM AND R. R. ERNST, *J. Magn. Reson.* **39**, 163 (1980).
18. R. FREEMAN, "A Handbook of Nuclear Magnetic Resonance," p. 14, Wiley, New York, 1988.
19. F. DELAGLIO, D. A. TORCHIA, AND A. BAX, *J. Biomol. NMR* **1**, 439 (1991).
20. G. ZHU AND A. BAX, *J. Magn. Reson.* **90**, 405 (1990).
21. M. IKURA, L. E. KAY, M. KRINKS, AND A. BAX, *Biochemistry* **30**, 5498 (1991).
22. A. J. SHAKA, C. J. LEE, AND A. PINES, *J. Magn. Reson.* **77**, 274 (1988).