Relative Orientation of $C^{\alpha}H^{\alpha}$-Bond Vectors of Successive Residues in Proteins through Cross-Correlated Relaxation in NMR

Elisabetta Chiarparin,[†] Philippe Pelupessy,[†] Ranajeet Ghose,[†] and Geoffrey Bodenhausen*[,†,‡]

Contribution from the Section de Chimie, Université de Lausanne, BCH, 1015 Lausanne, Switzerland, and Département de Chimie, associé au CNRS, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France

Received September 13, 1999. Revised Manuscript Received December 14, 1999

Abstract: Cross-correlation between the fluctuations of $^{13}C^{\alpha}-H^{\alpha}$ interactions affects the relaxation behavior of two-spin coherences (zero- and double-quantum coherences) involving the $^{13}C^{\alpha}$ nuclei of two successive amino acid residues. The cross-correlation rates are shown to depend on a dihedral angle Σ defined by two planes subtended by the atoms $\{H^{\alpha}(i-1),^{13}C^{\alpha}(i-1),^{13}C^{\alpha}(i)\}$ and $\{^{13}C^{\alpha}(i-1),^{13}C^{\alpha}(i),H^{\alpha}(i)\}$. This dihedral angle is related to the secondary structure of a protein, and can be used as a constraint in various protein structure calculation protocols.

Introduction

Nuclear magnetic resonance (NMR) has proven to be a reliable method to determine the structure of proteins in solution.[1] Nuclear Overhauser effects (NOE's) provide a measure of the spatial separation between two nuclei,[2] but the scarcity of long-range NOE's generally makes it difficult to completely determine the structure of a protein from NOE's alone. Long-range scalar coupling constants can be empirically related to backbone dihedral angles to provide further structural constraints,[3−5] although there is a lack of information about the backbone ψ angle. Residual dipolar couplings in dilute liquid-crystalline environments (bicelles or suspensions of bacteriophage) yield the orientation of bond vectors with respect to a common director, and hence with respect to each other.[6,7] Interference effects between two relaxation mechanisms furnish a direct measure of backbone dihedral angles in proteins. Reif et al.[8] used cross-correlation between the $^{15}N-H^{N}$ and $^{13}C^{\alpha}-H^{\alpha}$ dipolar interactions to estimate hitherto inaccessible ψ angles. Yang et al.[9−11] demonstrated that this angle can also be obtained from cross-correlation between the chemical shift anisotropy (CSA) of carbonyl $^{13}C'$ and the $^{13}C^{\alpha}-H^{\alpha}$ dipolar interaction. These methods were extended by our group.[12,13] We have also shown that the backbone angle ϕ can be estimated in a similar way.[14]

In this paper, cross-correlation between pairs of $^{13}C^{\alpha}-H^{\alpha}$ dipolar interactions is measured to estimate the angle θ subtended between two consecutive $^{13}C^{\alpha}-H^{\alpha}$ bond vectors in a protein. This angle depends on a dihedral angle, which we call Σ, which is defined by two planes subtended by the atoms $\{H^{\alpha}(i-1),^{13}C^{\alpha}(i-1),^{13}C^{\alpha}(i)\}$ and $\{^{13}C^{\alpha}(i-1),^{13}C^{\alpha}(i),H^{\alpha}(i)\}$. The angle Σ is related to the secondary structure, and information about Σ allows one to refine structure calculation protocols.

Theory and Design of Experiments

Assuming overall isotropic tumbling, the cross-correlation rate $R(\theta)$ is given by

$$R(\theta) = \left(\frac{\mu_0\hbar}{4\pi}\right)^2 \frac{\gamma_H{}^2\,\gamma_C{}^2}{r_{CH}{}^6} \frac{(3\cos^2\theta - 1)}{2} \frac{2S^2\tau_C}{5} \qquad (1)$$

where θ is the angle subtended between successive $^{13}C^{\alpha}-H^{\alpha}$ bond vectors, r_{CH} is the $C^{\alpha}H^{\alpha}$ internuclear distance (assumed to be 1.12 Å),[15] S^2 is the generalized Lipari−Szabo order parameter, τ_C is the overall correlation time for molecular rotational diffusion, and all the other symbols have their usual meaning.

A pulse sequence which allows one to determine the angle θ is depicted in Figure 1. The magnetization of the amide proton H^N is transferred to ^{15}N and then to the $^{13}C^{\alpha}$'s of the same (i) and of the previous ($i-1$) residue in the fashion of an HNCA experiment,[16] utilizing $^1J(NC^{\alpha}) \sim 7-11$ Hz and $^2J(C^{\alpha}N) \sim 4-9$ Hz. At this stage of the sequence, there are four relevant components of the density operator: $N_x(i)$, $2N_y(i)C^{\alpha}{}_z(i)$, $2C^{\alpha}{}_z(i-1)N_y(i)$, and $4C^{\alpha}{}_z(i-1)N_x(i)C^{\alpha}{}_z(i)$. To maximize the amplitude of the last term a transfer delay $2\tau_3 = 50$ ms is chosen, so that $\sin[\pi^1J(NC^{\alpha})2\tau_3]\ \sin[\pi^2J(C^{\alpha}N)2\tau_3] \approx 0.8-0.9$. The offspring of the second and third operators is eliminated by alternating ϕ_3 while keeping the receiver phase constant. For the sake of simplicity, we will omit the indices in the following, e.g.

* Address correspondence to this author.
† Université de Lausanne.
‡ Ecole Normale Supérieure.

(1) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(2) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Oxford University Press: Oxford, 1987.
(3) Karplus, M. *J. Phys. Chem.* **1959**, *30*, 11.
(4) Wang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 2483.
(5) Branden, C.; Twose, J. *Introduction to Protein Structure*, 2nd ed.; Garland Publishing: New York, 1999.
(6) Bax, A.; Tjandra, N. *J. Biomol. NMR* **1997**, *10*, 289.
(7) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111.
(8) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230.
(9) Yang, D.; Konrat, R.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 11938.
(10) Yang, D.; Gardner, K.; Kay, L. E. *J. Biomol. NMR* **1998**, *11*, 213.
(11) Yang, D.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 9880.
(12) Pelupessy, P.; Chiarparin, E.; Ghose R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *13*, 375.
(13) Chiarparin, E.; Pelupessy, P.; Ghose R.; Bodenhausen, G. *J. Am. Chem. Soc.* **1999,** *21*, 6876.
(14) Pelupessy, P.; Chiarparin, E.; Ghose R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *14*, 277.
(15) Ottiger, M.; Bax, A *J. Am. Chem. Soc.* **1998**, *120*, 12334.
(16) Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1990**, *89*, 496.

10.1021/ja9933184 CCC: $19.00
Published on Web 02/08/2000

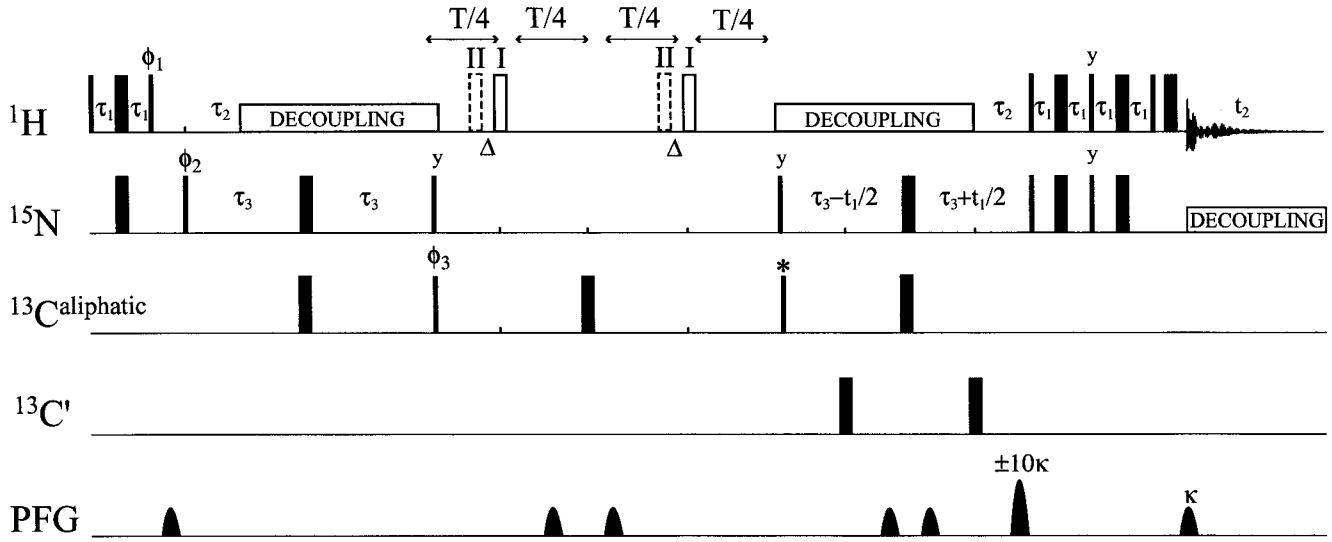

Figure 1. Pulse sequences employed for the measurement of cross-correlation rates $R(\Sigma)$. Narrow and wide rectangles indicate $\pi/2$ and π pulses. The fixed delays are set to $\tau_1 = 2.5$ ms, $\tau_2 = 5.4$ ms, $\tau_3 = 25$ ms, and $\Delta = [8J(C^\alpha H^\alpha)]^{-1} = 0.89$ ms. The bold-outlined rectangles in experiment I are placed so that the scalar couplings are refocused. In experiment II these pulses (dashed rectangles) are shifted to allow evolution under $J(C^\alpha H^\alpha)$. The relaxation delay T must be either very small or a multiple of $1/J(C^\alpha C^\beta)$ to prevent losses due to evolution under this coupling. The evolution in the ^{15}N dimension occurs in t_1 in a constant-time manner during the conversion of $4C^\alpha_z N_x C^\alpha_z$ into $2N_y H^N_z$. Unless specified otherwise, all pulses are applied along the x-axis. The phase cycle is as follows: $\phi_1 = y, -y, y, -y, y, -y, y, -y$; $\phi_2 = x, x, -x, -x, x, x, -x, -x$; $\phi_3 = x, x, x, x, -x, -x, -x, -x$, receiver phase $\phi_{rec} = x, -x, -x, x, x, -x, -x, x$. The pulse marked with an asterisk is applied in the first 8 scans and omitted in the second series of 8 scans, while the receiver phase is reversed between the first and second series of eight scans, to distinguish $4C^\alpha_y N_z C^\alpha_y$ from N_z.

$4C^\alpha_z(i-1)N_x(i)C^\alpha_z(i)$ is simply denoted $4C^\alpha_z N_x C^\alpha_z$. All nuclei placed before the symbol N belong to the previous residue while the nuclei placed after N belong to the same residue. Two-spin coherences are then excited $4C^\alpha_z N_x C^\alpha_z \rightarrow 4C^\alpha_y N_z C^\alpha_y$ by the application of a ^{13}C $(\pi/2)_x$ pulse and ^{15}N $(\pi/2)_y$ pulse. A set of two complementary two-dimensional experiments has been designed to measure the conversion process $4C^\alpha_y N_z C^\alpha_y \rightarrow 16C^\alpha_x H^\alpha_z N_z C^\alpha_x H^\alpha_z$ due to cross-correlation.[12−14,17] Experiment I allows one to measure signals proportional to the initial term $S_I = <4C^\alpha_y N_z C^\alpha_y>$, while experiment II provides signals $S_{II} = <16C^\alpha_x H^\alpha_z N_z C^\alpha_x H^\alpha_z>$. These experiments differ only in the evolution under scalar couplings in the relaxation period T of Figure 1. In experiment I all scalar couplings are refocused, while in experiment II, the two $J(C^\alpha H^\alpha)$ couplings are active during $4\Delta_1 = [2J(C^\alpha H^\alpha)]^{-1}$. Since both ZQ- and DQ-coherences contained in the operator $4C^\alpha_y N_z C^\alpha_y$ have to be retained, phase cycling cannot be used to separate the remaining terms N_z and $4C^\alpha_y N_z C^\alpha_y$. Therefore, the carbon $\pi/2$-pulse (marked with an asterisk in Figure 1) that converts $4C^\alpha_y N_z C^\alpha_y \rightarrow 4C^\alpha_z N_z C^\alpha_z$ is omitted in a second series of scans while the receiver phase is reversed. The cross-correlation rate of eq 1 can be determined from the ratio of the signal intensities:

$$\frac{S_{II}}{S_I} = \frac{\exp(R(\theta)T) - \exp(-R(\theta)T)}{\exp(R(\theta)T) + \exp(-R(\theta)T)} = \tanh[R(\theta)T] \qquad (2)$$

The π pulses are positioned so that the effects of CSA/DD cross-correlations are averaged out in both experiments I and II.[13] Evolution of the ^{15}N chemical shift in a constant-time fashion in t_1 during the reconversion processes $4C^\alpha_z N_x C^\alpha_z \rightarrow 2N_y H^N_z$ allows one to obtain a dispersion of resonances in the manner of heteronuclear single-quantum correlation (HSQC) spectroscopy.

However, there is a complication that might hamper the analysis. The zero- and double-quantum spectra associated with the coherence $4C^\alpha_y N_z C^\alpha_y$ both consist of a doublet-of-doublets split by two different $^1J(C^\alpha H^\alpha)$ couplings. Since the values of these two coupling constants are nearly equal, the two central lines overlap, so that the secular approximation is inappropriate. As a consequence, the cross-correlation rate measured through eq 2 can deviate from $R(\theta)$ of eq 1 by an amount that has a maximum given by the off-diagonal element in the Liouville matrix,[13] which is equal to $^1/_4(R_{ii} - R_{ai} - R_{ia} + R_{aa})$, where R_{ii}, R_{ai}, R_{ia}, and R_{aa} are the auto-relaxation rates of density operator terms $4C_y N_z C_y$, $8C_x H_z N_z C_y$, $8C_y N_z C_x H_z$, and $16C_x H_z N_z C_x H_z$ which are in-phase and antiphase with respect to the $H^\alpha(i-1)$ and $H^\alpha(i)$ protons. To first approximation, these rates can be written as $R_{ia} \approx R_{ii} + 1/T_1\{H^\alpha(i)\}$, $R_{ai} \approx R_{ii} + 1/T_1\{H^\alpha(i-1)\}$, $R_{aa} \approx R_{ii} + 1/T_1\{H^\alpha(i)\} + 1/T_1\{H^\alpha(i-1)\}$, where $1/T_1\{H^\alpha\}$ denotes the spin−lattice relaxation rate of the relevant H^α nucleus due to the proton−proton dipole−dipole interactions. Hence, the deviation from eq 1 is very small.

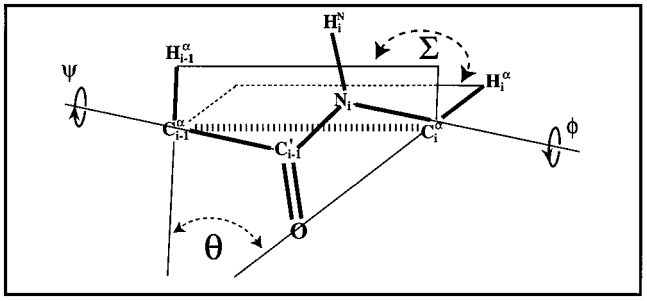

Figure 2. Fragment of the backbone of a protein. The θ angle is subtended between two bond vectors $^{13}C^\alpha(i-1)$-$H^\alpha(i-1)$ and $^{13}C^\alpha(i)$-$H^\alpha(i)$ in successive amino acids. The Σ angle is the dihedral angle between the planes defined by the atoms $\{H^\alpha(i-1),^{13}C^\alpha(i-1),^{13}C^\alpha(i)\}$ and $\{^{13}C^\alpha(i-1),^{13}C^\alpha(i),H^\alpha(i)\}$.

Results and Discussion

The angle θ depends on the dihedral angle Σ between two planes subtended by the atoms $\{H^\alpha(i-1),^{13}C^\alpha(i-1),^{13}C^\alpha(i)\}$ and $\{^{13}C^\alpha(i-1),^{13}C^\alpha(i),H^\alpha(i)\}$. The angles Σ, ψ, and ϕ are drawn in Figure 2. Assuming that the four atoms $H^N(i)$, $N(i)$, $C'(i-1)$, and $O(i-1)$ lie in a plane (i.e. $\omega = 180°$) we can derive the following relationship

$$\cos\theta_{CH,CH} = -0.106 + 0.894\cos(\psi + \phi + 180°) \qquad (3)$$

The angle Σ is given by $[\psi(i-1) + \phi(i)]$ ($\Sigma = 180°$ when the two H^α atoms lie in the same plane as the two C^α atoms

(17) Felli, I. C.; Richter, C.; Griesinger, C.; Schwalbe, H. *J. Am. Chem. Soc.* **1999**, *121*, 1956.

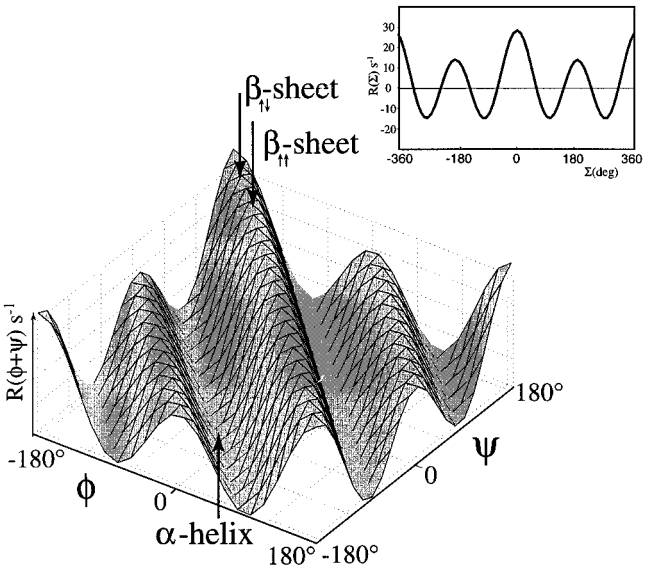

Figure 3. Plot of cross-correlation rate $R(\phi+\psi)$ in eq 1 according to eq 3. The regions of α-helices, parallel β-sheets, and antiparallel β-sheets are indicated. The rates are negative for residues in α-helices and positive for residues in β-sheets. In the top-right corner, a cross-section along the diagonal of the grid plot shows $R(\Sigma) = R(\phi+\psi)$.

and on the same side with respect to the $C^{\alpha}(i-1)$-$C^{\alpha}(i)$ axis). Knowledge of the relative orientation of consecutive $^{13}C^{\alpha}H^{\alpha}$ vectors allows one to define the backbone structure. A concerted change of ϕ and ψ in opposite directions is possible, without affecting Σ. This degree of freedom defines the orientation of the peptide plane with respect to the $C^{\alpha}H^{\alpha}$ vectors.

In Figure 3 the theoretical rate derived from eq 1 is plotted as a function of the two angles ϕ and ψ. The regions corresponding to α-helices and β-strands are outlined. Note that the angles ϕ and ψ refer to consecutive amino acids, in contrast to the angles represented in Ramachandran plots.

In the derivation of eq 3, an ideal geometry was assumed. Deviations from ideality may become important when the vectors are separated by many bonds. To evaluate the errors due to nonideal structures, the relationship between cos θ and cos Σ was statistically investigated in Figure 4 for eighteen proteins from the Protein Data Bank for which crystal structures were available with a resolution better than 2.2 Å. Where absent, protons were inserted into the structures using the program MolMol.[18] Linear regression of the data yielded a correlation coefficient of 0.99 (see Figure 4a) and the following relationship between θ and Σ was obtained:

$$\cos\theta = -0.09 + 0.894\cos(\Sigma + 180°) \qquad (4)$$

The coefficients in eqs 3 and 4 are very similar, suggesting that the assumption underlying the derivation of eq 3 (notably the planarity of the peptide bond) is largely justified. To correlate the angle Σ with secondary structural elements, the proteins of the database mentioned above were examined. The secondary structure was evaluated using the algorithm of Kabsch and Sander[19] which utilizes hydrogen-bonding patterns. Residues belonging to α-helices and β-sheets were then collected into a database. Only residues for which the previous residue also belong to the same segment of secondary structure were included. For example, a residue is classified as representative of an α-helix only if the previous residue also belongs to the same α-helix. The databases corresponding to α-helices and β-sheets were further restricted by excluding residues which did not fulfill the Ramachandran most-favored angles ($100° \le \psi \le 180°$ for β-sheets, and $-20° \le \psi \le -80°$ for α-helices).[5] This yielded a total of 430 residues in β-sheet regions and 539 in α-helical regions. The frequency of occurrence of the dihedral angles Σ in this database is shown in Figure 4b. The mean values of the dihedral angle Σ were found to be $21.0° \pm 22.9°$ in the β-sheets and $-104.7° \pm 9.0°$ in α-helices. The spread of the Σ angles is larger in β-sheets than in α-helices. Two consecutive $^{13}C^{\alpha}-H^{\alpha}$ bond vectors are oriented approximately antiparallel in β-sheets and roughly orthogonal in α-helices.

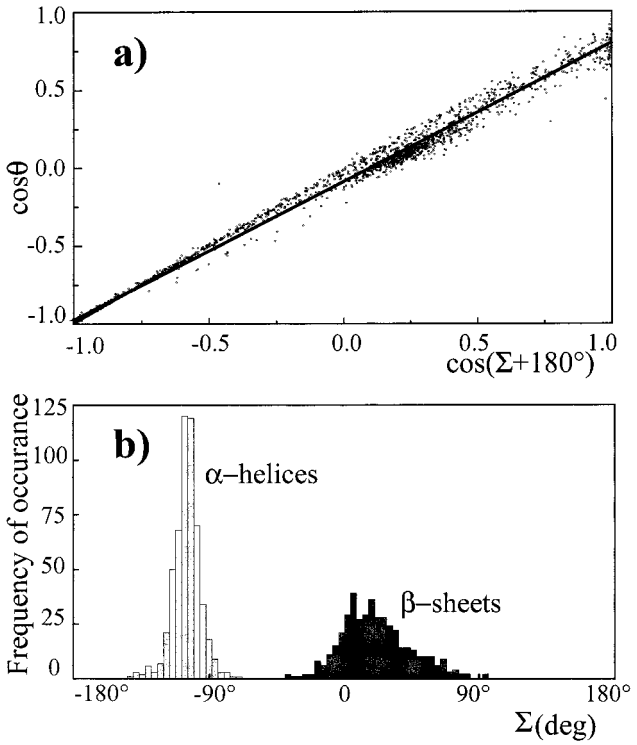

Figure 4. (a) Linear regression of cos θ against cos(Σ + 180°). When the two H^{α} atoms lie in the same plane as the two C^{α} atoms and on the same side with respect to the $C^{\alpha}(i-1)$-$C^{\alpha}(i)$ axis, $\Sigma = 180°$. The points represent angles calculated from the X-ray structures for 18 proteins from the Protein Data Bank (see Supporting Information). (b) The frequency of occurrence of dihedral angles Σ of α-helices and β-sheets, selected as described in the text.

The experiments of Figure 1 were applied to ^{15}N,^{13}C-labeled human ubiquitin and the rates $R(\theta)$ were estimated by fitting eq 2 to the experimental data with $T = 25$ ms. Figure 5 (top) shows a comparison between the theoretical and experimental rates $R(\theta)$. The abcissae of the experimental points have been determined by deriving the angles Σ appropriate for the relevant pairs of neighboring amino acids from the X-ray structure of ubiquitin.[20] All glycine residues and the residues immediately following them were removed, since the relaxation in these systems is complicated due to the presence of two H^{α} nuclei in each glycine. The rates are negative (-13.5 ± 3.6 s^{-1}) in α-helices, whereas those in β-sheets are positive (16.9 ± 7.1 s^{-1}), as expected from Figure 3. Thus, a series of consecutive negative rates $R(\Sigma)$ are indicative of the presence of an α-helix whereas a series of consecutive positive rates indicates the presence of a β-strand.

(18) Konradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51.

(19) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.

(20) Vijay-Kumar, S.; Bugg, C. E.; Cook, C. J. *J. Mol. Biol.* **1987**, *194*, 531.

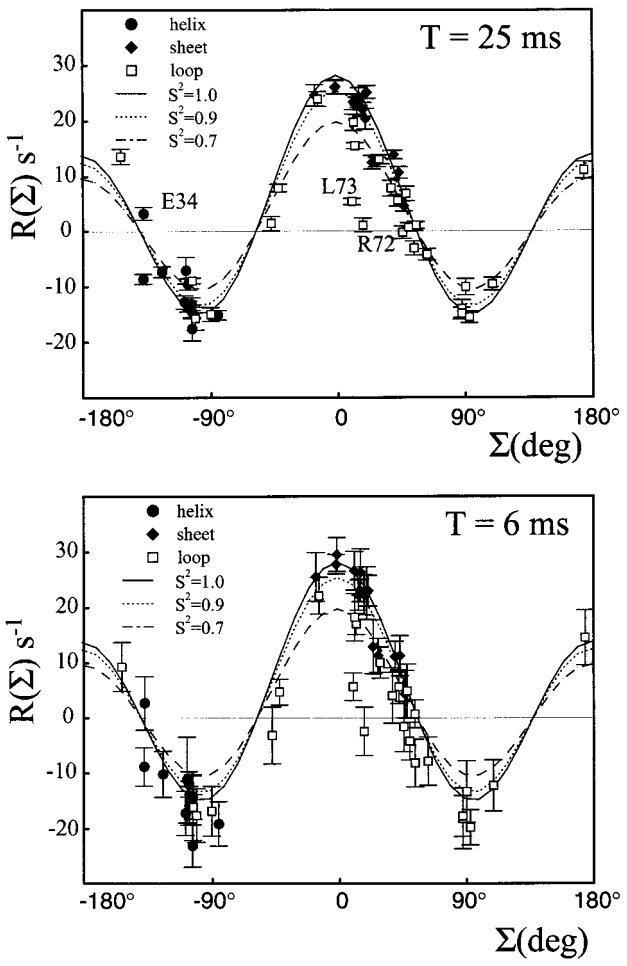

Figure 5. Cross-correlation rates in ubiquitin at 303 K and 14 T. The experimental rates $R(\theta)$, derived with eq 2 from data recorded with the experiments of Figure 1 using a Bruker Avance 600 MHz spectrometer, are plotted as a function of the dihedral angle Σ derived from the X-ray structure of ubiquitin, for $T = 6$ (bottom) and 25 ms (top). The curves represent the theoretical dependence predicted from eqs 1−3 for order parameters $S^2 = 1$, 0.90, and 0.76. The overall correlation time was assumed to be $\tau_c = 4.1$ ns and $r_{CH} = 1.12$ Å. The amino acids leucine-73 (L73) and arginine-72 (R72) are near the highly mobile C-terminus. Glutamic acid-34 (E34) is the last residue of an α-helix. For the assignments of the other data points to individual amino acids, see the Supporting Information.

The long delay $T \approx 1/J(C^\alpha C^\beta) \approx 25$ ms might cause significant signal losses in larger proteins. To overcome this problem, one can insert a band-selective C^α refocusing pulse in the middle of T to prevent evolution under $J(C^\alpha C^\beta)$.[10] However, this approach brings about undesirable complications. On the one hand, one has to account for cross-correlation effects that occur during the band-selective pulse, which may typically have a duration of 2−3 ms. Since cross-correlation rates involving two $^{13}C^\alpha H^\alpha$ interactions are $\gamma_C/\gamma_N \approx 2.5$ times larger than those involving one $^{13}C^\alpha H^\alpha$ and one $^{15}NH^N$ interaction, the effects of relaxation during a band-selective pulse are much more pronounced than in previous experiments.[9−14] On the other hand, it is not possible to refocus all C^α spins without affecting at least some of the C^β spins. Indeed, the chemical shifts of C^β belonging to threonines and serines lie in the C^α region. Furthermore, the effects of imperfections of band-selective pulses are exacerbated since each two-spin coherence involves two C^α spins, in contrast to previous experiments. For larger proteins we therefore propose to use small delays $T \ll 1/J(C^\alpha C^\beta)$, so that losses due to evolution under these scalar couplings are minimized. Since the proton π pulses need to be shifted in experiment II to allow evolution of $J(C^\alpha H^\alpha)$, the delay T cannot be made shorter than $[2J(C^\alpha H^\alpha)]^{-1} = 3.6$ ms. The signal intensities are attenuated by a factor of $\cos^2\{\pi J(C^\alpha C^\beta)T\}$, which implies a loss of 17 or 34% for $T = 4$ or 6 ms. The rates shown in Figure 5 (bottom) have been measured with $T = 6$ ms.

At $T = 6$ ms the errors are larger than for $T = 25$ ms. The errors are almost entirely due to experiment II. At the beginning of the build-up (i.e. at $T = 6$ ms) the amplitude of $16C_xH_zN_zC_xH_z$ is proportional to $R(\theta)T$. Since the cross-correlation rate is proportional to the correlation time τ_C, the build-up rate increases with the size of the molecules. However, the conversion $N_xH^\alpha{}_z \rightarrow 4C^\alpha{}_zN_xC^\alpha{}_z$ in the interval $2\tau_3$, which requires about 50 ms, leads to significant signal losses in large proteins (the duration of the sequence, excluding T, is $6\tau_1 + 4\tau_3 \approx 115$ ms). At high magnetic fields, the experiments can be modified by using spin-state selective methods to benefit from the so-called TROSY effect.[21] Another source of error might be caused by variations of $J(C^\alpha H^\alpha)$. Since the rates measured for $T = 6$ and 25 ms were found to be the same within experimental error, these errors are believed to be small.

Conclusions

We have developed a method to measure the effects of cross-correlation of the fluctuations of two $^{13}C^\alpha-H^\alpha$ dipolar interactions on the relaxation of two-spin coherences involving two consecutive $^{13}C^\alpha$ nuclei. These effects allow us to determine the dihedral angle Σ between the planes defined by the atoms $\{H^\alpha(i-1),^{13}C^\alpha(i-1),^{13}C^\alpha(i)\}$ and $\{^{13}C^\alpha(i-1),^{13}C^\alpha(i),H^\alpha(i)\}$. This angle is indicative of the secondary structure in proteins and may be used in conjunction with other dihedral angles to refine structure calculation protocols.

Acknowledgment. This work was supported by the Fonds National de la Recherche Scientifique (FNRS) and the Commission pour la Technologie et l'Innovation (CTI) of Switzerland and by the Centre National de la Recherche Scientifique (CNRS) of France. We would like to thank Dr. Catherine Zwahlen for stimulating discussions.

Supporting Information Available: Table of proteins used in the database for the analysis of the Σ angles and a table giving the cross-correlation rates $R(\Sigma)$ measured at $T = 6$ and 25 ms for all suitable residues in ubiquitin (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA9933184

(21) Pervushin, K. V.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, 94, 12366.