A Comprehensive Analysis of Multifield ^{15}N Relaxation Parameters in Proteins: Determination of ^{15}N Chemical Shift Anisotropies

Daniel Canet,*,† Philippe Barthe,‡ Pierre Mutzenhardt,† and Christian Roumestand*,‡

Contribution from the Laboratoire de Méthodologie RMN, Université H. Poincaré, Nancy I, B.P. 239, 54506 Vandoeuvre-lès-Nancy Cedex, France, and Centre de Biochimie structurale, Université de Montpellier I, Faculté de Pharmacie, 15 Avenue Charles Flahault, 34060 Montpellier Cedex, France

Received November 3, 2000. Revised Manuscript Received February 13, 2001

Abstract: This study deals with the exploitation of the three classical ^{15}N relaxation parameters (the longitudinal relaxation rate, R_1, the transverse relaxation rate, R_2, and the ^{1}H−^{15}N cross-relaxation rate, σ_{NH}) measured at several magnetic fields in uniformly ^{15}N-labeled proteins. Spectral densities involved in R_1, R_2 and σ_{NH} are analyzed according to the functional form $A + B/(1 + \omega^2\tau_s^2)$, where τ_s is the correlation time associated with slow motions sensed by the *NH* vector at the level of the residue to which it belongs. The coefficient *B* provides a realistic view of the backbone dynamics, whereas *A* is associated with fast local motions. According to the "model free approach", *B* can be identified with $2\tau_s S^2$ where *S* is the generalized order parameter. The correlation time τ_s is determined from the field dependency of the relaxation parameters while *A* and *B* are determined through linear equations. This simple data processing is needed for obtaining realistic error bars based on a statistical approach. This proved to be the key point for validating an extended analysis aiming at the determination of nitrogen chemical shift anisotropy. The protein C12*A*−*p*8^{MTCP1} has been chosen as a model for this study. It will be shown that all data (obtained at five magnetic field strengths corresponding to proton resonance of 400, 500, 600, 700, and 800 MHz) are very consistently fitted provided that a specific *effective* correlation time associated with slow motions is defined for each residue. This is assessed by small deviations between experimental and recalculated values, which, in all cases, remain within experimental uncertainty. This strategy makes needless elaborate approaches based on the combination of several slow motions or their possible anisotropy. Within the core of the protein τ_s fluctuates in a relatively narrow range (with a mean value of 6.15 ns and a root-mean-square deviation of 0.36 ns) while it is considerably reduced at the protein extremities (down to ∼3 ns). To a certain extent, these fluctuations are correlated with the protein structure. *A* is not obtained with sufficient accuracy to be valuably discussed. Conversely, order parameters derived from *B* exhibit a significant correlation with the protein structure. Finally, the multi-field analysis of the evolution of longitudinal and transverse relaxation rates has been refined by allowing the ^{15}N chemical shift anisotropy (*csa*) to vary residue by residue. Within uncertainties (derived here on a statistical basis) an almost constant value is obtained. This strongly indicates an absence of correlation between the experimental value of this parameter obtained for a given residue in the protein, the nature of this residue, and the possible involvement of this residue in a structured area of the protein.

Introduction

Heteronuclear spin relaxation is being used increasingly to study the dynamics of proteins.[1] An incentive to these studies is the search for correlation between structure, dynamics, and function.[2,3] A large number of such heteronuclear relaxation studies have focused on amide ^{15}N−^{1}H spin system in isotopically enriched protein samples,[4,5] thereby allowing the local dynamics along the protein backbone to be explored residue by residue. Typically, dynamical information is derived from the longitudinal and transverse relaxation rates, R_1 and R_2 respectively, and also from the cross-relaxation (^{1}H−^{15}N) rate (denoted by σ_{NH} in the following).

Such studies are aimed at characterizing (i) the protein overall tumbling, (ii) local fast motions, (iii) restrictions of these local motions defined by a generalized order parameter *S*, (iv) possibly more complicated slow motions, which would be substituted to an isotropic overall tumbling. Moreover, the so-called chemical shift anisotropy (*csa*) at the ^{15}N nucleus is in principle accessible through measurements of relaxation parameters at several magnetic field strengths, and should possibly provide further information. It is indeed well-known that multi-field data has been extensively used in the past for studying complex systems (such as surfactant aggregates[6]) and made it possible an accurate description of the different types of motion involved. This is of course also appropriate for proteins.[7] The goal of this paper is to delineate the potentiality of this approach by considering critically the reliability of each parameter and by assessing its role for the determinations listed above. Relaxation parameters have been determined at five different magnetic field strengths (9.4, 11.75, 14.1, 16.45, and 18.8 T) on a ^{15}N-labeled

† UPRESA CNRS 7042, INCM-FR CNRS 1742.
‡ UMR- CNRS 9955, INSERM-U414.

(1) Wagner, G. *Curr. Opin. Struct. Biol.* **1993**, *3*, 748.
(2) Karplus, M.; McCammon, J. A. *Annu. Rev. Biochem.* **1983**, *53*, 263.
(3) Frauenfelder, H.; Farak, F.; Young, R. D. *Annu. Rev. Biophys. Chem.* **1988**, *17*, 451.
(4) Peng, J. W.; Wagner, G. *Methods Enzymol*. **1994**, *239*, 563.
(5) Palmer, A. G.; Williams, J.; McDermott, A. *J. Phys. Chem.* **1996**, *100*, 13293.
(6) Lindman, B.; Olsson, U.; Söderman, O. In *Dynamics of Solutions and Fluid Mixtures by NMR*; Delpuech, J. J., Ed.; John Wiley & Sons: New York, 1995.

10.1021/ja0038676 CCC: $20.00
Published on Web 04/21/2001

sample of C12$A-p8^{MTCP1}$. The 3D solution structure of this small 68-residue protein has been solved[8] and recently refined in our laboratory.[9] It turns out that defining an *effective* correlation time for each residue provides a remarkably consistent analysis of *all* data, supported by the evaluation of realistic confidence intervals. In addition to the above-mentioned correlation times and order parameters, we focus here on the determination of the ^{15}N chemical shift anisotropy which, in principle, can be accessed through multifield determinations. In contrast with recent publications,[10] the value of the ^{15}N chemical shift anisotropy obtained through our analysis for a given residue appears not to be correlated either with the nature of this residue or with an element of the secondary structure to which it belongs.

Theory

The three relaxation parameters will be expressed by assuming that the following relaxation mechanisms are dominant: the dipolar ^{15}N$-^{1}$H interaction (d), the ^{15}N chemical shift anisotropy (csa), and possible other contributions (e.g., exchange) prone to affect the transverse relaxation rate. We denote by $\tilde{J}(\omega)$ a spectral density function which involves only dynamical parameters and whose simplest form would be $2\tau_c/(1+\omega^2\tau_c^2)$, τ_c being an effective correlation time. As explained below, more elaborate forms will be needed, while $\tilde{J}_d$ and $\tilde{J}_{csa}$ may be different. These spectral density functions are multiplied by factors defining the amplitude of the considered mechanism. One has for the dipolar and *csa* mechanisms, respectively (the various symbols having their usual meaning. ($\omega_N/2\pi$): ^{15}N resonance; d_{NH}: N$-$H bond length; $\Delta\sigma_N$: ^{15}N shielding anisotropy)

$$K_d = \frac{1}{20}\left(\frac{\mu_0}{4\pi}\right)^2\left(\frac{\gamma_H\gamma_N\hbar}{d_{NH}^3}\right)^2 \quad (1)$$

($K_d = 0.25957 \times 10^9$ assuming $d_{NH} = 1.02$ Å)

$$K_{csa} = \frac{1}{15}(\Delta\sigma_N)^2\omega_N^2 \quad (2)$$

(when not extracted from experimental data, $\Delta\sigma_N$ may be taken as -170 ppm; this average value has emerged from experimental measurements of ^{15}N amide shielding tensor using various methods[11−15]).

According to those conventions, we can write[16]

$$R_1 = K_d[6\tilde{J}_d(\omega_H+\omega_N) + 3\tilde{J}_d(\omega_N) + \tilde{J}_d(\omega_H-\omega_N)] + K_{csa}\tilde{J}_{csa}(\omega_N) \quad (3)$$

$$R_2 = K_d\left[3\tilde{J}_d(\omega_H+\omega_N) + \frac{3}{2}\tilde{J}_d(\omega_N) + 3\tilde{J}_d(\omega_H) + \frac{1}{2}\tilde{J}_d(\omega_H-\omega_N) + 2\tilde{J}_d(0)\right] + K_{csa}\left[\frac{1}{2}\tilde{J}_{csa}(\omega_N) + \frac{2}{3}\tilde{J}_{csa}(0)\right] + R_{2,others} \quad (4)$$

$$\sigma_{NH} = K_d[6\tilde{J}_d(\omega_H+\omega_N) - \tilde{J}_d(\omega_H-\omega_N)] \quad (5)$$

One way to exploit those relaxation parameters measured at different values of the magnetic field consists of inverting eqs 3−5 so as to obtain the spectral density as a function of ω (spectral density mapping[17,18]). This procedure is model-independent but provides only a qualitative interpretation of backbone dynamics. Moreover it generally implies some assumptions; for example, $\tilde{J}(\omega_H+\omega_N) \approx \tilde{J}(\omega_H-\omega_N)$.[19,20] Rather, we shall directly exploit these relaxation rates according to the following functional form

$$\tilde{J}(\omega) = A + \frac{B}{1+\omega^2\tau_s^2} \quad (6)$$

τ_s is an effective correlation time associated with slow motions. The parameter A includes the extreme narrowing contribution, associated with local fast motions. As a matter of fact, the two terms of eq 6 can be identified with the spectral density function of the model-free approach:[21]

$$A = (1-S^2)(2\tau_e) \quad (7)$$

$$B = S^2(2\tau_s) \quad (8)$$

Here τ_e is the effective correlation time describing the fast local motions (in fact $1/\tau_e = 1/\tau_f + 1/\tau_s$ where τ_f is associated with fast local motions, in practice very close to τ_e), whereas S is an order parameter specifying the restriction of these motions with respect to a local director, reorienting itself according to τ_s. One should note that the concept of a local director is reminiscent of organized systems (e.g., liquid crystals) and that, as far as a spherical object is concerned, the local director would be along the direction passing through the sphere center and the atom involved in the relaxation study. Clearly, in the case if a protein having approximately a spherical shape, τ_s would be the overall tumbling correlation time. Of course, the order parameter depends on the orientation of the relaxation vector with respect to the local director. This relaxation vector is the N$-$H bond for dipolar spectral densities, or the symmetry axis of the nitrogen shielding tensor (supposed to be of axial symmetry) for *csa* spectral densities. As a consequence, because these two vectors are not collinear,[22,23] we should have one order parameter for the dipolar interaction and another one for the *csa* mechanism. This means that the parameters A and B are not only site-dependent but also, in principle, mechanism-dependent. Nevertheless, the angle between the N$-$H vector and the shielding tensor symmetry axis is small[14] ($\sim$13−16°) so that considering that order parameters for the N$-$H dipolar interaction and the

(7) Peng, J.; Wagner, G. *Biochemistry* **1995**, *34*, 16733−16752; Vis, H.; Vorgias, C. E.; Wilson, K. S.; Kaptein, R.; Boelens, R. *J. Biomol. NMR* **1998**, *11*, 265−277.
(8) Barthe, P.; Yang, Y.-S.; Chiche, L.; Hoh, F.; Strub, M. P.; Guignard, L.; Soulier, J.; Stern, M. H.; van Tilbeurgh, H.; Lhoste, J.-M.; Roumestand, C. *J. Mol. Biol.* **1997**, *274*, 80.
(9) Barthe, P.; Chiche, L.; Declerck, N.; Delsuc, M.-A.; Lefèvre, J.-F.; Malliavin, T.; Mispelter, J.; Stern, M.-H.; Lhoste, J.-M.; Roumestand, C. *J. Biomol. NMR* **1999**, *15*, 271.
(10) Fushman, D.; Tjandra, N.; Cowburn, D. *J. Am. Chem. Soc.* **1999**, *121*, 8577−8582.
(11) Oas, T. G.; Hartzell, C. J.; Dahlquist, F. W.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5966.
(12) Hiyama, Y.; Niu, C. H.; Silverton, J. V.; Ravoso, A.; Torchia, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 2378.
(13) Tjandra, N.; Wingfield, P.; Stahl, S.; Bax, A. *J. Biol. NMR* **1996**, *8*, 273.
(14) Ottiger, M.; Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 9825.
(15) Fushman, D.; Tjandra, N.; Cowburn, D. *J. Am. Chem. Soc.* **1998**, *120*, 10947.
(16) See, for instance: Canet, D. *Nuclear Magnetic Resonance. Concepts and Methods*; Wiley: Chichester, 1996.
(17) Peng, J. W.; Wagner, G. *Biochemistry* **1992**, *31*, 8571.
(18) Peng, J. W.; Wagner, G. *J. Magn. Reson.* **1992**, *98*, 308.
(19) Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E. *J. Biomol. NMR* **1995**, *6*, 153.
(20) Ishima, R.; Nagayama, K. *Biochemistry* **1995**, *34*, 3162.
(21) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.
(22) Boyd, J.; Redfield, C. *J. Am. Chem. Soc.* **1998**, *120*, 9692.
(23) Fushman, D.; Cowburn, D. *J. Biomol. NMR* **1999**, *13*, 139.

^{15}N *csa* mechanism are identical may constitute a reasonable assumption (see below). On the other hand, owing to their fairly high values, order parameters should not change by a significant amount.

It appears that a unique correlation time may not be sufficient for describing the slow motions undergone by the *NH* vector. This leads, for instance, to an extension of the model-free approach[24] implying the introduction of an additional correlation time and of an additional order parameter.

Another approach comes from the Lipari−Szabo equations[21] in the case of an anisotropic overall motion. Tjandra et al.[25] inserted, in place of the simple exponential describing the slow motion, a correlation function of the Woessner type[26] related to the whole protein.

However, none of these approaches was found adequate for treating the whole set of data (results not shown). We therefore turned to another concept involving the definition of *a single correlation time* τ_s *per residue*. τ_s thus becomes an effective correlation time describing the slow motions sensed by the relevant *NH* vector. Of course, these motions are certainly very complicated and not limited to the overall tumbling for which, anyway, some anisotropy should be taken into account. In fact, relying on an effective correlation time amounts to consider that the superimposition of different motions lead *approximately to an exponential correlation function* which may be different from one residue to another. To further assess this point, we can recall that the sum of two exponential functions with relatively close time constants does not differ significantly from a single-exponential function with an effective time constant. In fact, the concept of an effective correlation time for each residue is not new. It has been invoked in several instances but always limited to the characterization of anisotropic rotational diffusion.[27−30] It is taken here in a broader sense as it includes pragmatically the superposition of different motions as well as any anisotropy of the overall tumbling.

Data Analysis

If the three parameters (R_1, R_2, and σ_{NH}) are available at a single frequency, only *A*, *B*, and τ_s can be extracted by assuming that $R_{2,others}$ is negligibly small (this could actually be achieved by correcting the R_2 values from exchange contributions implying that multifield data are available) and that $\tilde{J}_d = \tilde{J}_{csa}$. As soon as data at two frequencies are at hand, it can be envisaged to extract some further information concerning (i) the exchange contribution to the R_2 values (see below) or (ii) the *csa* mechanism, in addition to the parameters discussed above. Owing to the low frequencies involved in *csa* spectral densities (see 3 and 4), we can approximate the *csa* contributions in R_1 and R_2 as

$$R_{1,csa} = u_{csa}\frac{\omega_N^2}{1+\omega_N^2\tau_s^2} \tag{9}$$

$$R_{2,csa} = u_{csa}\left(\frac{1}{2}\frac{\omega_N^2}{1+\omega_N^2\tau_s^2} + \frac{2}{3}\omega_N^2\right) \tag{10}$$

The simplified form of eqs 9 and 10 comes from the fact that, in spectral densities at low frequencies (zero frequency or nitrogen frequency), the constant term *A* (relating to fast motions) in eq 6 is negligible with respect to the second term $B/1 + \omega\tau_s^{22}$ (relating to slow motions). Assuming that the shielding tensor is axially symmetric and that its symmetry axis possesses the same dynamical properties as the *NH* vector, we can express u_{csa} from K_{csa} as follows:

$$u_{csa} = \frac{1}{15}(\Delta\sigma_N)^2 B \tag{11}$$

and thus hope to determine $\Delta\sigma_N$, the nitrogen-shielding anisotropy.

A computer program based on the above considerations has been written. If τ_s is known (for instance from the procedure described below), *A*, *B*, and possibly u_{csa} are found by solving a system of linear equations, (this capability can be recognized by examining eqs 3−6), thus avoiding the unnecessary use of a general search algorithm with the risk of divergence. Such a treatment does not proceed from a new model but represents a new strategy which of course is consistent with the usual procedures. It possesses the unique capability of calculating realistic errors through a statistical approach. More explicitely, this set of linear equations can be written as

$$\sum_{j=1}^{m} a_{ij}x_j = R_i^{exp} \tag{12}$$

with the unknowns $x_1 = A$, $x_2 = B$, and possibly $x_3 = u_{csa}$ (so that $m = 2$ or 3), R_i^{exp} being one of the *n* experimental relaxation parameters used in the fit (so that the subscript *i* runs from 1 to *n*). The coefficients a_{ij} are calculated from eqs 1−5. For example, if R_i^{exp} is a longitudinal relaxation rate, one has

$$a_{i1} = 10K_d$$

$$a_{i2} = K_d\left[\frac{6}{1+(\omega_H+\omega_N)^2\tau_s^2} + \frac{3}{1+\omega_N^2\tau_s^2} + \frac{1}{1+(\omega_H-\omega_N)^2\tau_s^2}\right]$$

$$a_{i3} = \frac{\omega_N^2}{1+\omega_N^2\tau_s^2}$$

Whenever it is decided not to determine u_{csa}, the *j* subscript takes only the values 1 and 2 and a_{i2} has to be appended by $K_{csa}/(1 + \omega_N^2\tau_s^2)$. Because the number *n* of linear eq 12 is generally larger than the number of unknowns, one has to recourse to the general regression procedure[31] which is briefly summarized below. Equation 12 can be recast in a matricial form

$$\mathbf{AX} = \mathbf{R} \tag{13}$$

A being the rectangular matrix of coefficients (of dimension $n \times m$), **X** and **R** the columns representing respectively the unknowns and the experimental relaxation parameters (of dimensions *m* and *n*, respectively). The least-squares condition

(24) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989.
(25) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562.
(26) Woesnner, D. T. *J. Chem. Phys.* **1962**, *252*, 7740.
(27) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269.
(28) Schurr, J. M.; Babcock, H. P.; Fujimoto, B. S. *J. Magn. Reson.* **1994**, *B105*, 211.
(29) Brüschweiler, R.; Liao, X.; Wright, P. E. *Science* **1995**, *268*, 886.
(30) Andrec, M.; Inman, K. G.; Weber D. J.; Levy, R. M.; Montelione, G. T. *J. Magn. Reson.* **2000**, *146*, 66.
(31) Draper, N. R.; Smith, H. *Applied Regression Analysis*; Wiley: New York, 1998.

amounts to multiply eq 13 on the left by $\mathbf{A^t}$, the transpose of $\mathbf{A}$.

$$\mathbf{A}^t\mathbf{A}\mathbf{X} = \mathbf{A}^t\mathbf{R}$$

or

$$\mathbf{C}\mathbf{X} = \mathbf{Z} \tag{14}$$

where $\mathbf{C} = \mathbf{A^tA}$ is a square matrix of dimension $m \times m$ and $\mathbf{Z} = \mathbf{A^tR}$ a column of dimension m. Thus, eq 14 represents a set of m equations for m unknowns, easily solved by standard numerical procedures. This treatment yields also statistical uncertainties[31]

$$\Delta x_j = t\sigma'\sqrt{\frac{C_{jj}^{-1}}{n-m}} \tag{15}$$

t is related to the confidence interval and can be extracted from Student-Fisher tables, σ' is the standard deviation

$$\sigma' = \sqrt{\sum_{i=1}^{n}\frac{(R_i^{\exp} - R_i^{\mathrm{calc}})^2}{n}}$$

and C_{jj}^{-1} is the jth diagonal element of the inverse of matrix $\mathbf{C}$. Besides the obvious advantages of linear calculations, the interest of this treatment arises from the fact that τ_s is determined solely from the frequency dependence of the available relaxation parameters.

However the determination of τ_s must be dealt with by non linear methods (in this computer program, the Simplex algorithm was used in order to minimize the χ^2 quantity defined as

$$\chi^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{R_i^{\exp} - R_i^{\mathrm{calc}}}{\Delta R_i}\right)^2$$

R_i being one of the relaxation parameter used in the fit and ΔR_i its experimental uncertainty); this implies however, at each step of the search algorithm, the calculation of A, B (and possibly u_{csa}) by the linear procedures just mentioned. The uncertainty on τ_s is obtained by considering that it is proportional to the standard deviation σ' defined above. The proportionality factor is deduced from the variation of σ' with τ_s.

Results and Discussion

Experimental data for most residues of the ^{15}N-labeled *C*12*A*-$p8^{MTCP1}$ have been obtained at 9.4 T, 11.75 T, 14.1, 16.45, and 18.8 T. The human protein $p8^{MTCP1}$ is a 68-residue protein encoded by the MTCP-1 oncogene. Its biological function is unknown, but its 3D solution structure has been recently solved.[8] As shown in Figure 1, it revealed an original scaffold consisting of three α-helices, associated with a new cysteine motif. The core of the protein mainly consists of two helices (helix I: residues 7−20, and helix II: residues 29−40) which are covalently paired by two disulfide bridges (Cys38-Cys7 and Cys17-Cys28), forming an α-hairpin. A relatively well-defined loop (residues 41−47) connects helix II to helix III, which spans residues 48−63. The third disulfide bridge (Cys39-Cys50) links the top of helix III to the tip of helix II. Helix III is oriented roughly parallel to the plane defined by the α-antiparallel motif and appears less defined. Except for the first N-terminal turns, few nOe contacts were found between the third helix and the α-hairpin, suggesting that helix III is loosely bound to the core of the protein. In the recombinant C12A- $p8^{MTCP1}$ mutant protein, the "free" cysteine residue at position 12 in the native sequence has been replaced by an alanine residue, to improve the expression yields in *Escherichia coli*.

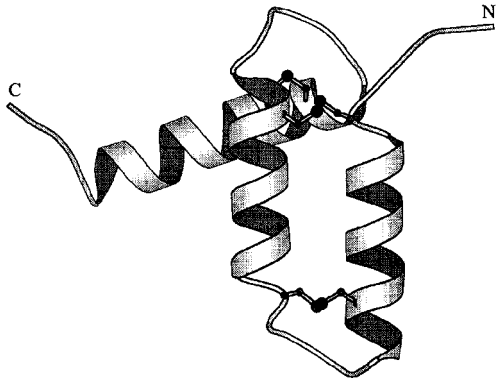

Figure 1. Ribbon diagram (MOLSCRIPT[48]) of the structure of $p8^{MTCP1}$ showing the backbone and disulfide bonds.

Experimental Section

The protein, at a final concentration of 0.4 mM, was prepared with a 50 mM phosphate buffer (pH adjusted at 6.5) and a 25 mM NaCl solution; ^2H_2O is present in a proportion of 10% for field-frequency stabilization purposes. Temperature was carefully adjusted at 20 °C using a calibration sample (80% glycol in d_6-DMSO). Complete assignment of ^{1}H−^{15}N cross-peaks can be found in a previous work.[9] In all experiments, the ^{1}H carrier was centered on the water resonance in order to use the WATERGATE sequence[32,33] for suppressing the solvent resonance. Quadrature detection in the indirect dimension was achieved using the States-TPPI scheme;[34] the initial t_1 (t_1: evolution period) was set at exactly half the chosen dwell time in order to remove baseline distortions and to optimize aliasing characteristics in the resulting two-dimensional spectra.[35] The pulse sequences used to determine R_1, R_2, and σ_{NH} (derived from the corresponding NOE factor) were similar to those described in the litterature.[17,18,36] To minimize artifacts, pulse field gradients were inserted during the intervals where the spin system is in a longitudinal spin-order state.[37] R_1 data sets are recorded in such a way that the signal intensity decays exponentially to zero as a function of the relaxation delay, thus enabling a simple two-parameter fit. A recycle time of 4 s was employed, and ^{15}N decoupling during proton acquisition was performed using a WALTZ sequence. R_1 experiments were performed with 13 relaxation delays (18, 54, 102, 150, 210, 258, 306, 402, 498, 606, 810, 1002, and 1506 ms) during which a train of 180° ^{1}H pulses (separated by 3ms) is applied so as to remove dipolar (^{15}N−^{1}H)−*csa*(^{15}N) cross-correlation effects. R_2 experiments were performed employing a Carr−Purcell−Meiboom−Gill (CPMG) pulse train[38,39] with an interval of 1 ms between two consecutive ^{15}N 180° pulses; ^{1}H decoupling 180° pulses are applied between two ^{15}N pulses every 4 ms. Thirteen experiments were run with the following duration of the CPMG train: 16, 32, 48, 80, 112, 128, 160, 192, 256, 320, 384, 512, 768 ms. All R_1 and R_2 experiments were performed with relaxation delays arbitrarily chosen in the relevant list (and not in an increasing or decreasing order) so as to prevent any bias that could arise from possible degradation of the main magnetic field homogeneity. The noise level was estimated through duplicate experiments for relaxation delays of 18 ms and 150 ms (R_1 measure-

(32) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661.
(33) Sklenar, V. *J. Magn. Reson.* **1995**, *A114*, 132.
(34) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393.
(35) Bax, A.; Ikura, M.; Kay, L. E.; Zhu, G. *J. Magn. Reson.* **1991**, *91*, 174.
(36) Kay, L. E.; Nicholson, L. K.; Delaglio, F.; Bax, A.; Torchia, D. A. *J. Magn. Reson.* **1992**, *97*, 359.
(37) Bax, A.; Pochapsky, S. S. *J. Magn. Reson.* **1992**, *99*, 638.
(38) Carr, H. Y.; Purcell, E. M. *Phys. Rev.* **1954**, *94*, 630.
(39) Meilboom, S.; Gill, G. *Rev. Sci. Instrum.* **1958**, *29*, 688.

ments) and of 16 ms and 112 ms (R_2 measurements). All experiments were recorded with a time domain (t_1,t_2) data size of 96 × 2K complex points and 4−16 transients per t_1 increment, depending on the spectrometer sensitivity.

The NOE factor, leading to the cross-relaxation rates σ_{NH}, is deduced from the comparison of ^{15}N intensities without and with proton saturation. The latter is achieved with a train of 120° pulses[40] separated by 20 ms and of a total duration of 3 s. In the experiment without saturation a carefully optimized water flip-back pulse was added before the last ^1H 90° pulse in order to avoid undesirable effects related to exchange between water and amide protons. A recycle time of 6 s between scans was used for obtaining a complete recovery of water magnetization and for reducing exchange effects. Moreover, the two experiments (with and without proton saturation) were acquired in an interleaved manner for each t_1 increment with a waiting time of 30 s between them. These NOE experiments were carried out with the same data size as for the R_1 and the R_2 experiments: 32−128 transients were accumulated per t_1 increment, again depending on the spectrometer sensibility.

NMR spectra were processed with the Gifa (version 4.22)[41] software. Cross-peak intensities were determined from peak heights[42] using the Gifa peak-picking routine. The relaxation rate constants R_1 and the R_2 were obtained from nonlinear fits to monoexponential functions.[43] A comparative study of decays characterized by both peak heights and peak volumes was found to yield essentially the same results. The uncertainties due to random errors in the measured heights were deduced from 500 Monte Carlo simulations. The root-mean-square (rms) values of noise was evaluated in free-peak regions and used to estimate the standard deviation of the peak intensities. The duplicate spectra, mentioned above, assess the validity of this estimate.

Transverse Relaxation Rates and Chemical Exchange. R_2 values have been corrected of exchange effects, by considering that the exchange contribution, R_{ex}, depends on the nitrogen Larmor frequency, ω_N, and can be written as[44] $R_{ex} = \Phi\omega_N^2$. The factor Φ is dependent on the intrinsic rate constant of the exchange process, on the chemical shift differences, on the population of participating sites, and on the applied ^{15}N spin−echo period in the CPMG pulse train. The presence of chemical exchange for a residue may be noted by a transverse relaxation rate that is high compared to that of nearby residues. However, if the relevant *NH* bond vector experiences both fast local motions, which reduce the value of R_2, and, simultaneously, slow conformational exchange, the interference of these two opposite contributions to R_2 may hamper the identification of possible exchange processes. With relaxation data sets available from more than one field, R_{ex} can be obtained from the ω_N^2 linear dependency of the quantity of $(2R_2 - R_1)$ according to[45]

$$2R_2 - R_1 = K_d[6\tilde{J}(\omega_H) + 4\tilde{J}(0)] + \left[\frac{4}{45}(\Delta\sigma_N)^2\tilde{J}(0) + 2\Phi\right]\omega_N^2 \quad (16)$$

where the notations of eqs 1−4 have been used and where $\tilde{J}_d = \tilde{J}_{csa} = \tilde{J}$ has been assumed. Assuming further that $\tilde{J}(\omega_H)$ is negligible with respect to $\tilde{J}(0)$, we notice that $\tilde{J}(0)$ is obtained from the intercept of the linear representation of $(2R_2 - R_1)$ as a function of ω_N^2 while the slope provides Φ if $\Delta\sigma_N$ is known (taken here as −170 ppm). From this determination of Φ for each residue, R_2 can be corrected for the exchange contribution (see Supporting Information). Obviously, such a correction which will proves essential requires multifield measurements.

Analysis without Considering Possible *csa* Variations. A first global fit (assuming $\tilde{J}_{csa} = \tilde{J}_d$) involving all relaxation parameters was attempted with a single (common to all residues) correlation time for describing the slow motions, *A* and *B* being determined for each residue according to the procedures described above (this approach will be denoted as "mode 1" in the following). The result (yielding a τ_s value of 6.26 ns, consistent with a previous work[9]) was rather disappointing with regard to recalculated values when compared to the experimental ones; this is quantified by the standard deviation (denoted as sd in the following) equal here to 0.241 s^{-1}. The situation is much improved by ignoring all transverse relaxation rates (this will be denoted as "mode 2" in the following) and under these circumstances we arrive at a tumbling correlation time, τ_s, equal to 4.90 ns ($sd = 0.071$ s^{-1}, obtained from experimental and calculated values of only R_1 and σ_{NH}). However, the agreement between recalculated and experimental values remained rather poor for residues 1, 3, 4, 5 (2 and 6 are proline residues)—the disordered N-terminus of the protein—and from residue 54—ending the very first turn of helix III—down to the C-terminal end of the protein.

Therefore, at this stage, only data between residue 7 and residue 53 were considered, still keeping a global description of slow motions (the same τ_s for all residues) while fitting the coefficients *A* and *B* for each residue. From the 3D structure, it can be seen that these residues belong to the core of the protein, including the α-hairpin and the first N-terminal turn of helix III. None of these residues can accommodate spectral densities involving two correlation times associated with slow motion, that is, with a spectral density more complicated than eq 6; in that case, unrealistic or unacceptable results are obtained. Nevertheless we observed severe inconsistencies between cross-relaxation and transverse relaxation rates. This is illustrated in Figure 2 where relative deviations between experimental and recalculated values are displayed for R_1, R_2, and σ_{NH}. When transverse relaxation rates are used in the fitting procedure (mode 1), a value of 6.28 ns is found for τ_s ($sd = 0.25$); differences between experimental and recalculated σ_{NH} are considerably scattered and can be very important (relative deviations appear quite large due to the fact that σ_{NH} is around zero). By contrast, when transverse relaxation rates are not used (mode 2), we arrive at $\tau_s = 4.98$ ns, and the agreement between experimental and recalculated values ($sd = 0.042$ calculated again from only R_1 and σ_{NH}) is greatly improved for σ_{NH}, whereas recalculated R_2 undergo a shift toward lower values. It can be noted that Lee and Wand reached recently similar conclusions.[46] The correlation time is seen to be significantly lower than the one obtained with all R_2 values (mode 1), and this can be ascribed to the spectral densities at zero frequency which are involved only in the R_2 data. Finally, interpreting R_1 and σ_{NH} experimental values of the protein extremities necessitated to append the spectral density (eq 6) with a term of the form $C/1 + \omega^2\tau_s'^2$; the relevant residues are located either in the disordered N(residues 1−6)- and C(residues 64−68)-termini of the protein or below the first turn of helix III (residues 54−63). However as R_2 values were still excluded and owing to the difficulty in interpreting *C* and τ_s', this analysis was not pursued.

A radical change occurred, making questionable the above analyses, when, still using a spectral density such as eq 6, τ_s was allowed to be different from one residue to the other (this type of analysis will be denoted as "mode 3"). The whole set of data was used in the fitting procedure (including all of the R_2 values and residues outside the protein core). The remarkable quality of the results can be appreciated from the right column of Figure 2, confirmed by a *sd* value of 0.068 s^{-1}. Relative deviations (between experimental and recalculated values) are now similarly scattered for the three parameters, lying in a very reasonable range and without any bias. These observations suggest that the concept of effective correlation times seems perfectly adequate and allows for a realistic analysis of ^{15}N relaxation parameters in proteins. Raw parameters involved in eq 6 are displayed in Figure 3 for the whole protein. *A*, as already mentioned, is associated with fast local motions and is consequently very small. Owing to the large uncertainties which affect this parameter, any discussion about its variations would be illusory; the only point worth noting is an increase toward the protein extremities (from residue 7 to the N-terminal on one hand, and from the beginning of helix III to the C-terminal on the other hand). Conversely, *B* and τ_s seem to be good indicators of protein backbone dynamics and appear to be related to the secondary structure. τ_s is as expected: smaller at the protein extremities (where a slow motion yet faster than the overall tumbling is likely to occur) and, in the protein

(40) Greziek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593.
(41) Pons, J. L.; Malliavin, T. E.; Delsuc, M. A. *J. Biomol. NMR* **1996**, *8*, 445.
(42) Skelton, N. J.; Palmer, A. G., III; Akke, M.; Kördel, J.; Rance, M.; Cahzin, W. J. *J. Magn. Reson.* **1993**, *B102*, 253.
(43) Press, W. H.; Flannery, B. P.; Teulkolsky, S. A.; Vetterling, W. T. *Numerical Recipies*; Cambridge University Press: Cambridge, 1986.
(44) Peng, J. W.; Wagner, G. *Biochemistry* **1995**, *34*, 16733.
(45) Habazettl, J.; Wagner, G. *J. Magn. Reson.* **1995**, *B109*, 100.
(46) Lee, A. L.; Wand, A. J. *J. Biomol. NMR* **1999**, *13*, 101.

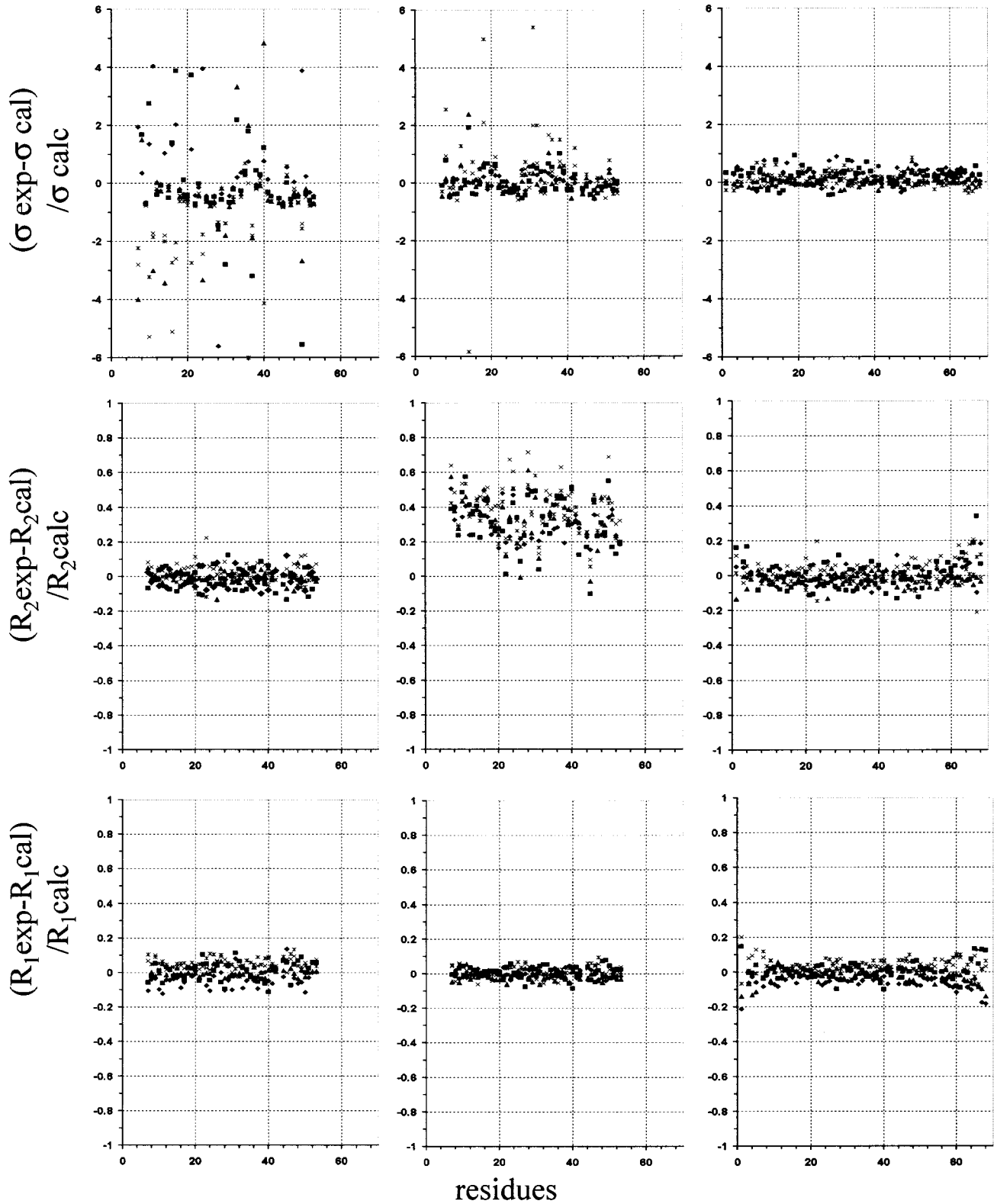

Figure 2. Relative deviations between experimental and recalculated values of R_1, R_2, and σ_{NH}. (Left) τ_s fitted from all relaxation parameters (mode 1 and residues belonging to the protein core only; extremities not used in the fit). (Middle) τ_s fitted without the transverse relaxation rates (mode 2 and residues belonging to the protein core only; extremities not used in the fit). (Right) τ_s fitted *residue-by-residue* from all relaxation parameters (mode 3). (Diamonds) 9.4 T data; (squares) 11.75 T; (triangles) 14.1 T data; (crosses) 16.45 T; (stars) 18.8 T. In this figure and in the following figures, 2, 6, and 43 do not appear as they are proline residues whereas data are not available for residues 55, 56, 61, 67 because of peak overlaps.

core, fluctuating in a narrow range corresponding to 6 ± 0.36 ns (these fluctuations presumably reflect the different orientations of *NH* vectors with respect to the overall rotation-diffusion tensor). The way in which slow motions affect each residue can therefore be quantified. Rather than discussing the parameter B itself, it may be more judicious to rely on the order parameter derived from B (eq 8). The relevant profile is shown in Figure 4 (top) and is globally sound. Despite large uncertainties at the two protein extremities, the square of the order parameter, S^2, is seen to decrease sharply in these two regions while it remains at a high level within the protein core (residues 7 to 53). Its

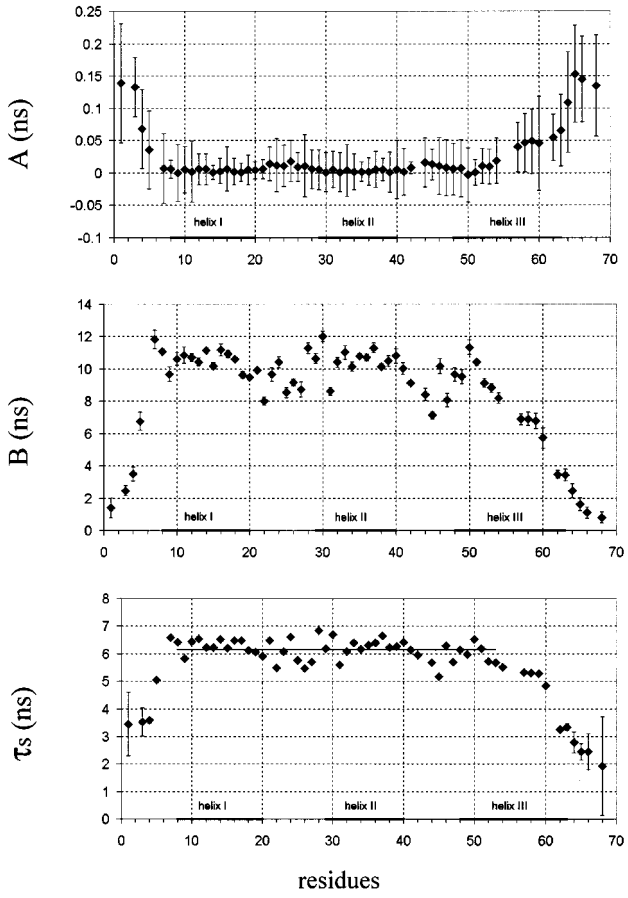

Figure 3. The raw parameters extracted from the analysis, residue-by-residue, of all the data: *A*, *B*, and τ_s (see eq 6).

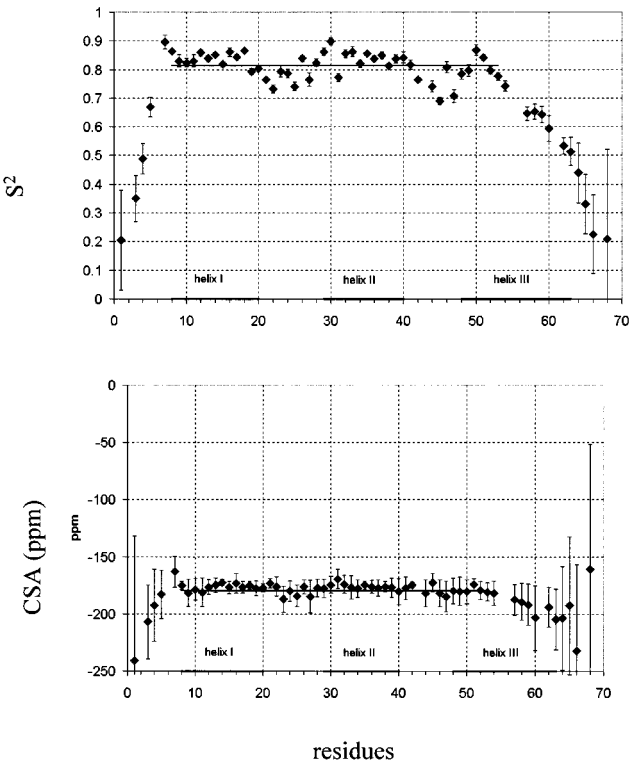

Figure 4. (Top) Profile of the order parameters S^2 determined from eq 8. The horizontal line corresponds to the mean value calculated over the protein core results. (Bottom) ^{15}N chemical shift anisotropy values determined from k_{csa} and *B* (both quantities being derived directly from experimental data; see eq 11). The horizontal line corresponds to the mean value calculated over the protein core.

variations in this region, albeit weak, are significant. The horizontal line in Figure 4 (top) represents its mean value over all of the residues in the protein core. It can be recognized that S^2 is systematically above this mean value for helices I and II, confirming their high degree of organization. S^2 decreases between residues 20 and 29, a region which corresponds to the loop which links the two helices. This occurs again between residues 41 and 47 for the link between helices II and III. It can also be observed that that a small part of helix III is well organized (up to residue 53) and that S^2 drops rapidly for the rest of helix III. These results are quite consistent with the protein structure and therefore support the present analysis.

Analysis Including Possible *csa* Variations . As explained in the Data Analysis section, the abundance of experimentally determined parameters makes possible the determination of ^{15}N *csa* values provided that some simplifying assumptions are made, that is, that an axially symmetric shielding tensor is assumed and that we postulate similar dynamical properties for the *NH* vector and the symmetry axis of the shielding tensor. It turns out that the values found for the other parameters, namely *A*, *B*, and τ_s, are only marginally modified when the *csa* is allowed to vary, even though the agreement between recalculated and experimental relaxation rates is slightly improved (*sd* $= 0.058\ s^{-1}$). The *csa* values obtained through the analysis based on effective correlation times are shown in Figure 4 (bottom). Its examination reveals that no significant variation of the ^{15}N *csa* can be detected within the displayed uncertainties (calculated according to the statistical treatment explained above; eq 15). It can be especially noted that large uncertainties occur for those residues presenting an unusual *csa* value. Still in Figure 4 (bottom), the horizontal line stands for the mean value calculated over the protein core (−180 ppm). It can be appreciated that this line virtually intersects *all* of the error bars. Now, this value of −180 ppm may appear too large. It can however be borne in mind that *csa* values are derived by dividing u_{csa} (directly determined from the fit) by *B* (see eq 11), the actual value of which depends on the chosen NH distance through the constant K_d (see eqs 3−6 and 1). It is easy to see that a 1% modification of d_{NH} entails a 3% modification of *csa* values. Thus, assuming for instance 1.04 Å for d_{NH} instead 1.02 Å decreases the *csa* mean value to ∼−170 ppm. Variations of the *csa* value can also be ascribed to the hypothesis made about the order parameter (which has been assumed to be identical to the one of the *NH* direction). To a rough approximation, these variations can be accounted for by a factor equal to $(3\cos^2\beta - 1)/2$ where β is the angle between the *NH* direction and the symmetry axis of the shielding tensor. As β cannot exceed 20°, the error on the *csa* value associated with the order parameter is less than 7%.

For the sake of completeness, one may wonder if the determination of *csa* in addition to *A*, *B*, and τ_s would not improve the quality of the fit when one relies on a single correlation time for describing the slow motion. Calculations were thus repeated for the other two modes of analysis (restricted to the core of the protein): mode 1 (a global τ_s with all of the relaxation data used in the fitting process), and mode 2 (a global τ_s without the R_2 relaxation rates in the fitting process). Again, the parameters *A*, *B*, and τ_s are almost unchanged ($\tau_s = 6.26$ ns for mode 1 and 4.79 ns for mode 2), whereas standard deviations drop significantly (0.08 and $0.036\ s^{-1}$ for modes 1 and 2, respectively). However, relative deviations between experimental and recalculated values exhibit the same trends as before (see Figure 2). This is shown in Figure 5: for mode 1, σ_{NH} relative deviations are badly scattered whereas, for mode 2, R_2 relative deviations are not only scattered but also displaced from the zero mean value. The only satisfactory fit corresponds again to the treatment based on an effective correlation time for each residue (mode 3; Figure 5, right).

Csa values deduced from these three modes are displayed in Figure 6 left (the bottom of which is identical to the bottom of Figure 4, restricted to the protein core). It can be observed that modes 1 and 3 (all data used, including transverse relaxation rates) yield virtually the same results yet with larger error bars for mode 1. For mode 2, important variations of *csa* values, seemingly related to the protein secondary structure, are observed in agreement with the conclusions of Fushman et al.[10] who recently performed a similar study (using only

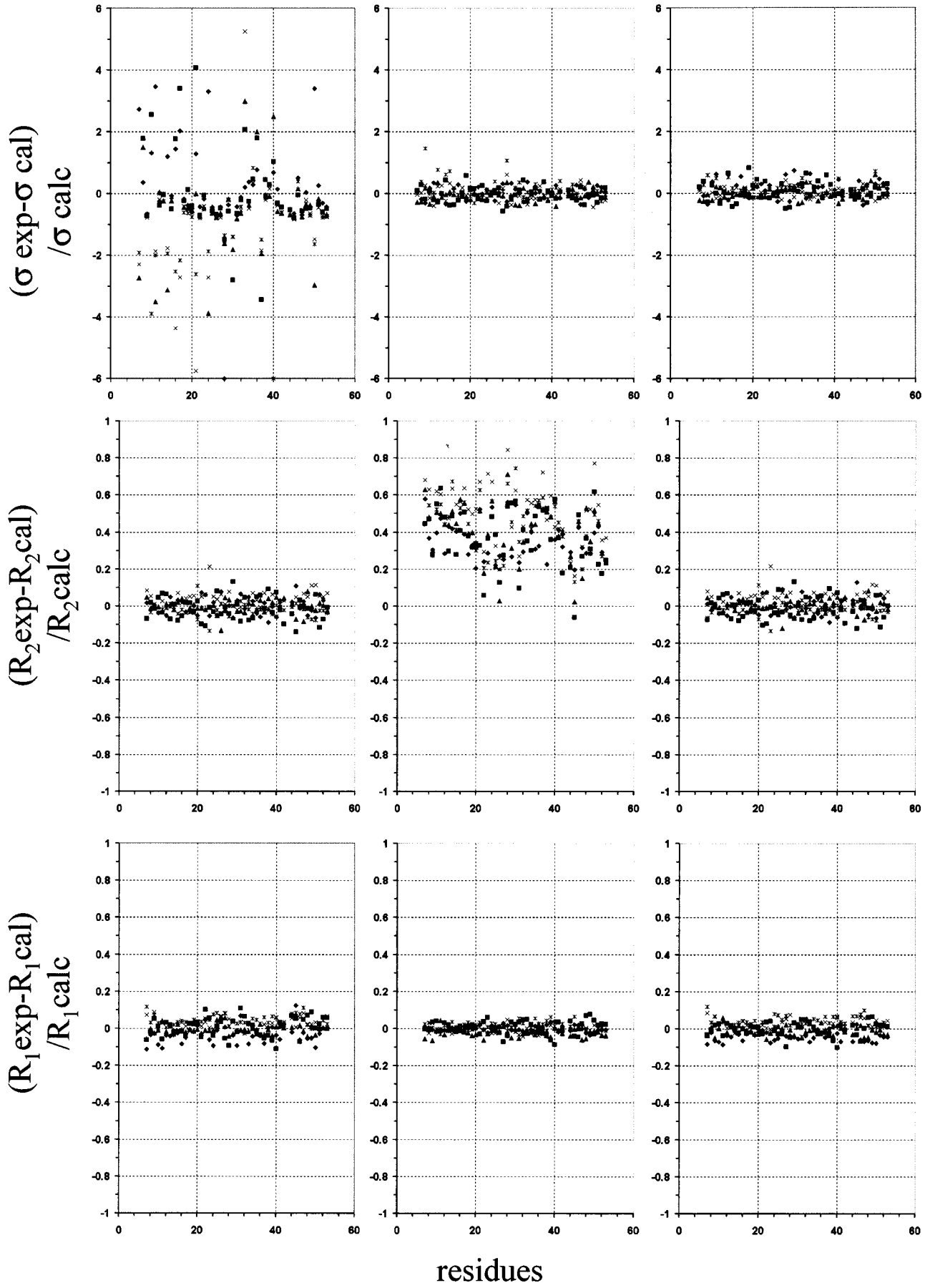

Figure 5. Same as Figure 2 when *csa* is allowed to vary for each residue. The representation is limited to the protein core.

the longitudinal relaxation rates for determining the *csa* value, ranging from −166 ppm to −231 ppm) on human ubiquitin. However, if we take into account the error bars (it can be reminded that they are derived here on a statiscal basis), results of mode 2 are seen to agree to a large

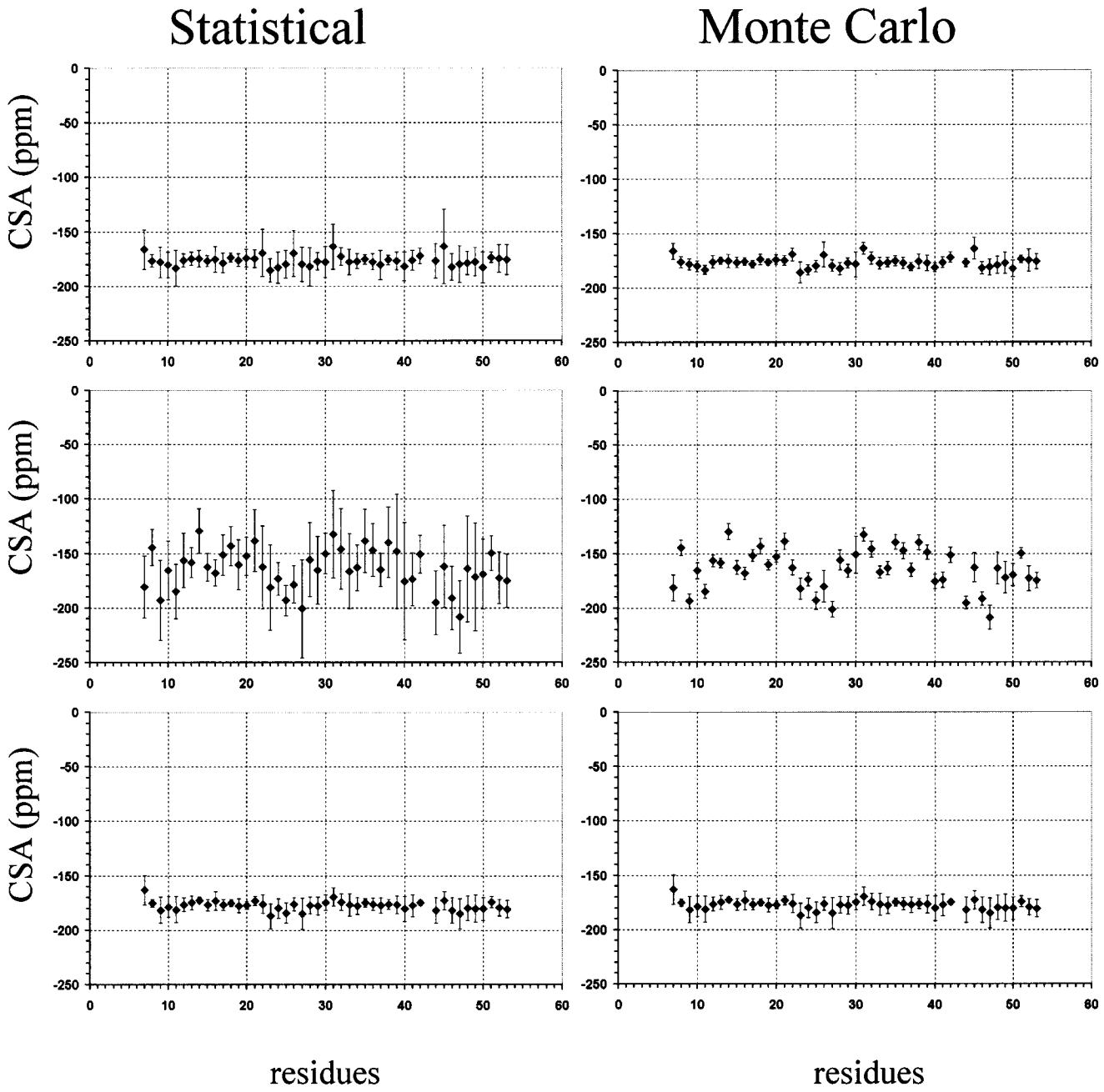

Figure 6. (Left) *csa* values (limited to the protein core) as obtained with the three mode of analysis. Error bars are deduced from a statistical treatment (see eq 15). (Top) All data with a single correlation time for describing the slow motion (mode 1). (Middle) R_1 and σ_{NH} data with a single correlation time for describing the slow motion (mode 2). (Bottom) All data with an effective correlation time for describing the slow motions at the level of each residue (mode 3). (Right) Same as left with error bars deduced from Monte Carlo calculations.

extent with those of mode 3 and this observation makes questionable any significant variation of the nitrogen *csa* value along the protein backbone.

To further investigate the validity of the above conclusions, all the calculations were repeated according to the Monte Carlo method. In each case, 500 data sets were generated by randomly perturbing the observables within their experimental uncertainties (plus or minus). The mean value of the parameters of interest was then calculated along with its error bar taken as the standard deviation of the corresponding (Gaussian) distribution (over the 500 data sets). First, it is reassuring (though not surprising) to obtain a mean value identical to the value derived from the original (not perturbed) data set and also error bars similar to the ones derived from the statistical treatment (eq 15), provided the latter are not too large. When this is not the case, error bars arising from the Monte Carlo treatment turned out to be much weaker. This is illustrated in Figure 6, where such a comparison is shown for the *csa* values. Of course, it would be tempting to rely rather on Monte Carlo error bars. However, it must be appreciated that such calculations ignore systematic deviations between recalculated and experimental values as well as the number of observables used in the fitting procedures, in contrast with the statistical treatment (see eq 15). For instance, the large discrepancies between recalculated and experimental cross-relaxation rates in mode 1 (see Figure 5) are not accounted for by the Monte Carlo method; likewise, in mode 2, the *csa* determinations involve a limited number of observables (the longitudinal relaxation rates) whereas, in mode 3, both longitudinal and transverse relaxation rates are used. Thus, error bars derived on a statistical basis are likely to be much more realistic and trustable than those obtained through a Monte Carlo procedure (see for instance figure 4 of ref 10).

Conclusions

The main findings of the present study can be summarized as follows: (i) defining a specific (effective) correlation time

for the slow motions sensed by each residue allows for a consistent interpretation of *all* relaxation rates (Figures 2 and 5) provided that transverse relaxation rates have been corrected for possible exchange contributions; it is remarkable that such an agreement is obtained for as many as five values of the magnetic field (although two values would be sufficient, exchange corrections to R_2 are better evaluated with additional measurements. It can be remembered that, as explained above, an effective correlation time is not in contradiction with other models such as the extended Lipari−Szabo model[24] or a treatment involving anisotropic overall reorientation;[25] (ii) the order parameter, determined in accordance with the "effective correlation time" concept, is well correlated with the protein structure; (iii) still relying on this concept, we were able to determine nitrogen chemical shift anisotropies which, in contrast with previous works,[10,15] do not show up any significant change along the protein backbone. It can be noted that such a conclusion is reached regardless of the mode of data analysis, provided that errors are properly evaluated (that is, according to a realistic statistical approach). This result can be related to recent theoretical studies[47] and would tend to indicate that, as expected, the secondary structure induces only small changes (if any) in nitrogen *csa* values.

Acknowledgment. We are grateful to Dominique Marion (IBS, Grenoble) for the use of the 800 MHz Varian spectrometer and to Jean-Pierre Simorre and Bernhard Brutscher for their help in running this machine. 700 MHz and 500 MHz data were obtained by Michael Salzmann (Bruker, Fällanden) and by Philippe Sizun (Sanofi-Synthelabo, Montpellier), respectively. We gratefully appreciate their cooperation.

Supporting Information Available: Table containing measured and exchange corrected R_2 values, diagram showing the Φ values involved in the exchange correction procedure, Table containing all experimental (exchange-corrected R_2) and recalculated relaxation parameters for each residue and quantities of interest (correlation times, order parameters, nitrogen shielding anisotropy) derived according to mode 3 (see text) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0038676

(47) Scheurer, C.; Skrynnikov, N. R.; Lienin, S. F.; Straus, S. K.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 8577−8582.

(48) Kraulis, P. J. *J. Appl. Crystallog.* **1991**, *24*, 946.