J|A|C|S
ARTICLES
Published on Web 05/24/2006

Variability of the ^{15}N Chemical Shielding Tensors in the B3 Domain of Protein G from ^{15}N Relaxation Measurements at Several Fields. Implications for Backbone Order Parameters

Jennifer B. Hall and David Fushman*

Contribution from the Department of Chemistry and Biochemistry, Center for Biomolecular Structure and Organization, University of Maryland, College Park, Maryland 20742

Received January 19, 2006; E-mail: fushman@umd.edu

Abstract: We applied a combination of ^{15}N relaxation and CSA/dipolar cross-correlation measurements at five magnetic fields (9.4, 11.7, 14.1, 16.4, and 18.8 T) to determine the ^{15}N chemical shielding tensors for backbone amides in protein G in solution. The data were analyzed using various model-independent approaches and those based on Lipari−Szabo approximation, all of them yielding similar results. The results indicate a range of site-specific values of the anisotropy (CSA) and orientation of the ^{15}N chemical shielding tensor, similar to those in ubiquitin (Fushman, et al. *J. Am. Chem. Soc.* **1998**, *120*, 10947; *J. Am. Chem. Soc.* **1999**, *121*, 8577). Assuming a Gaussian distribution of the ^{15}N CSA values, the mean anisotropy is −173.9 to −177.2 ppm (for 1.02 Å NH bond length) and the site-to-site CSA variability is ±17.6 to ±21.4 ppm, depending on the method used. This CSA variability is significantly larger than derived previously for ribonuclease H (Kroenke, et al. *J. Am. Chem. Soc.* **1999**, *121*, 10119) or recently, using "meta-analysis" for ubiquitin (Damberg, et al. *J. Am. Chem. Soc.* **2005**, *127*, 1995). Standard interpretation of ^{15}N relaxation studies of backbone dynamics in proteins involves an a priori assumption of a uniform ^{15}N CSA. We show that this assumption leads to a significant discrepancy between the order parameters obtained at different fields. Using the site-specific CSAs obtained from our study removes this discrepancy and allows simultaneous fit of relaxation data at all five fields to Lipari−Szabo spectral densities. These findings emphasize the necessity of taking into account the variability of ^{15}N CSA for accurate analysis of protein dynamics from ^{15}N relaxation measurements.

Introduction

The chemical shielding tensor (CST) reflects the local electronic environment of a nucleus under nuclear magnetic resonance observation and therefore contains valuable information on the local chemical structure and conformation of a molecule. Fast random molecular tumbling in solution averages the individual components of the tensor, so that only its trace, reflected in the isotropic chemical shift, is directly observed in high-resolution NMR spectra. Site-specific variations in the isotropic chemical shifts allow separation of NMR signals from various sites in macromolecules, and deviations of chemical shifts from their random coil values are widely used for secondary[1,2] and tertiary[3] structure predictions in proteins. The anisotropic components of the tensor contribute directly to nuclear spin relaxation; their knowledge is therefore essential for NMR applications to protein dynamics,[4−8] especially at higher field strengths, for the development of TROSY techniques to study large molecules,[9] and for the use of chemical shielding anisotropy for structure refinement.[10]

Understanding of the relationship between the chemical shielding tensor and protein structure is likely to facilitate the development of new approaches to structure prediction and to refine the theoretical models for chemical shielding calculations. Amide ^{15}N CSTs in proteins present a particular challenge, because they are susceptible to a variety of factors, including conformations (torsion angles) of both current and preceding residues, hydrogen bonding, solvent accessibility, and long-range electrostatics.[11−14]

The complete chemical shielding tensor could, in principle, be measured directly by solid-state NMR methods, and such studies provided valuable information on ^{15}N CSTs in short peptides.[15−21] However, applications of these techniques to uniformly labeled proteins are still in development. Recent

(1) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171−180.
(2) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(3) Cornilescu, G.; Delaglio, F.; Bax, A. *J Biomol. NMR* **1999**, *13*, 289−302.
(4) Tjandra, N.; Szabo, A.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6986−6991.
(5) Fushman, D.; Tjandra, N.; Cowburn, D. *J. Am. Chem. Soc.* **1998**, *120*, 10947−10952.
(6) Fushman, D.; Tjandra, N.; Cowburn, D. *J. Am. Chem. Soc.* **1999**, *121*, 8577−8582.
(7) Kroenke, C. D.; Loria, J. P.; Lee, L. K.; Rance, M.; Palmer, A. G. I. *J. Am. Chem. Soc.* **1998**, *120*, 7905−7915.
(8) Kroenke, C. D.; Rance, M.; Palmer, A. G. I. *J. Am. Chem. Soc.* **1999**, *121*, 10119−10125.
(9) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366−12371.
(10) Lipsitz, R. S.; Tjandra, N. *J. Magn. Reson.* **2003**, *164*, 171−176.
(11) Glushka, J.; Lee, M.; Coffin, S.; Cowburn, D. *J. Am. Chem. Soc.* **1989**, *111*, 7716−7722.
(12) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217−225.
(13) Sitkoff, D.; Case, D. *Prog. NMR Spectrosc.* **1998**, *32*, 165−190.
(14) Poon, A.; Birn, J.; Ramamoorthy, A. *J. Phys. Chem. B* **2004**, *108*, 16577−16585.

solution NMR approaches based on orientation-dependent changes in ^{15}N resonances in weakly aligned protein solutions[22−24] are very promising, although the accuracy and precision of these measurements do not yet allow site-specific determination of the CST values.

It has been demonstrated[4−6,8,25,26] that the anisotropy (CSA) of the ^{15}N CST can be directly obtained from ^{15}N relaxation measurements in proteins in solution. Measurements in ubiquitin[5,6] revealed a range of site-specific backbone ^{15}N CSA values, from approximately −120 to −220 ppm, with a mean of −157 ppm and a standard deviation (not the site-to-site variability) of 19 ppm. This range includes data for both conformationally well-defined amides and those located in the flexible regions. The angle between the unique axis of the ^{15}N CST and the NH bond was found to vary from 6° to 26°, with the mean of 15.7° and standard deviation (std) of 5°.[5,6] These findings were confirmed by independent relaxation studies in ubiquitin.[27] A higher in absolute value average CSA of −173 ppm (converted to an NH distance of 1.02 Å) with site-to-site variation of up to ±17 ppm was derived from shifts in peak positions in weakly aligned solutions of ubiquitin,[24] while recent MAS studies[28] of aligned ubiquitin in a similar medium yielded −162.0 ± 4.3 ppm for the mean ^{15}N CSA and 18.6° ± 0.5° for the angle, in agreement with those from previous ^{15}N relaxation data.[5,6] A similar range of site-specific ^{15}N CSA values (−129 to −213 ppm) was reported for ribonuclease H,[8] although with a somewhat different mean (−172 ppm), for a selection of well-ordered amides. For this subset of residues, the site-to-site variability in CSA was estimated to be ±5.5 ppm (upper limit ±9.6 ppm at 95% confidence), assuming a Gaussian distribution for ^{15}N CSA values. This number is relatively small, given the ∼30 ppm range of variation in the isotropic chemical shifts, and could be a result of the limited experimental precision in the CSA data, as the experimental uncertainties (±13 ppm) in the individual ^{15}N CSA values in that paper are noticeably larger than the reported variability. A recent study[29] combining new experimental measurements in ubiquitin with the literature data[5,27] resulted in an even more extreme mean ^{15}N CSA of −179.6 ppm (converted to NH distance of 1.02 Å) and a CSA variability comparable to that in ribonuclease H. However, the results of another recent study based on a combination of 14 auto- and cross-correlation rates in ubiquitin[30] agree with the earlier data[5,6] and give average CSAs ranging from −146.4 to −164.0 ppm and angles from 17.5° to 18.9°, depending on the choice of local motional model.

While the existence of some site-specific variability in ^{15}N CSA is now established, it still remains to be understood whether the differences between the reported data reflect some protein specificity of the CSA distribution or differences in the experimental approaches and/or in data analyses. Measurements in other proteins and with higher experimental precision are therefore required in order to address this issue.

It is even more important to understand the effect of site-specific variations in ^{15}N CSA on the motional characteristics of proteins derived from ^{15}N relaxation data, to improve the accuracy of NMR approaches to protein dynamics. Although computer simulations[31] show that ignoring the variability in CSA values could significantly affect the NMR-derived picture of the backbone dynamics, a direct experimental analysis of this issue has not been at hand.

Here we apply a combination of NMR relaxation and cross-correlation measurements at several magnetic fields to determine the ^{15}N chemical shielding tensors in a 56-amino acid protein, the third immunoglobulin-binding domain of protein G (further called GB3). We use several model-independent methods of data analysis to derive the ^{15}N CSA values and compare them with the values obtained assuming the Lipari−Szabo spectral densities. We then analyze the effect of these site-specific CSA values on the LS analysis of the backbone dynamics and on the derived order parameters in GB3.

Materials and Methods

Sample Preparation and NMR Measurements. The GB3 domain construct (56 a.a.) in these studies was the same as in ref 32. The protein was a generous gift from Dr. Ad Bax, NIH. All measurements were performed on the same protein sample containing 1.8 mM of uniformly ^{15}N-enriched GB3 dissolved in 30 mM phosphate buffer (pH 5.8) containing 9% D_2O. Sample temperature was set to 24 °C using a glycerol temperature calibration sample, with each spectrometer being individually calibrated.

Relaxation measurements included rates of ^{15}N longitudinal (R_1) and transverse (R_2) relaxation and the rate of ^{15}N−^{1}H cross-relaxation measured via steady-state ^{15}N{^{1}H} nuclear Overhauser effect (NOE). These experiments were performed at five magnetic fields of 9.4, 11.7, 14.1, 16.4, and 18.8 T and used standard pulse sequences (e.g., ref 33). The NOEs were determined using a flip-back measurement scheme[34] for water suppression, and the recycling delay was 4−5 s (see Supporting Information Table 1 for complete list of relaxation delays). The R_1, R_2, and NOE measurements at 9.4 T were performed twice, on different instruments (at UMD and at CERM), yielding similar results. ^{15}N CSA/^{15}N−^{1}H dipolar cross-correlation measurements were performed using the method described in refs 35 and 36. Transverse cross-correlation rates (η_{xy}) were measured at 9.4, 11.7, 14.1, and 18.8 T, while longitudinal cross-correlation (η_z) was measured at 9.4, 11.7, and 14.1 T. To verify that the η_z values were not affected by dipolar cross-relaxation of proton magnetization, the η_z measurements at 14.1 T were repeated on a perdeuterated ^{15}N-labeled GB3 sample and yielded the same results as for the protonated sample (Hall and Fushman,

(15) Harbison, G. S.; Jelinski, L. W.; Stark, R. E.; Torchia, D. A.; Herzfeld, J.; Griffin, R. G. *J. Magn. Reson.* **1984**, *60*, 79−82.
(16) Hartzell, C. J.; Whitfield, M.; Oas, T. G.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5966−5969.
(17) Oas, T. G.; Hartzell, C. J.; Dahlquist, F. W.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5962−5966.
(18) Hiyama, Y.; Niu, C.; Silverton, J.; Bavoso, A.; Torchia, D. *J. Am. Chem. Soc.* **1988**, *110*, 2378−2383.
(19) Shoji, A.; Ozaki, T.; Fujito, T.; Deguchi, K.; Ando, S.; Ando, I. *Macromolecules* **1989**, *22*, 2860−2863.
(20) Mai, W.; Hu, W.; Wang, C.; Cross, T. A. *Protein Sci.* **1993**, *2*, 532−542.
(21) Wu, C. H.; Ramamoorthy, A.; Gierasch, L. M.; Opella, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 6148−6149.
(22) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836−6837.
(23) Boyd, J.; Redfield, C. *J. Am. Chem. Soc.* **1999**, *121*, 7441−7442.
(24) Cornilescu, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 10143−10154.
(25) Fushman, D.; Cowburn, D. *J. Am. Chem. Soc.* **1998**, *120*, 7109−7110.
(26) Damberg, P.; Jarvet, J.; Allard, P.; Graslund, A. *J. Biomol. NMR* **1999**, *15*, 27−37.
(27) Kover, K. E.; Batta, G. *J. Magn. Reson.* **2001**, *150*, 137−146.
(28) Kurita, J.; Shimahara, H.; Utsunomiya-Tate, N.; Tate, S. *J. Magn. Reson.* **2003**, *163*, 163−173.
(29) Damberg, P.; Jarvet, J.; Graslund, A. *J. Am. Chem. Soc.* **2005**, *127*, 1995−2005.
(30) Loth, K.; Pelupessy, P.; Bodenhausen, G. *J. Am. Chem. Soc.* **2005**, *127*, 6062−6068.
(31) Fushman, D.; Cowburn, D. In *Methods in Enzymology*; James, T., Schmitz, U., Doetsch, V., Eds.; 2001; Vol. 339, pp 109−126.
(32) Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, *27*, 261−275.
(33) Fushman, D.; Cahill, S.; Cowburn, D. *J. Mol. Biol.* **1997**, *266*, 173−194.
(34) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593−12594.
(35) Hall, J. B.; Dayie, K. T.; Fushman, D. *J. Biomol. NMR* **2003**, *26*, 181−186.
(36) Hall, J. B.; Fushman, D. *Magn. Reson. Chem.* **2003**, *41*, 837−842.

manuscript in preparation). Note also that the η_{xy} values at 11.7 and 14.1 T were also measured using the spin-state selection method, yielding similar results.[37]

The spectra were recorded in an interleaved fashion, as detailed in ref 32 and then processed using XWINNMR. Further analysis including automatic peak picking and integration, relaxation curve fitting, and data analysis was performed using an in-house suite of Matlab programs. The program DYNAMICS[32,33] was modified to include the site-specific ^{15}N CSA as an additional fitting parameter.

Auto-relaxation and cross-correlation rates were obtained from least-squares fitting of peak intensities in the corresponding series of 2D spectra to a monoexponential decay. The heteronuclear NOE values were obtained from the ratio of peak intensities in the NOE and NONOE experiments.[33] Experimental errors in peak intensities were estimated in two ways:[38] by integrating regions of spectra containing no cross-peaks or, where applicable, from repeated (quadruplicate) measurements, using the method of ref 39. The errors in the rates were estimated using a Monte Carlo simulation of 500 experimental data sets per residue and assuming a normal distribution of experimental errors in peak intensities. The experimental errors in relaxation rates were around 1% on average: 1.16%, 0.83%, 1.43%, 1.09%, and 1.37% for R_1; 1.21%, 1.21%, 1.33%, 0.96%, and 1.30% for R_2, and 1.13%, 1.14%, 1.05%, 1.00%, and 1.06% for NOE values measured at 9.4, 11.7, 14.1, 16.4, and 18.8 T, respectively. The average errors in η_{xy} were 1.37%, 1.50%, 1.67%, and 1.47% at 9.4, 11.7, 14.1, and 18.8 T, respectively; the errors in η_z were 1.27%, 1.16%, and 1.52% at 9.4, 11.7, and 14.1 T.

Determination of ^{15}N CSA and the Backbone Dynamics from the Experimental Data. The ^{15}N chemical shielding anisotropies were derived from the measured relaxation and cross-correlation rates using five different methods, outlined below.

(A) Model-Independent Methods. (1) The R/η Method. This method is a generalization of that of ref 5 and is based on the fact that the ratio of the corresponding cross-correlation and auto-relaxation rates is independent, to a good approximation, of the spectral densities $J(\omega)$:[7,25]

$$\frac{R_2'}{\eta_{xy}} = \frac{R_1'}{\eta_z} = \frac{d^2 + c^2}{2dc_g} \tag{1}$$

Here $d = -\mu_0\gamma_H\gamma_N\hbar/(8\pi r_{HN}^3)$ is the strength of the $^{15}N-^1H$ dipolar coupling and $c = \gamma_N B_o \Delta\sigma/3 = -\omega_N\cdot\Delta\sigma/3$ and $c_g = -\omega_N\cdot\Delta\sigma_g/3$ represent the ^{15}N CSA contributions to auto-relaxation and cross-correlation rates, respectively, where[40]

$$|\Delta\sigma| = [\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - (\sigma_{xx}\sigma_{yy} + \sigma_{xx}\sigma_{zz} + \sigma_{yy}\sigma_{zz})]^{1/2} \tag{2}$$

$$\Delta\sigma_g = (\sigma_{zz} - \sigma_{yy})P_2(\cos\beta_z) + (\sigma_{xx} - \sigma_{yy})P_2(\cos\beta_x) \tag{3}$$

σ_{ii} are the principal values of the ^{15}N CST. $P_2(x)$ is the Legendre polynomial, and β_z, β_x are the intervening angles between the principal axes (z, x) of the ^{15}N CST and the N−H bond vector. r_{HN} is the internuclear distance (here assumed to be 1.02 Å for all backbone amides), γ_H, γ_N, and ω_H, ω_N are the gyromagnetic ratios and the absolute values of the Larmor frequencies, respectively, of ^1H and ^{15}N, and μ_o is the permeability of vacuum. $\Delta\sigma$ has the meaning of the effective anisotropy of ^{15}N CST and will be referred to as the ^{15}N CSA throughout this paper; $\Delta\sigma_g$ has the meaning of a "projection" of the CSA tensor onto the NH vector and can be represented as $\Delta\sigma$ times an orientation factor. Note that here we use the convention that $\sigma_{zz} \le \sigma_{yy} \le \sigma_{xx}$ and define the principal axes of the ^{15}N CST such that its z-axis corresponds to the least shielded component (σ_{zz}), i.e., is close in orientation to the NH bond. The other two axes are then defined such that the y-axis is approximately orthogonal to the peptide plane, and the x-axis located approximately in-plane. Under the assumption of an axial symmetry of the ^{15}N CST ($\sigma_{xx} = \sigma_{yy} = \sigma_\perp$, $\sigma_{zz} = \sigma_\parallel$), eqs 2 and 3 simplify into their more "conventional" form (e.g., ref 25):

$$\Delta\sigma = \sigma_\parallel - \sigma_\perp \text{ and } \Delta\sigma_g = \Delta\sigma\cdot P_2(\cos\beta_z) \tag{4}$$

The primes in eq 1 and throughout this paper indicate "reduced" relaxation rates: the contributions from high-frequency components of the spectral density were subtracted as follows:[6]

$$R_1' = R_1[1 - 1.249|\gamma_N/\gamma_H|(1 - \text{NOE})] = 3(d^2 + c^2)J(\omega_N) \tag{5}$$

$$R_2' = R_2 - 1.079|\gamma_N/\gamma_H|R_1(1 - \text{NOE}) = 0.5(d^2 + c^2)[4J(0) + 3J(\omega_N)] \tag{6}$$

Equation 1 can be recast to yield a linear dependence on ω_N^2,

$$\frac{2\omega_N R_2'}{\eta_{xy}} = \frac{2\omega_N R_1'}{\eta_z} = -\frac{3d}{\Delta\sigma_g} - \frac{(\Delta\sigma)^2}{3d\cdot\Delta\sigma_g}\omega_N^2 \tag{7}$$

which can then be fit to a straight line, $m\cdot x + b$ (where $x = \omega_N^2$), using a simple linear regression. This form allows a direct determination of $\Delta\sigma_g$ and $\Delta\sigma$ from the intercept b and slope m of this line:

$$\Delta\sigma_g = -3d/b \tag{8}$$

$$\Delta\sigma = -3d(m/b)^{1/2} \tag{9}$$

The choice of the sign in eq 9 reflects negative ^{15}N CSA, according to solid-state NMR data. For an axially symmetric ^{15}N CST this gives[5,6] (cf. eq 4) $\sigma_\parallel - \sigma_\perp = -3d(m/b)^{1/2}$ and $P_2(\cos\beta_z) = (m\cdot b)^{-1/2}$.

It should be mentioned here that the CSA parameters ($\Delta\sigma$, $\Delta\sigma_g$) obtained using this method are *independent* of the motional characteristics of the molecule, as discussed in ref 25.

(2) The $2R_2 - R_1$ Method. This method is based on a quadratic field dependence of the following combination of the auto-relaxation rates (e.g., ref 6),

$$2R_2' - R_1' = 4d^2J(0) + (4/9)J(0)(\Delta\sigma)^2\omega_N^2 \tag{10}$$

which allows a direct determination of $J(0)$ and $\Delta\sigma$ from the intercept b and the slope m of the line $m\cdot\omega_N^2 + b$ representing a linear dependence of $2R_2' - R_1'$ on ω_N^2:

$$J(0) = b/(4d^2) \tag{11}$$

$$\Delta\sigma = -3d(m/b)^{1/2} \tag{12}$$

In this method, the spectral density $J(0)$ is determined solely from the intercept of the fitting line and, therefore, is *independent* of the ^{15}N CSA. We assume throughout this paper that the conformational exchange contribution to R_2 is negligible, which holds for all residues in GB3 except possibly Val39.[32] When present, a conformational exchange contribution (in the case of fast exchange) has the same field dependence as the $(\Delta\sigma)^2$ term (e.g., eqs 7, 10), and special care is required in order to separate them.[6,7]

(3) The $2\eta_{xy} - \eta_z$ Method. This method utilizes a linear field dependence of the following combination of the cross-correlation rates:

$$2\eta_{xy} - \eta_z = -(8/3)d\Delta\sigma_g\cdot J(0)\omega_N = m\cdot\omega_N \tag{13}$$

which allows determination of the product, $\Delta\sigma_g\cdot J(0)$, directly from the slope m of the fitting line with zero intercept:

$$\Delta\sigma_g\cdot J(0) = -m\cdot 3/(8d) \tag{14}$$

(37) Vasos, P. R.; Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2005**, *31*, 149−154.
(38) Note that the results previously presented in ref 32 were obtained using experimental uncertainties estimated from noise integration.
(39) Skelton, N.; Palmer, A.; Akke, M.; Kordel, J.; Rance, M.; Chazin, W. *J. Magn. Reson.* **1993**, *B102*, 253−264.
(40) Canet, D. *Concepts Magn. Reson.* **1998**, *10*, 291−297.

Table 1. Statistics of the ^{15}N CSA Values in GB3 Determined Here Using Various Methods

method	analyzed set of residues	number of residues	max.$(\Delta\sigma)^a$ (ppm)	min.$(\Delta\sigma)^b$ (ppm)	$\langle\Delta\sigma\rangle^c$ (ppm)	μ^d (ppm)	median^e (ppm)	$\langle\delta\Delta\sigma\rangle^f$ (ppm)	std $(\Delta\sigma)^g$ (ppm)	Λ^h (ppm)
$2R_2 - R_1$	all[i]	47	−111.3	−241.0	−174.2	−173.9	−175.4	6.0	22.2	21.4
	χ^2/df_{fit} < 95% cutoff	32	−154.0	−207.0	−178.1	−178.2	−178.9	7.0	12.9	10.6
	α-helix[j]	11	−140.4	−198.2	−175.8	−176.4	−177.0	7.6	18.1	14.1
	β-strands[j]	19	−154.0	−241.0	−180.3	−180.2	−177.5	7.3	19.1	16.3
R/η	all[i]	44	−127.9	−237.9	−177.4	−177.2	−178.4	7.5	19.5	17.6
	χ^2/df_{fit} < 95% cutoff	33	−155.7	−203.5	−178.2	−178.2	−178.3	7.8	12.5	10.2
	α-helix[j]	11	−141.6	−203.5	−177.6	−179.3	−178.3	9.2	16.7	8.3
	β-strands[j]	19	−159.2	−237.9	−181.1	−180.7	−178.5	7.5	18.3	14.7
LS-CSA	all[i]	32	−126.0	−243.4	−176.9	−176.9	−176.8	3.1	20.0	19.2
	χ^2/df_{fit} < 95% cutoff	25	−158.1	−201.9	−178.3	−178.3	−177.2	3.3	12.6	11.9
	α-helix[j]	11	−126.0	−196.9	−174.3	−174.3	−180.5	3.4	21.3	19.9
	β-strands[j]	16	−159.3	−243.4	−180.7	−180.6	−175.9	3.1	20.9	19.6
average of all 3 methods	all[i]	50	−111.3	−240.8	−174.2	−173.8	−175.9	7.1	22.2	21.2
	χ^2/df_{fit} < cutoff	35	−155.7	−203.4	−177.7	−177.2	−178.3	7.9	11.9	9.1
	α-helix[j]	11	−136.0	−196.3	−176.0	−177.3	−184.6	9.2	18.1	12.0
	β-strands[j]	20	−159.8	−240.8	−180.3	−179.9	−177.8	8.2	18.6	14.5

[a] Smallest absolute value of the ^{15}N CSA. [b] Largest absolute value of the ^{15}N CSA. [c] Arithmetic mean of measured values of the ^{15}N CSA. [d] Value of μ that maximizes the likelihood function $p(\mu,\Lambda)$ (eq 19); μ is an estimate of the true mean of the CSA distribution. [e] Median of measured values of the ^{15}N CSA. [f] Arithmetic mean of experimental uncertainties in the ^{15}N CSA. [g] Standard deviation of the measured values of the ^{15}N CSA. [h] Value of Λ that maximizes the likelihood function $p(\mu,\Lambda)$; Λ is an estimate of the true site-to-site variability in the CSA distribution. [i] All residues with acceptable agreement of regression methods (out of 50 analyzable residues, see text). [j] The α-helix in GB3 extends from Ala23 to Asp36 with Thr25, Glu27, and Asn35 impossible to resolve in the spectra due to overlap (hence 11 analyzable residues). The β-strands comprise Tyr3-Ile7, Gly14-Lys19, Val42-Asp46, and Thr51−Thr55, with Glu15 excluded due to overlap (altogether 20 analyzable residues). Gln2 was excluded from the LS analyses (see text).

This approach has the advantage over the above-mentioned methods in that (1) it is not affected by the possible conformational exchange contribution to R_2 and (2) it does not require correction for the high-frequency components of the spectral density (cf. eqs 5, 6). The drawback is that it does not allow separate determination of $\Delta\sigma_g$ and $J(0)$. If one of these parameters is known (e.g., $J(0)$ from the $2R_2 - R_1$ method), then the other one (in this case, $\Delta\sigma_g$) can be directly obtained from eq 14.

(B) Analyses of the Relaxation Data Using the Lipari−Szabo Approximation. While the methods outlined above are truly independent of the model of local and overall motion, the following approaches use a specific, so-called "model-free" or Lipari−Szabo (LS) form of the spectral density function[41−43] that describes the backbone dynamics in terms of an order parameter S and a correlation time τ_{loc} of local motion.

(1) "Standard" Lipari−Szabo Approach (LS). The now standard, LS-type analysis of the relaxation data (R_1, R_2, NOE) (see e.g. refs 33, 44) was performed using the program DYNAMICS and assuming a uniform ^{15}N CSA value, as described in refs 32 and 33. Up to eight motional models (listed in ref 33) were considered per residue, depending on the number of available observables. The overall tumbling of GB3 was assumed anisotropic, described by the average diffusion tensor shown in Table 2. For amides in the loop regions, where the NH-vector orientation is less well defined than in the elements of secondary structure, we adopted a conservative approach, in that the overall tumbling was assumed isotropic, to avoid bias by a particular loop conformation captured in the crystal structure. Using the anisotropic diffusion model and crystal structure coordinates for residues in the loop regions resulted in slightly different values of the order parameters[32] but did not alter the conclusions of the analysis. The same approach was also adopted for the other LS-based models throughout this paper.

(2) Lipari−Szabo Approach Including CSA (LS-CSA). This approach is an extension of the "standard" LS analysis of the relaxation data (R_1, R_2, NOE) (see above) that here includes site-specific ^{15}N CSA ($\Delta\sigma$) as an additional adjustable parameter. The LS-CSA method, therefore, yields $\Delta\sigma$ and the conventional LS parameters (e.g., S^2, τ_{loc}) and possibly R_{ex}, depending on the model selection for local dynamics. Up to eight motional models (listed in ref 33) were considered per residue, depending on the number of available observables. For these purposes, the recent version of our computer program DYNAMICS[32] that already accounts for the overall rotational anisotropy was upgraded to include $\Delta\sigma$ as an additional fitting parameter in a simplex-based optimization. Previously, a similar type of inclusion of CSA in the derivation of the LS parameters has been used to assess the accuracy of overall rotational diffusion parameters[45] and for simultaneous analysis of single-field relaxation and cross-correlation data.[27]

The robustness of this procedure of deriving $\Delta\sigma$ was tested on 1000 sets (per model) of synthetic relaxation data (R_1, R_2, NOE at the five field strengths) containing 1% "experimental" noise. The range of the input parameters for the simple LS models was as follows: S^2 from 0.6 to 1, τ_{loc} from 0 to 100 ps (typical range of values for elements of secondary structure), and $\Delta\sigma$ from −100 to −300 ppm. The output order parameters and the $\Delta\sigma$ were within 4.38% (mean 0.004%, std 1.08%) and 6.68% (mean −0.012%, std 1.71%), respectively, from their input values, although only 94.9% of the data could be fit to within a 95% confidence level with this level of noise. In the case of the "extended model-free" model,[43] the fast dynamics were characterized by S_{fast}^2 from 0.7 to 1 (with $S^2 = S_{slow}^2 \cdot S_{fast}^2 < S_{fast}^2$) and τ_{fast} from 0 to 100 ps, while the slow motions had S_{slow}^2 from 0.6 to 1 and τ_{slow} from 200 to 500 ps. Here the output order parameters and the $\Delta\sigma$ were within 4.78% (mean 0.008%, std 1.11%) and 8.93% (mean −0.02%, std 1.86%), respectively, from their input values, and 95.9% of the data could be fit to within a 95% confidence level with this level of noise. No R_{ex} contributions to R_2 were included in the simulation. From these analyses, we concluded that the order parameter and CSA could be fit to within reasonable uncertainty with the existing errors in the experimental relaxation data.

(3) Lipari−Szabo Analysis of Spectral Densities Directly (LS-SDF). The CSA values were also derived by simultaneous fitting of the spectral densities measured at all five fields to a LS spectral density, $J_{LS}(\omega)$,[41] that describes local dynamics in terms of S^2 and τ_{loc} only. The $J_{LS}(\omega)$ values included the effect of the overall rotational anisotropy,[46,47] calculated from the diffusion tensor characteristics (Table 2) and the orientation of a given NH vector reconstructed according to

(41) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(42) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.
(43) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4936.
(44) Mandel, A. M.; Akke, M.; Palmer, A. G. I. *J. Mol. Biol.* **1995**, *246*, 144−163.
(45) Lee, A. L.; Wand, A. J. *J. Biomol. NMR* **1999**, *13*, 101−112.
(46) Woessner, D. *J. Chem. Phys.* **1962**, *37*, 647−654.

the crystal structure of GB3 (1IGD.pdb). For each residue, the experimental values of the spectral density function $J(\omega)$ at $\omega = 0$, ω_N, and $0.87\omega_H$ were directly derived from the relaxation data (R_1, R_2, NOE) at each field strength using the reduced spectral density approximation,[48,49] as follows:

$$J(0.87\omega_H) = |\gamma_N/\gamma_H|(1 - \mathrm{NOE})R_1/(5d^2) \quad (15)$$

$$J(\omega_N) = \frac{R_1 - 7(0.87/0.921)^2 d^2 J(0.87\omega_H)}{3(d^2 + c^2)} \quad (16)$$

$$J(0) = \frac{2R_2 - R_1 - 6(0.87)^2 d^2 J(0.87\omega_H)}{4(d^2 + c^2)} \quad (17)$$

Altogether this resulted in 15 values of $J(\omega)$ per residue, five of which were $J(0)$ values derived from different-field measurements and which are expected to be the same within experimental precision. The LS parameters (S^2, τ_{loc}) and the ^{15}N CSA value for each residue were obtained from an unconstrained nonlinear minimization of the following target function:

$$\chi_{LS}^2 = \sum_i \left[\frac{J(\omega_i) - J_{LS}(\omega_i)}{\delta J_i}\right]^2 \quad (18)$$

where the sum is over all available ω_i values for a given residue, and δJ_i represents the experimental error in $J(\omega_i)$. This method is analogous to the "classical" LS analysis except that reduced spectral densities are being used and the CSA is an additional fitting parameter.

Robust Analysis of Data. The methods described above usually rely on a least-squares fit of experimental data. Given the small number of available experimental data points per residue, the results of this fit are susceptible to experimental errors. Measures were taken to ensure that the conditions of each experiment were identical within practical limits; however, there are outlying data points in several residues, as can be seen, for example, from the linear regression plots (Supporting Information). These deviations do not seem to come from the random noise in the spectra, but rather are a result of spectral artifacts caused by baseline drift, water suppression problems, etc., the distribution of which is unknown and cannot be readily determined from the small sample of measurements. Least-squares fits (including linear regression) are particularly susceptible to outliers,[50−52] as their contributions to the target function increase as a square of the deviation from the fitting curve. In light of this, for each method of deriving the CSA, in addition to the "standard" least-squares regression analysis, two so-called "robust" regression methods[50,51] were used to obtain alternative values of the CSA and other pertinent parameters, with slightly different weights given to outlying data points. A least-squares regression involves minimization of the target function $\rho(z) = (1/2)z^2$, where z is given by $z = (y_i^{meas} - y^{pred}(x_i))/\delta y_i$, where y_i^{meas} and $y^{pred}(x_i)$ are the measured and predicted data, respectively, for a given residue, and δy_i is the experimental uncertainty in y_i^{meas}. For this type of $\rho(z)$, the more deviant the point from the model, the greater the weight that this point is given in the minimization. Robust regression methods involve minimization of alternative functions of z. Here we use two such functions as the target of the minimization:[50,51] (1) the absolute value of z ($\rho(z) = |z|$), in which all deviant points are given the same relative weight, and (2) $\rho(z) = \log(1 + (1/2)z^2)$, where the relative weight given to deviant points initially increases with deviation (while $z < \sqrt{2}$) and then decreases so that those points that are the farthest from the fitting curve are given the least relative weight.

For the majority of residues in GB3 the results of the least-squares regression and the two robust methods agreed within their estimated uncertainties. For these residues the average of the parameters from the three types of regression is reported. As the experimental uncertainties in the derived parameters we report the biggest of the errors from the least-squares fit (using standard equations[50] for uncertainties in linear regression parameters or Monte Carlo simulations) and from the robust methods (using Monte Carlo simulations), estimated by propagating the experimental errors in relaxation and cross-correlation rates.

For those few residues where the three methods disagreed (i.e., where using a different weight function $\rho(z)$ for the same data set resulted in significant changes in the derived fitting parameters) no CSA is reported, except those cases where the deviation in the results of the least-squares regression can be unambiguously ascribed to undue weight given to a single clearly outlying data point (see examples in the Supporting Information). For these residues, the average of the two robust methods is reported. All three fits (least-squares and the two robust methods) for each model-independent method for every amide are shown in the Supporting Information.

Separation of True Variability in the CSA from Experimental Uncertainty. The observed range of site-specific ^{15}N CSA values reflects both true CSA variability and random statistical errors in the measured parameters.[8] To address the actual variability of the CSA tensor, we adopted the same statistical approach as in refs 8 and 29 that assumes that the CSA values in proteins follow a Gaussian distribution. Assuming that the experimentally determined uncertainties are correct, the "true" values of the mean CSA (μ, in ppm) and site-to-site CSA variability (Λ, also in ppm) can be determined by maximizing the following likelihood function:[29,50]

$$p(\mu,\Lambda) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi(\Lambda^2 + (\delta\Delta\sigma_i)^2)}} \exp\left(-\frac{(\mu - \Delta\sigma_i)^2}{2(\Lambda^2 + (\delta\Delta\sigma_i)^2)}\right) \quad (19)$$

Here N is the number of residues probed in the measured distribution, and $\Delta\sigma_i$ and $\delta\Delta\sigma_i$ are the measured CSA value and its experimental uncertainty for residue i. The confidence limits for μ and Λ were estimated from the boundaries of a 95% bivariate confidence region determined from the following equation: $p(\mu,\Lambda)/\max\{p(\mu,\Lambda)\} = \exp(-0.5\chi^2_{0.95,2})$, where $\chi^2_{0.95,2} = 5.99$ is the 95th percentile point of the chi-square distribution with 2 degrees of freedom.

Results

The transverse (R_2) and longitudinal (R_1) ^{15}N relaxation rates and the steady-state $^{15}N\{^1H\}$ NOEs in GB3 were measured at five magnetic fields, 9.4, 11.7, 14.1, 16.4, and 18.8 T. The transverse (η_{xy}) and longitudinal (η_z) ^{15}N CSA/dipolar cross-correlation measurements were performed at four fields (9.4, 11.7, 14.1, and 18.8 T) for η_{xy} and at three fields (9.4, 11.7, and 14.1 T) for η_z. The experimental details are outlined in Materials and Methods; the actual data and errors are listed in Supporting Information Table 2. Overall, 50 out of 55 amides were analyzed; residues Glu15, Thr25, Glu27, and Asn35 were excluded because of signal overlap, and Val39 was excluded due to conformational exchange.[32] Gln2 was excluded from LS analyses since the atom coordinates for this residue (which is a mutation) were not available from the crystal structure.

Model-Independent Determination of ^{15}N CSA. The use of data measured at multiple field strengths is expected to

(47) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.
(48) Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E. *J. Biomol. NMR* **1995**, *6*, 153−162.
(49) Ishima, R.; Nagayama, K. *Biochemistry* **1995**, *34*, 3162−3171.
(50) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C*; Cambridge University Press: New York, 1992.
(51) Rousseeuw, R.; Leroy, A. M. *Robust Regression and Outlier Detection*; John Wiley & Sons: New York, 2003.
(52) Draper, N. R.; Smith, H. *Applied Regression Analysis*, 3rd ed.; John Wiley & Sons: New York, 1998.

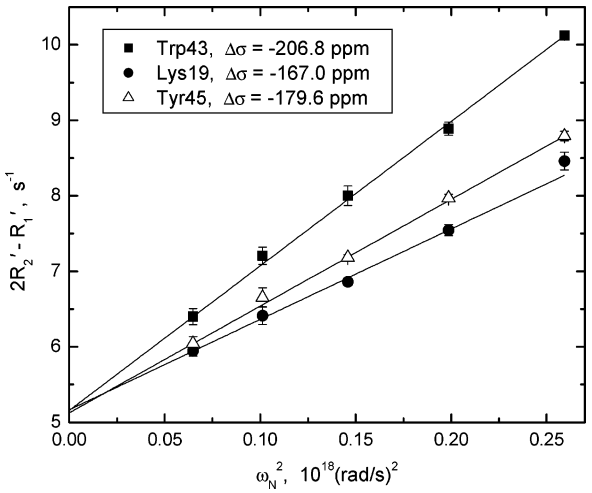

Figure 1. Representative fits of the dependence of $2R_2' - R_1'$ on ω_N^2. Shown are fits from the $2R_2 - R_1$ method for three residues in GB3. This plot also illustrates the variation in the ^{15}N CSA values between these residues. The amides shown here have very similar values of $J(0)$, as evidenced by the fact that they have the same intercept b (cf. eq 11), but exhibit strikingly different slopes, reflecting the difference in their CSA values (eq 12). The plots of $2R_2' - R_1'$ versus ω_N^2 for all residues in GB3 can be found in the Supporting Information. The error bars here and in all other figures represent standard errors (corresponding to 68.3% confidence intervals).

improve the accuracy of the derived picture of protein dynamics by allowing direct and independent determination of the spectral densities and the ^{15}N CSA.[31] The ^{15}N CSA values for backbone amides in GB3 were obtained using three different model-independent methods, detailed in Materials and Methods. These methods are model-independent in the sense that they involve no assumption about a particular parametrization of the spectral density function.

(A) The $2R_2 - R_1$ Method. The ^{15}N $\Delta\sigma$ values and the spectral density $J(0)$ were determined directly from the observed field dependence of the combination of reduced auto-relaxation rates, $2R_2' - R_1'$ (Figure 1). Relaxation data (R_1, R_2, and NOE) at all five fields were used for each residue. The data were fitted to a linear dependence on ω_N^2 (eq 10) using the three linear regression methods (least-squares and two robust methods) as discussed in Materials and Methods; the quality of the fit for each residue is shown in the Supporting Information. All three regression methods had good agreement (both slope and intercept agreed within the experimental uncertainty) for 38 out of 50 residues in GB3. For an additional nine residues (Leu12, Ala20, Val21, Gly38, Asp40, Asp47, Ala48, Thr49, and the C-terminal residue, Glu56) the two robust methods agreed within their experimental uncertainties (68.3% confidence interval). Only for three residues (Lys10, Gly41, and Lys50, all of which are in the loops in GB3) can no definitive CSA be reported because all three regression methods disagree for the $2R_2' - R_1'$ fit.

The average site-specific ^{15}N CSA values from the three fits are presented in Figure 2 (solid squares), and the values of $J(0)$

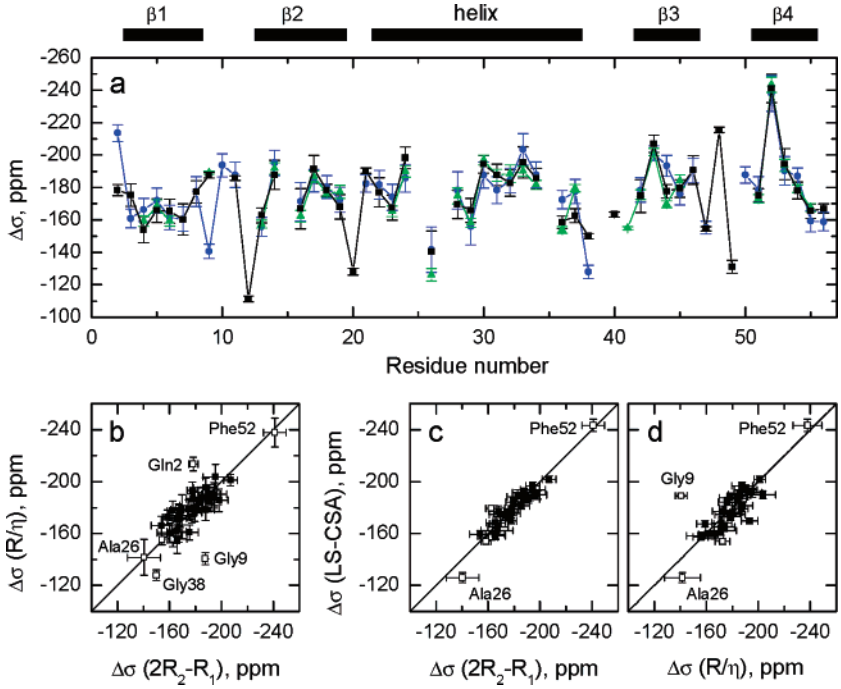

Figure 2. Site-specific ^{15}N CSA values in GB3 obtained using the three methods ($2R_2 - R_1$, R/η, and LS-CSA). (a) The site-specific ^{15}N CSAs, from the $2R_2 - R_1$ method (black squares), the R/η method (blue circles), and the LS-CSA method (green triangles) versus residue number. The secondary structure of GB3 is indicated at the top of the panel. (b) Correlation between ^{15}N CSA values measured using the model-independent methods, $2R_2 - R_1$ and R/η. Pearson's correlation coefficient r for these two data sets is 0.79; 81% of these CSA data agree within the experimental uncertainties. These values improve to $r = 0.80$ and 87% agreement if only those data (shown as solid squares) where the least-squares fits pass the 95%-confidence level χ^2/df cutoff are considered. (c) Correlation between the CSAs from $2R_2 - R_1$ and LS-CSA methods. The correlation coefficient is 0.95; it decreases to $r = 0.93$ if only those fits that pass the χ^2/df cutoff (solid squares) are included, though the percent agreement improves from 94% to 96%. (d) Correlation between the results from R/η and LS-CSA methods. The correlation coefficient is 0.80 and remains unchanged when the χ^2/df cutoff is applied (solid squares). The percent agreement increases from 84% for all considered residues to 88% for those residues with the χ^2/df below the cutoff value. In all correlation plots (panels b−d) the solid symbols represent values obtained for least-squares fits that passed the χ^2/df cutoff, while open symbols correspond to the remaining residues. Outliers and extreme values of the CSA are labeled. Note that those few residues that show significant differences in the CSA values between the methods are all located in the loops/termini. Also in the loops are all residues where only one out of the three methods resulted in an acceptable fit (panel a).

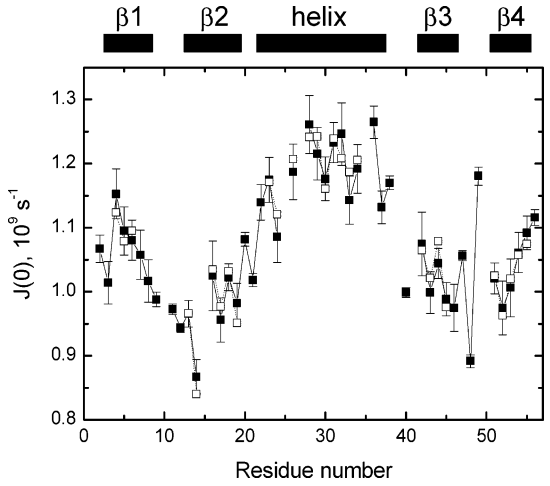

Figure 3. Agreement between the spectral density component, $J(0)$, measured using the $2R_2 - R_1$ method and reconstructed from the LS parameters. The spectral density component $J(0)$ obtained from the $2R_2 - R_1$ method directly (solid symbols) and calculated from the order parameters and local correlation times obtained in the LS-CSA method (open symbols). Throughout this paper, the factor 2/5 arising from the normalization of the spectral density of the overall rotational diffusion is explicitly included in the corresponding expression for $J(\omega)$.

are shown in Figure 3 (solid squares). The site-specific ^{15}N CSAs from this method range from -111.3 ± 1.7 ppm (Leu12) to -241.0 ± 8.7 ppm (Phe52), with a mean of $\langle\Delta\sigma\rangle = -174.2$ ppm and a standard deviation of 22.2 ppm. The median $\Delta\sigma$ is -175.4 ppm, in good agreement with the mean, indicating that the mean is not dominated by a small number of outliers (Table 1).

The observed site-specific ^{15}N CSA values were analyzed assuming a normal distribution of the true CSA values, as detailed in Materials and Methods. The average estimated relative uncertainty is 2.67% for $J(0)$ and 3.44% (or 6.0 ppm) for $\Delta\sigma$. From these data, the true CSA values in GB3 are characterized by a mean of $\mu = -173.9$ ppm and the site-to-site variability $\Lambda = 21.4$ ppm (see eq 19). We estimate a joint 95% confidence interval for μ from this method to range from -165.7 to -182.2 ppm and for Λ from 16.6 to 28.6 ppm (Figure 4). It is worth pointing out that site-to-site variability in the CSA in GB3 is evident from a comparison of the linear dependence of $2R_2' - R_1'$ on ω_N^2 (eqs 10−12) for three residues with similar $J(0)$ values (Figure 1).

(B) The *R*/*η* Method. This method is based on the field dependence of the ratio of the (reduced) auto-relaxation rate (R_2' or R_1') and the corresponding ^{15}N CSA/dipolar cross-correlation rate (η_{xy} or η_z, respectively), eqs 1−9. Both R_2'/η_{xy} and R_1'/η_z ratios are expected to have the same values (eq 1); therefore these data were analyzed together (see also below). The analysis included R_2'/η_{xy} data at four fields and R_1'/η_z at three fields for each residue. Using both R_2'/η_{xy} and R_1'/η_z data improves the accuracy of analysis by increasing the number of data points included in the fit. In addition, the R_1'/η_z values have the advantage of being free of any contribution from conformational exchange. The quality of the fit for each residue in GB3 is shown in the Supporting Information. All three regression methods had good agreement (both slope and intercept agreed within the experimental uncertainty) for 37 out of 50 amides in GB3. For an additional seven residues (Gly9, Thr11, Lys13, Ala26, Gly38, Phe52, and the C-terminal Glu56)

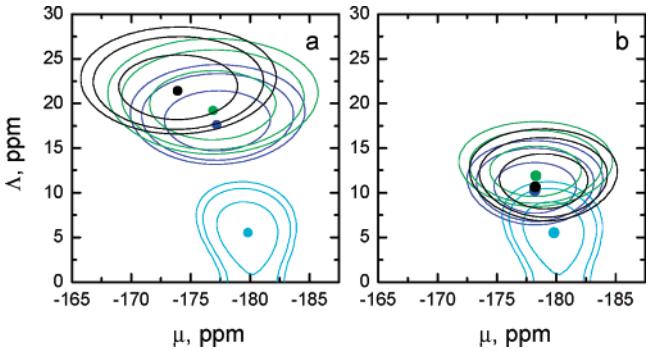

Figure 4. Likelihood functions (eq 19) obtained from different methods and sets of data that show significant site-so-site variability in the ^{15}N CSA values. Contour plots of the likelihood functions $p(\mu, \Lambda)$ (eq 19) corresponding to the ^{15}N CSA values from the three methods ($2R_2 - R_1$ (black), R/η (blue), and LS-CSA (green)) (a) for all analyzed residues in GB3 and (b) for only those residues where χ^2/df from the least-squares fits passed the goodness-of-fit test at a 95% confidence level. Also shown (in cyan), for comparison, is the analogous likelihood function obtained for the recently reported ^{15}N CSAs in ubiquitin,[29] scaled to a N−H bond length of 1.02 Å. The location of the maximum for each function is indicated by a dot (see also Table 1); the contour lines represent 68.3%, 90%, and 95% bivariate confidence regions for μ and Λ. In panel a, the 95% joint confidence intervals (in ppm) for μ and Λ are (−165.7, −182.2) and (16.6, 28.6) from $2R_2 - R_1$, (−169.9, −184.6) and (13.2, 24.3) from R/η, and (−168.0, −185.7) and (14.3, 27.3) from LS-CSA methods. For a subset of residues (panel b) that pass the χ^2/df cutoff, the corresponding confidence intervals for μ and Λ are (−172.7, −185.2) and (6.8, 17.1) from $2R_2 - R_1$, (−172.5, −183.9) and (6.4, 15.8) from R/η, and (−171.8, −184.7) and (8.5, 18.0) from LS-CSA methods.

the two robust methods agreed within their experimental uncertainties (68.3% confidence interval). For six residues (Leu12, Ala20, Asp40, Gly41, Ala48, and Thr49, all of which are in loop/turn regions of GB3), no CSA is reported here because all three regression methods disagree in the R/η fit. The ^{15}N CSA values ($\Delta\sigma$) obtained using this approach are shown in Figure 2, and the values of $\Delta\sigma_g$ are presented in Figure 5. These ^{15}N CSAs range from -127.9 ± 4.0 ppm (Gly38) to -237.9 ± 11.1 ppm (Phe52), with a mean value of -177.4 ppm and a standard deviation of 19.5 ppm. The median is -178.4 ppm. The average estimated level of the experimental errors is 4.23% (or 7.5 ppm) for $\Delta\sigma$. The maximization of the likelihood function (eq 19, Materials and Methods) yielded the true variability in $\Delta\sigma$ of $\Lambda = 17.6$ ppm and a true mean CSA of -177.2 ppm. We estimate a 95% confidence interval on μ from this method to be from -169.9 to -184.6 ppm and for Λ from 13.2 to 24.3 ppm (Figure 4).

The angles β_z derived from these $\Delta\sigma$ and $\Delta\sigma_g$ values assuming axial symmetry of the ^{15}N CST are shown in Figure 5c. The range of β_z values is from 7.5° (Val6) to 27.6° (Thr11) with a mean value of 19.9° and standard deviation of 4.5°, in agreement with the β_z values observed in ubiquitin.[5,24,28] Very similar β_z values were also determined from a combination of the $\Delta\sigma_g$ values from the $2\eta_{xy} - 2\eta_z$ method with the $\Delta\sigma$ values from $2R_2 - R_1$ (see below).

Note that using the mean of R_2'/η_{xy} and R_1'/η_z as the R/η value at a given field (where both data are available at 9.1, 11.7, and 14.1 T) resulted in the CSA values from -127.9 to -237.9 ppm with a mean CSA of -177.4 ppm and a standard deviation of 19.5 ppm. These results have a Pearson's correlation coefficient r of 0.97 to CSA values obtained using the individual measurements (see above). Fitting the R_2'/η_{xy} values alone gave ^{15}N CSA values in the range from -140.5 to -234.8 ppm, with

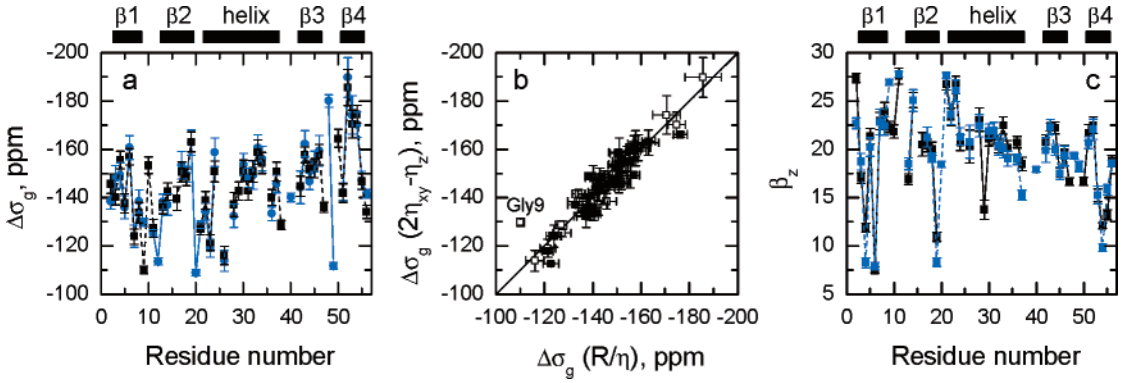

Figure 5. Values of $\Delta\sigma_g$ and the β_z angles from the R/η and $2\eta_{xy} - \eta_z$ methods. (a) Measured site-specific ^{15}N $\Delta\sigma_g$ values for GB3 from the R/η (black squares) and the $2\eta_{xy} - \eta_z$ methods (blue circles). The $\Delta\sigma_g$ values range from -108.9 ppm (Ala20, $2\eta_{xy} - \eta_z$) to -189.8 ppm (Phe52, $2\eta_{xy} - \eta_z$). (b) Correlation between $\Delta\sigma_g$ values measured using the R/η and $2\eta_{xy} - \eta_z$ methods. The correlation coefficient is 0.94 for all residues and 0.96 for only those fits that pass the χ^2/df cutoff. (c) β_z angles (in degrees) determined from the R/η method (black squares) and by combining the $\Delta\sigma_g$ values from the $2\eta_{xy} - \eta_z$ method with the $\Delta\sigma$ values from $2R_2 - R_1$ (blue circles). Pearson's correlation coefficient for the agreement of the β angles from these two measurements is 0.94. The derivation of β_z assumed axial symmetry of the ^{15}N chemical shielding tensor. The secondary structure of GB3 is indicated on the top of panels a and c.

a mean of -179.2 ppm and a standard deviation of 19.2 ppm, with a correlation coefficient of 0.91 to the CSAs derived from both transverse and longitudinal data. The R_1'/η_z data alone yielded a broader spread of CSAs, from -129.9 to -251.6 ppm, with somewhat larger absolute values of the mean (-185.5 ppm) and standard deviation (23.9 ppm). These data show a poor correlation ($r = 0.13$) with the CSAs obtained from both transverse and longitudinal data together, which likely reflects a lesser accuracy of the R_1/η_z data alone due to a narrower range of magnetic fields covered by the η_z measurements.

It is worth mentioning that ^{15}N CSAs obtained by the R/η method are expected to be independent of the magnitude of the spectral density function.[5,25,31] Indeed, the correlation coefficient between the $J(0)$ values derived from the $2R_2 - R_1$ method (these values are independent of $\Delta\sigma$) and the CSA values from the R/η approach was -0.23.

(C) Quality Control Using the $2\eta_{xy} - \eta_z$ Method. The field dependence of the cross-correlation data alone (eqs 13, 14) yields the product of $\Delta\sigma_g$ and $J(0)$. The quality of fit is shown in the Supporting Information. This analysis is totally independent of the auto-relaxation data. We then used the value of $J(0)$ derived from the $2R_2 - R_1$ method (this value is independent of $\Delta\sigma$) to obtain $\Delta\sigma_g$ (Figure 5). The $\Delta\sigma_g$ values thus obtained range from -107.2 (± 1.2) ppm for Leu12 to -186.1 (± 1.0) ppm for Ala34, with the mean value of -154.4 ppm and a median at -154.1 ppm. These values were then compared with the $\Delta\sigma_g$ values derived from the R/η approach, which are independent of $J(0)$. The excellent agreement ($r = 0.94$ to 0.96, Figure 5b) between the values of the same parameter determined independently from different sets of measurements thus provides strong quality control for our analysis.

Assuming axial symmetry of the ^{15}N CST, and using $\Delta\sigma$ values from the $2R_2 - R_1$ method, we determined the angle β_z between the unique (least shielded) component of the tensor and the NH bond vector (Figure 5c). These β_z values are in very good agreement ($r = 0.93$) with β_z derived from the R/η method described above.

Determination of ^{15}N CSA and Order Parameters in GB3 Using the Lipari−Szabo Approximation. The analysis of relaxation data was also performed assuming the so-called "model-free" form of the spectral density,[41−43] using both the conventional Lipari−Szabo approach (LS) and its modifications, LS-CSA and LS-SDF, described in Materials and Methods.

(A) Analysis of the Overall Tumbling. The importance of a correct treatment of the overall tumbling of a molecule for the accurate derivation of local motional parameters has been established in the literature.[32,45,53−55] The overall rotational diffusion tensor of GB3 was derived from the ^{15}N relaxation data (R_1, R_2, NOE) at each field using the program ROTDIF.[56] The diffusion tensor obtained by this method is independent, to a good approximation, of site-specific values of the $^1H-^{15}N$ dipolar interaction, ^{15}N CSA, and NH order parameters.[57,58] This follows from the fact that R_2' and R_1' are both proportional to $(d^2 + c^2)$ (see eqs 5, 6) and, for protein core residues with restricted backbone mobility, also to S^2 (because $J(0)$, $J(\omega_N) \propto S^2$). Thus, in the R_2'/R_1' ratio, analyzed in ROTDIF, all these factors unrelated to overall rotational diffusion cancel out. The results of the analyses are shown in Table 2. At all five field strengths an axially symmetric diffusion tensor was found to be a significant improvement over an isotropic model (evaluated by the statistical F-test[50]), whereas the use of a more complex, fully anisotropic diffusion tensor model was not statistically warranted (see also below).

The good agreement (within the experimental errors) between the diffusion tensors determined at different fields indicates that there is no significant difference in the experimental conditions (in particular, temperature) between the measurements on different spectrometers. This then justifies the simultaneous analysis of these relaxation data acquired at various fields for the purpose of extracting field-independent parameters, such as CSA, S^2, etc. Note also that there is practically no difference between the diffusion tensors derived using the crystal and

(53) Luginbuhl, P.; Pervushin, K. V.; Iwai, H.; Wuthrich, K. *Biochemistry* **1997**, *36*, 7305−7312.
(54) Korzhnev, D. M.; Orekhov, V. Y.; Arseniev, A. S. *J. Magn. Reson.* **1997**, *127*, 184−191.
(55) Fushman, D.; Cowburn, D. In *Structure, Motion, Interaction and Expression of Biological Macromolecules*; Sarma, R., Sarma, M., Eds.; Adenine Press: Albany, NY, 1998; pp 63−77.
(56) Walker, O.; Varadan, R.; Fushman, D. *J. Magn. Reson.* **2004**, *168*, 336−345.
(57) Fushman, D.; Varadan, R.; Assfalg, M.; Walker, O. *Prog. NMR Spectrosc.* **2004**, *44*, 189−214.
(58) Fushman, D.; Cowburn, D. In *Protein NMR for the Millenium (Biological Magnetic Resonance Vol 20)*; Krishna, N. R. L. B., Ed.; Kluwer: Dordrecht, 2002; pp 53−78.

Table 2. Characteristics of the Overall Rotational Diffusion Tensor of GB3 Derived from ^{15}N Relaxation Data at Different Magnetic Fields[a]

magnetic field (T)	^1H resonance frequency (MHz)	$D_\perp$[b] (10^7 s^{-1})	$D_\parallel$[b] (10^7 s^{-1})	Φ^{oc}	Θ^{oc}	τ_c[d] (ns)	anisotropy[e]	χ^2/df[f]	P[g]
From Auto- and Cross-Relaxation Rate Measurements									
9.4	400	4.40(0.19)	6.13(0.62)	89(18)	66(23)	3.35(0.20)	1.39(0.13)	0.64	6×10^{-11}
11.7	500	4.45(0.31)	6.20(1.12)	95(15)	68(19)	3.31(0.32)	1.39(0.24)	0.69	4×10^{-13}
14.1	600	4.45(0.15)	6.05(0.44)	90(8)	70(10)	3.34(0.14)	1.36(0.09)	0.72	2×10^{-13}
16.4	700	4.44(0.14)	6.24(0.41)	99(7)	63(11)	3.31(0.13)	1.41(0.08)	0.88	6×10^{-19}
18.8	800	4.46(0.08)	6.15(0.27)	100(7)	67(10)	3.32(0.08)	1.38(0.06)	0.74	3×10^{-14}
averaged tensor		4.44	6.14	99	66	3.33	1.38		
global-fit tensor		4.44	6.14	95	66	3.33	1.38	0.72	6×10^{-15}
From Cross-Correlation Rate Measurements									
9.4	400	4.50(0.16)	6.00(0.52)	101(9)	77(13)	3.33(0.16)	1.33(0.11)	0.66	9×10^{-11}
11.7	500	4.38(0.12)	6.14(0.40)	90(6)	59(9)	3.36(0.12)	1.40(0.08)	0.96	1×10^{-12}
14.1	600	4.40(0.06)	6.20(0.19)	93(4)	65(6)	3.33(0.06)	1.41(0.04)	0.51	3×10^{-17}

[a] The NH vectors for this analysis were taken from the original crystal structure of GB3 (PDB entry 1IGD); similar results were obtained for GB3 structures refined using residual dipolar couplings (PDB entries 1P7E and 1P7F[70]) (Supporting Information Table 3). These diffusion tensor characteristics are in good agreement with those theoretically predicted from the shape of the molecule.[32] Also shown are the diffusion tensors derived from the cross-correlation rates, η_{xy} and η_z. Numbers in parentheses represent standard deviations. [b] Principal values of the rotational diffusion tensor. [c] Polar and azimuthal angles $\{\Theta, \Phi\}$ (in deg) describe the orientation of the diffusion tensor axis with respect to protein coordinate frame. [d] Overall rotational correlation time of the molecule, $\tau_c = 1/[2\ \mathrm{tr}(\underline{\underline{D}})]$. [e] Degree of anisotropy of the diffusion tensor, $D_\parallel/D_\perp$. [f] Residuals of the fit divided by the number of degrees of freedom. [g] Probability that the reduction in χ^2 compared to the isotropic diffusion model occurred by chance.

solution structures of GB3 (cf. Table 2 and Supporting Information Table 3). Also there is no significant difference between the diffusion tensor obtained from a simultaneous (global) fit of all the data and the result of averaging the diffusion tensors obtained at each field (Table 2). Therefore for our LS analyses, we used the diffusion tensor resulting from the simultaneous fit of all data.

The overall rotational diffusion tensor can also be derived from cross-correlation measurements,[7] independently of the auto- and cross-relaxation measurements, when both η_{xy} and η_z data are available. This approach has the advantage of being essentially free of any effects of conformational exchange contributions to R_2 and also does not require correction for high-frequency components of the spectral density. The diffusion tensors obtained from η_{xy} and η_z measurements at 9.4, 11.7, and 14.1 T (shown in Table 1) are in excellent agreement with those derived from the auto-relaxation rates and NOE.

(B) Backbone Order Parameters: Assuming a Uniform ^{15}N CSA. When relaxation data (R_1, R_2, NOE) at several fields are available, order parameters for a given NH vector can be obtained from the data at each field separately or from a simultaneous fit of the relaxation data for all available field strengths. Because the LS backbone dynamics should not depend on the applied magnetic field, all these order parameters are expected to agree with each other.

We first analyzed the relaxation data at each field separately using a standard LS approach[32] assuming a uniform value of ^{15}N CSA of -160 ppm. In all these analyses the quality of fit was very good: the residuals of the fit for the majority of residues (96% at 9.4T, 96% at 11.7 T, 98% at 14.1 T, 94% at16.4 T, 84% at 18.8 T, and 94% overall) were within the acceptance level for a 95%-confidence goodness-of-fit test,[50] which indicates that the uncertainties in the experimental data are correct or overestimated, but not underestimated. The results, however, show a striking discrepancy between the derived order parameters corresponding to different field strengths (Figure 6a,f): for most residues in GB3 the observed variation in the derived S^2 values among the fields exceeds their experimental uncertainties. Even in the well-ordered parts of the protein, the difference in derived S^2 values between 800 and 400 MHz data exceeded 0.10 for some residues. Similar results were obtained when using a ^{15}N CSA of -170 ppm (suggested in ref 59) or the mean CSA of -174.2 ppm (the mean CSA from the three determination methods, $2R_2 - R_1$, R/η, and LS-CSA, Figure 6b,g). The observed disagreement between the derived S^2 values obtained for the same NH group from the measurements at different fields thus raises significant concern about the accuracy of the order parameters derived by the standard analysis.

We also attempted to analyze simultaneously the relaxation data at all five fields using a uniform CSA of -160 ppm and the average diffusion tensor and NH vector orientations from the crystal structure. This analysis indicated serious problems of fitting: only 8 out of 51 (Tyr3, Lys4, Leu5, Val6, Thr16, Ala23, Lys28, and Ala29) amides had residuals of the fit (χ^2) that passed the goodness-of-fit test at the 95% confidence level.[50] Using a uniform CSA of -170 ppm did not significantly improve the fit: here only 12 residues (Tyr3, Lys4, Leu5, Thr16, Thr18, Lys19, Ala23, Lys28, Gln32, Asp46, Thr51, and Thr55) had acceptable χ^2 values. Using the mean CSA value of -174.2 ppm gave only 14 residues (Gln2, Tyr3, Leu5, Thr16, Thr18, Lys19, Ala23, Lys28, Gln32, Ala34, Val42, Tyr45, Asp46, and Thr51) with acceptable χ^2 values. These results from multiple approaches clearly indicate that the conventional LS treatment assuming a uniform ^{15}N CSA fails totally to describe the multifield experimental data in GB3.

A similar problem was previously noted by Farrow et al.,[48] who observed that order parameters obtained from fitting relaxation data measured at several field strengths have low precision (although they are more accurate than order parameters obtained from data at one field strength) due to poor fits of multifield data to a LS spectral density function. Other examples of discrepancies in the LS parameters derived from relaxation measurements at several fields can be found elsewhere.[45,59,60]

It is noteworthy that for most residues in GB3 the observed difference in the order parameters appears systematic; that is, it increases with the field strength (Figure 6a,f). This tendency

(59) Tjandra, N.; Wingfield, P.; Stahl, S.; Bax, A. *J. Biomol. NMR* **1996**, *8*, 273−284.
(60) Korzhnev, D. M.; Billeter, M.; Arseniev, A. S.; Orekhov, V. Y. *Prog. NMR Spectrosc.* **2001**, *38*, 197−266.

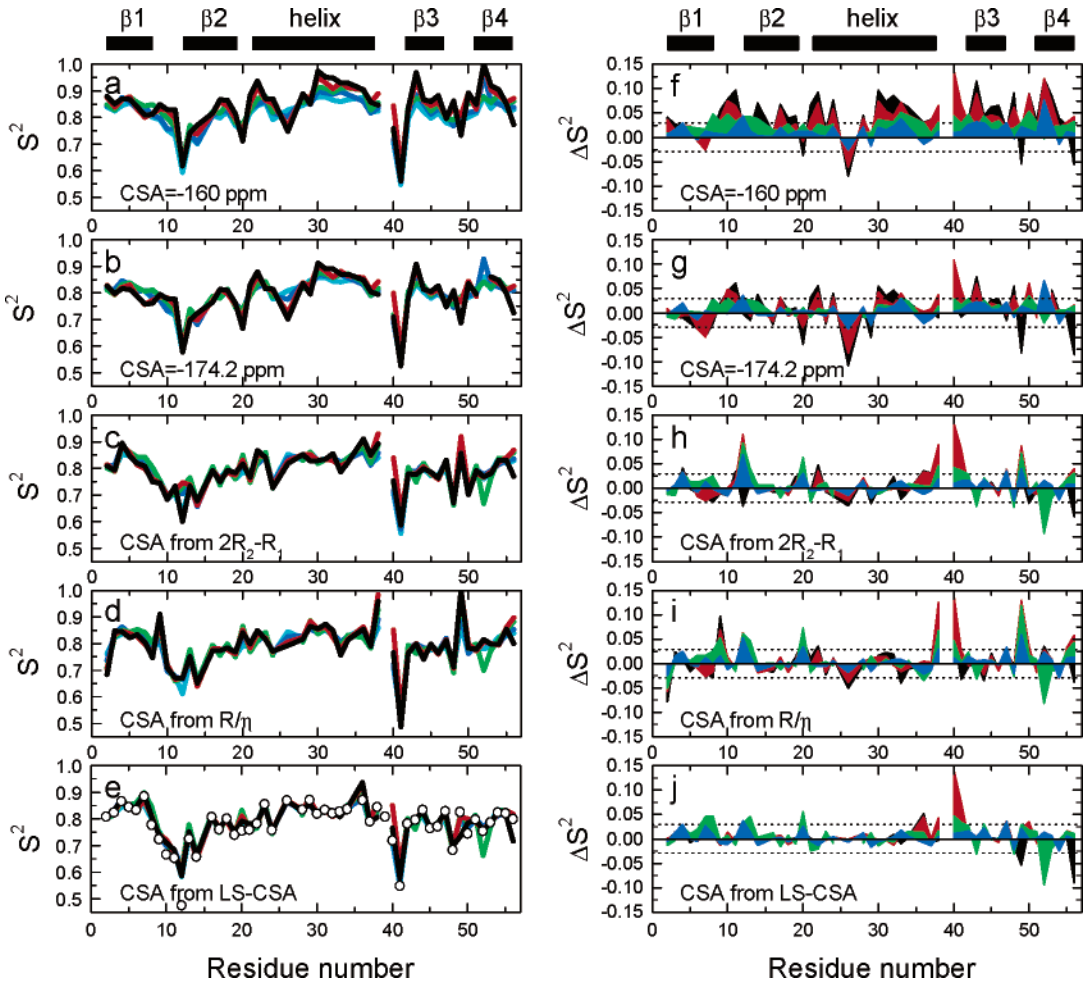

Figure 6. Backbone order parameters determined from ^{15}N relaxation data at each field using different CSA models. Shown are backbone order parameters in GB3 derived from a LS analysis of the ^{15}N relaxation data (R_1, R_2, NOE) at different fields (left panels). Right panels represent the differences, $\Delta S^2 = S^2 - S^2$(9.4 T), between the S^2 values at a particular field and at 9.4 T, where the ^{15}N CSA contribution to ^{15}N relaxation rates is the weakest. (a, f) LS analysis was performed in a conventional way, i.e., assuming a uniform CSA of −160 ppm for all residues. (b, g) LS analysis was performed assuming a uniform CSA of −174.2 ppm (the average of the site-specific CSAs in GB3, see Table 1) for all residues. (c, h) Site-specific ^{15}N CSA values from the $2R_2 - R_1$ method were used as input parameters. (d, i) Site-specific ^{15}N CSA values from the R/η method were used as input parameters. (e, j) LS analysis was performed for each field separately using the site-specific CSAs derived from the global fit (LS-CSA) of all five fields. Also shown as open circles in panel d are the order parameters from the global fit. The color scheme is as follows: the 18.8 T data are shown in black, 16.4 T in red, 14.1 T in green, 11.7 T in blue, and 9.4 T in cyan. The dashed horizontal lines represent the average estimated level (±0.029) of the experimental uncertainty in ΔS^2. Val39 has been removed from all panels because of the conformational exchange contribution.[32] To exclude deviations in S^2 due to a change in the model selection for different fields in a few residues, all data presented here were obtained assuming a model of local motion (model 2 in ref 44, model "B" in ref 33) that includes S^2 and τ_{loc} as fitting parameters. Our model-selection analysis showed that for the majority of residues in the secondary-structure elements of GB3 this was the preferred model.[32] Allowing freedom in the model selection led to even greater discrepancies between the order parameters from different fields, which, however, exhibit the same behavior as shown here (Supporting Information Figure 2). As a measure of the discrepancy in order parameters, the rmsd from the average (over all five fields) S^2 value for each method is 0.024 (panel a), 0.015 (b), 0.010 (c), 0.012 (d), and 0.009 (e), calculated for the secondary structure elements only.

is present even in the data obtained using the average CSA of −174.2 ppm (Figure 6b,g). This behavior could arise from (1) conformational exchange contributions to ^{15}N R_2 not accounted for in the analysis or (2) deviations in the site-specific values of ^{15}N CSA from their assumed values. Site-specific deviations in the ^{1}N−^{15}H bond length from a uniform value of 1.02 or 1.04 Å could also result in erroneous order parameters; however, the currently available experimental data on variations in the NH bond length in proteins are insufficient in order to rigorously address this issue. A failure of the LS spectral density model to accurately represent data at multiple fields cannot be excluded (e.g., refs 45, 54), particularly with regard to the uncoupling of local and global motions; however our analysis (below) indicates that a modified LS model (using site-specific CSAs) nicely fits the observed spectral densities in GB3.

Several lines of evidence suggest that conformational exchange is not the source of the observed discrepancy in the order parameters in GB3. As shown earlier,[32] conformational exchange contributions are negligible for most of amides in GB3, except Val39. This conclusion is also confirmed by the agreement (Supporting Information Figure 1) between the measured R_2's and their reconstructed "exchange-free" values,[7] $R_{2\text{free}}' = R_1' \cdot \eta_{xy}/\eta_z$. The exclusion of conformational exchange as a possible cause of the observed discrepancy between the S^2 values is further supported by the results of a LS analysis of the data at the individual fields. Here, 12 (excluding Val39) residues (Tyr3, Leu5, Ile7, Thr16, Ala23, Tyr33, Asp36, Asn37, Asp40, Thr44, Ala48, and Thr51) required a R_{ex}-containing model of local motion[33] at 18.8 T, where the R_{ex} contribution is expected to be the strongest. These R_{ex} values were relatively small

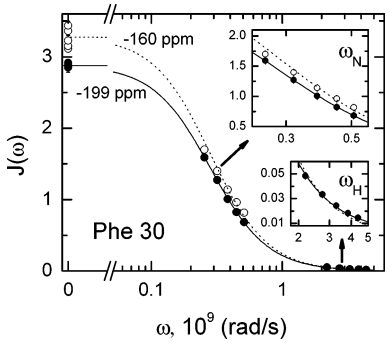

Figure 7. Illustration of the LS fit of the spectral density components determined at all five fields. Representative LS fit of all spectral density components from the five-field measurements for Phe30. Symbols depict the $J(\omega)$ values for $\omega = 0$, ω_N, and $0.87\omega_H$ derived from relaxation data for each field separately (eqs 15−17) assuming a CSA of −160 ppm (open circles) or the CSA value of −199.1 ppm for Phe30 that optimizes the fit (solid circles). The corresponding fitting curves are shown as dashed and solid lines, respectively. Shown in the insets is a blowup of the regions corresponding to $\omega = \omega_N$ and $0.87\omega_H$, indicated as "ω_N" and "ω_H". The values of S^2 and τ_{loc} were 0.93 and 3.0 ps when using a CSA of −160 ppm, and 0.81 and 10.3 ps for the fit CSA values. A 35-fold decrease in χ^2/df was observed when using the CSA and the LS parameters from the LS-SDF fit. The $\Delta\sigma$ value derived using the $2R_2 - R_1$ method (−194.3 ± 5.4 ppm for Phe30) resulted in a fit that was practically indistinguishable from the LS-SDF fit shown here, as does the use of the CSA value ($\Delta\sigma$ = −196.9 ± 2.9 ppm) from the LS-CSA fit for Phe30. For comparison, the result of this fit when the mean site-specific CSA of −174.2 ppm is used is shown in Supporting Information Figure 7.

(maximum 0.53 ± 0.10 s^{-1} for Asp36 at 18.8 T) and likely reflect errors in LS model selection, because the only residue that systematically showed conformational exchange at all five fields was Val39. In addition, excluding R_2s from the simultaneous analysis of the five-field data (hence using only R_1s and NOEs, as suggested in ref 45) did not improve the quality of fit for CSA = −160 ppm: only nine residues passed the goodness-of-fit test (Tyr3, Leu5, Lys13, Thr16, Lys19, Ala23, Ala29, Thr51, and Thr5) in this case. Note also that in terms of spectral densities, the presence of the R_{ex} contribution will affect $J(0)$ but not the $J(\omega_N)$ values (eqs 16, 17); therefore, the introduction of the R_{ex} terms might force the $J(0)$ values from different fields to converge, but will not improve the fit of spectral densities at $\omega = \omega_N$ (Figure 7, Supporting Information Figure 7) derived assuming a uniform CSA of −160 ppm or even −174.2 ppm (see below). Finally, the R_{ex}-free values of the overall diffusion tensor obtained solely from the cross-correlation measurements are in excellent agreement with those from the R_2/R_1 ratio (Table 2).

(C) Backbone Order Parameters: The Effect of Site-Specific ^{15}N CSAs. To verify that the observed field dependence in the order parameters (Figure 6a) could reflect site-specific variations in the ^{15}N CSA (Figure 2) unaccounted for in the conventional LS analysis, we performed the same derivation as above, this time using as input the site-specific ^{15}N CSA values measured using the model-independent approaches. As shown in Figure 6c,d,h,i, the inclusion of the site-specific ^{15}N CSA has dramatically reduced the variation in the order parameters among the fields, which is now within the level of experimental noise for most residues.

We therefore modified the LS analysis by including the CSA as an additional fitting parameter (LS-CSA method, Materials and Methods). This resulted in a significant improvement in the quality of fit of the five-field data for the majority of residues in GB3. For example, when the ^{15}N CSA was allowed to vary in the LS-CSA method, the mean χ^2/df for residues in the secondary structure dropped from 5.12 (for a uniform CSA of −160 ppm) to a value of 0.92. All of the secondary structure residues except for Ala26 and Phe52 now have χ^2/df low enough to pass the goodness-of-fit test at a 95% confidence level. Altogether, 47 out of 49 analyzed residues exhibited a decrease in χ^2 of the LS fit, and in 40 residues there is also a decrease in χ^2/df. The residues where the χ^2/df is not improved (Asn8, Leu12, Lys13, Thr16, Gly38, Asp40, Gly41, Asp47, and Thr49) are all in flexible regions of GB3 except for Thr16, for which the resulting CSA (−162.3 ppm) is very close to −160 ppm and the residuals of fit were already sufficiently low: χ^2/df = 0.56 and 0.67 for the LS and LS-CSA methods, respectively.

For those residues where a reduction in χ^2 was accompanied by an increase in the number of fitting parameters (33 residues in GB3), a statistical F-test was performed[50] to determine if the improvement in the χ^2 was significant. For 31 (94%) of these residues, the reduction in the χ^2 is statistically justified at a 95% significance level or higher (i.e., the probability, P, that the reduction in χ^2 occurred by chance is $P < 0.05$). For 25 (76%) of these residues the significance level is higher than 99% (i.e., $P < 10^{-2}$), and for 22 (67%) of these residues the significance level is even higher than 99.9% (i.e., $P < 10^{-3}$).

The order parameters derived from a simultaneous (global) fit of data from all five fields using the LS-CSA method are shown as open symbols in Figure 6e. All three regression methods (the least-squares and two robust methods) had good agreement (within the experimental uncertainty for both S^2 and the CSA) for 28 out of 49 amides in GB3 (Gln2 not included here because its coordinates are unavailable from the crystal structure). For an additional four residues (Gly9, Asp36, Asn37, and Gly41) the two robust methods agreed within their experimental uncertainties (68.3% confidence interval). For 17 residues (Tyr3, Ile7, Asn8, Lys10, Thr11, Leu12, Ala20, Val21, Asp22, Gly38, Asp40, Asp46, Asp47, Ala48, Thr49, Lys50, and the C-terminal Glu56), all of which are either in the loops/termini or at the edges of secondary structure elements, no CSA is reported here for the LS-CSA method because all three regression methods disagreed for either S^2 or $\Delta\sigma$.

The site-specific ^{15}N CSA values from the LS-CSA method were in the range from −126.0 ± 3.9 ppm (Ala26) to −243.4 ± 4.7 ppm (Phe52), with a mean of −176.9 ppm, a median of −176.8 ppm, and standard deviation of 20.0 ppm. The average estimated level of the experimental errors is 1.76% (or 3.1 ppm) for the CSA, which gives a true site-to-site CSA variability Λ of 19.2 ppm and a true mean of −176.9 ppm. We estimate 95% confidence limits from this method to be from −168.0 to −185.7 ppm for μ and from 14.3 to 27.3 ppm for Λ (Figure 4).

Using these site-specific ^{15}N CSA values as input for the LS analyses at separate fields resulted in a further reduction in the spread of the order parameters among the fields (Figure 6e,j). These results clearly indicate that the discrepancy in the order parameters in Figure 6a is caused by site-specific variations in the ^{15}N CSA.

(D) LS Fit of the Spectral Densities Directly. A direct analysis of the spectral densities produced similar results. For a uniform CSA of −160 ppm, the χ^2/df of the fit of the spectral density functions at all five fields for the secondary structure elements of GB3 ranges from 0.46 (Tyr16) to 20.6 (Trp43) with

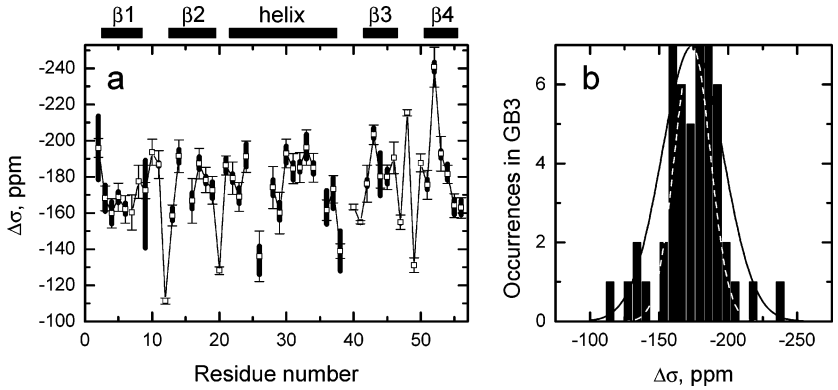

Figure 8. Site-specific ^{15}N CSA values, averaged over all three methods, show significant CSA variability in GB3. (a) Range of ^{15}N CSAs for each backbone amide in GB3 from the three methods ($2R_2 - R_1$, R/η, and LS-CSA) shown as solid vertical bars. The open symbols represent the average site-specific CSA, $\Delta\sigma$, from the three methods; the error bars represent the maximum error from the three methods for each residue. (b) Histogram of the average site-specific CSA values shown in panel a. Including these average site-specific CSA values into the analysis of the derivation of the true CSA values (eq 19) resulted in the true mean $\mu = -173.8$ ppm and the site-to-site variability $\Lambda = 21.2$ ppm (Table 1). The black curve represents a Gaussian distribution with the mean of -174.2 ppm and the standard deviation of 22.2 ppm. The dashed curve is also a Gaussian, with the same mean but with a standard deviation of 13.0 ppm; this curve corresponds to the case when all seven outliers in panel b are taken out.

a mean value of 4.73. The quality of the fits of the spectral density functions is illustrated in Figure 7; a similar comparison for a CSA of -174.2 ppm can be found in Supplementary Information Figure 7. Overall, major discrepancies between the experimental data and the LS model were for $\omega = 0$, due to the spread in the $J(0)$ values derived at various fields, and at $\omega = \omega_N$, where the experimental $J(\omega_N)$ values noticeably deviate from the theoretical curve. There is a good agreement for the high-frequency components (which are CSA-independent), particularly taking into account the reduced spectral density approximation[48,49] (eq 15) made when deriving $J(0.87\omega_H)$ from the experimental data.

The inclusion of CSA as a third fitting parameter (in addition to S^2 and τ_{loc}, see LS-SDF in Materials and Methods) resulted in the reduction of the residuals of fit for 29 out of 35 residues (or 83%) in the secondary structure elements; the χ^2/df with CSA as an additional adjustable parameter ranged from 0.3 (Thr18) to 6.1 (Phe52) with a mean of 1.25. The LS-SDF method resulted in a significantly better convergence of $J(0)$ values from different fields and, at the same time, in a better fit of the $J(\omega_N)$ values (Figure 7). A similar improvement in the fit was obtained when using site-specific CSA values from the $2R_2 - R_1$ method, resulting in reduced χ^2/df for 27 amides in the secondary structure.

Discussion

Agreement between the ^{15}N CSA Values in GB3 Derived from Various Methods. There is an excellent agreement between the results of the LS-CSA and LS-SDF methods: for the residues in the secondary structure, the CSAs from the two methods agree within their errors and have a correlation coefficient of 0.99. The order parameters and τ_{loc} values derived using these methods agree within their respective errors for all but two residues (Ala23 and Lys28) in the secondary structure. For those residues where there is good agreement, this indicates that the use of reduced spectral densities does not significantly alter the values of these parameters.

Furthermore, the CSA values from these two approaches based on the LS form of the spectral density function are in good agreement with the results of the model-independent approaches (Figure 2c,d). For all residues in GB3, the Pearson's correlation coefficient is 0.95 between the CSAs from the LS analyses and the $2R_2 - R_1$ method and 0.80 between the CSA values from the LS analyses and those measured using the R/η method. To validate the characteristics of the backbone dynamics (S^2, τ_{loc}) derived simultaneously with site-specific ^{15}N CSAs (LS-CSA method), we compared the spectral density $J(\omega)$ at $\omega = 0$ reconstructed from these data with $J(0)$ values obtained directly from the $2R_2 - R_1$ method (recall that this latter $J(0)$ is *independent* of the ^{15}N CSA). The good agreement between the two values of $J(0)$ (Figure 3) for the secondary structure elements of GB3 thus validates the LS parameters derived using the LS-CSA method.

Distribution of Site-Specific ^{15}N CSA Values. The range of ^{15}N CSAs obtained from all above-mentioned methods for each residue in GB3 is shown in Figure 8, together with a histogram of the average CSA values (from the three determination methods) for each residue. The likelihood functions $p(\mu,\Lambda)$ (eq 19) generated from the results of each of the three CSA determination methods are shown in Figure 4. The true mean CSA values (μ) from these methods are on average slightly higher in absolute value than those observed earlier in ubiquitin (mean CSA $= -157$ ppm)[5,6,27] and in Rnase H ($\mu = -172$ ppm)[8] and slightly lower than those recently reported for ubiquitin [29] ($\mu = -179.6$ ppm when scaled to a NH bond length of 1.02 Å), although within the average uncertainty of these measurements. These site-specific ^{15}N CSA values were then combined with the isotropic chemical shift data in order to reconstruct the individual components of the ^{15}N CST in GB3, assuming axial symmetry of the tensor (Supporting Information Tables 5a,b).

We observed no significant correlation between CSA values and secondary structure or amino acid type. There is no obvious correlation with the isotropic chemical shifts (Supporting Information Figure 8), although some residues with large $|\Delta\sigma|$, in particular Phe52 and Trp43, do show large isotropic shifts, while Asp49 and Gly38 have both small isotropic chemical shifts and $|\Delta\sigma|$. The mean CSAs of residues in the α-helix and β-strands are shown in Table 1. There is a weak correlation between the β_z angles and secondary structure, with slightly smaller angles in the β-strands (mean angle 18.9°) and turns (mean angle 19.1°) than in the helix (where the mean angle is

21.0°). Both the CSAs and β_z angles show smaller variation in the α-helix (where the standard deviations in the CSA and the angle are 18.1 ppm and 3.1°, respectively) compared to the β-strands (18.6 ppm and 4.7°), and even larger variations were observed in the loops/turns (26.3 ppm and 7.5°), possibly consistent with significantly different electronic arrangement in the secondary structures.

Site-to-Site ^{15}N CSA Variability in GB3: Comparison with Literature Data. The true site-to-site variability Λ in ^{15}N CSA obtained here is comparable to the standard deviation of the CSA values in ubiquitin[5,6] but significantly larger than the Λ values reported for Rnase H[8] and recently for ubiquitin.[29] The CSA distribution in ubiquitin, reconstructed from the individual CST components reported in ref 30, is in better agreement with our data for GB3: the standard deviations in these CSAs range from 10.1 to 13.7 ppm, and the site-to-site variability, Λ, from 7.8 to 10.5 ppm, depending on the model of local motion.

The value of Λ extracted from the observed site-specific CSA values, naturally, depends on the experimental uncertainties in CSA. Therefore, at least in principle, higher Λ values in GB3 could be a result of an underestimation of the experimental errors in the CSA. However, several lines of evidence suggest that this is not the case here. First of all, the residuals of fit from the diffusion tensor analyses (Table 2, column χ^2/df) are smaller than the ideal value of $\chi^2/df \approx 1$. This suggests that the errors in the relaxation and cross-correlation rates were possibly overestimated rather than underestimated. Second, the residuals of fit in the LS analysis (uniform CSA of −160 ppm) of the auto-relaxation data and NOEs at each field separately passed the goodness-of-fit test for the overwhelming majority of residues in GB3 (98%, 96%, 100%, 98%, and 84% of residues passed the 95% confidence test at 9.4, 11.7, 14.1, 16.4, and 18.8 T, respectively, and 97% overall), also suggesting that the errors in the relaxation data were not underestimated. Similar results were obtained for a CSA of −174.2 ppm. Third, to reduce Λ to the 5 ppm level reported in refs 8 and 29, we had to scale-up significantly the experimental errors in CSA (by a factor of 3 for the R/η method, 4 for the LS-CSA method, and >6.5 for the $2R_2 - R_1$ method) assuming that all errors are uniformly underestimated. This scaling factor is too large, given the reasonable χ^2/df values in all other fits presented here.

In addition, to further explore this issue, we introduced a certain χ^2/df cutoff level (determined here by a 95% confidence level for the goodness-of-fit test[50]) as a highly conservative criterion for eliminating fits from consideration here. This cutoff excludes those residues where the robust regressions were acceptable but the χ^2/df of the least-squares fit was too high due to an outlier that was effectively ignored by the robust methods: there are nine such exclusions from the $2R_2 - R_1$ method, six from R/η, and four from the LS-CSA fit. If only those residues with the χ^2/df of the least-squares fit lower than its 95% confidence limit are considered (32 amides from the $2R_2 - R_1$ method, 33 from R/η, and 25 from LS-CSA, represented by the filled symbols in Figure 2b,c,d, and Figure 5b), the CSA variability from each method is reduced to what could probably be considered its lower bound in GB3: $\Lambda_{2R_2-R_1} = 10.6$ ppm, $\Lambda_{R/\eta} = 10.2$ ppm, and $\Lambda_{\mathrm{LS-CSA}} = 11.9$ ppm. These estimates of the site-to-site CSA variability are still, consistently, almost a factor of 2 higher than those reported for Rnase H[8] or recently for ubiquitin.[29]

The results obtained here also differ from the ^{15}N CSA statistics in short peptides, where for a set of 39 solid-state NMR data (summarized in ref 60) we estimate a mean CSA of −155.8 ppm and a standard deviation of the distribution of 5.8 ppm. The larger range of CSA variability in GB3 compared to peptides could reflect greater internal structural heterogeneity in proteins.

To explore the effect of outliers as a possible source of the higher CSA variability observed here, we excluded from the set of residues for which $p(\mu, \Lambda)$ was generated for each method the extrema of the corresponding CSA range (Figure 2a, Figure 8b). The mean CSA values were largely unchanged ($\mu = -174.0$, −177.4, and −176.3 ppm for $2R_2 - R_1$, R/η, and LS-CSA, with Leu12 and Phe52, Ala26 and Phe52, and Gly38 and Phe52 excluded, respectively), and the measures of the site-to-site variability Λ were reduced to 17.2, 14.1, and 13.3 ppm, respectively. Restricting the CSA distribution even further by excluding all seven outliers in Figure 8 (Leu12, Ala20, Ala26, Gly38, Ala48, Thr49, and Phe52), thus effectively reducing the distribution to that contained within the dashed Gaussian curve shown in Figure 8, reduced the calculated site-to-site CSA variability Λ to 11.5, 13.8, and 13.1 ppm (for $2R_2 - R_1$, R/η, and LS-CSA, respectively), while the values of the true mean μ were only slightly affected (−176.5, −177.9, and −176.3 ppm, respectively). These exclusions also resulted in similar changes for the distribution function generated from the average CSAs of the three methods (Figure 8): $\mu = -175.1$ ppm and $\Lambda = 13.5$ ppm. Note that all these reduced estimates of the site-to-site variability in ^{15}N CSA are still significantly larger than those reported in refs 8 and 29.

In summary, all these data then suggest that the site-to-site variability in ^{15}N CSA reported here for GB3 is most probably correctly estimated or underestimated. This conclusion has important implications for the analysis of protein dynamics, since this degree of variability in the ^{15}N CSA means that the assumption of a uniform ^{15}N CSA value could result in significant errors in LS parameters.

Is There a Correlation between the Individual Components of the ^{15}N Chemical Shielding Tensor? It is instructive to discuss the CSA variability obtained here in relationship to the spread in the isotropic chemical shifts in GB3. The isotropic chemical shift (δ_{iso}) and the CSA are both combinations of the principal values of the ^{15}N CST: $\delta_{\mathrm{iso}} = (\delta_{xx} + \delta_{yy} + \delta_{zz})/3 \approx \sigma_{\mathrm{ref}} - (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3$; $\Delta\sigma \approx \sigma_{zz} - (\sigma_{xx} + \sigma_{yy})/2$, where σ_{ref} is the isotropic shielding of the reference compound, and the equation for $\Delta\sigma$ used here is an approximate form of eq 2, which is exact in the case of the axial symmetry of the CST, eq 4. Assuming a random model, when all three components of the ^{15}N CST are allowed to vary from site to site and are normally distributed with equal variances,[61] one can easily obtain from

(61) The standard deviations of the individual components of the ^{15}N chemical shift tensor for a collection of 39 solid-state measurements in short peptides (see p 221 in ref 60 for individual references) are approximately equal: 5.7, 7.3, and 6.5 ppm for δ_{zz}, δ_{yy}, and δ_{xx}, respectively. The standard deviations of the individual components of the ^{15}N CST derived from solution NMR measurements in ubiquitin (ref 28) are also approximately similar, ranging (depending on the model of local motion) from 6.8 to 9.1 ppm, 11.3 to 13.2 ppm, and 7.3 to 8.8 ppm for δ_{zz}, δ_{yy}, and δ_{xx}, respectively. Recall that the ^{15}N CST components are defined here such that $\sigma_{zz} \leq \sigma_{yy} \leq \sigma_{xx}$, i.e., σ_{zz} is the least shielded component.

these equations the following relationship between the standard deviations in the CSA (here referred to as the variability Λ) and in the isotropic chemical shift ($\Delta\delta_{iso}$):

$$\Lambda = \frac{3}{\sqrt{2}}\kappa\cdot\Delta\delta_{iso} \quad (20)$$

where κ is a numeric coefficient reflecting the interrelationship between the individual components of the CST:

$$\kappa = \sqrt{\frac{3 - 2R_{zx} - 2R_{zy} + R_{xy}}{3 + 2R_{zx} + 2R_{zy} + 2R_{xy}}} \quad (21)$$

Here R_{ij} is the correlation coefficient between σ_{ii} and σ_{jj}. In a particular case when all three CST components vary completely independently, $\kappa = 1$. Given that the standard deviation of the isotropic chemical shift in GB3 is 6.5 ppm, the expected value of Λ in this case is 13.8 ppm. This number is smaller than the CSA variability obtained for all residues in GB3 ($\Lambda_{2R_2-R_1} = 21.4$, $\Lambda_{R/\eta} = 17.6$, and $\Lambda_{LS-CSA} = 19.2$ ppm) but slightly larger than the values ($\Lambda_{2R_2-R_1} = 10.6$, $\Lambda_{R/\eta} = 10.2$, and $\Lambda_{LS-CSA} = 11.9$ ppm) obtained when considering only those residues with χ^2/df below the 95% goodness-of-fit cutoff. The deviation in the value of κ from 1 suggests that the individual components of the ^{15}N CST tensor are not independent from each other; however, it is impossible at this stage to draw a more definitive conclusion about the correlation coefficients between the individual components, and further studies are required to address this issue.

For example, it follows from eq 20 that a positive correlation between σ_{xx} and σ_{yy}, both being independent of σ_{zz}, will give $\kappa < 1$ (with the lower bound at $\kappa = 2/\sqrt{5}\}$), while an anticorrelation of these two components will result in $\kappa > 1$ (up to $\sqrt{2}$) with the upper bound on the CSA variability at $\Lambda = 3\Delta\delta_{iso}$ (or 19.5 ppm for GB3). It has been suggested [22] that σ_{xx} and σ_{yy} possibly vary in an anticorrelated manner; this would be consistent with the CSA variability in GB3 larger than 13.8 ppm. However, if the ^{15}N CST is truly axially symmetric (i.e., $\sigma_{xx} = \sigma_{yy}$; hence $R_{xy} = 1$), then the Λ value is expected to be smaller, $\Lambda = 3\sqrt{2/5}\Delta\delta_{iso}$, which gives the CSA variability around 12.3 ppm for GB3, again assuming that σ_{xx} and σ_{zz} (or $\sigma_{\perp}$ and $\sigma_{||}$ in this case) are normally distributed and vary independently. A positive correlation between $\sigma_{\perp}$ and $\sigma_{||}$ will further reduce the Λ values, down to zero at full correlation, while the anticorrelation will result in greater Λ's, with an upper bound at $\Lambda = 6\ \Delta\delta_{iso} = 39$ ppm. Using the correlation coefficients calculated from a collection[60] of 39 solid-state NMR data on short peptides, $R_{zx} = 0.06$, $R_{zy} = 0.43$, $R_{xy} = -0.12$, one would expect a Λ of 14 ppm in GB3. Inserting into eq 20 the correlation coefficients between the individual components of the ^{15}N CST recently measured in ubiquitin,[30] we estimate Λ to range from 9.6 to 13.3 ppm in ubiquitin (where the standard deviation in the isotropic chemical shift is 5.9 ppm) and from 10.5 to 14.6 in GB3.

Possible Sources of Systematic Errors in ^{15}N CSA Determination from Multiple-Field Data. In addition to the imprecision in the CSA values caused by random noise associated with the measurements, there could be systematic errors—largely inaccuracy—stemming from the underlying assumptions in the analysis. Here we focus on some of them; a detailed analysis can be found elsewhere.[31]

(A) The N−H Bond Length. As it is clear from eqs 9, 12, and 14, the ^{15}N CSA values are determined via the dipolar term d, hence depend on our knowledge of the N−H bond length. Two aspects are of importance here. First, a uniform value of the N−H bond length is usually assumed. Site-to-site variations in r_{HN} will necessarily affect the $\Delta\sigma$ values. Thus, a small, unaccounted for, deviation in the bond length by δr_{HN} will introduce an error in the CSA value on the order of $3(\delta r_{HN}/r_{HN})$. However, the currently available information on the variations in the N−H bond length is insufficient for a rigorous analysis of this issue. Second, the CSA values derived here were obtained assuming the N−H bond length of 1.02 Å. For comparison with the CSA data obtained for $r_{HN} = 1.04$ Å, our results should be uniformly scaled by $(1.02/1.04)^3 = 0.94$ (see also ref 31). Thus, the mean ^{15}N CSA and the site-to-site variability (average of all three methods) obtained here correspond to −164.3 and 20.0 ppm, respectively, if r_{HN} is 1.04 Å.

(B) Spectral Densities. The usual assumption made when analyzing ^{15}N relaxation data, be it the LS approach or the model-independent analyses, is to neglect the difference between the spectral densities describing the effect of motion on the contributions to the spin Hamiltonian from the ^{15}N−^{1}H dipolar interaction ($J_{DD}(\omega)$) and from the ^{15}N CSA ($J_{CSA}(\omega)$), i.e., $J_{DD}(\omega) = J_{CSA}(\omega) = J(\omega)$. In general, however,[31,62] $J_{DD}(\omega) \neq J_{CSA}(\omega)$, and a correction for the difference between the spectral densities can be included as

$$\Delta\sigma^{correct} = \Delta\sigma\cdot f \quad (22)$$

where f is the correction factor: $f = [J_{DD}(0)/J_{CSA}(0)]^{1/2}$ for the $2R_2 - R_1$ method, $f = \{[4J_{DD}(0) + 3J_{DD}(\omega_N)]/[4J_{CSA}(0) + 3J_{CSA}(\omega_N)]\}^{1/2}$ for R_2/η_{xy}, and $f = [J_{DD}(\omega_N)/J_{CSA}(\omega_N)]^{1/2}$ for R_1/η_z. There are several reasons why the spectral densities $J_{DD}(\omega)$ and $J_{CSA}(\omega)$ are not the same.[31]

First, the very nature of the chemical shielding suggests that it should fluctuate when the local environment of a nucleus changes as a result of internal motions in a protein. Here not only the orientation (as usually assumed in the equations relating relaxation rates to the spectral densities) of the CST but also the principal values themselves are expected to fluctuate. In contrast, the N−H bond length is less likely to change, except when transient hydrogen bonding occurs in the course of protein dynamics. Note also that the changes in local environment that modulate the CST do not necessarily have to affect the orientation of the N−H bond. A detailed analysis of the "breathing" of the ^{15}N CST requires molecular dynamics simulations (e.g., ref 63) and is beyond the scope of this paper.

Second, even when neglecting the differences in the mechanisms of modulation of these two tensors by motions within a protein, the difference between the spectral densities is expected to arise from the fact that the CSA and dipolar tensors are not collinear. As follows from our data (Figure 5c), the average angle β_z between the NH vector and the z-axis of the CSA tensor is 19.9°. The effect of CSA−dipolar noncollinearity on the contribution to the spectral density from anisotropic overall tumbling has been analyzed in detail in ref 62. Our calculations (not shown) using the average site-specific CSAs from the three

(62) Fushman, D.; Cowburn, D. *J. Biomol. NMR* **1999**, *13*, 139−147.
(63) Scheurer, C.; Skrynnikov, N. R.; Lienin, S. F.; Straus, S. K.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 4242−4251.

methods and the β_z angles (from R/η, Figure 5c) for GB3 resulted in the contributions from the noncollinearity to relaxation and cross-correlation rates that were on average within their respective experimental errors. As a result, the inclusion of these corrections in the model-independent and LS methods outlined above had no significant effect on the derived CSA values. In addition, because of the anisotropic character of backbone motion in proteins,[64,65] where the principal mode of motion is rocking of the peptide plane about the $C_\alpha-C_\alpha$ axis, the CSA−dipolar noncollinearity will result in different amplitudes (and associated order parameters) of the NH vector and CSA tensor motions. To investigate the effect of noncollinearity due to anisotropic backbone motions, we explored the difference in the order parameters for the NH vector and for a vector (representing the σ_{zz} axis) tilted by 20° toward the carbonyl atom in the peptide plane in a model system undergoing angular fluctuations about the $C_\alpha-C_\alpha$ axis. We found that the maximum difference in the squared order parameters for these vectors was 5%, with S_{CSA}^2 always smaller than S_{NH}^2, for a rotational angle of 40°, which is well above the maximum amplitude of Gaussian angular fluctuations about this axis recently reported for GB3.[66] Assuming that the correlation time of GAF motion is similar to that of the LS model, and the order parameters are close to 1; eq 22 gives $f \approx S_{NH}/S_{CSA} < 1.03$. This difference in the order parameters is insufficient to account for the large variability in the CSA that we observe in GB3. For example, if we assume for the sake of argument that the CSA in GB3 has a uniform mean value of −174.2 ppm, the factor f would have to range from 0.7 to 1.6 (hence $J_{DD}(0)/J_{CSA}(0)$ from 0.5 to 2.6) to account for the observed range of CSAs from the $2R_2 - R_1$ method. Similarly, to account for all the variability in the R/η measurements with respect to the average, f would have to vary from 0.7 to 1.4.

(C) Assumption of Axial Symmetry of the Overall Tumbling. The order parameters and the ^{15}N CSA values derived from the LS-based methods (but not those from the model-independent approaches) are sensitive to the model of overall tumbling used for the analysis. As demonstrated earlier[32] and further supported by the data presented here (Table 2, Supporting Information Tables 3, 4), the overall tumbling of GB3 in solution is anisotropic. While the axially symmetric and fully anisotropic tumbling models both provide a significant improvement in the fit over the isotropic diffusion model, the axially symmetric model for the overall tumbling was assumed here, based on several lines of evidence.

(1) The molecular shape of GB3 to a good approximation is axially symmetric. The ratio of the principal values of the inertia tensor of the molecule is 1.80:1.79:1.00. Moreover, theoretical predictions for GB3[32] based on hydrodynamic calculations using the HYDRONMR program[67] give a rotational diffusion tensor with the ratio of the principal components of 1.00:1.05:1.43, which suggests a high degree of axial symmetry.

(2) The fully anisotropic diffusion tensor derived from the relaxation data (Supporting Information Table 4) also shows a high degree of axial symmetry, with the principal values of the tensor, D_{xx} and D_{yy}, within their mutual errors at all fields. Also a global fit of the relaxation of data at all five fields resulted in a diffusion tensor with near zero rhombicity (0.08). This is also reflected in the large experimental uncertainties in the orientation of the x- and y-axes of the fully anisotropic tensor (angle Ψ, Supporting Information Table 4), indicating that the orientations of these axes of the diffusion tensor are not well defined.

(3) Based on the statistical F-test,[50] the probabilities that the observed reduction in χ^2 for the fully anisotropic model compared to the axially symmetric model occurred by chance (rightmost column in Supporting Table 4) at each field are not low enough (or the corresponding F-values are not high enough) in order to reject with certainty the null-hypothesis that both models fit the data similarly well.

(4) The fact that the spectral densities obtained using the axially symmetric model are in good agreement with the model-independent analyses (Figure 3) derived without any assumption about the overall tumbling further supports this conclusion.

All these observations support the conclusion that the axially symmetric diffusion tensor is an adequate model for rotational diffusion of GB3. (Note that the theoretical predictions mentioned above suggest that this is likely due to the overall shape of the protein, rather than a consequence of the quality or limited amount of experimental data.) However, to completely rule out the possibility that some ^{15}N CSA values might be affected by the neglect of the deviation of the diffusion tensor from axial symmetry and therefore might appear site-specific due to the individual orientations of amide bonds with respect to the protein's diffusion tensor, we also performed the LS-CSA analysis using the globally fit fully anisotropic diffusion tensor (Supporting Information Table 4). These ^{15}N CSAs are in excellent agreement (Supporting Information Figure 6) with the above-reported CSA values derived for the axially symmetric diffusion tensor. It should be noted here that the model-independent methods for CSA determination ($2R_2 - R_1$, R/η, and $\eta_{xy} - \eta_z$) presented in the text do not depend on the overall diffusion model. Therefore, the good agreement between the ^{15}N CSAs determined using the LS-CSA and LS-SDF-CSA methods for an axially symmetric tumbling model with the results of the model-independent analyses (Figure 2c,d) is also a strong indication that the results presented in this paper are not affected by the axially symmetric diffusion tensor representation.

^{15}N CSAs and the Order Parameters: What Errors in the Order Parameters Are Expected? As shown in this study, relaxation data at multiple fields allowed an accurate assessment of the site-specific ^{15}N CSAs, and these values, in turn influence the order parameters extracted from the data. Because measurements at multiple fields (particularly higher fields) are not always available to a general NMR user, it is instructional to estimate the level of uncertainties in the order parameters expected from the use of a constant CSA instead of the true CSA values. A comparison of the order parameters obtained from the LS-CSA analysis of all five-field data with those obtained for a typical field of 14.1 T, assuming a constant CSA, gave pairwise rmsd values of 0.06 (or 6.5%, range of deviations from −0.06 to 0.11) for −160 ppm and 0.04 (or 4.1%, range from −0.09 to 0.07) for −174.2 ppm. The corresponding numbers for 11.4 T were, naturally, smaller: rmsd = 0.04 (4.9%, range from −0.04 to 0.09) for −160 ppm and 0.03 (3.2%, range

(64) Bremi, T.; Bruschweiler, R. *J. Am. Chem. Soc.* **1997**, *119*, 6672−6673.
(65) Lienin, S. F.; Bremi, T.; Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1998**, *120*, 9870−9879.
(66) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Bruschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885−13890.
(67) García de la Torre, J.; Huertas, M. L.; Carrasco, B. *J. Magn. Reson.* **2000**, *B147*, 138−146.

−0.07 to 0.06) for −174.2 ppm. This comparison included only residues in the secondary structure of GB3; the deviations in the loop regions could be larger. Thus, even at low fields, the errors in the order parameters might not be negligible, particularly for those applications where quantitative changes in order parameters are of importance (such as entropy changes monitored by ^{15}N relaxation).

Conclusions

Here we presented a comprehensive study of the ^{15}N chemical shielding anisotropy in a protein based on a combination of ^{15}N relaxation and CSA/dipolar cross-correlation measurements at five static magnetic fields. The analysis was performed using various combinations of the experimental data and using model-independent approaches as well as methods based on the Lipari−Szabo approximation. The results indicate significant site-to-site variations in the principal values and the orientation of the ^{15}N CSA, similar to those observed earlier in ubiquitin.[5,6] Our estimates of the true variability in the ^{15}N CSA in GB3 depend to some degree upon which method for determining the CSA was used and which subset of residues is considered. These estimates range from 10.2 ppm (for 33 residues that pass the χ^2/df cutoff from the R/η method) to 21.4 ppm for all 47 residues from the $2R_2 - R_1$ method. Although this range of values could be a result of limited statistics, all of these estimates are still larger than the derived variability in the ^{15}N CSA from studies of ribonuclease H[8] or recently of ubiquitin using a subset of the methods used here.[29] The true mean CSA values range from −173.9 ppm ($2R_2 - R_1$) to −177.2 ppm (R/η).

Our data show that using the site-specific values of the ^{15}N chemical shielding anisotropy obtained here significantly improves the agreement between LS order parameters measured at different fields and allows simultaneous fit of the ^{15}N relaxation data at five fields to LS spectral densities. These findings emphasize the necessity of taking into account the variability of the ^{15}N chemical shielding tensor for accurate analysis of protein dynamics from ^{15}N relaxation measurements. This can be achieved by including CSA as an additional fitting parameter in the LS analysis of multiple-field data, provided the sample temperature and other experimental conditions were the same at all fields/spectrometers. These analyses also show that the Lipari−Szabo form of the spectral density provides a satisfactory approximation for the experimental spectral densities obtained using the reduced spectral density approach.

Significant variation in the true CSAs from their assumed values will affect several applications of NMR relaxation that depend on ^{15}N CSA. These include, in addition to the backbone order parameters and local correlation times from Lipari−Szabo analysis, the spectral density components (specifically, $J(0)$ and $J(\omega_N)$, but not $J(\omega_H)$) obtained from spectral density mapping, conformational exchange contributions derived from the field dependence of the ^{15}N relaxation rates (see discussion in ref 6), and local molecular geometries and order parameters determined from cross-correlation measurements involving the CSA mechanism. Among other characteristics that could be influenced by CSA variability are changes in the local conformational entropy (for example, accompanying ligand binding) estimated from the differences in order parameters. As mentioned above, characteristics of the overall tumbling are not expected to be sensitive to the ^{15}N CSA values, when determined from the ratio of cross-correlation or reduced relaxation rates.

In contrast to other heteronuclei (e.g., carbonyl ^{13}C[68,69]) the ^{15}N shifts (and CSAs) in proteins are not yet predictable, indicating that the subtleties of noncovalent bonding forces are still poorly understood in proteins. The site-specific ^{15}N CSA values presented above provide experimental data for testing and calibration of theoretical methods for shielding tensor predictions.

Acknowledgment. This work was supported by NIH grant GM65334 to D.F. We thank Dr. Ad Bax for providing us with the protein sample and for the encouragement and Dr. David Cowburn for useful suggestions concerning the manuscript. The measurements at 9.4, 16.4, and 18.8 T were performed at CERM (University of Florence), and we thank Prof. Ivano Bertini for the NMR time and Dr. Rainer Kümmerle (Bruker) for help with the experimental setup. The modified version of our program DYNAMICS that includes ^{15}N CSA as an additional fitting parameter is available from the authors upon request.

Supporting Information Available: Eight figures depicting the agreement between R_2 and R_{2free} values, the order parameters obtained from LS analysis with no restriction on the model selection, and, for each residue, fits of the experimental data for all three model-independent methods used here, a comparison of CSAs obtained from the LS-CSA analysis using the axially symmetric and fully anisotropic rotational diffusion tensors, the LS-SDF fit with a CSA of −174.2 ppm, and a comparison of ^{15}N CSAs with isotropic chemical shifts. Five tables listing the experimental delays, all experimental data (relaxation and cross-correlation rates and hetero-NOEs), the results of the axially symmetric rotational diffusion tensor analysis using GB3 solution structure, the fully anisotropic rotational diffusion tensor analysis, and the individual components of the ^{15}N CSTs for GB3. This material is available free of charge via the Internet at http://pubs.acs.org.

JA060406X

(68) Markwick, P. R.; Sattler, M. *J. Am. Chem. Soc.* **2004**, *126*, 11424−11425.
(69) Markwick, P. R.; Sprangers, R.; Sattler, M. *Angew. Chem. Int. Ed.* **2005**, *44*, 3232−3237.
(70) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179−9191.