Amplitudes of Protein Backbone Dynamics and Correlated Motions in a Small α/β Protein: Correspondence of Dipolar Coupling and Heteronuclear Relaxation Measurements†

G. Marius Clore*,‡ and Charles D. Schwieters§

Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, and Division of Computational Bioscience, Building 12A, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892-5624

Received April 1, 2004; Revised Manuscript Received May 14, 2004

ABSTRACT: Backbone residual dipolar coupling (N−H, Cα−Hα, N−C′, and Cα-C′) data collected in five different media on the B3 IgG binding domain of streptococcal protein G (GB3) have been analyzed by simultaneous refinement of the coordinates and optimization of the magnitudes and orientations of the alignment tensors using single and multiple structure representations. We show, using appropriate error analysis, that agreement between observed and calculated dipolar couplings at the level of experimental uncertainty is obtained with a two-structure ($N_e = 2$) ensemble representation which represents the simplest equilibrium description of anisotropic motions. The data permit one to determine the magnitude of the anisotropic motions along the four different backbone bond vectors in terms of $\langle S^2(\text{jump})\rangle$ order parameters. The order parameters, $\langle S^2_{NH}(\text{jump})\rangle$, for the N−H bond vectors are in qualitative agreement with the generalized order parameters, $\boldsymbol{S}^2_{NH}$(relaxation), derived from ^{15}N relaxation measurements, with a correlation coefficient of 0.84. $\boldsymbol{S}^2_{NH}$(relaxation) can be regarded as the product of an anisotropic order parameter, corresponding to $\langle S^2_{NH}(\text{jump})\rangle$ derived from the residual dipolar couplings, and an axially symmetric order parameter, S^2_{NH}(axial), corresponding to bond librations which are expected to be essentially uniform along the polypeptide chain. The current data indicate that the average value of S^2_{NH}(axial) is ∼0.9. The close correspondence of $\langle S^2_{NH}(\text{jump})\rangle$ and $\boldsymbol{S}^2_{NH}$(relaxation) indicates that any large-scale displacements from the mean coordinate positions on time scales longer than the rotational correlation time are rare and hence do not perturb the observed dipolar couplings. Analysis of a set of 100 $N_e = 2$ ensembles reveals the presence of some long-range correlated motions of N−H and Cα−Hα vectors involving residues far apart in the sequence but close together in space. In addition, direct evidence is obtained for ubiquitous crankshaft motions along the entire length of the polypeptide backbone manifested by the anticorrelation of the backbone torsion angles ϕ_i and ψ_{i-1}.

Protein dynamics are crucial to protein function and have been the subject of extensive investigation over many years using a variety of experimental (spectroscopic and crystallographic) and theoretical (molecular dynamics simulations and normal-mode analysis) approaches (*1*−*16*). In addition, correlated motions, which are extremely difficult to observe experimentally, may play an important role in recognition and binding processes, enzyme catalysis, and protein folding and unfolding.

The principal experimental approach that has been traditionally used to study protein dynamics in solution has involved the application of heteronuclear NMR relaxation measurements which, in principle, can provide information for almost all heavy atom positions in a protein (*1*−*7*). Motions on time scales from picoseconds to nanoseconds and from microseconds to milliseconds can be probed using laboratory and rotating frame measurements, respectively. The majority of published work has focused on the application of ^{15}N relaxation measurements to examine the dynamics of backbone N−H bond vectors, although ^{13}C and ^{2}H relaxation measurements have recently come to the fore in the investigation of side-chain dynamics.

Residual dipolar couplings (RDCs)[1] measured in weak alignment media are highly sensitive to internuclear vector orientations relative to the external alignment tensor and yield extremely powerful orientational restraints for structure determination (*17*−*22*). RDCs are also sensitive to motion (*20*−*22*). Thus, axially symmetric motion will scale the magnitude of the observed RDC by the generalized order parameter $\boldsymbol{S}$ (as opposed to $\boldsymbol{S}^2$ in the case of relaxation measurements). Anisotropic motions have a more complex

† This work was supported in part by the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G.M.C.).

* To whom correspondence should be addressed. E-mail: mariusc@intra.niddk.nih.gov. Tel: (301) 496-0782. Fax: (301) 496-0825.

‡ Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases.

§ Division of Computational Bioscience, Center for Information Technology.

[1] Abbreviations: RDC, residual dipolar coupling; GB3, B3 IgG binding domain of streptococcal protein G.

10.1021/bi049357w
Published on Web 07/29/2004

effect on the observed RDCs. A number of studies have indicated that RDCs measured in multiple alignment media can potentially yield information relating to amplitudes of both axially symmetric and anisotropic motions on a time scale ranging from picoseconds to about 1 ms (*23*−*30*). Earlier work suggested that the amplitudes of motions derived from RDC data are considerably larger than those probed by relaxation measurements (time scale picoseconds to nanoseconds), which led to the suggestion that proteins generally exhibit large-scale concerted motions that are present at all times over the nanosecond to millisecond range (*23*−*30*). More recently, we have developed an intuitively simple approach for deriving both magnitudes of motions and ascertaining the presence of correlated motions from RDC data recorded in multiple alignment media that makes use of ensemble averaging coupled with simultaneous refinement of the coordinates and optimization of the magnitude and orientation of the alignment tensor (*31*). Using previously published N−H RDC data on ubiquitin in 11 different alignment media (*28*−*30*), we were able to show that, with only a few exceptions, the data can be accounted for by rather small amplitude anisotropic motions [with S^2(jump) values ≥ 0.8], and that the structural impact of those few residues that undergo larger scale motions is small and can be accommodated by minimal shifts in the atomic coordinates (*31*). The ubiquitin data, however, were limited in both accuracy and extent, insofar that the RDCs in all but 2 of the 11 media were confined to the N−H bond vectors.

Recently, Ulmer et al. (*32*) published a set of highly accurate backbone RDC data (N−H, Cα−Hα, N−C′, and Cα−C′) with well-defined experimental uncertainties in five alignment media for the small (56-residue) B3 IgG binding domain (GB3) of streptococcal protein G. Analysis of these data, excluding regions of known mobility as assessed from ^{15}N relaxation data (*33*), has been used to demonstrate the existence of small deviations from peptide bond planarity (*32*). In addition, the structure of GB3 has been solved at a resolution of 1.1 Å (*34*), enabling highly accurate generalized order parameters for the N−H bond vectors to be derived from ^{15}N relaxation data (*33*). GB3, therefore, provides an ideal system for probing, for the first time, motion (both magnitude and correlation effects) along four different backbone bond vectors using our previously developed methodology. We show that the $\langle S^2_{NH}(\text{jump})\rangle$ order parameters for the N−H bond vectors derived from the RDC data are in good agreement with the general order parameters, $\boldsymbol{S}^2_{NH}$(relaxation), derived from ^{15}N relaxation measurements. In addition, our analysis reveals the presence of extensive anticorrelated crankshaft motions along the backbone, as well as some long-range correlations between residues that are distant in primary sequence but in close spatial proximity.

METHODS

Residual Dipolar Couplings. Backbone N−H, Cα−Hα, N−C′, and Cα−C′ residual dipolar couplings, measured in five different alignment media, on the B3 domain of streptococcal protein G (GB3) were taken from Ulmer et al. (*32*). The five media consist of the following: (i) 4.2% bicelles (w/v) comprising a 30:10:1 ratio of ditetradecylphosphatidylcholine, dihexylphosphatidylcholine, and cetyltrimethylammonium bromide; (ii) 4.3% (w/v) C12E5 alkyl poly(ethylene glycol) and 1-hexanol in a ratio of 0.96, 11 mg/mL bacteriophage pf1, and axially stretched positively and negatively charged polyacrylamide gels. The errors in the measured N−H, Cα−Hα, N−C′, and Cα−C′ RDCs were 0.26, 0.58, 0.10, and 0.10 Hz, respectively (*32*).

Structure Refinement. All simulated annealing and minimization calculations were carried out using the IVM internal variable module (*35*) within the molecular structure determination package Xplor-NIH (*36*), incorporating the various new features related to ensemble averaging and RDCs described previously (*31*). Structures were visualized using the molecular graphics package VMD-XPLOR (*37*).

Calculations were carried out using ensemble sizes of $N_e = 1$, 2, 3, and 8. A total of 100 ensembles were calculated in each case. The protocol employed for structure refinement was identical to that described in ref *8*. The target function comprised terms for the experimental NMR restraints, covalent geometry, and nonbonded contacts. The experimental NMR restraints comprise 1031 RDCs with final reference force constants of 15, 3, 100, and 100 kcal·mol^{-1}·Hz^{-2} for the N−H, Cα−Hα, N−C′, and Cα−C′ RDCs, respectively. This weighting scheme is designed to achieve agreement between observed and calculated RDCs at the level of experimental uncertainty. Since the target function for the RDC term is represented by a quadratic harmonic potential (*8*), the force constants, k_q, employed for the different RDC types are readily derived from their respective experimental errors, Err_q, using the equation $k_q = k_{NH}(\text{Err}_{NH}/\text{Err}_q)^2$, where k_{NH} and Err_{NH} are the force constant and experimental error, respectively, for the N−H RDCs. In addition, the RDC force constants employed for the different media, m, are weighted relative to those of the reference medium (in this instance PEG/hexanol) by a scale factor given by $[D_a^{NH}(\text{PEG/hexanol})/D_a^{NH}(\text{m})]^2$, where D_a^{NH}(m) is the magnitude of the alignment tensor for medium m, thereby ensuring that each medium contributes equally to the target function. (Empirically, to attain an observed experimental error of 1 Hz, the force constant for the RDCs should be set to 1 kcal·mol^{-1}·Hz^{-2}. Hence, to approximate an experimental error of 0.1 Hz, the force constant should be set to 100 kcal·mol^{-1}·Hz^{-2}.)

The RDC restraints were supplemented by 68 distances for 34 backbone hydrogen bond restraints [derived from $^{3h}J_{NC'}$ couplings (*38*)] with a final force constant of 30 kcal·mol^{-1}·Å^{-2} (for each backbone hydrogen bond there are two distance restraints: $r_{NH-O} \leq 2.3$ Å, $r_{N-O} \leq 3.3$ Å) and 143 torsion angle restraints with a final force constant of 200 kcal·mol^{-1}·rad^{-2} (consisting of 53 ϕ and 51 ψ backbone torsion angle restraints with minimum ranges of ±30°) and 30 χ_1 and 9 χ_2 internal side-chain torsion angle restraints with minimum ranges of ±50° and ±20°, respectively, taken from the previously published data on the GB1 domain (*39*, *40*) which is 86% sequence identical to the GB3 domain (note all torsion angle restraints are consistent with the crystal structure of GB3; the torsion angle restraints on ϕ and ψ serve simply to prevent the potentially rare occurrence of local mirror images during the course of simulated annealing, while the side-chain torsion angle restraints for internal side chains ensure that unwarranted rotamer flips do not occur; the presence or absence of these torsion angle restraints does not impact the results of the current analysis). The force

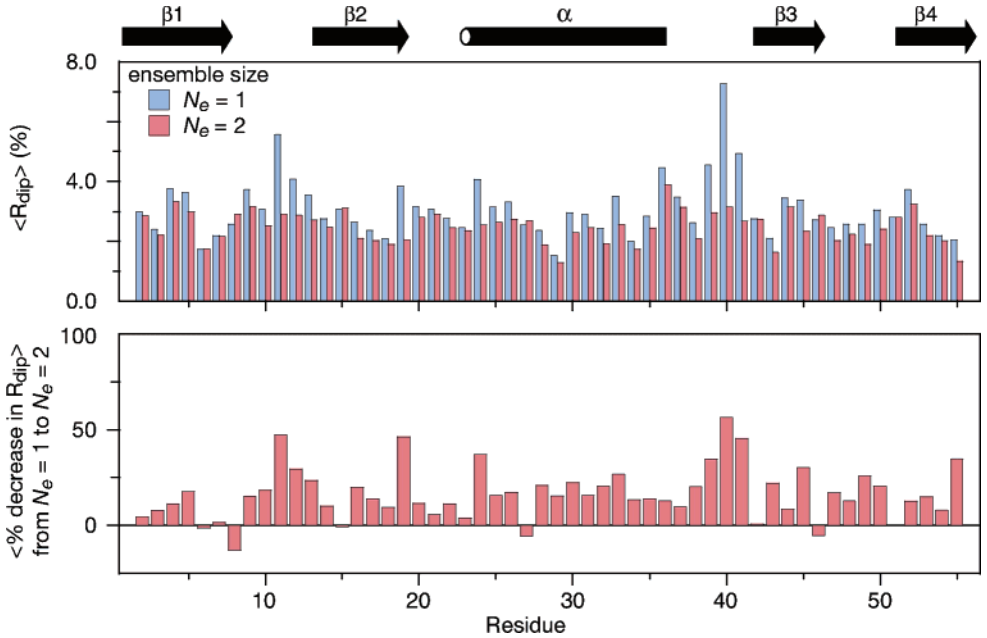

FIGURE 1: Overall agreement between observed and calculated dipolar couplings for GB3 as a function of residue for structures refined with ensemble sizes of $N_e = 1$ and 2. The top panel displays the overall average dipolar coupling *R*-factor (R_{dip}) (i.e., the average for the N−H, Cα−Hα, N−C′, and Cα−C′ RDCs over all five media) for the $N_e = 1$ (blue) and 2 (red) calculations. The bottom panel displays the percentage decrease in the overall R_{dip} upon increasing the ensemble size from $N_e = 1$ to 2 and is given by $100(R_{dip}^{N_e=1} - R_{dip}^{N_e=2})/R_{dip}^{N_e=1}$. The angle brackets denote averaging over all 100 calculated ensembles.

constants for the bonds, angles, and improper torsions (excluding those related to the peptide bond) were set to 1000 kcal·mol^{-1}·Å^{-2}, 500 kcal·mol^{-1}·rad^{-2}, and 500 kcal·mol^{-1}·rad^{-2}, respectively. The three improper torsion angles related to the planarity of the peptide bond and the planarity of the carbonyl oxygen atom and HN atom relative to the peptide plane were set to 250 kcal·mol^{-1}·rad^{-2} to permit deviations from ideality of the magnitude observed by Ulmer et al. (*32*) to occur. The nonbonded contacts were represented by the following five terms: a quartic van der Waals repulsion term with a final force constant of 4 kcal·mol^{-1}·Å^{-4} and a van der Waals radius scale factor of 0.8 (*41*); a torsion angle database potential of mean force with a final weighting factor of 1.0 (*42*); a radius of gyration restraint (*40*) with a force constant of 100 kcal·mol^{-1}·Å^{-2} and a target value of 10.6 Å, corresponding to the radius of gyration of the crystal structure used to ensure optimal packing (*34*) (the presence or absence of this term does not impact the results of the current analysis); and an empirical hydrogen-bonding term incorporating both distance and angular dependencies with a weighting factor of 500 (*43*). For calculations with an ensemble size $N_e \geq 2$, four additional terms were included: the relative atomic position term (eq 14 in ref *31*) applied to the Cα atoms with a force constant of 100 kcal·mol^{-1}·Å^{-2} and a width of 0.6 Å; the shape term (eq 16 of ref *31*) applied to atoms outside a sphere of 5 Å from the center of the molecule defined by the Cα atoms of residues 1−56 with a force constant of 1 kcal·mol^{-1}·Å^{-2} for the size component and 1000 kcal·mol^{-1}·rad^{-2} for the orientation component; and harmonic spread terms for D_a (eq 7 of ref *31*) and rhombicity (eq 8 of ref *31*) with weighting factors of 10^5 kcal·mol^{-1}·Hz^{-2} and 10^5, respectively.

The cooperative nature of the backbone motions was examined using two correlation functions related to bond vector orientations and torsion angles. The bond vector correlation function C_{kl} between residues k and l (given by eqs 19 and 20 in ref *8*), averaged over all ensembles, is a measure of how closely the various bond vectors of constituent members of each ensemble line up with those from all other calculated ensembles and over what range of residues. C_{kl} scales between 0 and 1, with 0 indicating no correlation and 1 perfect correlation. The normalized correlation function of the joint fluctuations in backbone torsion angles (*44*, *45*), $\sigma(\alpha_k,\beta_l)$, is given by $\langle\Delta\alpha_k\Delta\beta_l\rangle/(\langle(\Delta\alpha_k)^2\rangle\langle(\Delta\beta_l)^2\rangle)^{1/2}$, where α_k is a ϕ or ψ backbone torsion angle of residue k and β_l is a ϕ or ψ torsion angle of residue l, and the average is over all ensembles. Δα is the difference of the torsion angle α from the ensemble mean value of this angle. $\sigma(\alpha_k,\beta_l)$ has values ranging from −1 to +1, with −1 indicating perfect anticorrelation, 0 no correlation, and +1 perfect correlation.

RESULTS AND DISCUSSION

Ensemble Refinement of GB3 against Backbone Dipolar Couplings in Five Alignment Media. The results of simultaneous refinement using ensemble sizes of $N_e = 1$ and 2 against N−H, Cα−Hα, N−C′, and Cα−C′ one-bond RDCs recorded in five alignment media with significantly different alignment tensors are summarized in Figure 1 and Table 1. The resulting values of the magnitude of the alignment tensors in the five media are provided in Table 2.

As discussed in detail previously (*31*), the two-structure ensemble ($N_e = 2$) formulation provides the simplest and most direct description of the equilibrium magnitudes of anisotropic motions. The overall calculated RDCs in the $N_e = 2$ representation are given by the average of the calculated RDCs of the individual members of the ensemble (i.e., averaging occurs at the level of the RDCs and not the coordinates). It is important to stress that the $N_e = 2$ representation does not involve the assumption of an overall two-state model since the majority of local motions are uncorrelated; that is to say, correlations tend to only extend

Table 1: Refinement of GB3 against All Residual Dipolar Couplings in All Five Media[a]

dipolar couplings[b]	X-ray[c]		refined[d]			
			$N_e = 1$[e]		$N_e = 2$	
	rms (Hz)	R_{dip} (%)	rms (Hz)	R_{dip} (%)	rms (Hz)	R_{dip} (%)
N−H (246)	1.98 ± 0.45	12.7 ± 1.7	0.50 ± 0.06	3.2 ± 0.5	0.30 ± 0.08	1.8 ± 0.3
Cα−Hα (269)	3.51 ± 0.75	11.2 ± 0.9	0.82 ± 0.26	2.5 ± 0.4	0.47 ± 0.17	1.4 ± 0.3
N−C′ (259)	0.35 ± 0.10	18.3 ± 3.5	0.12 ± 0.02	6.2 ± 1.1	0.11 ± 0.02	5.5 ± 1.0
Cα−C′ (268)	0.41 ± 0.07	13.2 ± 0.7	0.15 ± 0.04	4.7 ± 0.7	0.13 ± 0.04	4.0 ± 0.6

[a] The dipolar coupling *R*-factor, R_{dip}, is given by the ratio of the rms difference between observed and calculated dipolar couplings and the expected value of the rms difference if the vectors are randomly oriented (*47*). The latter is given by $[2\lambda D_a^{NH}(4 + 3\eta^2)/5]^{1/2}$, where D_a^{NH} is the magnitude of the axial component of the alignment tensor for the N−H dipolar couplings, η is the rhombicity, and λ is a normalization factor for the other dipolar couplings given by $\gamma_m\gamma_n r_{NH}^3/\gamma_N\gamma_H r_{mn}^3$ where γ are the gyromagnetic ratios for the difference nuclei (proton, nitrogen, or carbon), and the vibrationally corrected, effective bond distances r_{mn} are taken from refs *62* and *69*. The experimental errors of 0.26, 0.58, 0.1, and 0.1 Hz for the N−H, Cα−Hα, N−C′, and Cα−C′ RDCs, respectively, correspond to R_{dip} values of 1.56%, 1.73%, 5.16%, and 3.13%, respectively. Thus, the relative errors, when taking the scaling of the D_a's and consequent range of RDCs for the different bond vectors into account, are substantially higher for the N−C′ (3-fold) and Cα−C′ (∼2-fold) RDCs than for the NH and Cα−Hα RDCs. [b] The number of measured dipolar couplings is given in parentheses. The values of χ^2/N (where $N = 1042$, the total number of measured RDCs) are 30.51, 2.33, and 1.22 for the X-ray structure, the $N_e = 1$ ensemble, and the $N_e = 2$ ensemble, respectively. (The latter two values are mean values obtained by averaging over all 100 calculated ensembles.) Using the statistical *F* test, the probability *P* that the reduction in χ^2 for the $N_e = 2$ ensemble relative to the $N_e = 1$ ensemble could occur by chance is less than 10^{-16} (*70*). [c] The 1.1 Å resolution X-ray coordinates (PDB accession code 1IGD) are taken from ref *34*. Protons were added using Xplor-NIH (*36*) with standard covalent geometry. [d] The mean values and standard deviations reported are obtained by averaging over all 100 calculated ensembles. [e] Calculations with $N_e = 1$ were also carried out in which the experimental dipolar couplings were scaled by $1/\boldsymbol{S}$(relaxation), where $\boldsymbol{S}_{NH}^2$(relaxation) are the generalized N−H order parameters derived from ^{15}N relaxation measurements (*33*). [Note that differences in $\boldsymbol{S}$(relaxation) for different types of interactions are already taken into account by using vibrationally corrected, effective internuclear distances in the calculation of the RDCs (*62*, *69*).] The resulting rms values for the N−H, Cα−Hα, N−C′, and Cα−C′ RDCs are 0.64 ± 0.12, 1.13 ± 0.32, 0.13 ± 0.02, and 0.21 ± 0.06 Hz, respectively; the corresponding dipolar coupling *R*-factors are 3.8 ± 0.5%, 3.4 ± 0.5%, 6.6 ± 1.3%, and 6.2 ± 1.3%, respectively. With the exception of N−C′, which is only minimally affected, the agreement between scaled and calculated dipolar couplings for the other three bond vectors is 30−40% worse than that obtained using the raw (i.e., unscaled) experimental dipolar couplings.

Table 2: D_a^{NH} and Rhombicity (η) of the Alignment Tensors for GB3 in the Five Liquid Crystalline Media

no.	medium[a]	X-ray[b]		refined[c]			
				$N_e = 1$		$N_e = 2$	
		D_a^{NH} (Hz)	η	D_a^{NH} (Hz)	η	D_a^{NH} (Hz)	η
1	bicelle/CTAB	−15.80	0.26	−15.91 ± 0.01	0.23 ± 0.00	−16.44 ± 0.05	0.22 ± 0.00
2	PEG/hexanol	−8.74	0.23	−8.85 ± 0.01	0.21 ± 0.00	−9.02 ± 0.04	0.21 ± 0.00
3	pf1	13.13	0.11	13.67 ± 0.01	0.09 ± 0.00	13.96 ± 0.05	0.09 ± 0.00
4	positive gel	10.17	0.66	10.16 ± 0.01	0.66 ± 0.00	10.47 ± 0.04	0.67 ± 0.00
5	negative gel	10.70	0.46	11.43 ± 0.01	0.43 ± 0.00	11.68 ± 0.04	0.42 ± 0.00

[a] The values of the scalar tensor products between medium 1 and media 2−5 are 0.80, −0.39, 0.86, and 0.33, respectively; between medium 2 and media 3−5 are −0.60, 0.58, and 0.23, respectively; between medium 3 and media 4−5 are −0.60 and 0.64, respectively; and between medium 4 and medium 5 is −0.14. [b] The 1.1 Å resolution X-ray coordinates (PDB accession code 1IGD) are taken from ref *34*. [c] The mean values and standard deviations reported are obtained by averaging over all 100 calculated ensembles.

to adjacent neighbors. In the context of the present calculations, the extent of correlated motions can only be quantified by examining the distribution of bond vectors within a large collection of calculated ensembles (in this instance 100).

The backbone atomic rms difference between the 100 calculated $N_e = 1$ structures and the X-ray coordinates (*34*) is 0.56 ± 0.01 Å, and the corresponding value for the 100 ensemble means of the $N_e = 2$ structures is 0.60 ± 0.02 Å. Refinement using an ensemble size of $N_e = 1$ results in a large (3−4-fold) reduction in RDC *R*-factors (*46*) relative to the X-ray coordinates. However, while the agreement between observed and calculated N−C′ and Cα−C′ RDCs is essentially at the level of the error in the measurement (∼0.1 Hz), the rms difference between observed and calculated values for the N−H and Cα−Hα RDCs is about 40% larger than the experimental error. Refinement using an ensemble size of $N_e = 2$ results in an ∼40% reduction in the *R*-factor for the N−H and Cα−Hα RDCs such that the rms difference between observed and calculated couplings is at the level of the experimental uncertainty [0.26 and 0.58 Hz for N−H and Cα−Hα RDCs, respectively (*32*)]. There is a small decrease in the RDC *R*-factors for the N−C′ and Cα−C′ RDCs, but this is not significant. Figure 1 displays the overall residue-based RDC *R*-factor for the $N_e = 1$ and 2 refinements together with the percentage decrease in the overall residue-based RDC *R*-factor from $N_e = 1$ to 2 as a function of residue number. Although the largest decreases involve residues in the turns between strands β1 and β2, and in the loops preceding and following the α-helix, significant decreases in the residue-based RDC *R*-factor are observed throughout the protein. Refinement using larger ensemble sizes ($N_e = 3$ and 8) results in no further significant reductions in RDC *R*-factors for any of the bond vectors.

As indicated in the Methods section the force constant for the improper torsion angles related to the planarity of the peptide bond was reduced by two relative to the others to permit deviations in peptide bond planarity to occur. The average rms deviation from peptide bond planarity for both the $N_e = 1$ and 2 calculations is 3.8 ± 0.3°, and the peptide bond torsion angle ω ranges from approximately 174° to −175°. This is fully consistent with the range of deviations observed in very high resolution protein and peptide crystal

Table 3: Refinement of GB3 with Cross-Validation[a,b]

dipolar couplings[c]	$N_e = 1$ rms(free) (Hz)	$N_e = 1$ R_{dip}(free) (%)	$N_e = 2$ rms(free) (Hz)	$N_e = 2$ R_{dip}(free) (%)
N−H (246)	1.80 ± 0.43	11.3 ± 1.5	1.79 ± 0.43	11.0 ± 1.4
Cα−Hα (269)	2.89 ± 0.56	9.2 ± 0.9	2.71 ± 0.50	8.4 ± 0.7
N−C′ (262)	0.21 ± 0.04	11.2 ± 0.8	0.22 ± 0.04	11.3 ± 1.1
Cα−C′ (254)	0.29 ± 0.07	9.1 ± 0.8	0.27 ± 0.06	8.4 ± 0.8

[a] Only the rms and free *R*-factors, R_{dip}(free), for the cross-validated RDC terms are shown. (Note that the free *R*-factor refers to the calculated cross-validated *R*-factor for RDCs *excluded* from the refinement; the working *R*-factor refers to the *R*-factor for the RDCs *included* in the refinement.) Four sets of cross-validation calculations were carried out, leaving out in turn the N−H, Cα−Hα, N−C′, and Cα−C′ dipolar couplings. The mean values and standard deviations reported are obtained by averaging over all 100 calculated ensembles. The agreement with the working set of RDCs that are included in the calculations is the same as that for the calculations reported in Table 1 where refinement is carried out against all RDCs. [b] The mean values and standard deviations reported are obtained by averaging over all 100 calculated ensembles. [c] The number of measured RDCs is given in parentheses.

Table 4: Error Analysis of GB3 Dipolar Coupling Refinement[a]

dipolar couplings[b]	rms between target and calculated RDCs (Hz): noise × 1, $N_e = 1$	noise × 1, $N_e = 2$	noise × 2, $N_e = 1$	noise × 2, $N_e = 2$
(A) Random Noise Added to Theoretical RDCs for a Single Structure ($N_e = 1$)				
N−H (246)	0.21 ± 0.03	0.20 ± 0.04	0.40 ± 0.06	0.35 ± 0.07
Cα−Hα (269)	0.51 ± 0.11	0.50 ± 0.10	1.00 ± 0.23	0.89 ± 0.20
N−C′ (262)	0.09 ± 0.01	0.09 ± 0.01	0.17 ± 0.01	0.17 ± 0.02
Cα−C′ (254)	0.09 ± 0.01	0.09 ± 0.01	0.17 ± 0.03	0.17 ± 0.02
(B) Random Noise Added to Theoretical RDCs for a Two-Structure Ensemble ($N_e = 2$)				
N−H (246)	0.41 ± 0.05	0.21 ± 0.05	0.53 ± 0.09	0.35 ± 0.10
Cα−Hα (269)	0.77 ± 0.21	0.43 ± 0.06	1.05 ± 0.20	0.74 ± 0.13
N−C′ (262)	0.11 ± 0.02	0.10 ± 0.02	0.20 ± 0.03	0.18 ± 0.03
Cα−C′ (254)	0.12 ± 0.03	0.10 ± 0.02	0.19 ± 0.04	0.17 ± 0.03

[a] The experimental errors in the measurements (i.e., noise × 1) are 0.26, 0.58, 0.1, and 0.1 Hz for the N−H, Cα−Hα, Cα−C′, and N−C′ RDCs, respectively. Theoretical RDCs were calculated from a single refined $N_e = 1$ structure and from a single refined two-structure $N_e = 2$ ensemble using the alignment tensors obtained from refinement against the experimental RDC data. Random Gaussian noise (×1 or ×2) was then added to the theoretical RDCs, which were subsequently employed for refinement using $N_e = 1$ or 2 ensembles. The mean values and standard deviations reported are obtained by averaging over all 100 calculated ensembles. [b] The number of measured dipolar couplings is given in parentheses.

structures (*47*, *48*). The values of ω observed in the structures from the $N_e = 1$ and 2 calculations are highly correlated to those obtained using different procedures and a single structure representation by Ulmer et al. (*32*) with a slope of ~1 and correlation coefficients of 0.85 and 0.86, respectively (see Supporting Information). The correlation coefficients for the average peptide bond angles observed in the $N_e = 1$ and 2 ensembles is 0.97. The average rms differences between ω peptide bond torsion angles for the $N_e = 1$ versus Ulmer et al. structures, the $N_e = 2$ versus Ulmer et al. structures, and the $N_e = 1$ versus $N_e = 2$ structures are 1.84°, 1.80°, and 0.82°, respectively.

Error Analysis. The question that arises is whether the improvement in the *R*-factors for the N−H and Cα−Hα RDCs upon increasing the ensemble size from $N_e = 1$ to 2 reflects the presence of anisotropic motions or is simply due to fitting noise in the experimental data.

One simple method of assessment involves cross-validation (*31*, *32*, *46*). To this end we carried out a series of calculations, leaving out all the RDC data for each bond vector type in turn (Table 3). In each case, the agreement with the RDCs included in the refinement remained unchanged relative to the calculations in which all RDCs are included in the refinement (cf. Table 1). The free *R*-factors (i.e., involving RDCs excluded from refinement) range from about 8% to 12%. The free *R*-factors for the N−C′ and Cα−C′ RDCs are significantly less than those for the X-ray coordinates; in contrast, those for the NH and Cα−Hα RDCs are only slightly less than those for the X-ray coordinates. From the results shown in Table 3, one can ascertain that there is no significant decrease in the free RDC *R*-factor upon increasing the ensemble size from $N_e = 1$ to 2. In regard to the N−C′ and Cα−C′ RDCs this is hardly surprising, since the working *R*-factors (for the RDCs included in the refinement) obtained with inclusion of these RDCs do not decrease significantly upon increasing the ensemble size (Table 1). In the case of the N−H and Cα−Hα RDCs, the absence of a decrease in free *R*-factor may at first appear surprising. However, when the structures are refined without inclusion of the N−H or Cα−Hα RDCs, the H_N or Hα protons will be placed in a position corresponding to idealized covalent geometry, and hence any distortions in peptide plane geometry in the case of the H_N vector or tetrahedral geometry about the Cα atom will not be taken into account (*32*). Thus the attainable values of the free *R*-factors for the N−H and Cα−Hα vectors are limited by the uncertainties in the placement of the respective protons. These uncertainties place a lower limit of approximately 10% on the attainable free *R*-factor. Since this lower limit is already reached in the $N_e = 1$ calculations, increasing the ensemble size to $N_e = 2$ does not result in any significant change in the free *R*-factors. The cross-validation results for the N−H and Cα−Hα RDCs also suggest that the motions of N−H and Cα−Hα bond vectors for a given residue are to a large extent independent.

We therefore resorted to an alternative, more rigorous method of error analysis. Two sets of calculations were carried out, and the results are summarized in Table 4. Theoretical RDCs were calculated from a refined $N_e = 1$ structure and from a refined two-structure $N_e = 2$ ensemble using the alignment tensors obtained from refinement against the experimental RDCs. Random Gaussian noise at one and two times the level of the experimental error was then added to the theoretical RDCs, which were subsequently used for refinement with $N_e = 1$ and 2 ensembles. (Note that several sets of calculations were carried out using theoretical RDCs computed from different structures and ensembles, and in each case the same results were obtained.)

When random noise is added to the theoretical RDCs calculated for a single structure, the results obtained with $N_e = 1$ and 2 refinement are essentially identical; that is, no improvement in RDC *R*-factors is observed upon increasing the ensemble size from $N_e = 1$ to 2, and in both cases the rms differences between observed and calculated values are equal to the noise level (Table 4A). Thus, one can conclude that the decrease in RDC *R*-factor observed upon refinement

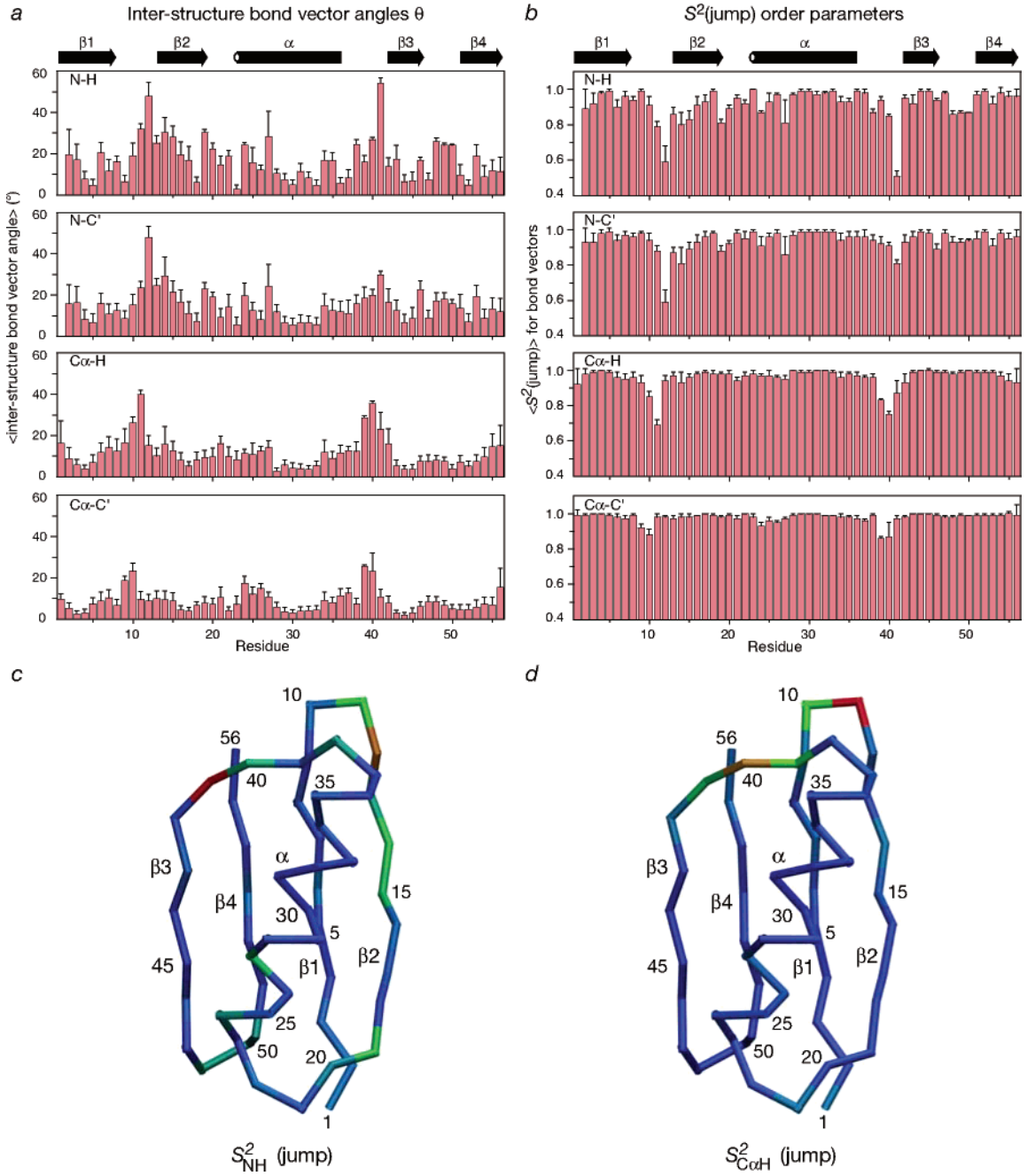

FIGURE 2: Amplitudes of backbone motions for GB3 deduced from the ensemble size $N_e = 2$ structures. The average interstructure N−H, N−C′, Cα−Hα, and Cα−C′ bond vector angles, $\langle\theta\rangle$, between members of an ensemble and the corresponding calculated average order parameters, $\langle S^2(\text{jump})\rangle$ (where the angle brackets denote averaging over all 100 calculated ensembles), are displayed as a function of residue in panels a and b, respectively. The error bars are the standard deviations from the 100 calculated ensembles. The secondary structure elements are shown above the figure: strands β1, β2, β3, and β4 comprise residues 1−8, 13−20, 42−46, and 51−56, respectively; the single α-helix extends from residue 23 to residue 36. Panels c and d display a Cα backbone trace of the GB3 domain color coded according to $\langle S^2(\text{jump})\rangle$ for the N−H and Cα−Hα bond vectors, respectively. The values of $\langle S^2(\text{jump})\rangle$ vary linearly through the color spectrum from red to blue, corresponding to the minimum and maximum values of $\langle S^2(\text{jump})\rangle$, respectively.

against all the experimental RDCs when the ensemble size is increased from $N_e = 1$ to 2 is not a result of fitting noise.

When random noise is added to the theoretical RDCs derived from a two-structure $N_e = 2$ ensemble, however, the rms differences between observed and calculated N−H and Cα−Hα RDCs are reduced by 45−50% with the noise set to one times the experimental error and by 30−35% with the noise set to two times the experimental error; in addition, the rms differences obtained for the $N_e = 2$ calculations are equal to the noise level (Table 4B). In contrast, the rms difference between observed and calculated N−C′ and Cα−C′ RDCs remains unchanged upon increasing the ensemble size from $N_e = 1$ and 2 and, in both cases, is equal to the noise level (Table 4B). This is simply due to the fact that the experimental error for the N−C′ and Cα−C′ RDCs is too large (relative to their respective D_a's) to detect any improvement upon increasing the ensemble size beyond $N_e = 1$. These results exactly reproduce the results observed with the experimental RDCs. One can therefore conclude that the substantial (ca. 40%) improvement in agreement observed for the experimental N−H and Cα−Hα RDCs upon increasing the ensemble size from $N_e = 1$ to 2 is real

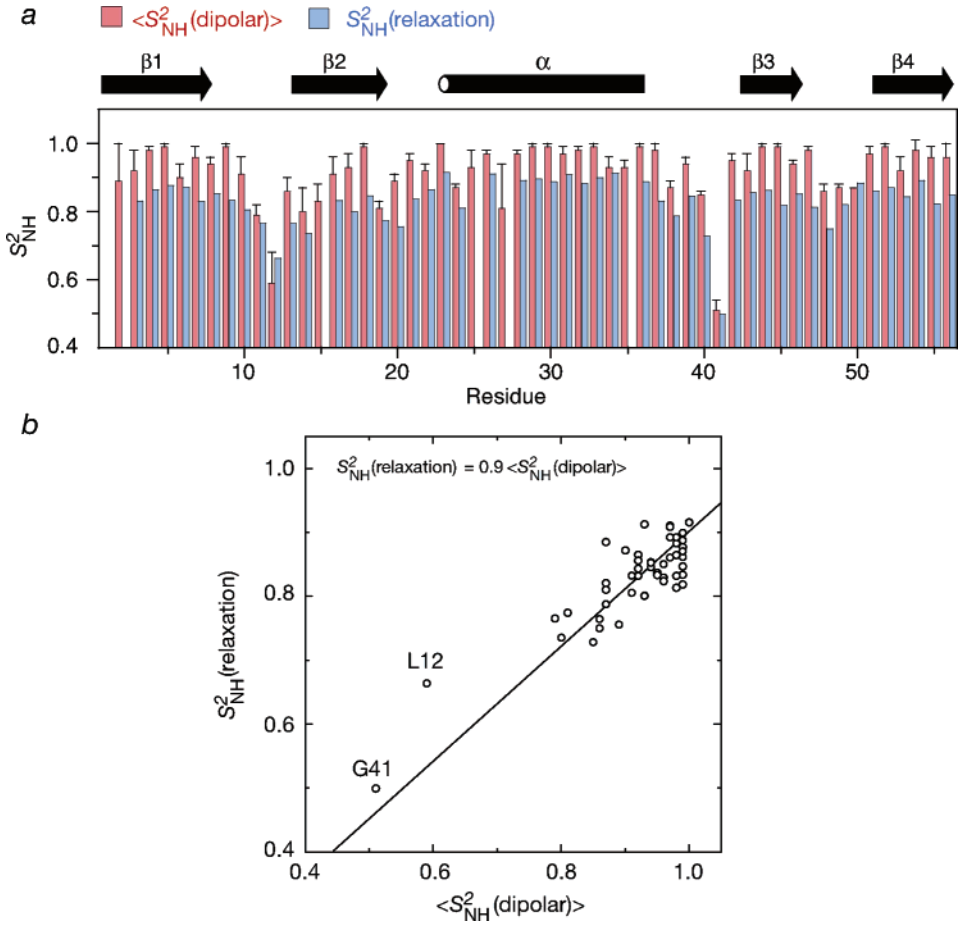

FIGURE 3: Comparison of the N−H order parameters derived from dipolar couplings and ^{15}N heteronuclear relaxation studies on GB3. (a) Plot of order parameters as a function of residue with $\langle S^2_{NH}(\text{dipolar})\rangle$ shown in red and $S^2_{NH}(\text{relaxation})$ in blue. $S^2_{NH}(\text{dipolar})$ corresponds to $\langle S^2_{NH}(\text{jump})\rangle$ for the N−H vectors shown in Figure 2, derived from the 100 calculated $N_e = 2$ ensembles. (b) Correlation plot of $\langle S^2_{NH}(\text{dipolar})\rangle$ versus $S^2_{NH}(\text{relaxation})$; the correlation coefficient is 0.82. The angle brackets denote averaging over all 100 calculated $N_e = 2$ ensembles. The values of $S^2_{NH}(\text{relaxation})$ are taken from ref *33* using the results obtained with the axially symmetric model of overall motion. Note that a full anisotropic diffusion model of overall tumbling motion only results in a small, albeit statistically significant, improvement, in agreement with the experimental ^{15}N relaxation data (*33*), and by its very nature must result in some reduction in the precision of the derived $S^2_{NH}(\text{relaxation})$ values. Moreover, the values of $S^2_{NH}(\text{relaxation})$ for the fully anisotropic model are essentially identical to those obtained for the axially symmetric diffusion model, and the rms difference between the two sets of values is only 0.019. The values, however, of $S^2_{NH}(\text{relaxation})$ for the two residues, Leu12 and Gly41, with the lowest order parameters are slightly reduced for the fully anisotropic model (0.61 and 0.47, respectively) relative to those for the axially symmetric model (0.66 and 0.50, respectively), resulting in even closer agreement with the corresponding values of $\langle S^2_{NH}(\text{dipolar})\rangle$.

and reflects unambiguously the presence of anisotropic motions along the backbone.

Amplitudes of Anisotropic Backbone Motions. The average interstructure bond vector angles, $\langle\theta\rangle$, and corresponding average $\langle S^2(\text{jump})\rangle$ order parameters [given by $\langle(3\cos^2\theta + 1)/4\rangle$] (*49*) derived from the $N_e = 2$ calculations are plotted as a function of residue in panels a and b of Figure 2, respectively, and the $\langle S^2(\text{jump})\rangle$ values for the N−H and Cα−Hα bond vectors, color coded on a Cα backbone trace of GB3, are displayed in Figure 2c. We note that the values of $\langle S^2(\text{jump})\rangle$ obtained from the $N_e = 2$ calculations are essentially unaffected as the ensemble size is increased further: thus, for example, the correlation coefficients for $\langle S^2_{NH}(\text{jump})\rangle$ between the $N_e = 2$ and 3 calculations, the $N_e = 2$ and 8 calculations, and the $N_e = 3$ and 8 calculations are 0.99, 0.96, and 0.99, respectively, with slopes of 0.98, 0.96, and 0.98, respectively (see Supporting Information).

The two regions exhibiting the largest amplitude motions are located in the $\beta1-\beta2$ hairpin turn (residues 9−12) and in the loop connecting the single α-helix to strand $\beta3$ (residues 37−41). Larger than average motions are also seen in the short turn connecting strand $\beta2$ to the α-helix (residues 21−22) and in the $\beta3-\beta4$ turn (residues 47−50). Within the regions of regular secondary structure, evidence for larger than average motions is seen for the first three residues (residues 13−15) of strand $\beta2$ and in the first turn (residues 23−27) of the α-helix. Within regions of secondary structure, $S^2_{NH}(\text{jump}) < S^2_{NC'}(\text{jump}) < S^2_{C\alpha H\alpha}(\text{jump}) < S^2_{NC'}(\text{jump})$ with corresponding average values for the interstructure bond vector angles of ∼14°, ∼13°, ∼9°, and ∼7°, respectively.

A comparison of the $\langle S^2_{NH}(\text{jump})\rangle$ values with the $S^2_{NH}(\text{relaxation})$ values derived from ^{15}N relaxation measurements (*33*) is shown in Figure 3. There is a close correspondence of the two order parameters. The trends as a function of residue are almost identical (Figure 3a), and the two order parameters are highly correlated (Pearson's correlation coefficient $r = 0.82$) (Figure 3c). Interestingly,

Table 5: Average Interstructure Bond Vector Angles, $\langle\theta\rangle$, and $\langle S^2(\text{jump})\rangle$ Order Parameters for GB3 Derived from the $N_e = 2$ Calculations

	$\langle\theta\rangle$ (deg)[a]	$\langle S^2(\text{jump})\rangle$[a]	$\boldsymbol{S}^2$(relaxation)[b]
	All Residues		
N−H	16.8 ± 10.3	0.92 ± 0.09	0.83 ± 0.07
Cα−Hα	11.3 ± 7.4	0.96 ± 0.06	
N−C′	14.7 ± 7.6	0.94 ± 0.06	
Cα−C′	8.6 ± 5.2	0.98 ± 0.03	
	Residues in Secondary Structure Only[c]		
N−H	14.4 ± 7.6	0.94 ± 0.06	0.85 ± 0.05
Cα−Hα	9.1 ± 3.9	0.97 ± 0.02	
N−C′	13.4 ± 6.1	0.95 ± 0.04	
Cα−C′	7.2 ± 3.6	0.98 ± 0.01	

[a] The mean values are obtained by averaging over all 100 calculated ensembles. [b] The average values of the generalized order parameter $\boldsymbol{S}^2_{NH}$(relaxation) for the N−H bond vectors derived from ^{15}N relaxation measurements are taken from the data in ref *33*. [c] Strands β1 (residues 1−8), β2 (residues 13−20), β3 (residues 42−46), and β4 (residues 51−56) and the single helix (residues 23−36).

$\boldsymbol{S}^2_{NH}$(relaxation) is on average ~0.9 times smaller than $\langle S^2_{NH}(\text{jump})\rangle$ derived from RDCs (Figure 3c and Table 5). This is readily rationalized since $\boldsymbol{S}^2_{NH}$(relaxation) can be considered to be the product of axially symmetric and anisotropic order parameters (*49*, *50*): $S^2_{NH}(\text{axial})\cdot S^2_{NH}(\text{anisotropic})$. S^2_{NH}(anisotropic) can be equated to $\langle S^2_{NH}(\text{jump})\rangle$ obtained from the RDC data. This suggests that S^2_{NH}(axial) is fairly uniform over the whole length of the polypeptide chain and, to a first approximation, originates from bond librations. S^2_{NH}(anisotropic), on the other hand, arises as a consequence of fluctuations in the ϕ and ψ backbone torsion angles.

Correlation plots for the interstructure bond vector angles, $\langle\theta\rangle$, are shown in Figure 4. Good correlations between $\langle\theta_{NH}(i)\rangle$ and $\langle\theta_{NC'}(i)\rangle$ (Figure 4a) and between $\langle\theta_{CH}(i)\rangle$ and $\langle\theta_{C\alpha C'}(i)\rangle$ (Figure 4b) are observed with correlation coefficients of 0.86 and 0.85 (excluding Thr11), respectively. The anomaly for Thr11 in the $\langle\theta_{CH}(i)\rangle$ versus $\langle\theta_{C\alpha C'}(i)\rangle$ correlation may possibly be attributed to more complex motions around this residue (see discussion below on correlated motions). No correlation, however, is observed between $\langle\theta_{NH}(i)\rangle$ and $\langle\theta_{C\alpha H}(i)\rangle$ (Figure 4c). However, a weak correlation ($r \approx 0.6$) is observed between $\langle\theta_{NH}(i)\rangle$ and $\langle\theta_{C\alpha H}(i-1)\rangle$.

Correlated Motions. From the perspective of overall backbone atomic positions, the amount of global motion is small (Figure 5a). The average rms difference between the structures within an ensemble is 0.48 ± 0.09 Å. The structures, however, reveal clear evidence of correlated backbone motions. In this regard it should be emphasized that the correlated nature of the motions derived from analysis of the calculated $N_e = 2$ ensembles arises from the RDCs acting in concert with the restraints imposed upon the system by covalent geometry and stereochemistry. In the

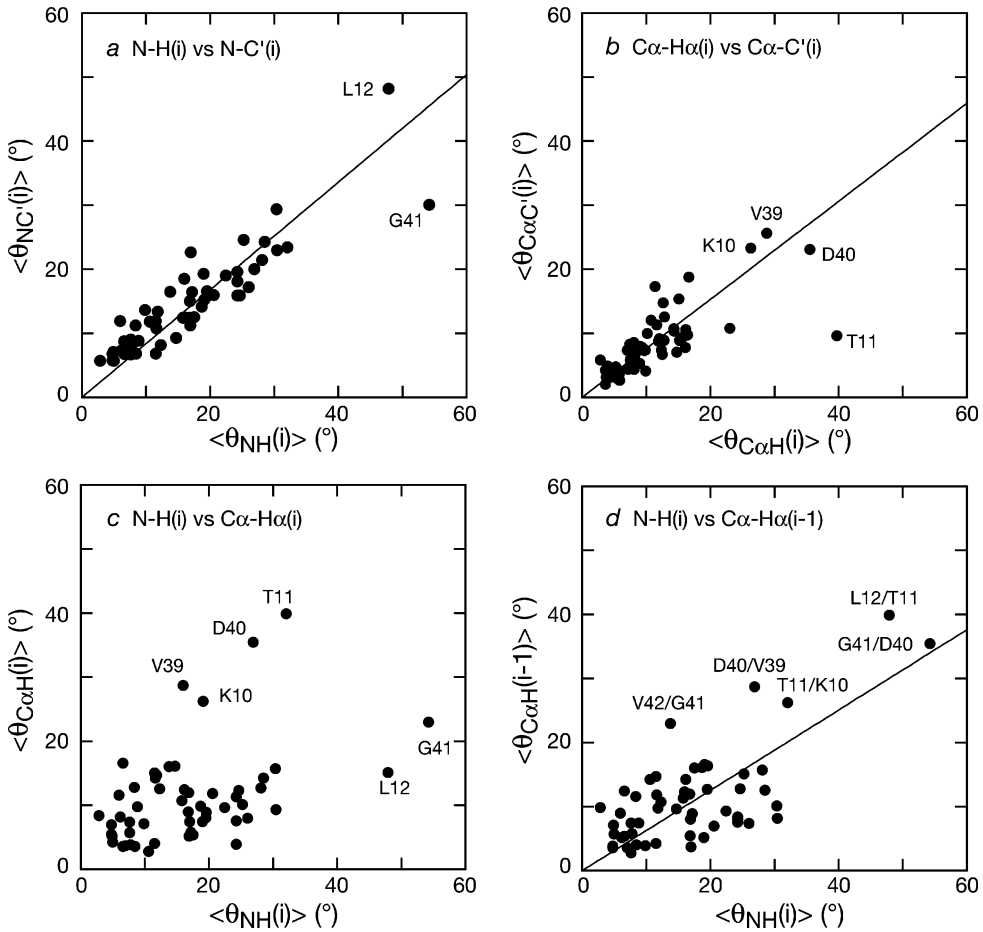

FIGURE 4: Correlation plots between various interstructure bond vector angles, $\langle\theta\rangle$, for GB3 derived from the 100 calculated $N_e = 2$ ensembles: (a) N−H(i) versus N−C′(i), (b) Cα−Hα(i) versus Cα−C′(i), (c) N−H(i) versus Cα−Hα(i), and (d) N−H(i) versus Cα−Hα(i−1). The angle brackets denote averaging over all 100 calculated $N_e = 2$ ensembles. The correlation coefficients for (a), (b), and (d) are 0.86, 0.85 (omitting the single outlier, T11), and 0.55, respectively.

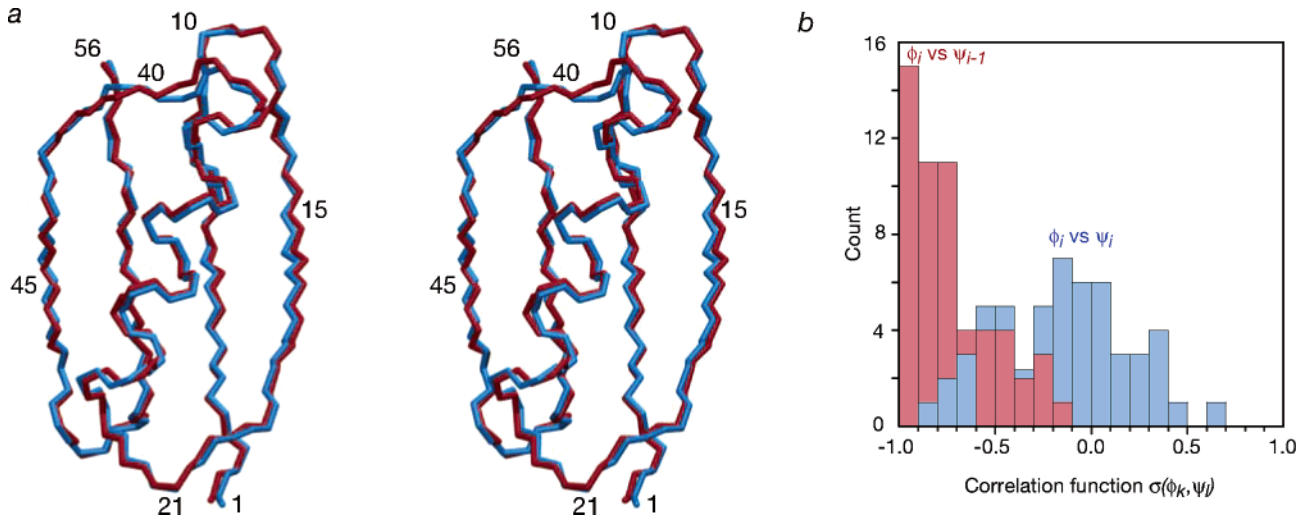

FIGURE 5: Global structural implications of backbone motions for GB3 derived from dipolar coupling refinement using an ensemble size of $N_e = 2$. (a) Stereoview showing the complete backbone (N, Cα, C atoms) of two members (displayed in red and blue) of a typical $N_e = 2$ ensemble. (b) Histogram of ϕ_i/ψ_{i-1} (red) and ϕ_i/ψ_i (blue) torsion angle correlations derived from the 100 calculated $N_e = 2$ ensembles.

absence of the latter, the RDCs would only provide information on the magnitudes of the motions.

The presence of extensive crankshaft motions along the polypeptide backbone involving anticorrelated fluctuations between the backbone torsion angles ϕ_i and ψ_{i-1} is observed (Figures 5b and 6d). The distribution is highly skewed, with mean and mode values for the angular correlation coefficient, $\sigma(\phi_i,\psi_{i-1})$, of 0.73 and 0.93, respectively (Figure 5b). This finding is consistent with the correlation of S^2(jump) for the N−H bond vector of residue i and the Cα−Hα of residue $i - 1$ (Figure 4d). In contrast, the distribution of $\sigma(\phi_i,\psi_i)$ is approximately normal with mean and standard deviations of −0.17 and 0.32, respectively. Thus, the motions of ϕ and ψ torsion angles within a residue are only rarely correlated.

The existence of crankshaft motions in helices was first proposed on theoretical grounds from normal-mode analysis (*51*) and was subsequently postulated to account for the observation that the generalized order parameters for the Cα−Hα bond vectors derived from heteronuclear relaxation measurements tend to be slightly larger than those for the N−H bond vectors (*52*, *53*). This is mirrored in the present study where the calculated values of $\langle S^2(\text{jump})\rangle$ for the Cα−Hα bond vectors are also slightly larger than those for the N−H bond vectors (Figure 2 and Table 5). In addition, crankshaft motions with anticorrelated ϕ_i and ψ_{i-1} torsion angles have been observed in molecular dynamics simulations (*44*, *45*, *54*−*56*), but the true physical extent of such correlated motions was difficult to ascertain. The present results (Figures 5b and 6d) clearly establish the extensive and pervasive nature of anticorrelated crankshaft motions involving ϕ_i and ψ_{i-1} along the full length of the polypeptide backbone.

The NH−NH bond vector correlations are highly localized, and no significant long-range correlations between distal elements in the sequence are observed (Figure 5a). The most prominent short-range correlations are seen in four regions: the $\beta1-\beta2$ hairpin turn, the beginning of the $\beta2$ strand, the first turn of the helix, and residues at the end of the loop connecting the helix with strand $\beta4$ and the N-terminal end of strand $\beta4$. Interestingly, in the latter case, the N−H bond vector of Gly41, which exhibits the lowest $\langle S^2_{NH}(\text{jump})\rangle$ order parameter, is not correlated to the motions of the adjacent residues.

The Cα−Hα bond vectors exhibit short-range CαH−CαH correlations similar to the N−H bond vectors, but they tend to be more extensive and more pronounced; in addition, short-range correlations are observed at the C-terminus of strand $\beta1$ and from the last turn of the helix to the first few residues of the subsequent loop (Figure 6b). Moreover, a number of long-range correlations between distal secondary structure elements that are in close spatial proximity are also observed: between the antiparallel strands $\beta1$ and $\beta2$, between the parallel strands $\beta1$ and $\beta4$, and between the loop preceding strand $\beta3$ and the C-terminus of strand $\beta4$.

The NH−CαH bond vector correlations are also predominantly local although there are some weaker long-range correlations that mirror the CαH−CαH correlations. The local NH−CαH correlations are asymmetric, and in those regions where local correlations are observed, the correlations for the N−H bond vector of a given residue to the Cα−Hα bond vectors of the preceding residues are generally larger than those to the Cα−Hα bond vectors of the following residues.

Detailed views of the two regions exhibiting the largest amplitude backbone motions (residues 9−12 and residues 39−42) within a typical $N_e = 2$ ensemble are provided in Figure 7. Relatively large interstructure bond vector angles (40−50°) are readily absorbed by appropriate changes in the ϕ and ψ backbone torsion angles which remain in the most favored region of the Ramachandran map. In the case of residues 9−12, $\phi_i-\psi_{i-1}$ anticorrelated fluctuations characteristic of crankshaft motions are not prominent, except for residues Leu12/Thr11. Instead, a number of other correlated compensatory fluctuations of the backbone torsion angles are observed. For example, ϕ and ψ of Lys10 are highly anticorrelated, ϕ of Leu12 is positively correlated with ϕ of residues 9−11, and ψ of Lys10 and Thr11 are positively correlated. In the case of residues 39−42, $\phi_i-\psi_{i-1}$ anticorrelations are significant and accompanied by a number of positive correlations involving the ϕ angles of Asn37 and Gly38, the ϕ angles of Val39 and Asp40, and the ψ angles of Val39 and Gly41.

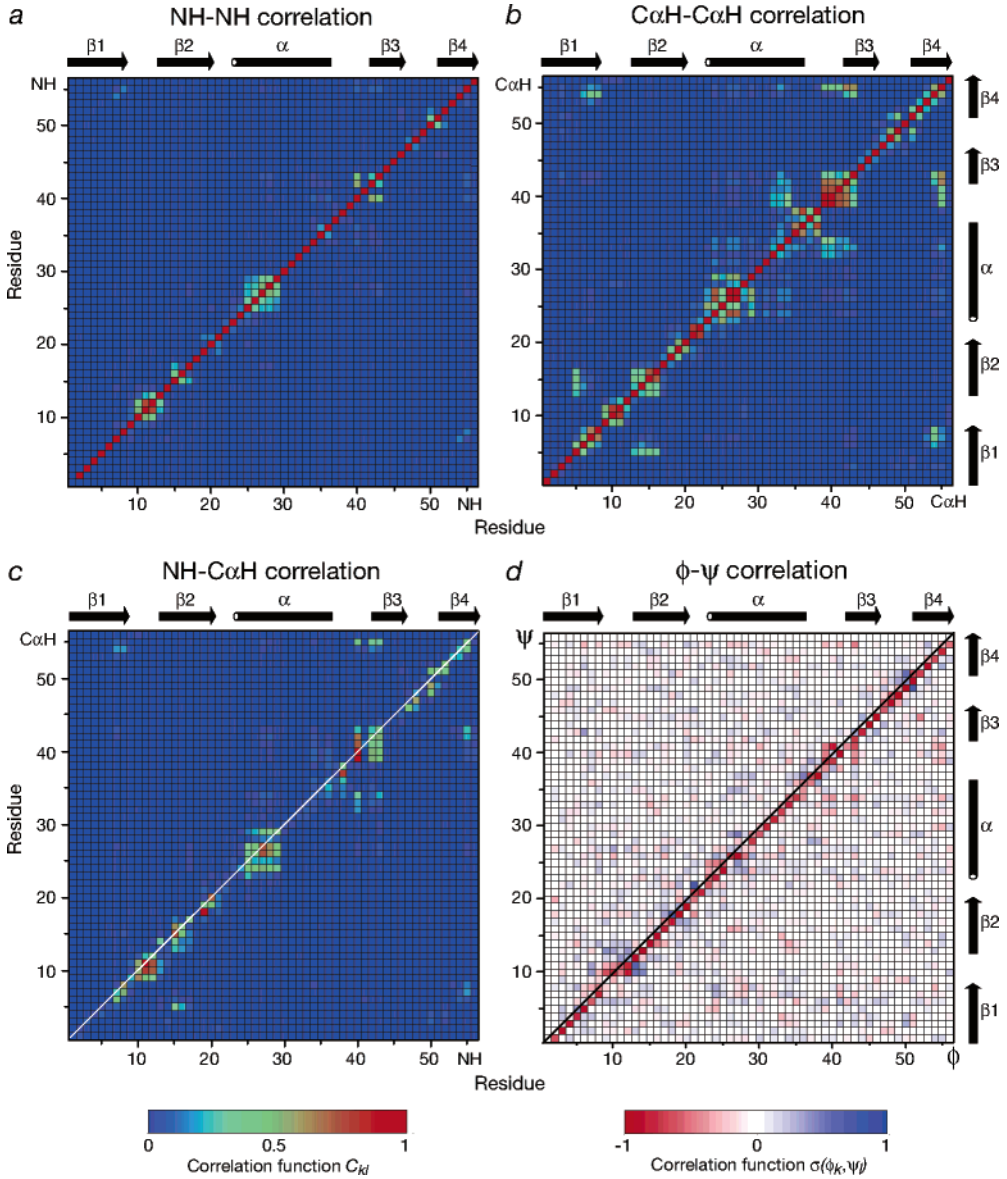

FIGURE 6: Correlated motions of N−H and Cα−Hα bond vectors and ϕ/ψ backbone torsion angles for GB3 derived from the 100 calculated $N_e = 2$ ensembles: (a) NH−NH, (b) CαH−CαH, and (c) NH−CαH bond vector correlations; (d) $\phi-\psi$ torsion angle correlations. The correlation function for the bond vectors varies between 0 (uncorrelated; blue) and 1 (fully correlated; red) with light green representing a value of 0.5. The torsion angle correlation function for ϕ/ψ, $\sigma(\phi_k,\psi_l)$, varies between −1 (fully anticorrelated, red) and +1 (fully correlated, blue) with a value of zero (white) indicating the absence of any correlation.

CONCLUDING REMARKS

Analysis of the extensive and highly accurate backbone N−H, Cα−Hα, N−C′, and Cα−C′ RDC data set available for GB3 in five alignment media (*32*) by simultaneous refinement of the coordinates and alignment tensors reveals the presence of anisotropic backbone motion. Specifically, large improvements (ca. 40%) in agreement between observed and calculated N−H and Cα−Hα RDCs are observed when the ensemble size used in refinement is increased from $N_e = 1$ to 2, such that the N−H and Cα−Hα RDCs (Table 1) are fit to the level of the experimental error (0.26 and 0.58 Hz, respectively). With a few exceptions, the amplitudes of the motions are small. This finding is entirely consistent with a recent study in which a small but statistically significant improvement in the agreement between observed N−H RDCs and those calculated from crystal structures was observed relative to a static representation by including a Gaussian axial function (GAF) model to represent small amplitude anisotropic peptide plane motion (*59*). Regions with significant mobility are identified in loops and turns (Figure 2) and display evidence of significant localized correlated motion extending over a range of four to six residues (Figure 6). In addition, crankshaft motions, characterized by anticorrelated fluctuations of ϕ_i and ψ_{i-1}, are ubiquitous throughout the polypeptide backbone (Figures 5b and 6d). Some evidence for weak long-range correlated motions between the strands of both antiparallel and parallel β-sheets is also observed (Figures 6b).

From the perspective of protein folding/unfolding, it is tempting to speculate that areas exhibiting significant local correlated motions may constitute hot spots that unfold cooperatively. Two such hot spots have been identified in the highly homologous GB1 domain (*60*, *61*): the first constitutes the $\beta1/\beta2$ hairpin and the second the first turn of the α-helix. Both of these regions exhibit NH−NH, CαH−CαH, and NH−CαH bond vector correlations extending over

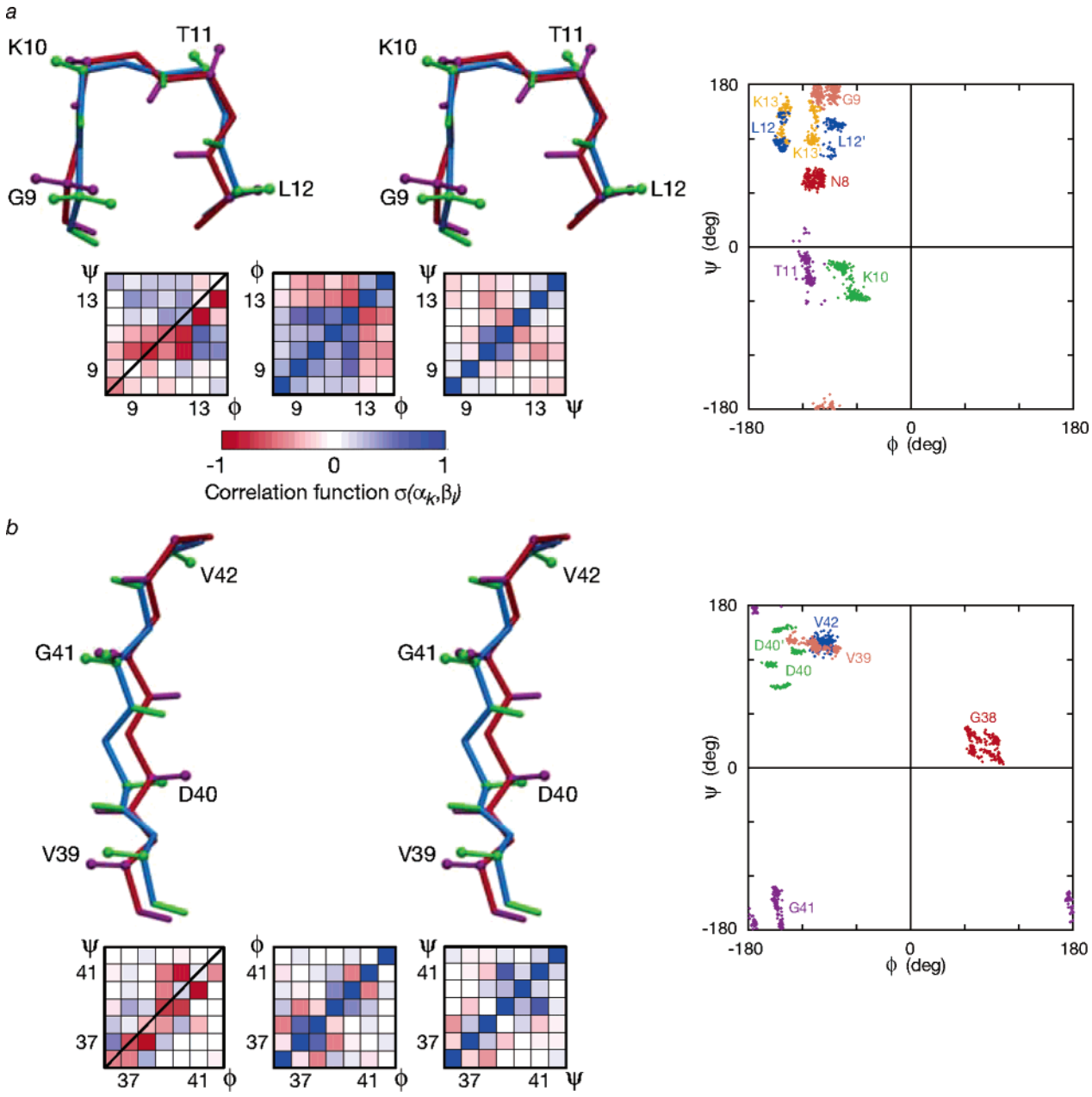

FIGURE 7: Local structural implications of backbone motions for GB3 derived from dipolar coupling refinement using an ensemble size of $N_e = 2$ for the two regions displaying the smallest S^2(jump) order parameters. (a) Residues 9−12 and (b) residues 39−42. The left-hand side of the figure provides (i) stereoviews with the backbone (N, Cα, C) atoms, N−H bonds, and Cα−Hα bonds shown in blue, green, and green, respectively, for one member of a typical $N_e = 2$ ensemble and in red, magenta, and magenta, respectively, for the other (the Hα atoms are represented by small spheres) and (ii) $\phi-\psi$, $\phi-\phi$, and $\psi-\psi$ torsion angle correlations derived from the 100 calculated $N_e = 2$ ensembles with the color coding for the torsion angle correlation function, $\sigma(\alpha_k,\beta_l)$ shown at the bottom of panel a, varying from red for $\sigma(\alpha_k,\beta_l) = -1$ (fully anticorrelated), through white for $\sigma(\alpha_k,\beta_l) = 0$ (uncorrelated), to blue for $\sigma(\alpha_k,\beta_l) = 1$ (fully correlated). Ramachandran ϕ/ψ plots for the 100 calculated $N_e = 2$ ensembles are displayed on the right-hand side of the figure.

three to five residues (Figure 6). In the case of the first turn of the single helix, examination of the $N_e = 2$ ensembles suggests that the backbone amide of Glu27 may be involved in oscillating bifurcated hydrogen bonds with the backbone carbonyls of the $i - 4$ (Ala23) and $i - 3$ (Glu24) residues. In contrast to the $\beta1/\beta2$ hairpin, which is disrupted in the folding transition state, the symmetrically disposed $\beta3/\beta4$ hairpin is stabilized (*60*, *61*). Interestingly, only minimal correlated motions are observed within the $\beta3/\beta4$ hairpin (Figure 6).

The present analysis indicates that overall the $\langle S^2_{NH}$(jump)$\rangle$ order parameters derived from the RDC data are qualitatively and quantitatively consistent with and highly correlated to the generalized $\boldsymbol{S}^2_{NH}$(relaxation) order parameters derived from relaxation measurements. The average value of $\boldsymbol{S}^2_{NH}$(relaxation) is ∼0.9 times less than that of $\langle S^2_{NH}$(jump)$\rangle$. Since $\boldsymbol{S}^2_{NH}$(relaxation) can be expressed as S^2_{NH}(axial)·S^2_{NH}(anisotropic) (*49*, *50*) where the anisotropic order parameter can be equated to $\langle S^2_{NH}$(jump)$\rangle$, it follows that S^2_{NH}(axial) is fairly uniform with an average value of ∼0.9. (This value is increased slightly as the ensemble size is increased: for ensemble sizes of $N_e = 3$ and 8, the value is 0.92 and 0.93, respectively.)

S^2_{NH}(axial) can be considered to arise from bond librations which would be expected to be essentially uniform along the polypeptide chain. The value of ~0.9 for S^2_{NH}(axial) that we derive from the current analysis is entirely consistent with previous experimental (*62*) and theoretical (*57*, *58*) work. Analysis of long time molecular dynamics trajectories on various proteins yields an average value of S^2_{NH}(libration) of ~0.9 with minimal variation (~1%) along the polypeptide chain (*57*, *58*). From a physical perspective, a reduction of S^2_{NH}(axial) from a value of 1.0 is equivalent to an increase in the effective bond length, r^{eff}_{NH}, relative to the equilibrium internuclear separation, R^{eq}_{NH}. This is because the effective bond length depends not only on harmonic and anharmonic vibrations but also on bond librations (bond bending) such that the trajectory of the NH vector arising from such motions essentially constitutes diffusion within a cone (*62*). In this representation S^2_{NH}(axial) = $(R^{eq}_{NH}/r^{eff}_{NH})^6$ (*63*, *64*). The value of R^{eq}_{NH} obtained from high-resolution neutron diffraction is 1.02 Å (*65*), yielding a value of ~1.04 Å for r^{eff}_{NH} if S^2_{NH}(axial) = 0.9. This value coincides with the vibrationally corrected, effective NH bond length of 1.041±0.006 Å derived from N−H and N−C′ RDC measurements in two alignment media (*62*) assuming a reference bond length of 1.329 Å for the N−C′ distance derived from statistical analysis of atomic resolution X-ray structures of small molecules (*66*). The scaling of the magnitude of the axial component of the alignment tensor for the Cα−H, N−C′, and Cα−C′ RDCs relative to the N−H RDCs makes use of these vibrationally corrected, effective bond distances derived from analysis of RDC data (*62*). The analysis of the ^{15}N relaxation data, on the other hand, makes use of the equilibrium N−H bond length of 1.02 Å. Thus, if the ^{15}N relaxation data employed a value of 1.04 Å for the N−H bond length, the slope of the correlation between $\boldsymbol{S}^2_{NH}$(relaxation) and $\langle S^2_{NH}(\text{jump})\rangle$ would be 1.

The $\boldsymbol{S}^2_{NH}$(relaxation) order parameters derived from ^{15}N relaxation data relate only to motions less than the rotational correlation time (which in the case of GB3 is ~3.3 ns). RDCs, on the other hand, potentially contain motional information on time scales up to ~1 ms. The excellent overall correlation between $\boldsymbol{S}^2_{NH}$(relaxation) and $\langle S^2_{NH}(\text{jump})\rangle$ therefore confirms the generally held view that large-scale collective motions away from the mean coordinate positions on time scales longer than the rotational correlation time are only present transiently (*1*−*16*, *67*) and, hence, do not significantly perturb the observed RDCs (*31*, *32*). In this regard, however, it is interesting to note that the $\langle S^2_{NH}(\text{jump})\rangle$ values for Leu12 (0.59 ± 0.09) and Gly41 (0.51±0.03) are slightly smaller and approximately equal, respectively, to the corresponding values for $\boldsymbol{S}^2_{NH}$(relaxation) (0.66 and 0.50, respectively). In addition, the value of $\langle S^2_{NH}(\text{jump})\rangle$ for Leu12 is further decreased when the ensemble size is increased beyond 2 (to 0.52 ± 0.05 for N_e = 3 and 0.47 ± 0.03 for N_e = 8), whereas that for Gly41 is barely affected (0.48 ± 0.00 for N_e = 3 and 0.47 ± 0.00 for N_e = 8). This suggests that a portion of the anisotropic motions involving Leu12 may occur on a time scale longer than the rotational correlation time.

The present study clearly shows that highly accurate RDC data for multiple backbone vectors in multiple alignment media can provide information relating to the magnitude of local backbone motions and the presence of correlated motions at a level of detail that was heretofore not possible. However, it must be emphasized that the effects of local motions on the observed RDCs are subtle and hence highly accurate data are required to be able to detect such relatively small scale motions using RDC data. The current analysis assumes that the scaling of the observed RDCs by S_{NH}(axial) is effectively uniform, and at the current level of accuracy and precision of the present data, it is not feasible to obtain residue-specific values of S^2_{NH}(axial). In addition, it is absolutely essential that data in multiple alignment media with significantly different alignment tensors are available. The dual requirement of highly accurate data and multiple alignment media presents a considerable technical challenge for the widespread application of this methodology. Finally, this study confirms that, from the perspective of the routine use of RDC data for structure refinement, the assumption of a single structure representation will in general be sufficient (*31*). This assumption is unlikely to have any significant untoward effect on coordinate accuracy, but it will result in significant increases in coordinate precision (i.e., the bundle of calculated structures will appear to be anomalously precise) (*47*). This is always important to bear in mind since the ratio of accuracy to precision for structures refined with RDC data using a unique structure representation (i.e., N_e = 1) will, in general, be considerably higher than for the corresponding structures refined without RDC data (*47*, *68*). In other words, providing there are no errors in the NOE restraints, the relative increase in coordinate precision achieved by incorporating RDC restraints will be significantly larger than the corresponding improvement in coordinate accuracy.

ACKNOWLEDGMENT

We thank Attila Szabo and Ad Bax for extensive discussions. We also thank Ad Bax and Tobias Ulmer for making their RDC data on GB3 available to us and David Fushman for providing us with the order parameters derived from ^{15}N relaxation measurements. This study utilized the high-performance computational capabilities of the Biowulf/LoBos3 cluster at the National Institutes of Health.

SUPPORTING INFORMATION AVAILABLE

Data relating to $\langle S^2_{NH}(\text{jump})\rangle$ order parameters obtained for ensemble sizes of N_e = 3 and 8. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wagner, G. (1983) Characterization of the distribution of internal motions in the basic pancreatic trypsin inhibitor using a large number of internal NMR probes, *Q. Rev. Biophys. 16*, 1−57.
2. Torchia, D. A. (1984) Solid-state NMR studies of protein internal dynamics, *Annu. Rev. Biophys. Bioeng. 13*, 125−144.
3. Kay, L. E. (1998) Protein dynamics from NMR, *Nat. Struct. Biol. 5*, S513−S517.
4. Fisher, M. W. F., Majumdar, A., and Zuiderweg, E. R. P. (1998) Protein NMR relaxation: theory, applications and outlook, *Prog. Nucl. Magn. Reson. Spectrosc. 33*, 207−272.
5. Ishima, R., and Torchia, D. A. (2000) Protein dynamics from NMR, *Nat. Struct. Biol. 7*, 740−743.
6. Wand, A. J. (2001) Dynamic activation of protein function: a view emerging from NMR spectroscopy, *Nat. Struct. Biol. 8*, 926−931.

7. Palmer, A. G., III (2001) NMR probes of molecular dynamics: overview and comparison with other techniques, *Annu. Rev. Biophys. Biomol. Struct. 30*, 129−155.
8. Ringe, D., and Petsko, G. A. (1985) Mapping protein dynamics by X-ray diffraction, *Prog. Biophys. Mol. Biol. 45*, 197−235.
9. Thüne, T., and Badger, J. (1995) Thermal diffuse X-ray scattering and its contribution to understanding protein dynamics, *Prog. Biophys. Mol. Biol. 63*, 251−276.
10. Karplus M., and McCammon, J. A. (1983) Dynamics of proteins: elements and function, *Annu. Rev. Biochem. 53*, 263−300.
11. Karplus, M., and Petsko, G. A. (1990) Molecular dynamics simulations in biology, *Nature 347*, 631−639.
12. Daggett, V., and Levitt, M. (1993) Realistic simulations of native-protein dynamics in solution and beyond, *Annu. Rev. Biophys. Biomol. Struct. 22*, 353−380.
13. van Gusteren, W. F., Luque, F. J., Timms, D., and Torda, A. E. (1994) Molecular mechanics in biology. from structure to function, taking account of solvation, *Annu. Rev. Biophys. Biomol. Struct. 23*, 847−863.
14. Wang, W., Donini, O, Reyes, C. M., and Kollman, P. A. (2001) Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein−ligand, protein−protein, and protein-nucleic acid noncovalent interactions, *Annu. Rev. Biophys. Biomol. Struct. 30*, 211−243.
15. Karplus, M., and McCammon, J. A. Molecular dynamics simulations of biomolecules, *Nat. Struct. Biol. 9*, 646−652.
16. Warshel, A. (2002) Molecular dynamics simulations of biological reactions, *Acc. Chem. Res. 35*, 385−395.
17. Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, *Science 278*, 1111−1114.
18. Clore, G. M., Starich, M. R., and Gronenborn, A. M. (1998) Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses, *J. Am. Chem. Soc. 120*, 10571−10572.
19. Hansen, M. R., Mueller, L., and Pardi, A. (1998) Tunable alignment of mactomolecules by filamentous phage yields dipolar coupling interactions, *Nat. Struct. Biol. 5*, 1065−1074.
20. Prestegard, J. H., Al-Hashimi, H. M., and Tolman, J. R (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings, *Q. Rev. Biophys. 33*, 371−424.
21. Tolman, J. R. (2001) Dipolar couplings as a probe of molecular dynamics and structure in solution, *Curr. Opin. Struct. Biol. 11*, 532−539.
22. Bax, A., Kontaxis, G., and Tjandra, N. (2001) Dipolar couplings in macromolecular structure determination, *Methods Enzymol. 339*, 127−174.
23. Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1997) NMR evidence for slow collective motions in cyanometmyoglobin, *Nat. Struct. Biol. 4*, 292−297.
24. Tolman, J. R., Al-Hashimi, J. M., Kay, L. E., and Prestegard, J. H. (2001) Structural and dynamic analysis of residual dipolar coupling data for proteins, *J. Am. Chem. Soc. 123*, 1416−1424.
25. Tolman, J. R. (2002) A novel approach to the retrieval of structural and dynamic information from residual dipolar coupling using several oriented media in biomolecular NMR, *J. Am. Chem. Soc. 124*, 12020−12030.
26. Briggman, K. B., and Tolman, J. R. (2003) De novo determination of bond orientations and order parameters from residual dipolar couplings with high accuracy, *J. Am. Chem. Soc. 125*, 10164−10165.
27. Meiler, J., Promper, J. J., Peti, W., Griesinger, C., and Brüschweiler, R. (2001) Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins, *J. Am. Chem. Soc. 123*, 6098−6107.
28. Peti, W., Meiler, J., Brüschweiler, R., and Griesinger, C. (2002) Model-free analysis of protein backbone motion from residual dipolar couplings, *J. Am. Chem. Soc. 124*, 5822−6833.
29. Hus, J.-C., Peti, W., Griesinger, C., and Brüschweiler, R. (2003) Self-consistency analysis of dipolar couplings in multiple alignments of ubiquitin, *J. Am. Chem. Soc. 125*, 5596−5597.
30. Meiler, J., Peti, W., and Griesinger, C. (2003) Dipolar couplings in multiple alignments suggest alpha helical motion in ubiquitin, *J. Am. Chem. Soc. 125*, 8072−8073.
31. Clore, G. M., and Schwieters, C. D. (2004) How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation, *J. Am. Chem. Soc. 126*, 2923−2938.
32. Ulmer, T. S., Ramirez, B. E., Delaglio, F., and Bax, A. (2003) Evaluation of backbone protein positions and dynamics in a small protein by liquid crystal NMR spectroscopy, *J. Am. Chem. Soc. 125*, 9179−9191.
33. Hall, J. B., and Fushman, D. (2003) Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G, *J. Biomol. NMR 27*, 261−275.
34. Derrick, J. P., and Wigley, D. B. (1994) The third IgG-binding domain from streptococcal protein G: an analysis by X-ray crystallography of the structure alone and in a complex with Fab, *J. Mol. Biol. 243*, 906−918.
35. Schwieters, C. D., and Clore, G. M. (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement, *J. Magn. Reson. 152*, 288−302.
36. Schwieters, C. D., Kuszewski, J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package, *J. Magn. Reson. 160*, 66−74.
37. Schwieters, C. D., and Clore, G. M. (2001) The VMD-XPLOR visualization package for NMR structure refinement, *J. Magn. Reson. 149*, 239−244.
38. Cornilescu, G., Ramirez, B. E., Frank, M. K., Clore, G. M., Gronenborn, A. M., and Bax, A. (1999) Correlation between $^{3h}J_{NC'}$ and hydrogen bond length in proteins, *J. Am. Chem. Soc. 121*, 6275−6279.
39. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. M. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G, *Science 253*, 657−661.
40. Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1999) Inproving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration, *J. Am. Chem. Soc. 121*, 2337−2338.
41. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints: application to crambin, potato carboxypeptiase inhibitor and barley serine proteinase inhibitor 2, *Protein Eng. 2*, 27−38.
42. Clore, G. M., and Kuszweski, J. (2002) χ_1 rotamer populations and angles of mobile surface side chains are accurately by a torsion angle database potential of mean force, *J. Am. Chem. Soc. 124*, 2866−2867.
43. Lipsitz, R. S., Sharma, Y., Brooks, B. R., and Tjandra, N. (2002) Hydrogen bonding in high-resolution protein structures: a new method to assess NMR protein geometry, *J. Am. Chem. Soc. 124*, 10261−10266.
44. McCammon, J. A., Gelin, B. R., and Karplus, M. (1977) Dynamics of folded proteins, *Nature 267*, 585−590.
45. Fadel, A. R., Jin, D. Q., Montelione, G. T., and Levy, R. M. (1995) Crankshaft motions of the polypeptide backbone in molecular dynamics simulations of human type-alpha transforming growth factor, *J. Biomol. NMR 6*, 221−226.
46. Clore, G. M., and Garrett, D. S. (1999) R-factor, free R and complete cross-validation for dipolar coupling refinement of NMR structures, *J. Am. Chem. Soc. 121*, 9008−9012.
47. Karplus, P. A. (1996) Experimentally observed conformation-dependent geometry and hidden strain in proteins, *Protein Sci. 5*, 1406−1420.
48. MacArthur, M. W., and Thorton, J. M. (1996) Deviations from planarity of the peptide bond in peptides and proteins, *J. Mol. Biol. 264*, 1180−1195.
49. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules: I Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546−4559.
50. Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of ^{15}N nuclear magnetic relaxation of proteins, *J. Am. Chem. Soc. 112*, 4989−4991.
51. Go, M., and Go, N. (1976) Fluctuations of an alpha-helix, *Biopolymers 15*, 1119−1127.
52. Dellwo, M. J., and Wand, A. J. (1989) Model dependent and model independent analysis of the global and internal dynamics of cyclosporin A, *J. Am. Chem. Soc. 111*, 4571−4578.
53. Palmer, A. G., III, and Case, D. A. (1992) Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide, *J. Am. Chem. Soc. 114*, 9059−9067.

54. Levitt, M. (1983) Molecular dynamics of native protein. II. Analysis and nature of motion, *J. Mol. Biol. 168*, 621−657.
55. Chandrasekhar, I., Clore, G. M., Szabo, A., Gronenborn, A. M., and Brooks, B. R. (1992) A 500 ps molecular dynamics simulation study of interleuin-1β in water: correlation with nuclear magnetic resonance spectroscopy and crystallography, *J. Mol. Biol. 226*, 239−250.
56. Brunne, R. M., Berndt, K. D., Güntert, P, Wüthrich, K., and van Gunsteren, W. F. (1995) Structure and internal dynamics of the bovine pancreatic trypsin inhibitor in aqueous solution from long-time molecular dynamics simulations, *Proteins 23*, 49−62.
57. Fushman, D., Ohlenschläger, O., and Rüterjans, H. (1994) Determination of the backbone mobility of ribonuclease T1 and the 2′GMP complex using molecular dynamics simulations and NMR relaxation data, *J. Biomol. Struct. Dyn. 11*, 1377−1402.
58. Pfeiffer, P., Fushman, D., and Cowburn, D. (2001) Simulated and NMR-derived backbone dynamics of a protein with significant flexibility: a comparison of spectral densities for the βARK1 PH domain, *J. Am. Chem. Soc. 123*, 3021−3036.
59. Bernado, P., and Blackledge, M. (2004) Anisotropic small amplitude peptide plane dynamics in proteins from residual dipolar couplings, *J. Am. Chem. Soc. 126*, 4907−4920.
60. McCallister, E. L., Alm, E., and Baker, D. (2000) Critical role of β-hairpin formation in protein G folding, *Nat. Struct. Biol. 7*, 669−673.
61. Ding, K., Louis, J. M., and Gronenborn, A. M. (2004) Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR, *J. Mol. Biol. 335*, 1299−1307.
62. Ottiger, M., and Bax, A. (1998) Determination of relative N-H^N, N-C′, C^α-C′ and C^α-H^α effective bond lengths in a protein by NMR in a dilute liquid crystalline phase, *J. Am. Chem. Soc. 120*, 12334−12341.
63. Henry, E. R., and Szabo, A. (1985) Influence of vibrational motion on solid-state line shapes and NMR relaxation, *J. Chem. Phys. 82*, 4753−4761.
64. Clore, G. M., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) Analysis of backbone dynamics of interleukin-1β using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy, *Biochemistry 29*, 7387−7401.
65. Jeffrey, G. (1992) in *Accurate Molecular Structures, Their Determination and Importance* (Domenicano, A., and Hargittai, Eds.) pp 270−298, Oxford University Press, Oxford.
66. Engh, R. A., and Huber, R. (1991) Accurate bond and angle parameters for X-ray structure refinement, *Acta Crystallogr. A47*, 392−400.
67. Bax, A., and Tjandra, N. (1997) Are proteins even floppier than we thought, *Nat. Struct. Biol. 4*, 254−256.
68. Clore, G. M., and Gronenborn, A. M. (1998) New methods of structure refinement for macromolecular structure determination by NMR, *Proc. Natl. Acad. Sci. U.S.A. 26*, 5891−5898.
69. Cornilescu, G., and Bax, A. (2000) Measurement of protein, nitrogen and carbonyl chemical shielding anisotropies in a protein dissolved in a dilute liquid crystalline phase, *J. Am. Chem. Soc. 122*, 10143−10154.
70. Bevington, P. R., and Robinson, D. K. (1992) in *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed., pp 205−209, McGraw-Hill, New York.

BI049357W