

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/5494463>

Simultaneous NMR Study of Protein Structure and Dynamics Using Conservative Mutagenesis †

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · JUNE 2008

Impact Factor: 3.3 · DOI: 10.1021/jp0772124 · Source: PubMed

CITATIONS

60

READS

13

4 AUTHORS, INCLUDING:



Beat Vögeli

ETH Zurich

46 PUBLICATIONS 814 CITATIONS

SEE PROFILE



Ad Bax

National Institutes of Health

485 PUBLICATIONS 78,298 CITATIONS

SEE PROFILE

Simultaneous NMR Study of Protein Structure and Dynamics Using Conservative Mutagenesis[†]

Lishan Yao,[‡] Beat Vögeli,[‡] Dennis A. Torchia,[§] and Ad Bax^{*,‡}

Laboratory of Chemical Physics, NIDDK, and National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, Maryland 20892-0520

Received: September 7, 2007; In Final Form: November 28, 2007

A novel iterative procedure is described that allows both the orientation and dynamics of internuclear bond vectors to be determined from direct interpretation of NMR dipolar couplings, measured under at least three orthogonal alignment conditions. If five orthogonal alignments are available, the approach also yields information on the degree of motional anisotropy and the direction in which the largest amplitude internal motion of each bond vector takes place. The method is demonstrated for the backbone ^{15}N – ^1H , $^{13}\text{C}^\alpha$ – $^1\text{H}^\alpha$, and $^{13}\text{C}^\alpha$ – $^{13}\text{C}'$ interactions in the previously well-studied protein domain GB3, dissolved in a liquid crystalline suspension of filamentous phage Pf1. Alignment variation is achieved by using conservative mutations of charged surface residues. Results indicate remarkably uniform backbone dynamics, with amplitudes that agree well with those of previous ^{15}N relaxation studies for most residues involved in elements of secondary structure, but larger amplitude dynamics than those found by ^{15}N relaxation for residues in loop and turn regions. In agreement with a previous analysis of dipolar couplings, the N–H bonds in the second β -strand, which is involved in antibody recognition, show elevated dynamics with largest amplitudes orthogonal to the chain direction.

Experimental NMR studies of protein structure traditionally have relied mostly on extracting interproton distances from extensive sets of nuclear Overhauser effects (NOEs), supplemented by dihedral angles obtained from J couplings and chemical shifts.^{1–3} NMR measurement of ^{15}N and ^{13}C relaxation rates complements the structural picture by providing quantitative information about the amplitudes and rates at which individual NH and CH bond vectors change their orientation relative to the time-averaged structure.^{4–6} These motions are often interpreted using the model-free approach of Lipari and Szabo,⁷ and report on angular excursions taking place on a time scale faster than the rotational correlation time of the protein, typically 10^{-8} s. However, many functionally important motions, such as seen in enzymatic catalysis or allosteric activation, involve rearrangement of large groups of atoms, which often takes place on a much slower time scale, largely invisible to conventional relaxation measurements. The changes in chemical shifts associated with such rearrangements result in resonance line broadening, whose measurement can yield the rates but usually not the amplitudes at which these slower processes are taking place.^{8,9}

With the advent of more recent technology to weakly align macromolecules in solution relative to an external magnetic field, residual internuclear dipolar couplings (RDCs) have become an additional source of NMR information on protein structure and dynamics.^{10–14} The RDCs represent the time-averaged second-order Legendre polynomials, $P_2(\cos \theta)$, where

θ is the time-dependent angle between a given internuclear vector and the magnetic field. The temporal behavior of θ is dominated by the Brownian diffusion of the molecule, but is also affected by internal motions of the internuclear vector relative to the molecular frame. Experimentally measured RDCs therefore carry information on the time-averaged orientation of the corresponding vectors relative to the molecular frame, as well as information on the dynamic behavior within this frame. When using RDCs for structure calculation and refinement purposes, the dynamic behavior is often neglected. However, it is widely recognized that the dynamic information encoded in RDCs intrinsically reports on both the amplitude and direction of the internal motions, integrated over the entire range of time scales faster than milliseconds, thereby providing a perfect complement to the time scale information obtained from relaxation measurements.^{10,14}

Weak alignment of proteins in a magnetic field is usually accomplished by dissolving the protein in a dilute liquid crystalline suspension of magnetically oriented particles, ranging from filamentous phages to lipid bilayers,^{15–18} or in anisotropically compressed hydrogels.^{19–21} Alternatively, attaching a paramagnetic metal to the system can induce magnetic alignment due to the magnetic properties of the metal-chelated molecule.^{22–26} Alignment of the molecule is defined by a traceless and symmetric 3×3 Saupe matrix, and contains five independent elements. This means that, at least in principle, up to five linearly independent alignment tensors can be generated, and measurement under such different alignment conditions can reveal important new information on the structure and internal dynamics of a molecule.^{27–29} Extracting dynamics information in a model-free manner for individual bond vectors in a protein relies on measurement of RDCs under these five different alignment orientations.^{10,11,30,31} However, in practice, save for paramagnetic tagging, alignment is dominated by steric and electrostatic

[†] Part of the “Attila Szabo Festschrift”.

^{*} Corresponding author. E-mail: bax@nih.gov. Address: Building 5, room 126, NIH, Bethesda, Maryland 20892-0520. Phone: 301 496 2848. FAX: 301 402 0907.

[‡] National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

[§] National Institute of Dental and Craniofacial Research.

(repulsive and attractive) forces between the protein and the medium.³² As a result, without changing the shape or charge of the protein, it has remained very difficult to generate alignment conditions that sufficiently map out this five-dimensional space.^{33,34} A second problem with extracting dynamic information from RDCs is that the impact of dynamics on the RDC value in practice is only second-order because, for small angular excursions, to first order, $\langle P_2(\cos \theta') \rangle = P_2(\cos \langle \theta' \rangle)$, where θ' is the angle between the internuclear vector and the x , y , or z -axis of the molecular frame. Therefore, even though only moderate measurement precision is needed to define the average orientation of a given internuclear vector, far more precise RDC measurements are required to extract dynamic information. Although a wide variety of alignment methods are available, many of these may not be compatible with high precision RDC measurement for any given protein, either owing to line broadening induced by the alignment method, or to instability of the alignment medium.

Our current study exploits a recently introduced alternate method for exploring five-dimensional alignment space: generation of mutant proteins that have a different surface charge distribution and therefore align differently in a liquid crystalline medium of negatively charged Pf1 phages.³⁵ The Pf1 medium is highly robust, commercially available, and has minimal adverse effect on the quality of the protein NMR spectrum.¹⁷ Measurement of RDCs for mutants of the third immunoglobulin binding domain of protein G (GB3) indicated no measurable effect of any of the mutations on the backbone amide N–H vector orientations for 6 mutants relative to the native domain.³⁵ Here, we report the measurement of RDCs for ^{15}N – ^1H , $^{13}\text{C}^\alpha$ – $^1\text{H}^\alpha$, and $^{13}\text{C}^\alpha$ – $^{13}\text{C}'$ vectors in these six mutants and use them to explore the dynamic properties of these vectors.

Before extracting the structure and dynamics parameters from RDCs, the alignment tensor has to be determined. If the structure of the molecule is known, a singular value decomposition (SVD) fit of the RDCs to the orientations of the corresponding bond vectors in the structure is used to calculate the alignment tensor elements.^{36,37} In the absence of the structure, determining the alignment tensor becomes more challenging. Approximate values for the magnitude and rhombicity of the alignment tensor may be obtained from the histogram of observed dipolar couplings,^{38–40} but this method decreases in accuracy when the bond vectors are not distributed uniformly or when fewer than ca. 100 RDCs are available. Provided that N–H, N–C', and C'–C $^\alpha$ data are available simultaneously under five different alignment conditions and under the assumption that these vectors associated with a single peptide group move as a rigid unit, the known relative orientations of these bond vectors can be exploited to extract simultaneous alignment and structure information.⁴¹ An elegant alternative approach, named *direct interpretation of dipolar couplings* or DIDC, has been introduced to extract the alignment tensors without the need for prior structural information or assumptions about the internal motions, but requires that each coupling is measured under alignment conditions that together map the entire five-dimensional alignment space.^{29,30,42}

Our present study describes a variant, iterative DIDC method, which we show to result in improved accuracy of the alignment tensors, and which yields reliable values for extracted asymmetric motion parameters when using simulated data, as well as reasonable numbers when using experimental RDCs. Importantly, our iterative DIDC method also allows determination of the alignment tensors if as few as three independent RDC data sets are measured.

The method is evaluated for the third IgG binding domain of protein G, GB3, a system previously characterized extensively by both X-ray crystallography and NMR spectroscopy.^{43–49} Application of the iterative DIDC method to RDCs measured for six different GB3 mutants is shown to yield a rather uniform view of backbone dynamics, but also highlights the presence of regions in the protein, in particular in its turns, where the amplitude of dynamics is larger than that observed by ^{15}N relaxation methods. In particular, for one of the edge strands, $\beta 2$, the amplitude of motions of the N–H vector around the C $^\alpha$ –C $^\alpha$ vector (the so-called γ motion)^{49,50} is of larger amplitude than motion within the peptide plane. The same motional anisotropy is also observed for C $^\alpha$ –H $^\alpha$ vectors in this region, whereas for C $^\alpha$ –C' vectors the limited accuracy at which these couplings can be measured (~ 0.1 Hz) prohibits quantitative analysis of motional anisotropy.

Theoretical Analysis

In the secular approximation, the dipolar coupling between nuclei I and S can be expressed as

$$D^{\text{IS}} = D_{\text{max}}^{\text{IS}} \langle (3 \cos^2 \theta - 1)/2 \rangle \quad (1)$$

where the angular brackets refer to the ensemble average, θ is the angle between the I–S internuclear vector and the magnetic field, and $D_{\text{max}}^{\text{IS}} = -\mu_0(h/2\pi)\gamma_I\gamma_S/(4\pi^2r_{\text{IS}}^3)$. Here, μ_0 is the magnetic permittivity of the vacuum, h is Planck's constant, γ_x is the magnetogyric ratio of spin X, and r_{IS} is the distance between nuclei I and S. For rigid molecules, eq 1 can be rewritten in the principal axis system of the alignment tensor as⁵¹

$$D^{\text{IS}} = D_a \left[(3 \cos^2 \theta - 1) + \frac{3}{2} R \sin^2 \theta \cos 2\phi \right] \quad (2)$$

where D_a is referred to as the magnitude of the residual dipolar coupling tensor, R is the rhombicity, and (θ, ϕ) are the polar coordinates defining the orientation of the vector in the alignment principal axis system. Below, we provide a slightly more compact formulation, applicable to the case where multi-alignment RDCs are measured, than originally introduced by Tolman.³⁰ Eq 1 can be rewritten in the molecular frame by using the spherical harmonics addition theorem:^{27–29}

$$D^{\text{IS}} = D_{\text{max}}^{\text{IS}} \langle b \cdot a \rangle \quad (3)$$

where $b \cdot a$ is the dot product between vectors b and a , and b is the five-dimensional unit vector,

$$b = \left\{ (3z^2 - 1)/2, \frac{\sqrt{3}}{2}(x^2 - y^2), \sqrt{3}xz, \sqrt{3}yz, \sqrt{3}xy \right\} \quad (4)$$

and x , y , and z are the Cartesian coordinates of vector IS in the molecular frame. The vector a has the same form as b , but with x , y , and z denoting the orientation of the magnetic field in the molecular frame. If L internuclear RDCs are measured under N different alignment conditions, eq 3 can be written in a matrix form,

$$\mathbf{D}_M = D_{\text{max}}^{\text{IS}} \langle \mathbf{B} \mathbf{A} \rangle \quad (5)$$

Here, \mathbf{D}_M is the $L \times N$ RDC matrix and \mathbf{B} is an $L \times 5$ matrix, each row of which is a vector b , and \mathbf{A} is a $5 \times N$ matrix, each column of which is a vector a . If bond and field fluctuations are uncorrelated, a premise confirmed by SECONDIA self-consistency analysis^{52,53} (*vide infra*), eq 5 can be written as

$$\mathbf{D}_M = D_{\max}^{\text{IS}} \langle \mathbf{B} \rangle \langle \mathbf{A} \rangle \quad (6a)$$

One can simplify eq 6a by incorporating D_{\max}^{IS} into $\langle \mathbf{A} \rangle$,

$$\mathbf{D}_M = \langle \mathbf{B} \rangle \langle \mathbf{A} \rangle \quad (6b)$$

Information content of RDC measurements is maximized when RDCs are measured in five independent alignments.^{10,30,34,42,52,53} However, because both $\langle \mathbf{B} \rangle$ and $\langle \mathbf{A} \rangle$ are unknown, to our knowledge it is not possible to solve eq 6 without any restriction on the bond motions. Tolman solved the problem by assuming that the most ordered vectors in a molecular system can be treated as rigid, i.e., $\langle \mathbf{B} \rangle = \mathbf{B}$, or $b \cdot b = 1$.⁴² In matrix form, this corresponds to

$$\text{Diag}\{\langle \mathbf{B} \rangle \langle \mathbf{B} \rangle^T\} = 1 \quad (7)$$

where $\langle \mathbf{B} \rangle^T$ is the transpose of $\langle \mathbf{B} \rangle$. Tolman used this condition and an elegant linear formulation to obtain the structure matrix \mathbf{B} and then the alignment matrix $\langle \mathbf{A} \rangle$. However, if a bond vector is rigid, besides eq 7, two additional constraints relate the elements of each rigid b -vector:

$$3b_3^2 = ((1 + \sqrt{3}b_2) - b_1) \times (1 + 2b_1) \quad (8a)$$

$$3b_4^2 = ((1 - \sqrt{3}b_2) - b_1) \times (1 + 2b_1) \quad (8b)$$

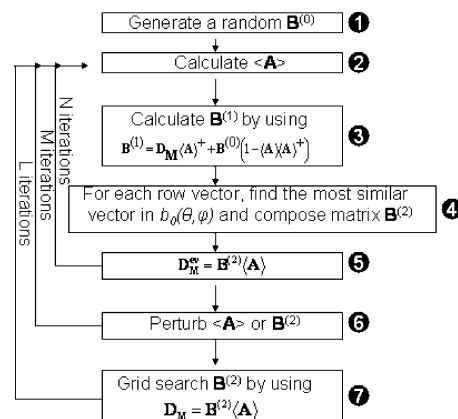
The five elements of each rigid b -vector are constrained by three equations because each such five-dimensional b -vector is a function of only two parameters: θ and φ . Explicitly,

$$b = \left\{ (3 \cos^2 \theta - 1)/2, \frac{\sqrt{3}}{2} \sin^2 \theta \cos 2\varphi, \frac{\sqrt{3}}{2} \sin 2\theta \times \cos \varphi, \frac{\sqrt{3}}{2} \sin 2\theta \sin \varphi, \frac{\sqrt{3}}{2} \sin^2 \theta \sin 2\varphi \right\} \quad (9)$$

It is anticipated that incorporating the two added constraints, eqs 8a and 8b, into the DIDC method will significantly improve the accuracy of the results obtained, and will also (under conditions discussed below) reduce the number of alignments needed to obtain unique solutions for $\langle \mathbf{A} \rangle$ and $\langle \mathbf{B} \rangle$. However, it is difficult to incorporate these two constraints directly into the DIDC method since eqs 8a and 8b are nonlinear. In this paper, we present a new algorithm that improves upon the DIDC approach by fully utilizing the three constraints for rigid b -vectors (eqs 7 and 8), thereby greatly increasing the accuracy of $\langle \mathbf{B} \rangle$ and $\langle \mathbf{A} \rangle$ derived from RDCs. It is important to note that, with the constraints of eqs 7 and 8 in place, there remain two unknown parameters for each N–H vector (θ and ϕ) as well as five elements describing $\langle \mathbf{A} \rangle$ of each mutant. When RDCs are available for K bond vectors in M mutants, the solution therefore becomes overdetermined when $K \times M > 2K + 2M + 3(M - 1)$, where $2M$ corresponds to the unknown magnitude and rhombicity of each $\langle \mathbf{A} \rangle$, and $3(M - 1)$ Euler angles define their relative orientations. A minimum of $K \geq 12$ RDC measurements is therefore required when using three independent alignment orientations.

We begin by using eq 9 to calculate an array of rigid b -vectors $b_0(\theta_i, \phi_j)$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$. The range is $0-90^\circ$ for θ and $0-360^\circ$ for ϕ . Since the inversion of a bond vector does not change b , half of the full range θ ($0-90^\circ$) suffices for generating b -vectors. The grid size is 0.2° for both θ and ϕ , fine enough to calculate accurate structure and dynamics parameters, as shown in the Results section. By construction, each b -vector in the array $b_0(\theta_i, \phi_j)$ satisfies the three rigid-vector

SCHEME 1: Flowchart of the Iterative DIDC Algorithm^a



^a For the simulated data using three alignments, the numbers of iterations used are $N = 10$, $M = 40$, and $L = 40$. For the calculations using five sets of simulated or six sets of experimental RDCs, L is reduced to 6 and M is reduced to 20 because increasing the number of iterations does not improve the fit quality.

constraints. Next, we construct an initial structure matrix, $\mathbf{B}^{(0)}$, whose N rows are N vectors randomly chosen from the array $b_0(\theta_i, \phi_j)$. Because N is typically much greater than 5, $\mathbf{B}^{(0)}$ will be full rank (5) enabling $\langle \mathbf{A} \rangle$ to be calculated according to $\langle \mathbf{A} \rangle = (\mathbf{B}^{(0)})^+ \mathbf{D}_M$, where the plus superscript refers to the Moore–Penrose inverse. We use $\langle \mathbf{A} \rangle$ to calculate an improved \mathbf{B} according to

$$\mathbf{B}^{(1)} = \mathbf{D}_M \langle \mathbf{A} \rangle^+ + \mathbf{B}^{(0)} [1 - \langle \mathbf{A} \rangle \langle \mathbf{A} \rangle^+] \quad (10)$$

This expression is similar to the one used by Tolman for calculating a refined structure matrix given a structural model $\mathbf{B}^{(0)}$ and an alignment tensor $\langle \mathbf{A} \rangle$.³⁰ The first term of the right side of eq 10 corresponds to the contribution to $\mathbf{B}^{(1)}$ made by the RDC data. Equation 10 is a solution of eq 6, but not a unique solution because it depends on the initial choice of $\mathbf{B}^{(0)}$. The degeneracy results from the fact that $\langle \mathbf{A} \rangle$ is not full ranked (smaller than 5). If $\langle \mathbf{A} \rangle$ were full ranked, the second term of eq 10 would equal zero, and would not contribute to driving convergence.

The aim of eq 10 is to find a solution satisfying eq 6, closest to the starting matrix $\mathbf{B}^{(0)}$. $\mathbf{B}^{(1)}$ is an improvement over $\mathbf{B}^{(0)}$ because $\mathbf{B}^{(1)}$ is a solution of eq 6 while $\mathbf{B}^{(0)}$ is not. Following the calculation of $\mathbf{B}^{(1)}$, we incorporate the three rigid-vector restraints for each b -vector (each row of $\mathbf{B}^{(1)}$) to obtain $\mathbf{B}^{(2)}$ in the following manner. First, we denote the rows of $\mathbf{B}^{(1)}$ as $\mathbf{B}^{(1)}_k$, $k = 1, \dots, N$, and denote b' as the vector in the array $b_0(\theta_i, \phi_j)$ that minimizes the norm $\|b_0(\theta_i, \phi_j) - \mathbf{B}^{(1)}_k\|$. Second, we construct $\mathbf{B}^{(2)}$ by taking as its rows the vectors b' . Note that $\mathbf{B}^{(2)}$ is constructed so that each of its row vectors satisfies the three rigid-bond constraints and also most closely represents the experimental data. The steps discussed above are summarized in the first four boxes of the flow chart in Scheme 1, which depicts the complete algorithm used to calculate $\langle \mathbf{A} \rangle$ and \mathbf{B} .

After obtaining $\mathbf{B}^{(2)}$ and $\langle \mathbf{A} \rangle$, a new RDC matrix \mathbf{D}_M^{ev} is calculated by using $\mathbf{D}_M^{\text{ev}} = \mathbf{B}^{(2)} \langle \mathbf{A} \rangle$, and the error ϵ is estimated by using

$$\epsilon = \text{Tr}\{(\mathbf{D}_M^{\text{ev}} - \mathbf{D}_M)(\mathbf{D}_M^{\text{ev}} - \mathbf{D}_M)^T\}$$

where \mathbf{X}^T is the transpose of matrix \mathbf{X} , and $\text{Tr}(\mathbf{X})$ is the trace of \mathbf{X} . The algorithm stops when the error is comparable to the

experimental error or when the maximum number of iterations is reached. Otherwise, $\mathbf{B}^{(2)}$ is used to calculate a new alignment matrix by $\langle \mathbf{A} \rangle = (\mathbf{B}^{(2)})^+ \mathbf{D}_M$, and the algorithm returns to box 2. Typically it takes 10–20 iterations to converge to a minimum, although this is not necessarily the global minimum. To ensure reaching the global minimum, a Monte Carlo method using the Metropolis criterion⁵⁴ is used to alternately perturb the \mathbf{B} and $\langle \mathbf{A} \rangle$ matrices (box 6 in Scheme 1, M iterations). A random walk was used for the \mathbf{B} (or $\langle \mathbf{A} \rangle$) matrix and the Metropolis rule is used to accept (or reject) the move. More details of the random walk are described in the Supporting Information. The final step is the grid search of \mathbf{B} , by using $\langle \mathbf{A} \rangle$ from the previous box, in order to minimize the RDC error for each vector. Note that, in all likelihood, more elegant and faster procedures can be developed for finding the best value of \mathbf{B} ; however, considering that the conceptually simple grid search was fully adequate for this purpose, development of such a more sophisticated algorithm was not pursued.

The iterative scheme of Flowchart 1 is carried out in multiple cycles. In the first cycle, we include all the vectors in the algorithm. In subsequent cycles, the residue with the largest RDC discrepancy between the measured and predicted value is excluded if this discrepancy exceeds the average fitting error by more than a factor of 2, and the whole process is repeated. This procedure excludes the flexible residues, and the thus obtained subset of RDCs is then used to find the best fitted alignment tensors, $\langle \mathbf{A} \rangle$. The essential idea behind this approach is to first find the best alignment matrix $\langle \mathbf{A} \rangle$ and structure matrix \mathbf{B} satisfying eqs 7 and 8, and which give computed RDCs that best match the experimental data. This procedure therefore yields better alignment tensors and structural information than would be obtained in the absence of using eqs 8, without requiring prior knowledge of vector orientations. The assumption implicit in eqs 7 and 8 (equivalent to $S = 1$) introduces errors in the (θ, ϕ) and amplitude of each NH $\langle b \rangle$ vector, which will be shown to be very small. By definition, the ratio between the norm of the true $\langle b \rangle$ and approximated $\langle b \rangle$ is less than 1. In addition, this norm ratio is nonuniform among different NH vectors and therefore introduces a scaling factor in the $\langle \mathbf{A} \rangle$ matrix, which cannot be determined from RDCs alone (in the absence of a motional model). For example, one can multiply $\langle \mathbf{B} \rangle$ by 2.0 and divide $\langle \mathbf{A} \rangle$ by 2.0 to obtain an identical RDC matrix \mathbf{D} . As confirmed by simulated data (*vide infra*), the scaling factor in the $\langle \mathbf{A} \rangle$ matrix falls very close to the average of the order parameters S of the residues included in the alignment determination. The nonuniform norm ratios and the orientation difference between the true $\langle b \rangle$ and approximated $\langle b \rangle$ also will introduce an error in $\langle \mathbf{A} \rangle$. However, as confirmed by simulated data (see Figure 3 in Results), this latter error is very small because many more than five $\langle b \rangle$ vectors are used in the alignment determination, and their random errors largely cancel out when determining $\langle \mathbf{A} \rangle$.

After having obtained the accurate alignment tensors, eq 6 is used again, but in the absence of the constraints implicit in eqs 7 and 8, in order to extract the structure and dynamics parameters. After obtaining the order parameters, any flexible residues not excluded from the above process are removed, and a new subset of RDCs corresponding to rigid residues are used to obtain a new $\langle \mathbf{A} \rangle$, which is subsequently used to extract the new $\langle \mathbf{B} \rangle$ matrix.

Although our above discussion tacitly assumes that data from five independent alignments are available, structural and dynamical information already can be obtained from the RDC data of only three independent alignments. This information can

be extracted because each rigid vector is a function of only two parameters (θ, φ) . Therefore, even when data is limited to only three alignments, the alignment matrix $\langle \mathbf{A} \rangle$ and (rigid bond) structure matrix \mathbf{B} can be determined using Scheme 1. In addition, order parameters can be determined for flexible bonds, but only under the assumption that their motions are axially symmetric. The performance of the algorithm will be discussed in the Results section.

Materials and Methods

Experimental Data. Six 1.2 mM samples of GB3 mutants—K19AD47K, K19ED40N, K19AT11K, K19EK4A-C-His₆, K19EK4A-N-His₆, and K19EK4A—were prepared in 20 mM phosphate, 50 mM NaCl buffer, pH 6–6.5, 93% H₂O, 7% D₂O, where N-His₆ and C-His₆ refer to N-terminal and C-terminal six-residue histidine tags, commonly used in protein purification. With the exception of K19EK4A-C-His₆ (pH 7.9 in pf1), the pH of all samples was in the 6–6.5 range.

All NMR spectra were recorded on a Bruker DMX500 spectrometer, operating at a ¹H resonance frequency of 500 MHz, and equipped with a cryogenic triple-resonance probehead. The C^α–C', C^α–H^α and N–H RDCs were derived from differences in splittings between the aligned samples and the corresponding isotropic samples. One-bond ¹⁵N–¹H and ¹³C^α–¹³C' splittings were extracted from three-dimensional HNCO experiments without ¹H decoupling during ¹⁵N evolution and with ¹³C^α decoupling omitted during C' evolution. The acquisition times were 112 (¹³C'), 90 (¹⁵N), and 64 ms (¹H), with the data matrix consisting of 56 × 128 × 512 complex points. In the 3D HNCO spectra, each correlation was split into four multiplet components, separated by ¹J_{HN} + ¹D_{HN} in the ¹⁵N dimension and by ¹J_{C'Ca} + ¹D_{C'Ca} in the ¹³C' dimension, yielding duplicates for each splitting and allowing error estimation (0.1 Hz for ¹D_{HN} and ¹D_{C'Ca}) from the root-mean-square (rms) value of the pairwise differences. One-bond C^α–H^α and ¹⁵N–¹H RDCs were extracted from 3D HN(CO)CA experiments with errors of ca. 0.4 Hz for ¹D_{HaCa} and 0.2 Hz for ¹D_{HN}. The HN(CO)CA acquisition times were 28 (¹³C^α), 68 (¹⁵N) and 64 ms (¹H), using data matrices of 87 × 96 × 512 complex points. All data were processed and analyzed using NMRPipe.⁵⁵

Generating Synthetic RDC Data in the Absence of Dynamics. Synthetic N–H RDCs were generated by using eq 6. Alignment tensors used were those obtained from an SVD fit of each mutant's experimental N–H RDCs to the GB3 NMR structure (PDB entry 2OED), previously determined by measurement of RDCs in five different alignment media,⁴⁷ and assuming $S = 1$. So, these synthetic RDCs fit perfectly to the 2OED structure, when using the experimentally determined alignment tensors.

Generating Synthetic RDC Data in the Presence of Dynamics. The synthetic N–H RDCs were generated in a manner similar to that described above for the rigid case, but with NH bond dynamics included. The N–H vector angles (θ, φ) were derived from the 2OED structure (implicitly assuming the Z axis of the NH bond tensor to be along the static NH bond direction in the 2OED structure), and the synthetic order parameters, S , mimic the ¹⁵N Lipari–Szabo relaxation order parameters:⁵⁶ $S = S_{LS}^2$. By setting $S = S_{LS}^2$, our simulated data correspond to much higher degrees of internal dynamics than observed by ¹⁵N relaxation, thereby allowing for motions taking place on a time scale longer than the protein's rotational correlation time. The asymmetric motion parameter η and rotational angle γ were generated randomly (Supporting Information). The five parameters are briefly described below. In a coordinate system where the principal axis of the

largest principal component of the bond tensor is parallel to the Z axis,

$$S = (3\langle z^2 \rangle - 1)/2 \quad (11)$$

$$\eta = (\langle y^2 \rangle - \langle x^2 \rangle)/S \quad (12)$$

with $\langle x^2 \rangle \leq \langle y^2 \rangle \leq \langle z^2 \rangle$, and x , y , and z are the Cartesian coordinates of the H atom, provided the N atom is at the origin, using a normalized bond length ($x^2 + y^2 + z^2 = 1$), and $\langle \rangle$ denotes ensemble averaging. If $(\theta, \varphi, \gamma)$ are the three Euler angles for rotating the coordinate system from the molecular frame to the principal axis system of the dynamically averaged bond tensor, a total of five parameters ($S, \eta, \theta, \varphi, \gamma$) are needed to generate each b vector and then the RDCs.

A total of 200 sets of RDCs each with random noise of 0.2, 0.4, and 0.8 Hz were generated to test the accuracy and precision of the algorithm and explore the effects of RDC uncertainties upon the extracted structure and dynamics parameters, with alignment strengths following those determined experimentally, corresponding to D_a values of 10.7, -7.2 , 8.6, 5.4, and 9.0 Hz for K19AD47K, K19ED40N, K19ET11K, K19EK4A-C-His-tag, and K19EK4A-N-His-tag, respectively.

Extracting Structure and Dynamics from Simulated RDCs. Because bond vector orientations very close to those of the 2OED structure can already be derived from RDC data using just three alignment frames, assuming a rigid structure, the otherwise time-consuming grid search procedure was greatly accelerated by taking the output of any of these initial rigid structure calculations as the starting structure for the iterative algorithm that aims to extract the full five-parameter $\langle b \rangle$ vectors, i.e., $(\theta, \varphi, S, \eta, \gamma)$, for each bond. After obtaining the alignment tensors, eq 6 is used to extract the structure (θ, φ) and dynamics information (S, η, γ) .

We also applied our approach to a simpler dynamic model, in which bond motion is assumed to be axially symmetric (isotropic), and S is simply a scaling factor applicable to the RDCs measured for any given bond vector. Therefore, in this case, each bond tensor is axially symmetric and defined by only three parameters (θ, φ, S) . With the alignment tensors known, S is now simply determined from a two-dimensional (θ, φ) grid search, where, at each point, the best value of S is determined from a linear fit between the RDCs predicted for the (θ, φ) grid point and the true RDCs.

The reduction in unknown bond tensor parameters from 5 to 3 makes it possible, at least in principle, to use Scheme 1 for extracting the dynamics and structure information when RDCs are measured for only three (rather than five) independent alignments. In this case, initially the $S = 1$ assumption is made, and this condition is relaxed only for residues that cannot be fit to within the RDC error. Note that, in the case of RDC noise, this latter procedure introduces a bias, because any small deviation of a given order parameter from the average S value that would alter the RDCs for that vector by less than the assumed RDC error remains undetected.

Extracting Structure and Dynamics from Experimental RDCs. Although both the orientation and dynamics of bond vectors can be derived from randomly selected starting orientations, this procedure yields poor convergence when aiming to fit all five sets of RDCs at once (see above). Therefore, a two-step procedure is again employed, where the fitting of the six sets of RDCs uses as a starting value the best fit obtained from an initial optimization of $\langle \mathbf{A} \rangle$ and \mathbf{B} for just three sets of RDCs (using 200 random starting structures). For each mutant, the ^{15}N – ^1H RDCs from HNCO experiments were fitted to the

corresponding values from HN(CO)CA in each alignment. The slope of the obtained fits, all having correlation coefficients of >0.995 , was then used to scale the RDCs from HNCO to match the amplitude of alignment strength in HN(CO)CA measurement, thereby accounting for minute changes in alignment strength with age of the sample. The N–H RDCs used for the alignment determination and dynamics extraction are the average of those from HNCO and HN(CO)CA experiments. The C^α – H^α RDCs were scaled by a factor of 2.08^{-1} ,⁵¹ and were combined with N–H RDCs to obtain the alignment tensors. Similarly, the C^α – C' RDCs were scaled by a factor 0.198^{-1} to match their amplitude to N–H couplings,⁵¹ but, considering their lower intrinsic precision, these were only used for deriving isotropic dynamics information of the C^α – C' bonds. Assuming a C^α – C' reference bond length of 1.329 \AA ,⁵⁷ these scale factors correspond to effective N–H and C–H bond lengths of 1.041 and 1.117 \AA ,⁵⁸ respectively, incorporating the effect of zero-point librations.⁵⁹ After obtaining $\langle \mathbf{A} \rangle$, a two-step procedure is used to derive the structure and dynamics parameters: (1) An isotropic model with three parameters (S, θ, φ) is used to fit the RDCs. (2) The error of the fit is compared with the measurement error. For example, assuming the measurement error of N–H couplings is 0.3 Hz , with six RDCs measured for each vector and three adjustable parameters, the fitting error threshold is $(6-3) \times 0.3^2 = 0.27 \text{ Hz}^2$. (3) Any vector with a fitting error larger than the threshold is fitted to the fully anisotropic model. Monte Carlo calculations were used to evaluate the errors in structure and dynamics parameters, adding Gaussian distributed uncertainties of 0.3 , 0.6 , and 0.15 Hz (before scaling) to the measured N–H, C^α – H^α and C^α – C' RDCs. As discussed below, the precision of parameters derived using the fully anisotropic model deteriorates quickly as the RDC error increases, effectively limiting the determination of reliable anisotropic motion parameters to measurements with RDC errors of $\leq 0.5 \text{ Hz}$ (for an alignment strength, $D_a \approx 8 \text{ Hz}$). Therefore, only the isotropic model, which has two fewer adjustable parameters and therefore is less sensitive to experimental noise, could be used for analysis of C^α – C' motions from the corresponding RDCs.

As described above for the simulated data, three sets of RDCs, in principle, suffice for extracting structure and dynamics from the RDC data when using the isotropic model. Following the same procedure, the $S = 1$ assumption was made initially, and only when the corresponding RDCs could not be fit to within experimental noise, the (θ, φ) grid was searched again, but optimizing S at each grid point.

Extracting H–N–C $^\alpha$ –H $^\alpha$ Dihedral Angles. Although the iterative DIDC method yields both dynamic and structural information, the latter is generated only as vector orientations and not as a complete three-dimensional model of the protein. To convert these orientations into H–N–C $^\alpha$ –H $^\alpha$ dihedral angle information, these vectors were added to the (non-protonated) backbone model of the 2OED NMR structure.

Results

Iterative DIDC Using Noise-Free Synthetic RDC Data. We first demonstrate our iterative DIDC method using three “error-free” synthetic RDCs data sets, generated as described in the previous section using the 2OED structure (without dynamics), to derive solutions for $\langle \mathbf{A} \rangle$ and \mathbf{B} . With the Monte Carlo parametrization used in our study, only 12 out of 200 randomly generated sets of N–H vector orientations, or “structures” converge to yield complete agreement with the input RDC data, and the bond vector orientations of these 12 “structures” fall

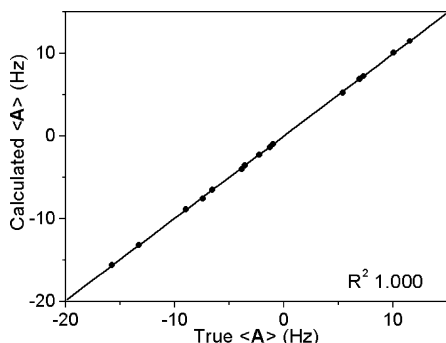


Figure 1. Correlation between converged alignment tensor components, calculated starting from random structures, and the true alignment components. The vertical axis represents 15 components of the predicted alignment matrix $\langle \mathbf{A} \rangle$ for three mutants of GB3, while the horizontal axis represents those of the true alignment matrices, used to generate the synthetic RDCs, without added noise, and without inclusion of the effect of internal dynamics.

very close to the 2OED structure used to generate the synthetic RDCs. Note that, as pointed out in the Theoretical Analysis section (below eq 9), the solutions are over determined by the RDCs, and obtaining a single converged structure already suffices in the absence of noise. However, in the presence of noise, we require only that a solution converges to within experimental error, and it therefore becomes desirable to find multiple solutions starting from random initial coordinates to ensure all solutions (converged structures) are essentially the same.

The total χ^2 error between synthetic and best-fitted RDCs, $\Sigma(\text{RDC}^{\text{syn}} - \text{RDC}^{\text{calc}})^2$, where the summation extends over all experimental couplings in the three mutants, equals $\leq 0.25 \text{ Hz}^2$. The average angular error for the resulting vector orientations equals 0.4° . Because the three alignment orientations employed in the present study are not perfectly orthogonal, not all RDCs are equally sensitive to small changes in vector orientation, and even in the converged sets, one residue (A48) shows angular errors as large as 1.3° , but these errors approach the limit resulting from the 0.2° resolution of the (θ, φ) grid. After rotating the alignment frames of the three mutants such that they correspond to the same molecular frame, Figure 1 compares the 15 matrix components of the thus-obtained predicted alignments with those of the corresponding true alignment matrices. The excellent agreement between the predicted and true alignment tensors confirms that the iterative DIDC method can accurately extract the three alignment tensors when three independent sets of RDCs are available, in the absence of prior structural information, albeit at relatively low convergence rates. Our results also show that it is possible to extract accurate structure information using just three sets of independent RDCs instead of five.

Higher convergence can be reached by increasing the Monte Carlo temperature T , and the iteration parameter M in Scheme 1, considerably lengthening the required calculation time. Instead, to ensure that the global minimum is reached, we prefer to increase the number of random starting structures, followed by verification that all structures that converge to agreement with the RDCs are essentially identical to one another.

By adding two more sets of synthetic RDCs without dynamics, with alignment tensors corresponding to those of K19EK4A-C-His₆ and K19EK4A-N-His₆, the number of convergent structures increases to 50 when starting from the 200 structures obtained above at the last stage of the analysis of the data of the three mutants. Interestingly, even though only 12 of those 200 had fully converged with three sets of RDCs, adding the

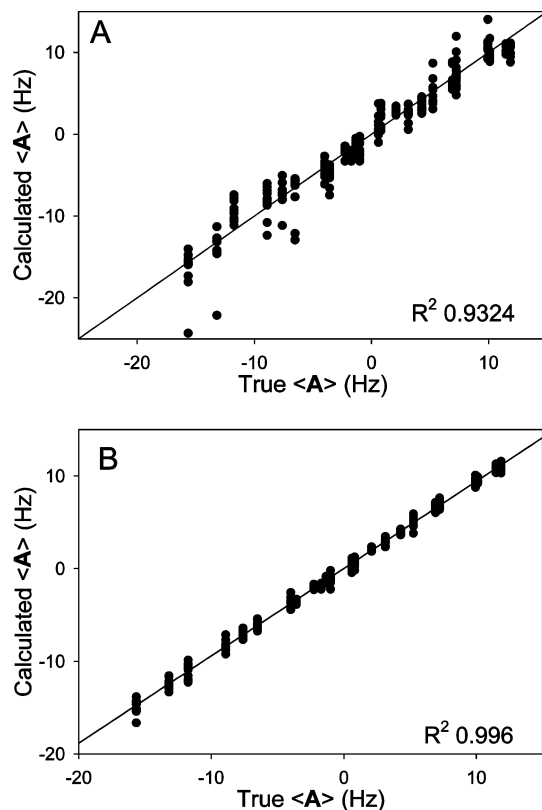


Figure 2. The predicted alignment matrix components calculated using (A) the direct (SVD) DIDC method and (B) the new, iterative DIDC method, for five sets of simulated RDCs, in the presence of noise. Synthetic RDCs were generated assuming fully anisotropic motion as described in the text, with the five alignments chosen to be those obtained from SVD fits of experimental N–H RDCs of GB3 mutants K19AT11K, K19ED40N, K19AD47K, K19EK4A-C-His₆, and K19EK4A-N-His₆ to the 2OED structure. Gaussian distributed noise of 0.8 Hz rms magnitude was added to the simulated RDCs, to generate 10 RDCs for each vector in each alignment. The small underestimate in $\langle \mathbf{A} \rangle$, manifested in a slope smaller than one in panel B, results from the initial approximation that $S = 1$ during iterative DIDC.

RDCs of the two additional mutants results in a considerable increase of the total number of converged structures. The main reason for this is that, of the 188 nonconverged structures, many were already close to the true solution but trapped in a local minimum. Instead, if completely random vector orientations are used as the initial guess when fitting the synthetic RDCs of five independent alignments, Scheme 1 yields poor convergence because either the \mathbf{B} or $\langle \mathbf{A} \rangle$ matrix has a tendency to become singular during the course of the iteration. This observation suggests that a reasonable starting structure is important for convergence when data from five or more alignments are utilized. In practice, we therefore use RDCs from three sets of alignments to obtain a reasonably good initial structure and then add the extra sets of RDCs to further refine the $\langle \mathbf{A} \rangle$ matrix, as done above. Alternatively, if available, an existing structure (e.g., X-ray or NMR structure) can be used.

Next, we demonstrate the use of the iterative DIDC method for five sets of RDCs in the presence of fully anisotropic motion, with synthetic RDCs generated as described in the Methods section. Because data from five alignments are used, the performance of SVD-DIDC and iterative DIDC can be compared (Figure 2). Results are shown for 10 sets of input RDCs, with Gaussian noise of 0.8 Hz. Eight flexible residues are excluded by using the procedure discussed in the Methods section. Although a close correlation between the true and best-fitted alignment tensor elements is observed for both methods,

TABLE 1: Comparison of rms Errors in the Structure and Dynamics Parameters from SVD-DIDC and Iterative DIDC^a

	orientation (°)	S	η	γ (°)
SVD-DIDC	0.88	0.015	0.040	62 (9.4 ^b)
Iterative DIDC	0.48	0.010	0.007	10.1 (2.7 ^b)

^a A total of 55 error-free RDCs simulated for each of five alignments were fitted. The rms error is defined as the rms deviation relative to the true value. The orientation error is defined as the rms angular deviation from the true orientation. ^b Only residues with η larger than 0.05 are included.

we note that the direct DIDC method is very sensitive to the presence of noise. The correlation coefficients are 0.932 for the direct DIDC results and 0.996 for iterative DIDC. As expected, owing to the enforcement of the two additional constraints of eq 8 during the iterative method, the alignment tensor accuracy obtained from iterative DIDC is much improved.

In both the direct and iterative DIDC methods, one implicit assumption is that the vectors selected for alignment determination are either rigid or have uniform order parameters. This assumption inevitably introduces some error when extracting the structure and dynamics parameters, and the magnitude of these errors can be estimated by using simulated RDC data (Table 1). The residues retained during iterative DIDC yield S order parameters ranging from 0.85–0.98. The rms error is 0.48° for the orientation, 0.01 for the order parameter S , 0.007 for η , and 2.7° for γ (for residues with $\eta > 0.05$). Although these small errors can be reduced by further restricting the evaluation to residues with a narrower variation in S , as discussed below, the error in practice is dominated by noise in the experimental RDCs. In the absence of experimental noise, excellent agreement is observed between S values used to generate the RDCs and those extracted from the RDCs using the iterative DIDC method (Figure 3). One intrinsic aspect of S values extracted from RDCs is the presence of an unknown but uniform scaling factor for S . In the simulated data, this scaling factor is determined by fitting the predicted alignments to the true alignments, while, in the experimental data, the S value of the most rigid vector is commonly set to 1.^{10,30}

Iterative DIDC Using Synthetic RDC Data in the Presence of Noise. In order to further explore the accuracy and precision of iterative DIDC and to gain general insights into the impact of experimental errors upon the accuracy of the extracted structure and dynamics parameters, Gaussian-distributed random errors of 0.2, 0.4, and 0.8 Hz rms magnitude are added to the simulated RDCs. The alignment tensors are calculated, and the structure and dynamics parameters are extracted as described above, with results summarized in Table 2. Compared to the direct DIDC method, iterative DIDC gives considerably improved structure and dynamics parameters (Table 2), with the improvement being most pronounced for the asymmetric motion parameters. Table 2 also shows that extraction of the asymmetry of the dynamics parameters is very sensitive to even small amounts of experimental noise.

Because the alignment tensor components from iterative DIDC are essentially the same as those obtained with error-free RDCs (data not shown), the errors in the extracted $\langle b \rangle$ vectors reflect the sensitivity of the extracted dynamics parameters to errors in the RDCs, rather than errors in the extracted alignment tensors. This conclusion is confirmed by using the known true alignment tensors, used to generate the synthetic data, when extracting $(\theta, \phi, S, \eta, \gamma)$ in the presence of noise (Table 2). It is remarkable that even a very small noise amplitude of only 0.2 Hz has a strong adverse impact on the accuracies of the asymmetric motion parameters, η and γ . Since simulated

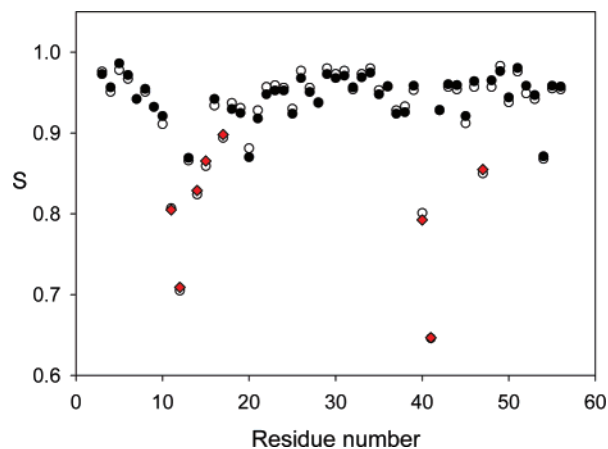


Figure 3. Comparison of the order parameters, S , extracted from synthetic RDCs without noise (filled symbols) and the true S used to generate synthetic RDCs (open symbols). The red diamonds correspond to residues excluded in the calculation of $\langle S \rangle$. The exclusion of residues is based on the fitting error (see text).

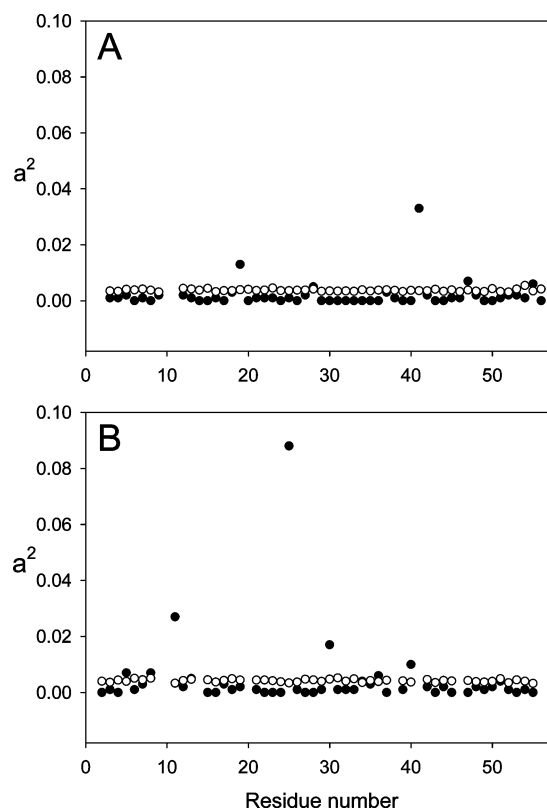


Figure 4. Self-consistency analysis of (A) ^{15}N – ^1H and (B) $^{13}\text{C}^\alpha$ – $^1\text{H}^\alpha$ RDCs in the 6 GB3 mutants. The two panels show the cumulative sum of heterogeneous modes for individual residues.⁵³ Filled symbols are the a^2 for experimental RDCs, while open symbols are the corresponding numbers for simulated RDCs to which 0.3 Hz Gaussian noise has been added. Elevated a^2 values are seen for K19 and G41 ^{15}N – ^1H RDCs, and for C^α – H^α RDCs of T11, T25, F30, and D40, suggesting a measurable impact of the mutations on these bond vectors, which therefore were excluded from dynamics analysis.

η values for most residues are smaller than 0.05, comparable to the error in the extracted values, the precise estimation of residue-specific η and γ angles from RDCs therefore is tenuous for such well-ordered residues, even in the presence of a complete set of five alignments and very small measurement errors.

As defined earlier, $\langle y^2 \rangle \geq \langle x^2 \rangle$, making η positive by definition. However, during the simulation the RDCs noise

TABLE 2: Comparison of rms Errors in the Structure and Dynamics Parameters from Iterative DIDC in the Presence of RDC Errors, with Alignments Obtained Three Different Ways

RDC error (Hz)	orientation(°)		<i>S</i>		anisotropy parameters	
	aniso	iso	aniso	iso	η	γ (°)
0.2 ^a	1.4		0.029		0.03 (0.03 ^b)	52 (24 ^b)
0.2 ^c	1.5	1.7 (1.5 ^d)	0.029	0.030 (0.027 ^d)	0.03 (0.03 ^b)	55 (24 ^b)
0.2 ^e	2.1		0.031		0.07 (0.05 ^b)	58 (34 ^b)
0.4 ^a	2.8		0.059		0.07	62
0.4 ^c	2.9	2.6 (2.4 ^d)	0.055	0.045 (0.043 ^d)	0.07	62
0.4 ^e	4.0		0.057		0.15	63
0.8 ^a	5.5		0.118		0.15	69
0.8 ^c	5.7	4.8 (4.6 ^d)	0.104	0.079 (0.077 ^d)	0.15	69
0.8 ^e	7.2		0.121		0.27	66

^a Perfect alignments are used. ^b When considering residues with $\eta > 0.05$. ^c Alignment from iterative DIDC. ^d When considering residues with $\eta < 0.1$. ^e Alignment from SVD–DIDC.

sometimes causes $\langle x^2 \rangle$ to exceed $\langle y^2 \rangle$, and to keep η positive, the x and y axes need to be switched. This manifests itself in a large jump of the γ angle. As a result, the average η from simulations would be uniformly higher than the true η , and the fluctuation of γ also would be artificially elevated. To solve this problem, the $\langle \mathbf{B} \rangle$ matrices are averaged first over all the simulated data sets, and then the average $(\theta, \phi, S, \eta, \gamma)$ are calculated. The error of each parameter is estimated as $\sqrt{\sum_i^n (\tau_i - \langle \tau \rangle)^2 / n}$, where n is the number of RDCs data sets, $\langle \tau \rangle$ is the average of the parameter τ calculated above, and τ_i is the parameter extracted from the i th RDCs data set. The average error in the γ angle is $\sim 55^\circ$ (Table 2), a rather large number considering that the full range of γ is 180° . When limiting the evaluation to residues with the largest asymmetric motions ($\eta \geq 0.05$), the error in γ decreases to 24° , while the error in η remains similar at 0.034. The accuracy of η and γ is limited by the RDC errors, and not by the iterative DIDC method itself. For example, when 0.2 Hz of RDC noise is present, the errors in S and orientation are about 3 times larger than those caused by the error in $\langle \mathbf{A} \rangle$ that results from the initial assumption of uniform order parameters, while the η and γ errors are about 5 times larger. Thus, for a typical ~ 0.2 Hz experimental measurement error in the N–H RDCs, in practice it will not be possible to derive quantitative information on the asymmetry of the internal motion parameters η and γ when this asymmetry is small ($\eta \leq 0.05$). For errors larger than ca. 0.4 Hz, the asymmetric motion parameters become obscured completely. As can be seen from Table 2, the errors in the extracted structure and dynamics parameters are roughly proportional to the amplitude of RDC errors, regardless of whether perfect or predicted alignment tensors are used to calculate the $\langle \mathbf{B} \rangle$ matrix. Taken together, these observations show that accuracies of the dynamics parameters, but not the alignment parameters, calculated using Scheme 1, are highly sensitive to the errors in simulated RDCs.

Application to Experimental RDC Data from At Least Five Alignments. Before performing the iterative DIDC algorithm on the experimental data, it is essential to evaluate whether the backbone structure and dynamics are the same in all mutants. Differences in structure and/or dynamics could result from the mutations themselves, or from differences in electrostatic interaction between Pf1 and the mutant proteins. An analysis method, SECONDA,^{52,53} is used to identify such residues. Figure 4 shows values of the a^2 parameter as a function of residue number for both N–H and C α –H α interactions, with elevated values reflecting inconsistency resulting either from problems with the experimental measurement, or from a difference in structure or dynamics between the mutants. Elevated a^2 values, above what is expected based on simply

measurement error, are observed for the N–H interactions of K19 and G41, as well as for the C α –H α RDCs of T11, T25, F30, and D40. Four of these six residues are at or immediately adjacent to the sites of mutation; for T25, the $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ nuclei resonate within 1 ppm from one another, violating the weak coupling approximation and compromising the accuracy of the $^{13}\text{C}\alpha$ – $^1\text{H}\alpha$ splitting measurement; we have no explanation for F30 being an outlier. Below, these six residues are excluded from both the N–H and C α –H α analysis of the structure and dynamics.

We first present the results of iterative DIDC for GB3 when using the experimental data obtained for all six mutants as input for deriving $\langle \mathbf{A} \rangle$ and \mathbf{B} matrices. As discussed above, a reasonable starting structure for the initial vector orientations greatly improves convergence of the iterative DIDC procedure, and for this purpose we used the output of an initial (θ, ϕ) search with only RDCs from three mutants as input. Note, however, that for this initial search when using data from three mutants but now in the presence of experimental noise, convergence to within experimental noise when starting from random vector orientations was very similar to what we found for simulated data (ca. 10–15 out of 200 starting structures). Flexible residues were excluded by using the procedure discussed in the Methods section. A total of 42 N–H vectors and 38 C α –H α vectors remain after the last cycle of the iterative DIDC method. After obtaining the $\langle \mathbf{A} \rangle$ matrix, both the three-parameter isotropic and the five-parameter fully anisotropic motion model were used to fit the RDCs. The GB3 NH order parameters, S , extracted from the experimental RDCs (Figure 5A) reflect the isotropic model when a fit within experimental error was obtained, but the fully anisotropic model when no satisfactory fit with just three adjustable parameters exists. As can be seen, elevated dynamics are found for residues 10–21, comprising the first loop and strand $\beta 2$, as well as residues 47 and 48. With the exception of two outliers (G38 and T44), these results are qualitatively consistent with GB3 NH generalized order parameters S_{LS} obtained from relaxation measurements⁵⁶ (Figure 5A). Inspection of Figure 5A shows that significant discrepancies between S_{LS} and S_{RDC} are seen only for residues that require the five-parameter fully asymmetric motion model. It cannot be excluded that several of these (e.g. G38 and T44) result from either outlier RDC measurement errors or small changes in the corresponding vector orientation in one of the mutants, undetected by SECONDA. In this respect, it is important to note that SECONDA will only identify an inconsistency in the set of six RDCs if it cannot be fit with any $(S, \theta, \phi, \eta, \gamma)$ model, regardless of how realistic a given S value may be.

Although the vast majority of residues involved in secondary structure yield similar order parameters in our iterative DIDC

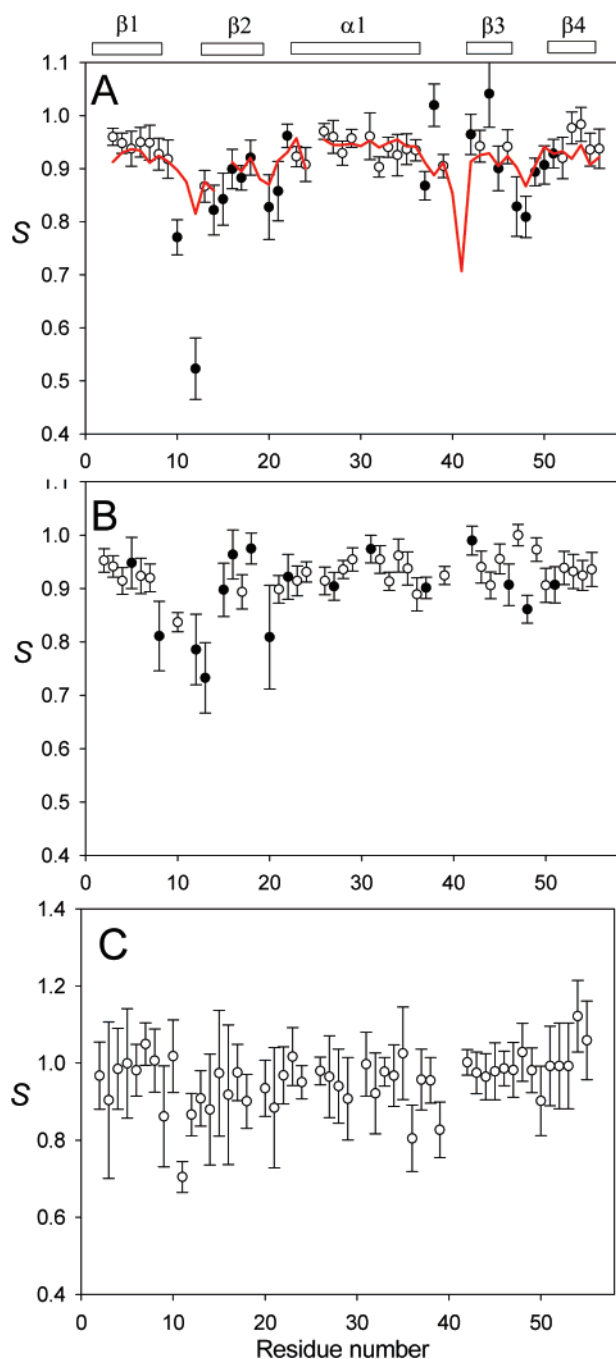


Figure 5. Experimental order parameters, S , of NH, $C^\alpha H^\alpha$, $C' C^\alpha$ bond vectors in GB3 derived from iterative DIDC using all six sets of RDCs. (A) N-H S values. The red line marks the order parameters derived from ^{15}N relaxation,⁵⁶ using $S = \sqrt{S_{\text{LS}}^2}$. (B) $C^\alpha-H^\alpha$ S values. RDCs of residue T11, A19, T25, F30, D40, and G41 are excluded in both panels A and B because of RDC inconsistency, as evaluated by SECONDA analysis (Figure 4). Filled symbols represent residues for which the fully anisotropic model was required to get a satisfactory fit to the data, while, for open symbols, the isotropic internal motion model was able to fit the RDCs to within the experimental noise. The order parameters S are scaled to yield $S = 1$ for the most rigid $C^\alpha-H^\alpha$ vector. The estimated error results from 100 Monte Carlo simulations, where Gaussian distributed noise with rms amplitudes of 0.3 Hz ($^1D_{\text{NH}}$), 0.6 Hz ($^1D_{\text{C}\alpha\text{H}\alpha}$) and 0.15 Hz ($^1D_{\text{C}'\text{C}\alpha}$) has been added to the experimental RDCs. (C) $C'-C^\alpha$ S values derived using the isotropic model, required because of the larger relative uncertainty of the RDCs.

analysis and in the relaxation study, residues found to be more dynamic in the relaxation study, such as G10, L12, A20, V21, D47, and A48, all involved in linkages between elements of

secondary structure, show even larger amplitude dynamics when evaluated by iterative DIDC (Figure 5A). This increased motion indicates the presence of internal dynamics of these residues on a time scale beyond the 4-ns cutoff applicable to the relaxation study of GB3. The order parameters presented in this study are in reasonable agreement with those extracted from seven sets of RDCs by Bouvignies et al.,⁴⁹ but, on average, are more uniform in both the α -helix and the β -sheet, with the exception of “recognition-strand” $\beta 2$. The $C^\alpha-H^\alpha$ order parameters (Figure 5B) confirm the flexibility of the various linkage residues, indicating that elevated motion is not simply restricted to γ -motions of the peptide plane, and point to a more heterogeneous behavior for the $\beta 2$ strand, with K13, E15, and T17 being flexible, but T16 and T18, which have side chains pointing to the interior of the protein, being well ordered. The S values for $C'-C^\alpha$ interactions, pointing more parallel to the main chain direction, do not reveal large variations in amplitude of the dynamics (Figure 5C). However, the small amplitude of $C'-C^\alpha$ RDCs results in fractional errors that are 3–5 times larger than those of the $C^\alpha-H^\alpha$ and N-H RDCs. Consequently, the $C'-C^\alpha$ order parameters show a large uncertainty and variance from residue to residue, possibly obscuring reductions in S caused by elevated dynamics.

A plot of the N-H asymmetry parameter η reveals significant asymmetric motions of flexible residues G10, L12, A20, V21, D47, and A48 (Figure 6A), a result that is consistent with the small S values obtained for these residues (Figure 5A). It is interesting that the $\beta 2$ strand also exhibits anisotropic motion, with the average N-H η being 0.066 ± 0.015 . When considering the average order parameters for the corresponding residues, 0.87 ± 0.04 , this yields fluctuations of $\langle y^2 \rangle \sim 0.072$ and $\langle x^2 \rangle \sim 0.015$. When considering this motion to take place in an ellipsoidal cone, the corresponding angular fluctuation is $\sim 15^\circ$ for the motion along the long (major) axis of the ellipsoid versus 7° for the motion along the short axis.

The orientation of the long axis is defined by the angle γ . However, it is not straightforward to interpret the meaning of γ , because it depends on the choice of the axis system. Therefore, we redefine the orientation of the long axis relative to the unit vector perpendicular to the peptide plane, with $\gamma = 0^\circ$ for out-of-peptide-plane motions, and $\gamma = 90^\circ$ for in-plane motions. As can be seen in Figure 6C, the long axis of the $\beta 2$ -strand asymmetric motion is approximately perpendicular to the peptide plane, pointing to out-of-plane motions that are about 8° larger than the in-plane motional amplitude. However, for the more flexible residues in regions that link secondary structural elements, this pattern does not persist, suggesting that motional processes for such amides are more complicated than simple “rocking” about the $C^\alpha_i-C^\alpha_{i+1}$ vector. The pattern of increased asymmetric motions observed for the N-H bonds of the $\beta 2$ strand is also reflected in elevated η values of the $C^\alpha-H^\alpha$ vectors (Figure 6B).

Validation of Structural Accuracy from $^3J_{\text{HNC}\alpha\text{H}\alpha}$ Couplings. To independently evaluate the accuracy of the vector orientations obtained using iterative DIDC, we evaluate how well they agree with recently reported $^3J_{\text{HNC}\alpha\text{H}\alpha}$ couplings. These 3J couplings were shown to agree considerably better (root-mean-square deviation (rmsd) 0.42 Hz) with a structure where the N-H vectors were allowed to deviate from an idealized in-peptide-plane orientation (PDB entry 2OED) than with the same backbone structure where the H^{N} and H^α protons were positioned in their idealized in-plane and tetrahedral positions, respectively (0.6 Hz).⁶⁰ In the structures where the protons were allowed to deviate from their idealized positions, they were

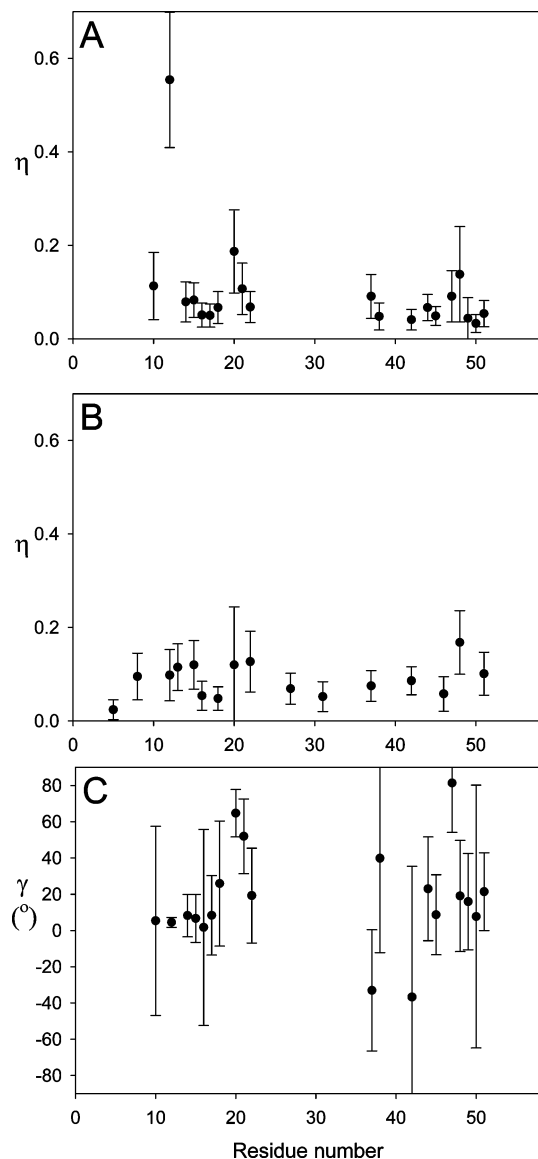


Figure 6. Asymmetric dynamics parameters η of (A) N-H and (B) C α -H α bond vectors, extracted from the six sets of experimental RDCs, using iterative DIDC. (C) Angle between the calculated long axis of the asymmetric motion bond tensor and the unit vector perpendicular to the peptide plane. 0° corresponds to the calculated long axis perpendicular to the peptide plane, while $\pm 90^\circ$ corresponds to the calculated long axis in the peptide plane. The error bars result from 100 simulations, where Gaussian distributed noise with rms magnitude of 0.3 Hz (N-H) and 0.6 Hz (C α -H α) was added to the experimental RDCs. Values are only shown for residues that could not be fit satisfactorily with the isotropic motional model.

nevertheless harmonically restrained by attenuated force constants to prevent them from moving unrealistically far from their idealized positions, a process that also improved their agreement with $^3J_{\text{HNC}\alpha\text{H}\alpha}$ couplings. The N-H and C α -H α vector orientations derived from iterative DIDC are completely unrestrained, and, in principle, could deviate very strongly from their idealized orientations. However, we find this not to be the case. The rmsd between the idealized orientation and that derived from iterative DIDC equals 5.2° for N-H, and 3.6° for C α -H α , and the previously reported $^3J_{\text{HNC}\alpha\text{H}\alpha}$ couplings fit very well to the dihedral angles obtained for these positions (0.39 Hz), confirming the high accuracy of the time-averaged vector orientations obtained with DIDC. Note that incorporating the dynamics information of the DIDC analysis in the Karplus curve fitting

is not straightforward, as DIDC provides no information on whether the N-H and C α -H α vectors move in a correlated, uncorrelated, or anticorrelated manner.

Our previous work⁶⁰ suggested that the effect of dynamics on $^3J_{\text{HNC}\alpha\text{H}\alpha}$ analysis is small, but not completely negligible, as evidenced by a slightly better fit to an ensemble of GB3 structures⁴⁶ (0.36 Hz rmsd) than to the best fitting individual structure (0.42 Hz). This dynamic ensemble of structures had been calculated on the basis of the same RDC data, supplemented by crystallographic B factors and ^{15}N S_{LS}^2 order parameters.⁴⁶ However, in this respect, it is also interesting to note that apparently the uncertainty in the average coordinates of any of these structures is not completely negligible, as reflected by an even slightly better fit (0.35 Hz) to a static set of dihedral angles derived from taking the average of the 2OED- and DIDC-derived dihedral angles.

Application to Experimental RDCs Data from Three Alignments. As demonstrated above for synthetic RDC data, the iterative DIDC algorithm can extract the alignment tensors in the absence of prior structural information, even if RDCs are measured for only three independent alignments. We first demonstrate this by using the experimental N-H RDCs obtained for the K19AD47K, K19ED40N, and K19AT11K mutants to calculate the three alignment tensors. Again, 200 random structures are used as the initial guess for iterative DIDC, and then the resulting structure with the smallest RDC error is selected. The flexible residues were excluded following the procedure described in the Methods section. After obtaining the alignment tensors, the dynamics and structure parameters are extracted. As discussed earlier, since there are three RDCs for each vector, only the orientation (θ , ϕ) and order parameter S can be determined for each bond vector. Initially, S is assumed to be uniform, and only the orientation (θ , ϕ) is extracted. At the second stage, S is allowed to vary if the fitting error when using the two-parameter model (θ , ϕ) is larger than the estimated 0.3 Hz experimental RDC error. We find that 36 out of 49 H-N vectors can be fitted to the two-parameter model, while the other 13 require fitting to all three parameters (S , θ , ϕ). Figure 7A shows the angle between the N-H vector extracted from just three RDCs in the above-described manner and the orientation obtained when using all six RDCs. As can be seen, with a pairwise average difference of 3.6°, the two sets of N-H orientations generally agree well with each other. However, several flexible residues, including K10, L12, D47, and A48, show larger discrepancies, which may in part result from the use of an isotropic motional model, whereas the true motions of these four residues appear quite asymmetric. As can be seen in Figure 7A for V39 and E42, and as was discussed earlier, some vector orientations are much more sensitive than others to small changes in the RDCs, resulting in substantial uncertainties in the extracted vector orientations when Monte Carlo noise is added to the input RDCs.

Figure 7A also shows that the uncertainty in the extracted vector orientation on average is considerably higher for the three-parameter fits than for the two-parameter fits, with $S = 1$. This increased structural uncertainty is caused by the fact that three observables are used to extract three parameters; i.e., the solution is no longer overdetermined and becomes quite sensitive to small errors in the input data or imperfections in the approximation of the motional model as being axially symmetric. Therefore it is generally preferable to use the two-parameter fit whenever possible.

The dynamics of the protein is revealed by the plot of S versus residue number (Figure 7B). Increased flexibility of K10, L12,

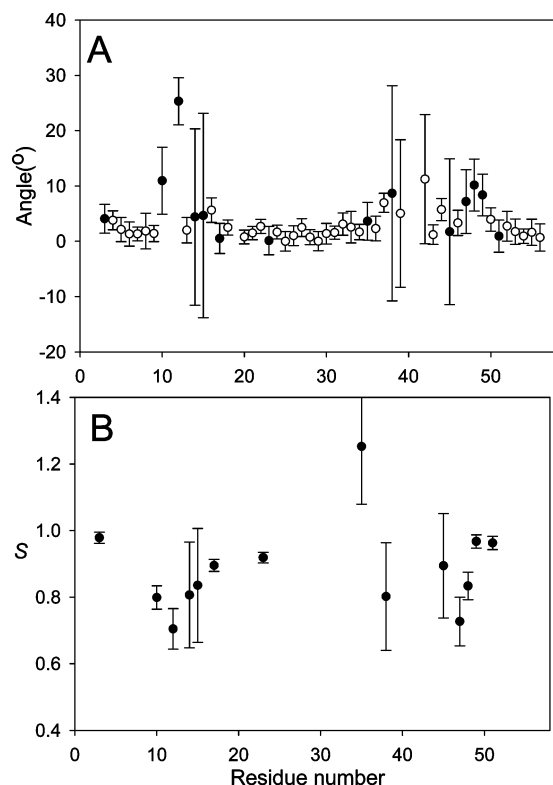


Figure 7. The structure and dynamics parameters extracted from just three sets of ^{15}N - ^1H RDCs, measured for K19AD47K, K19ED40N, and K19AT11K. (A) The average angles between vectors extracted from three and six sets of RDCs. Open symbols represent residues fit with $S = 1$, while the filled circle represents residues which required adjustment of S to obtain a satisfactory fit. The result of the six-RDC fit is used as a reference, as the vector orientations resulting from this analysis resulted in the best fit to $^3J_{\text{HNH}\alpha}$ couplings of any GB3 structure to date (see Concluding Remarks section). The average angle difference over all residues is 3.6° . (B) S values extracted from the three sets of RDCs, using the isotropic model. Error bars for both panels A and B result from 200 Monte Carlo simulations, with 0.3 Hz Gaussian noise added to the experimental RDCs.

strand β_2 , D47, and A48 is clearly evident, consistent with results obtained from six RDCs. However, the flexibility of A20 and V21 are not identified in the fitting. On the other hand, the amide of residue N35 now shows a non-physical $S = 1.25$, and its very large error margin indicates that S is poorly defined by the RDCs. In contrast, the error in orientation of this amide is only 3.6° , suggesting that S is more sensitive to RDC errors than is the structure.

Concluding Remarks

In this study, we introduced an iterative method for DIDC in terms of both structure and dynamics. When applying the method to the small protein GB3, the iterative DIDC provides a considerably more accurate determination of alignment tensors without reference to prior molecular structure information than does the direct, SVD-based DIDC method. The improved definition of the alignment tensors obtainable with iterative DIDC also results in more accurate structure and order parameters, and allows extraction of reasonably accurate asymmetric dynamics parameters (Table 1).

Evaluation of the iterative DIDC method using simulated N-H RDC data from five independent alignments shows that the alignment tensors obtained from iterative DIDC reproduce the dynamics and structure parameters as well as those from the true, error-free alignment tensors, even for RDC errors as

large as 0.8 Hz (for $D_a \approx 8$ Hz), confirming the robustness of this method (Table 2). The simulations also indicate that the errors in both the structure and dynamics parameters are proportional to the RDC errors, and that, for RDC errors larger than ca. 0.2 Hz, accurate values of the asymmetric motion parameters η and γ can no longer be determined when $\eta \leq \sim 0.05$. This simply is a consequence of the fact that asymmetry of internal motions tends to impact RDCs only very little (unless internal motion amplitude is large), causing a change in the RDC that tends to be less than the RDC error. However, under the assumption that asymmetric motion is very similar within a cluster of residues of a given secondary structure, α -helix, or β -sheet, it remains feasible to obtain characterization of the average behavior of such asymmetric motion in such regions because the error of the average value is inversely proportional to the square root of the number of vectors.

For cases where the intrinsic uncertainty in the experimental RDCs is larger, either caused by less precise RDCs or by error propagation when the five alignments are insufficiently orthogonal to one another, it becomes more realistic to simply employ the isotropic motion model for RDCs, which is less sensitive to experimental error and then provides more accurate structure and dynamics parameters than the fully asymmetric model (Table 2).

The N-H order parameters are in reasonable agreement with the order parameters S_{LS}^2 , derived from ^{15}N relaxation measurements (Figure 5). However, the linker residues between various elements of secondary structure, including G10, L12, A20, V21, D47, and A48 exhibit larger angular excursions than those observed in the NMR relaxation study, indicating that the time scale of these motions extends beyond the limit set by the rotational correlation time (~ 4 ns), where these motions impact ^{15}N relaxation. All these linker residues display asymmetric motions of NH vectors, but the orientations of the axes of motion relative to the peptide plane varies from residue to residue, suggesting the asymmetric motion of these residues is more complex than simple rocking of the peptide planes.

Our results also show elevated, asymmetric motion for the N-H vectors of strand β_2 , with the larger amplitude motion being orthogonal to the chain direction, and out-of-plane motion being larger than in-plane motions by about 8° . Interestingly, strand β_2 is also the site where GB3 is destined to interact with the IgG antibody,⁴³ a point also noted by Bouvignies et al.⁴⁹ Presumably, the flexibility of β_2 allows the residues to explore the conformational space to adapt and bind to IgG in an optimal fashion. Indeed, the asymmetric motion of the β_2 strand is consistent with a recent dynamics study by Bouvignies et al. also based on RDC analysis.⁴⁹ However, our study does not find evidence for the pattern of alternating large and small amplitude motions observed in β -sheets by these workers and interpreted as evidence for collective motion of hydrogen-bonded amide groups. Instead, our study suggests that, with the exception of edge strand β_2 , residues in the GB3 β -sheet appear to be rather uniform in their dynamic characteristics. The C^α - H^α order parameters confirm the presence of elevated dynamics for residues G10, L12, A20, V21, and D47, but display more diverse S values in the β_2 strand (Figure 5). The anisotropic motion analysis of C^α - H^α RDCs suggests that the majority of residues with N-H asymmetric motion also tend to have C^α - H^α asymmetric motion.

Application of iterative DIDC to six sets of experimental N-H, C^α - H^α , and C' - C^α RDCs for GB3 allowed determination of the structure and dynamics parameters of these vectors. The resulting vector orientations are in excellent agreement

(rmsd 3.9°) with the previously determined NMR structure (PDB entry 2OED),⁴⁷ and, with an rmsd of 0.39 Hz, the $^3J_{\text{HNCaH}\alpha}$ couplings agree very well with the DIDC-derived H–N–C α –H α dihedral angles. So, even while the main purpose of the iterative DIDC method is to provide a model-free interpretation of RDCs in terms of dynamics, it also proves to be remarkably robust in terms of defining structural information. Besides confirming that RDCs indeed can be exploited to extract dynamics information, our results also show that ignoring the effect of dynamics when interpreting RDCs in terms of a single, time-averaged structure does not lead to significant systematic errors for a well-structured, globular protein such as GB3. Considering that such a time-averaged structure already can provide an extremely good fit to the experimental RDCs ($R^2 > 0.99$), the effect of internal dynamics on RDCs, neglected in such a fit, can be viewed as second-order. Quantitative extraction of the amplitude and orientation of the internal dynamics from the very small discrepancy between experimental RDCs and a single, static model requires highly accurate experimental RDCs. Analysis of dynamics from RDCs can benefit from reasonable assumptions, such as the concerted movement of all RDC vectors associated with an entire peptide plane (C'–C α , C'–N, and N–H)⁴⁹ or, as first proposed by Tolman and demonstrated by our results, can also be carried out by DIDC in a model-free manner.

An appealing feature of the iterative DIDC is that it permits determination of alignment tensors from just three independent sets of RDCs, without recourse to prior structural information (Figure 1). This significantly reduces the difficulty in generating sufficient independent alignments, required for the SVD-based DIDC method. Our results confirm that it is possible to simultaneously extract both structure and dynamics with reasonable accuracy.

Acknowledgment. This work was supported in part by the Intramural Research Program of the NIDDK, NIH, and by the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH.

Supporting Information Available: Five tables with the parameters used for generating synthetic RDCs and order parameters derived from the RDCs, and a detailed description of the iterative DIDC algorithm, and a figure showing χ^2 trajectories during Scheme 1. This material is available free of charge via the Internet at <http://pubs.acs.org>. The MATLAB program is available for downloading at <http://spin.niddk.nih.gov/bax/software/>.

References and Notes

- Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
- Clore, G. M.; Gronenborn, A. M. *Curr. Opin. Chem. Biol.* **1998**, 2, 564–570.
- Ferentz, A. E.; Wagner, G. *Q. Rev. Biophys.* **2000**, 33, 29–65.
- Wagner, G. *Q. Rev. Biophys.* **1983**, 16, 1–57.
- Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol., Part B* **2001**, 339, 204–238.
- Bruschweiler, R. *Curr. Opin. Struct. Biol.* **2003**, 13, 175–183.
- Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, 104, 4546–4559.
- Akke, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1996**, 118, 911–912.
- Mulder, F. A. A.; Skrynnikov, N. R.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, 123, 967–975.
- Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Bruschweiler, R. *J. Am. Chem. Soc.* **2001**, 123, 6098–6107.
- Tolman, J. R.; Al-Hashimi, H. M.; Kay, L. E.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, 123, 1416–1424.
- Bax, A. *Protein Sci.* **2003**, 12, 1–16.
- Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. *Chem. Rev.* **2004**, 104, 3519–3540.
- Blackledge, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, 46, 23–61.
- Tjandra, N.; Bax, A. *Science* **1997**, 278, 1111–1114.
- Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, 120, 10571–10572.
- Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, 5, 1065–1074.
- Ruckert, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, 122, 7793–7797.
- Tycko, R.; Blanco, F. J.; Ishii, Y. *J. Am. Chem. Soc.* **2000**, 122, 9340–9341.
- Meier, S.; Haussinger, D.; Grzesiek, S. *J. Biomol. NMR* **2002**, 24, 351–356.
- Sass, H.-J.; Musco, G.; Stahl, S. J.; Wingfield, P. T.; Grzesiek, S. *J. Biomol. NMR* **2000**, 18, 303–309.
- Ma, C.; Opella, S. J. *J. Magn. Reson.* **2000**, 146, 381–384.
- Rodriguez-Castaneda, F.; Haber, P.; Leonov, A.; Griesinger, C. *Magn. Reson. Chem.* **2006**, 44, S10–S16.
- Bertini, I.; Del Bianco, C.; Gelis, I.; Katsaros, N.; Luchinat, C.; Parigi, G.; Peana, M.; Provenzani, A.; Zoroddu, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 6841–6846.
- Wohnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. *J. Am. Chem. Soc.* **2003**, 125, 13338–13339.
- Gaponenko, V. V.; Walsey, C. J.; Hoffman, B. M.; Rosevear, P. R. *Biophys. J.* **2000**, 78, 288A.
- Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, 120, 9106–9107.
- Al-Hashimi, H. M.; Valafar, H.; Terrell, M.; Zartler, E. R.; Eidsness, M. K.; Prestegard, J. H. *J. Magn. Reson.* **2000**, 143, 402–406.
- Tolman, J. R.; Ruan, K. *Chem. Rev.* **2006**, 106, 1720–1736.
- Tolman, J. R. *J. Am. Chem. Soc.* **2002**, 124, 12020–12030.
- Peti, W.; Meiler, J.; Bruschweiler, R.; Griesinger, C. *J. Am. Chem. Soc.* **2002**, 124, 5822–5833.
- Zweckstetter, M.; Hummer, G.; Bax, A. *Biophys. J.* **2004**, 86, 3444–3460.
- Ruan, K.; Tolman, J. R. *J. Am. Chem. Soc.* **2005**, 127, 15032–15033.
- Lakomek, N. A.; Carlomagno, T.; Becker, S.; Griesinger, C.; Meiler, J. *J. Biomol. NMR* **2006**, 34, 101–115.
- Yao, L.; Bax, A. *J. Am. Chem. Soc.* **2007**, 129, 11326–11327.
- Losonczi, J. A.; Andrec, M.; Fischer, M. W. F.; Prestegard, J. H. *J. Magn. Reson.* **1999**, 138, 334–342.
- Sass, J.; Cordier, F.; Hoffmann, A.; Rogowski, M.; Cousin, A.; Omichinski, J. G.; Lowen, H.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, 121, 2047–2055.
- Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1998**, 133, 216–221.
- Bryce, D. L.; Bax, A. *J. Biomol. NMR* **2004**, 28, 273–287.
- Warren, J. J.; Moore, P. B. *J. Magn. Reson.* **2001**, 149, 271–275.
- Bouvignies, G.; Markwick, P.; Bruschweiler, R.; Blackledge, M. *J. Am. Chem. Soc.* **2006**, 128, 15100–15101.
- Briggman, K. B.; Tolman, J. R. *J. Am. Chem. Soc.* **2003**, 125, 10164–10165.
- Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, 243, 906–918.
- Gronenborn, A. M.; Filpula, D. R.; Essiz, M. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, 253, 657–661.
- Stone, M. J.; Gupta, S.; Snyder, N.; Regan, L. *J. Am. Chem. Soc.* **2001**, 123, 185–186.
- Clore, G. M.; Schwieters, C. D. *J. Mol. Biol.* **2006**, 355, 879–886.
- Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, 125, 9179–9191.
- Meier, S.; Haussinger, D.; Jensen, P.; Rogowski, M.; Grzesiek, S. *J. Am. Chem. Soc.* **2003**, 125, 44–45.
- 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.
- Lienin, S. F.; Bremi, T.; Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1998**, 120, 9870–9879.
- Bax, A.; Kontaxis, G.; Tjandra, N. *Meth. Enzymol.* **2001**, 339, 127–174.
- Hus, J. C.; Bruschweiler, R. *J. Biomol. NMR* **2002**, 24, 123–132.
- Hus, J. C.; Peti, W.; Griesinger, C.; Bruschweiler, R. *J. Am. Chem. Soc.* **2003**, 125, 5596–5597.
- Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, 21, 1087–1092.
- Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, 6, 277–293.
- Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, 27, 261–275.
- Engh, R. A.; Huber, R. *Acta Crystallogr., Sect. A* **1991**, 47, 392–400.
- Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, 120, 12334–12341.
- Case, D. A. *J. Biomol. NMR* **1999**, 15, 95–102.
- Vogeli, B.; Ying, J. F.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2007**, 129, 9377–9385.