Characterization of protein dynamics from residual dipolar couplings using the three dimensional Gaussian axial fluctuation model

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · APRIL 2008

Impact Factor: 2.63 · DOI: 10.1002/prot.21703 · Source: PubMed

3 AUTHORS, INCLUDING:

Guillaume Bouvignies
Ecole Normale Supérieure de Paris
35 PUBLICATIONS **975** CITATIONS

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Characterization of protein dynamics from residual dipolar couplings using the three dimensional Gaussian axial fluctuation model

Guillaume Bouvignies, Phineus R. L. Markwick, and Martin Blackledge*

Institute de Biologie Structurale Jean-Pierre Ebel, CNRS; CEA; UJF; UMR 5075, 41 rue Jules Horowitz, F-38027 Grenoble, Cedex, France

ABSTRACT

Residual dipolar couplings are potentially very powerful probes of slower protein motions, providing access to dynamic events occurring on functionally important timescales up to the millisecond. One recent approach uses the three dimensional Gaussian Axial Fluctuation model (3D GAF) to determine the major directional modes and associated amplitudes of motions along the peptide chain. In this study we have used standard and accelerated molecular dynamics simulations to determine the accuracy of 3D GAF-based approaches in characterizing the nature and extent of local molecular motions. We compare modes determined directly from the trajectories with motional parameterization derived from RDCs simulated from the same trajectories. Three approaches are tested, that either suppose a known three-dimensional structure, simultaneously determine backbone structure and dynamics, or determine dynamic modes in the absence of a structural model. The results demonstrate the robustness of the 3D GAF analysis even in the presence of large-scale motions, and illustrate the remarkably quantitative nature of the extracted amplitudes. These observations suggest that the approach can be generally used for the study of functionally interesting biomolecular motions.

Proteins 2008; 71:353–363.

Key words: NMR; residual dipolar couplings; dynamics; molecular dynamics; slow motions; protein.

INTRODUCTION

Nuclear magnetic resonance (NMR) is now established as a key technique for the study of protein motions on a vast range of timescales.[1] Fast motions on picosecond to nanosecond timescales can be characterized using spin relaxation measurements,[2–4] while relaxation dispersion experiments are sensitive to chemical shift exchange effects occurring over the micro- to millisecond range, a timescale of particular interest because biological activity is expected to occur in this range.[5–7] Residual dipolar couplings (RDCs) measured under conditions of partial molecular alignment[8,9] can also be used to probe these slower molecular motions. RDCs are averages over all orientations of the magnetic dipolar interaction vector that are sampled up to the inverse of the alignment-induced coupling, thus reporting on averages up to the millisecond range.[10,11] These parameters are therefore highly complementary to the dynamic information derived from spin relaxation measurements. Comparison of motional averaging on the two time-scales can in principle provide information on dynamics in the nano- to millisecond range. This ability of RDCs to encode the details of local conformational fluctuation has been exploited by a number of groups in recent years to explore the nature of local dynamics in proteins over this time range.[12–24]

Methods exploiting specific geometric models to describe local motion have recently been developed. Initial studies, using a one-dimensional Gaussian Axial Fluctuation (GAF) model for peptide plane reorientation about the C^{α}_{i-1}—C^{α}_{i} axis, demonstrated that a common anisotropic component of protein backbone dynamics could be extracted from ^{15}N—^{1}H RDCs, and that this allowed a significantly more accurate description of the overall molecular alignment tensor.[18] Following on from this the presence of slow motions were investigated in protein GB3, by interpreting an extensive set of RDCs using the three-dimensional Gaussian Axial Fluctuation model (3D-GAF), a model of peptide plane dynamics allowing for stochastic motions around three orthogonal axes attached to the peptide plane,[25] assuming a fixed time- and ensemble-averaged model of the average structure. The study delivered a site-specific anisotropic motional description of each peptide plane in the protein, and provided a quantita-

Grant sponsor: EU; Grant number: EU-NMR JRA3; Grant sponsor: French Research Ministry; Grant number: ANR NT05-4_42781; Grant sponsor: CEA.
Guillaume Bouvignies and Phineus R. L. Markwick contributed equally to this work.
*Correspondence to: Martin Blackledge; Institute de Biologie Structurale Jean-Pierre Ebel, CNRS; CEA; UJF; UMR 5075, 41 rue Jules Horowitz, F-38027 Grenoble, Cedex, France. E-mail: martin.blackledge@ibs.fr
Received 27 April 2007; Revised 8 June 2007; Accepted 27 June 2007
Published online 23 October 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21703

tive estimate of the nature and extent of dynamics present on the protein backbone, identifying a heterogeneous distribution of slower motions in the protein in comparison to ^{15}N spin relaxation data.[24] Further progress was then achieved by the development of *Dynamic-Meccano*, an approach that allows a dynamically averaged conformation to be determined simultaneously with a description of the principal dynamic component of each peptide plane from RDCs alone, thereby alleviating the need for any *a priori* knowledge of the average solution conformation.[26] The average conformation showed very strong similarity to the RDC-refined crystal structure, suggesting that the shallow base of the conformational energy landscape of this particular protein was essentially harmonic. The aim of the current study is to determine the level of accuracy that such GAF-based approaches can be expected to deliver under conditions of the potentially large amplitude motions occurring over longer (up to the millisecond) timescales.

One source of possible artifact has its origin in the potential complexity of slower motions, for which the 3D GAF model may be an inappropriate model. This model was initially inspired from classical MD simulation lasting 12 ns, and applied to the characterization of motions occurring over this time range in a small cyclic polypeptide.[25] It is not clear how the model responds to motional modes of more complex nature that may occur on a longer time scale, and whether extracted amplitudes or order parameters would be accurately determined when the model is applied under these conditions. Another question we would like to address is whether dynamic amplitudes can be quantitatively estimated from residual dipolar couplings. Care was taken in the initial study to avoid absorbing motion into the alignment tensor determination, an effect that would have the direct result of reducing the detected motional amplitudes. To minimize these effects, vectors corresponding to dipolar couplings that were identified as being the least dynamically averaged in an initial analysis were used to determine the alignment tensors, and all dynamic amplitudes determined relative to these tensors. Using the *Dynamic-Meccano* approach, alignment tensors are estimated ab initio, in the absence of information about the fold of the protein. The accuracy of these approaches is also addressed here.

Molecular dynamics (MD) simulation enables the exploration of conformational energy landscapes accessible to protein molecules[27,28] and thereby provides invaluable tools for aiding our understanding and interpretation of experimental NMR data in terms of molecular motion.[4,29–31] In the current work we use MD simulations, including accelerated molecular dynamics (AMD),[32] an approach that extends the time scale available to molecular simulation by applying an additional bias potential which raises and flattens the complex potential energy landscape, thereby enhancing the escape rate from one potential energy basin (sub-state) to the next. Sampling from MD and AMD is used to investigate the behavior of the 3D GAF RDC analysis, and to test the capacity of this approach to detect motions of differing amplitude and geometry. Order parameters from AMD simulations have recently been compared to order parameters derived from the 3D GAF analysis of RDCs, and found to reproduce these order parameters very closely.[33] The advantage of MD simulation is that dynamic amplitudes and modes can be calculated directly from the vector correlation functions and compared with amplitudes extracted from a 3D GAF analysis.

METHODS

MD simulation

The X-ray structure of the third IgG binding domain in Streptococcal protein G (GB3, PDB code 1IGD) was placed in a periodically repeating box with 15,000 water molecules and two Na^+ counter ions. The system was brought to thermodynamic equilibrium at 300 K, 1 bar pressure using a weak coupling thermostat with a different random seed generator, before performing a 2 ns MD simulation.

In all simulations performed, bonds involving protons were constrained using the SHAKE algorithm[34] and a time-step of 2 fs was employed. All simulations were performed under periodic boundary conditions using weak coupling temperature and pressure conditions. Electrostatic interactions were treated using the Particle Mesh Ewald (PME)[35] method with a direct space sum limit of 10 Å. The TIP3P water force-field[36] was employed for solvent molecules and the newly developed ff99SB force field[37] was used for solute residues. All simulations were performed using a modified in-house version of the AMBER8 code.[38]

Accelerated MD simulation

The details of the application of accelerated MD (AMD) have been discussed previously in the literature, and a full description of the production of the trajectory analyzed here has recently been published.[33] Briefly the trajectory resulted from 10 ns of AMD applied using a boost energy set to 300 kcal mol^{-1} above the dihedral angle energy and an α parameter of 60 kcal mol^{-1}.

Simulation of dipolar coupling data

Two sets of 1000 structures representing sampling from the classical MD simulation and the accelerated MD were used in the analysis. For each structure N—H, $^{\alpha}C—C'$ and N—C′ one-bond dipolar couplings were simulated from each peptide plane. To reproduce experimental conditions as closely as possible the five alignment tensors determined from the experimental data

were imposed.[24] RDCs were then averaged over the 1000 structures and these data treated to determine $S^2_{RDC,NH}$ and σ_α, σ_β, and σ_γ (vide infra). Noise levels equal to those estimated experimentally were added using a randomly sampled Gaussian distribution centered on these simulated data (0.26, 0.1, and 0.1 Hz for the N—H, $^{\alpha}$C—C′, and N—C′ one-bond dipolar couplings, respectively).

Calculation of the order parameters

The internal dynamics present in the two simulations of GB3 were also characterized by directly calculating order parameters ($S^2_{MD,NH}$). Following superposition of all the structures representing the MD trajectories onto the heavy backbone atoms of all residues 6–61 of the average structures, the order parameters were calculated from:[39]

$$S^2 = \frac{1}{2}\left[3\sum_{i=1}^{3}\sum_{j=1}^{3}\left\langle \mu_i\mu_j\right\rangle^2 - 1\right] \qquad (1)$$

where m_i are the Cartesian coordinates of the normalized inter-nuclear vector of interest. We note that while the 1000 structures extracted across the AMD simulation give a representation of the range of variation of the order parameters extracted from a full AMD analysis, the numerical result is not exactly the same as that previously published due to the very small number of structures extracted in comparison to the previous study.[33]

Calculation of the 3D GAF amplitudes from MD trajectories

3D GAF amplitudes were extracted from the MD trajectories using previously described procedures.[25]

Determination of dynamic parameters from RDCs: 3D GAF analysis

3D GAF dynamic parameters were calculated using previously described procedures and some recently developed modifications. Three scenarios were adopted to determine the motional amplitudes from the simulated RDCs. In each case we assume that the five alignment tensors are unknown.

RDCs are described with respect to the 3D GAF motional amplitudes σ_α, σ_β, and σ_γ using Eqs. (2) and (3) in Ref. 24. Gaussian widths σ_α, σ_β, and σ_γ are then determined with respect to the simulated data.

In Scenario I we assume that the energy-minimized mean structure from the AMD simulation represents the average conformation in solution, and the amplitudes are determined relative to the orientation of the peptide planes as described above. The alignment tensors are initially determined by fitting all couplings to the X-ray conformation using the program Module. The motional amplitudes are determined relative to these tensors, and the least-mobile vectors are then used to determine the alignment tensors and the entire procedure repeated relative to these tensors.

In Scenario II we determine the structure of the average conformation using the Dynamic-Meccano approach.[26] The first step involves the elucidation of the components of all five tensors. This entails the determination of 22 independent parameters defining the tensors (five values of D_a and D_r, and the relative orientation of the tensors $\{\alpha, \beta, \gamma\}_i$), and four parameters defining the orientation of each considered plane $(\theta,\phi,\psi)_i$ and its associated dynamic amplitude. A total of 238 parameters are therefore optimized in this step. The most appropriate model from three orthogonal 1D GAF[40] models of peptide plane reorientation or a common scaling factor was used.

The effect of a 1D-GAF motion on a residual dipolar coupling can be analytically expressed as follows:

$$\left\langle D_{ij}\right\rangle_{1D-GAF} = -\frac{\mu_0}{8\pi^2}\frac{\gamma_i\gamma_j\hbar}{r_{ij}^3}\left\{\begin{array}{l}\frac{1}{4}D_a\left[s_1\left(3\cos^2\beta-1\right)+3\,s_2\sin 2\,\beta\cos\alpha+3\,s_3\sin^2\,\beta\cos 2\alpha\right]+\\ \\ \frac{3}{8}D_r\left[\begin{array}{l} s_1\sin^2\,\beta\cos 2\gamma-\\ 2s_2\sin\beta\left(\cos(\alpha+2\gamma)\cos^2\frac{\beta}{2}-\cos(\alpha-2\gamma)\sin^2\frac{\beta}{2}\right)+\\ 2s_3\left(\cos(2\alpha+2\gamma)\cos^4\frac{\beta}{2}+\cos(2\alpha-2\gamma)\sin^4\frac{\beta}{2}\right)\end{array}\right]\end{array}\right\}$$

with $s_1 = 2(3\cos^2\theta-1)$; $s_2 = 2\sin 2\theta\exp\left(-\frac{\sigma^2}{2}\right)$, $s_3 = 2\sin^2\theta\exp(-2\sigma^2)$, where r_{ij} is the internuclear distance, σ the motional amplitude, $\{\alpha, \beta, \gamma\}$ the Euler angles transforming from the alignment tensor principal axis system into the local peptide plane frame defined by its z-axis along the rotational axis and its x-axis so that the internuclear vector lies in the x-z-plane and θ the angle between the rotational axis and the internuclear vector.

The backbone coordinates of the protein are constructed by positioning peptide planes of fixed-internal

geometry and intervening tetrahedral junctions to best reproduce the experimental data. Plane orientation is accompanied by optimisation of a local motional amplitude, the most appropriate selected from the dynamic modes described above (amplitudes σ_α, σ_β, σ_γ of the GAF motions, or an order parameter *S*, of an axially symmetric motion). Finally a complete 3D GAF analysis of the dynamic disorder present along the chain is applied using the resulting structure, as described previously.

In Scenario III the only conformational information results from determination of the optimal orientation of each peptide plane, using the most appropriate dynamic model from the three 1D GAF models and the axially symmetric motion, as in the initial step of Scenario II. A complete 3D GAF analysis of the dynamic disorder present along the chain is again applied, the only difference here being that the peptide conformation is not constructed. This removes any dependence of dynamic parameters and peptide orientation, on the orientation of the previous plane that may be introduced due to the inclusion of the tetrahedral junction in the *Dynamic Meccano* molecular construction algorithm.[26]

RESULTS

To enable a bias-free analysis of potential artifacts we have employed MD calculations to simulate different modes and amplitudes of protein dynamics and subsequently to test the ability of GAF-based approaches to reproduce these motions accurately. This was achieved by employing either a short (2 ns) classical MD simulation, that essentially samples fast modes associated with subnanosecond dynamics, or accelerated MD simulation. Accelerated MD simulation enhances higher activation energy transition probabilities between conformational sub-states, thereby emulating larger-scale motions that would be expected to occur on the timescales relevant to RDC averaging. The availability of such methodologies allows us to address questions raised concerning the ability of 3D GAF analyses to extract meaningful amplitudes and modes from experimental RDCs.

A detailed analysis of the AMD approach applied to protein GB3 has recently been published,[33] so we restrict ourselves here to the observation that the AMD ensemble indeed samples more conformational space than the MD ensemble. Figure 1 shows the root mean square deviations of the two 1000-strong ensembles from their means. These values are found to be (0.54 ± 0.09)Å and (0.76 ± 0.16)Å for the standard MD and AMD ensembles, respectively. RDCs were simulated from the AMD and MD ensembles using alignment tensors that have been determined from experimental measurements made in five different alignment media,[16] and the 3D GAF analysis used to determine local dynamic amplitudes. Using this approach we implicitly assume that the overall shape of the molecule does not change as a function of

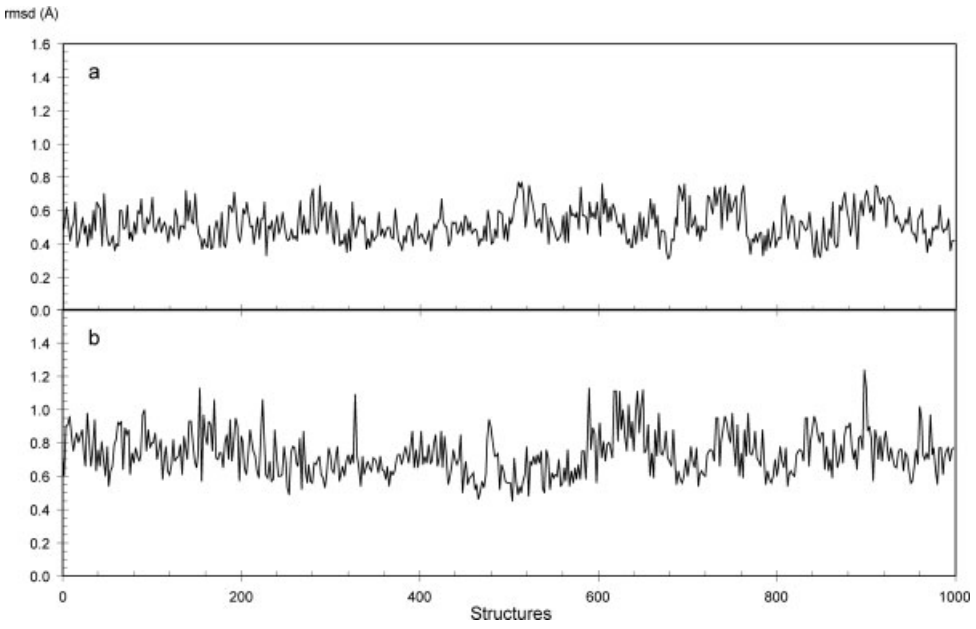

Figure 1

*Conformational Sampling of MD and AMD Ensembles. Root mean square deviation of the 1000 structures comprising the MD (**a**) and accelerated MD (**b**) ensembles. Lines show rmsd of backbone atoms (N, C^α, C') for amino acids 8–60 with respect to their means.*

local dynamics, and that contributions to measured RDCs from local dynamics and molecular alignment can be dissociated.

Alignment tensor determination

Accurate determination of the local motional modes and amplitudes present in the protein from RDCs alone requires a reasonable estimate of the molecular alignment tensor, normally entirely unknown at the start of the analysis. The alignment tensor is treated in two different ways depending on the scenario.

Scenario I

The first of these, Scenario I, uses a structural model to represent the average conformation in solution. In the case of this simulation the structure is represented by the average conformation from the MD trajectories. In practice an experimental technique is assumed to be available (X-ray crystallography or NMR) that can give us a reasonable model of this conformation. The alignment tensor is determined by fitting the vectors and RDCs presenting the least motion in combination with the static structure. Here the $C'—C^{\alpha}$ vectors with the lowest motional amplitudes were used.

Scenarios II and III

In the second and third scenarios the tensor is determined in the absence of structural information concerning the protein fold. A minimization is performed in which the eigenvalues of all five tensors, the orientation of each peptide plane, the type of dynamic model and the amplitude parameter associated with it, are simultaneously optimized (see Methods). As previously described[26] the motional amplitudes (σ_α, σ_β, and σ_γ) are all independently determined using Eq. (2) by minimizing a target function containing all RDCs associated with the plane in question. The dynamic parameter retained in the final target function corresponds to the amplitude of the best fitting motional mode on the basis of the fitting function χ^2, whether this is $\{\sigma_\alpha, \sigma_\beta, \text{or } \sigma_\gamma\}$ or a scaling factor for all vectors in the plane. RDC averaging under the latter motion requires only scaling of a single order parameter S common to all RDCs in the plane, while the three anisotropic models affect differently oriented vectors in the plane in different way. The peptide plane conformation is assumed to be the same throughout the molecule, and defined from a database of high-resolution crystal structures.

The resulting alignment tensors determined in each case are compared in Table I for RDCs simulated from AMD in the presence of experimental levels of random noise. Eigenvalues A_a and A_r, and the relative orientation of the tensors $\{\alpha, \beta, \gamma\}_i$ are shown for each tensor, and compared to the true tensors used to simulate the data. Eigenvalues are also shown when all simulated couplings are used to analyze the alignment tensors and no account is taken of dynamic averaging ("static"). Not surprisingly the "static" tensors are found to be systematically lower (by a factor of 0.9) than the true tensors. This is in line with previous observations, that a certain level of common motion will inevitably be absorbed into the tensor eigenvalues if it is

Table I
Alignment Tensor Determination Using Scenarios I–III

	Aa (10^{-4})	Ar (10^{-4})	α (°)	β (°)	γ (°)
Tensor 1 (true)[a]	13.6	1.7	−75.3	88.9	−35.1
Static[b]	12.3	1.6	−78.4	89.0	−34.8
SI[c]	13.2	1.7	−76.1	89.0	−35.1
SII/SIII[d]	13.3	1.6	−75.3	88.9	−35.1
Tensor 2 (true)	9.5	6.0	−52.6	73.0	54.1
Static	8.4	5.4	−54.7	73.4	54.5
SI	9.2	5.8	−52.8	72.7	54.0
SII/SIII	9.2	6.0	−53.1	70.8	53.9
Tensor 3 (true)	11.0	4.7	105.4	102.6	122.7
Static	9.9	4.6	104.2	101.8	122.4
SI	10.5	4.7	104.6	102.9	122.8
SII/SIII	10.8	4.8	104.3	102.1	122.9
Tensor 4 (true)	−8.4	−1.9	102.0	82.1	172.3
Static	−7.8	−1.7	101.7	82.8	172.6
SI	−8.1	−1.8	104.6	81.6	172.3
SII/SIII	−8.2	−1.7	102.8	82.6	172.7
Tensor 5 (true)	−15.0	−3.5	−87.7	122.7	−10.4
Static	−13.3	−3.3	−87.9	121.1	−10.1
SI	−14.5	−3.3	−88.6	122.9	−10.5
SII/SIII	−14.7	−3.1	−85.3	121.7	−9.1

[a]Tensors used to simulate the RDCs.
[b]Tensors determined using all couplings, and taking no account of dynamic averaging (static model).
[c]Tensors determined using coordinates from the structural model (Scenario I).
[d]Tensors determined using no structural model (Scenario II/III).

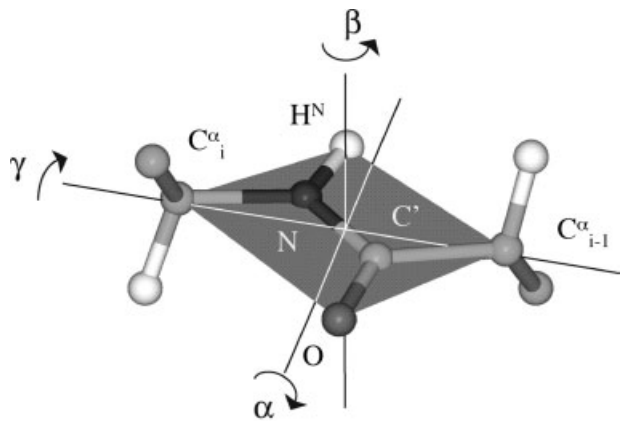

Figure 2
Three-Dimensional Gaussian Axial Fluctuation Model. Figure showing the axes of the 3D GAF model relative to one peptide plane. Diffusive reorientational motions around the three axes (α, β, γ) are described by Gaussian distributions with half-height amplitudes of {σ_α, σ_β, and σ_γ}.

Table II
Alignment Tensor Determination Using Linearly Independent Tensors

	Aa (10^{-4})	Ar (10^{-4})	α (°)	β (°)	γ (°)
Tensor 1 (true)[a]	20.00	13.00	60.00	60.00	60.00
Static[b]	17.65	11.5	61.7	60.6	59.3
SII/SIII[c]	19.2	12.6	60.00	60.00	60.00
Tensor 2 (true)	20.00	13.00	−60.00	−60.00	−60.00
Static	−17.8	11.7	−57.2	65.2	45.0
SII/SIII	19.2	12.9	−59.3	−59.1	−60.3
Tensor 3 (true)	20.00	13.00	0.00	45.00	0.00
Static	−17.5	−11.6	179.7	45.5	180.0
SII/SIII	19.4	12.6	1.3	45.3	−1.8
Tensor 4 (true)	20.00	13.00	45.00	0.00	45.00
Static	−17.1	−11.4	179	90.4	−90.1
SII/SIII	19.0	12.7	−1.3	−0.6	90.8

[a]Tensors used to simulate the RDCs.
[b]Tensors determined using all couplings, and taking no account of dynamic averaging (static model).
[c]Tensors determined using no structural model (Scenario II/III).

not explictly taken into account, and in fact coincides with the experimentally observed differences between tensors determined using a static model, or a model allowing for pervasive anisotropic GAF-like motions.[18] Tensors determined using Scenarios I and II/III are found to reproduce the actual eigenvalues much more closely. The structure-free approaches (II/III) are very slightly more accurate, but the average ratio of predicted and actual values of A_a are found to be very close to 1.0 in both cases (0.97 ± 0.01 for SI and 0.98 ± 0.01 for SII/III compared to 0.90 for the "static" tensors). Similar results were found for the classical MD simulation.

This result already provides key information concerning the capacity of these methods to quantitatively estimate the level of alignment, and indicates that both approaches are valid in providing alignment tensors against which reference framework dynamic amplitudes can be determined. According to these simulations only a near-negligible amount of motion is absorbed into the amplitude of the alignment tensors.

We have also tested the dependence of the tensor determination on tensor orientation, by simulating data for four orthogonal tensors. These results reproduce the features shown here and are summarized in Table II.

Determination of dynamic parameters from RDCs: comparison with MD and AMD

Once the alignment tensors have been determined as described above, the 3D GAF analysis is used to interpret the simulated RDCs as defined using the following three scenarios.

Scenario I

In the first case the average orientation is assumed to be well represented by an available model of the mean solution conformation, and the coordinates of this structure are used to define the axes of reorientation of each peptide plane. Again in this case this model is represented by the minimized average coordinates of the trajectories. The results of fitting the motional amplitudes $\{\sigma_\alpha, \sigma_\beta, \text{or } \sigma_\gamma\}$ are shown in Figures 3 and 4 for MD and AMD ensembles. The reproduction of all three 3D GAF amplitudes and generalized order parameters for the $N—H^N$ and $C'—C^\alpha$ vectors is clearly very precise for the MD ensemble, with rms values of S^2 of 0.06 and 0.04 respectively.

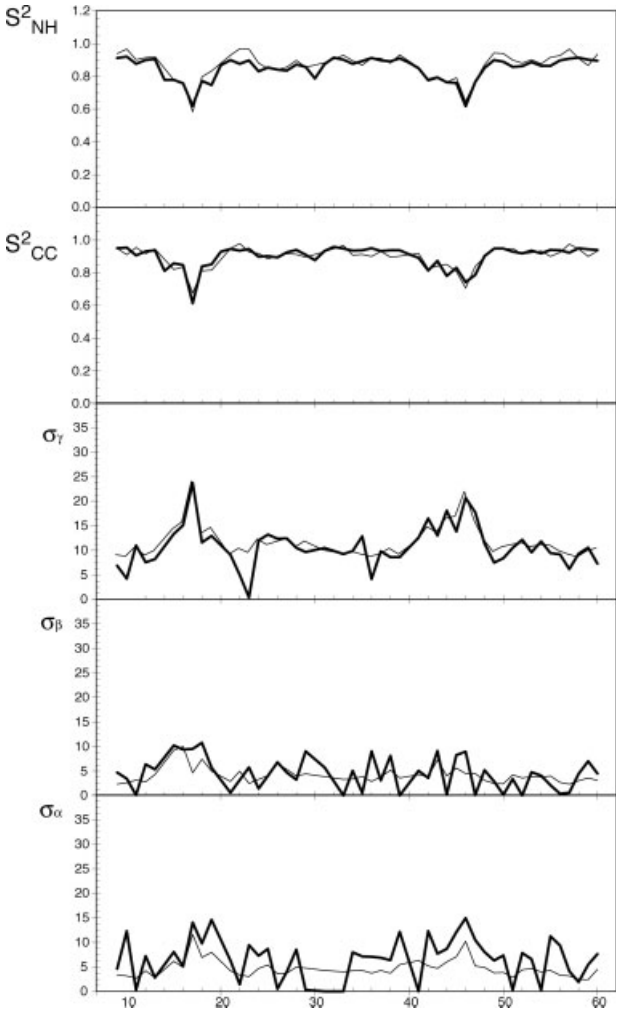

Figure 3
*Reproduction of motional parameters from MD ensemble using Scenario I. (**a**) Comparison of S^2_{NH} extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**b**) Comparison of $S^2_{C'C\alpha}$ values (same line thickness). (**c**) Comparison of σ_γ extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**d**) Comparison of σ_β values (same line thickness). (**e**) Comparison of σ_α values (same line thickness).*

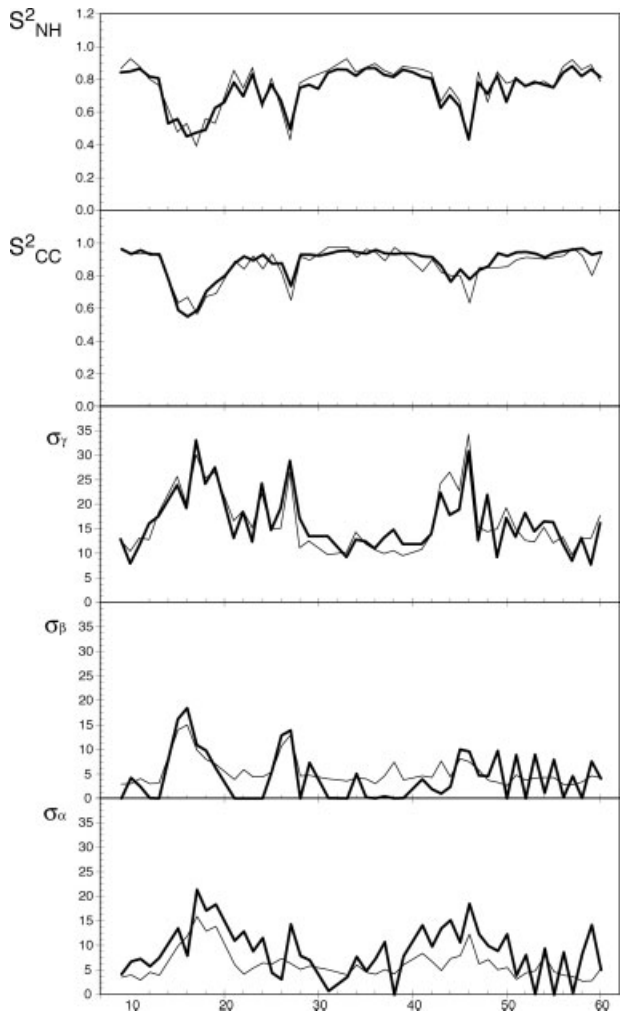

Figure 4

*Reproduction of motional parameters from AMD ensemble using Scenario I. (**a**) Comparison of S^2_{NH} extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**b**) Comparison of $S^2_{C'C\alpha}$ values (same line thickness). (**c**) Comparison of σ_γ extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**d**) Comparison of σ_β values (same line thickness). (**e**) Comparison of σ_α values (same line thickness).*

In the case of the AMD ensemble, where increased amplitude motions are present, the agreement between $\{\sigma_\alpha, \sigma_\beta, \text{or } \sigma_\gamma\}$ extracted directly from the MD simulation and those determined using the 3D GAF applied to the simulated RDCs, although less precise, is still remarkably good. Rms values for each amplitude are shown in Table III. We note that σ_α is apparently slightly, but systematically overestimated and σ_β similarly underestimated, although within the range $\{0^\circ < \sigma < 5^\circ\}$ GAF amplitudes are known to be quite insensitive to small angular fluctuations.

Finally we have compared generalized order parameters (S^2) calculated for different vectors in the peptide plane

Table III

3D GAF Reproduction of Data Extracted from MD and AMD

	RMS					
		AMD				
Scenario	MD SI	SI	SII	SIII	SII (4 tensors)[b]	SII (3 tensors)[c]
S^2_{NH}[a]	0.04	0.05	0.09	0.09	0.06	0.07
$S^2_{C'C\alpha}$	0.06	0.055	0.06	0.06	0.05	0.06
γ (°)	1.8	3.2	6.6	7.0	3.1	3.9
β (°)	2.2	3.3	3.2	3.8	3.0	3.7
α (°)	3.5	4.7	3.9	4.7	4.2	4.7

[a]Rms difference of order parameter or GAF amplitude between MD/AMD extracted parameters and values extracted from 3D GAF analysis of simulated RDCs. Values are averaged over the whole sequence.
[b]Using data simulated from 4 orthogonal tensors.
[c]Using data simulated from 3 orthogonal tensors.

from the 3D GAF amplitudes, with those calculated directly from the AMD derived ensemble. The results, again shown in Figure 4, indicate that these parameters are actually extremely robust, with rms values of 0.05 and 0.055 for S^2_{NH} and $S^2_{C'C\alpha}$, respectively.

Scenario II

In this case the mean conformation is constructed directly from the simulated RDCs using the *Dynamic-Meccano* protocol to determine the average conformation of the polypeptide chain assuming only a common structure of the peptide plane, and a weakly constrained tetrahedral junction. The conformation calculated from RDCs simulated from the AMD is shown in Figure 5, in comparison to the minimized average coordinates extracted from the AMD simulation (backbone atom rmsd: 0.56 Å).

The generalized order parameters extracted using this structure in the 3D GAF analysis of the RDC datasets are compared to values directly extracted from the AMD ensemble in Figure 6. Similar features are shown to those found using Scenario I. Although S^2_{NH} values are very slightly higher than the true values, the overall profile is again well reproduced despite fluctuations in σ. We note that the alignment tensor eigenvalues are in general closer to the true tensors. This small effect ($\Delta S^2_{NH} < 0.02$), may arise from a slight absorption of motion into the average conformations of the peptide planes.

Scenario III

In this case a three-dimensional structure is not constructed, rather the optimal average orientation of each peptide plane determined in Step A is used in the full 3D GAF analysis. The advantage of using such an approach is that there is no dependence on the determination of the orientation of the neighboring peptide planes. The

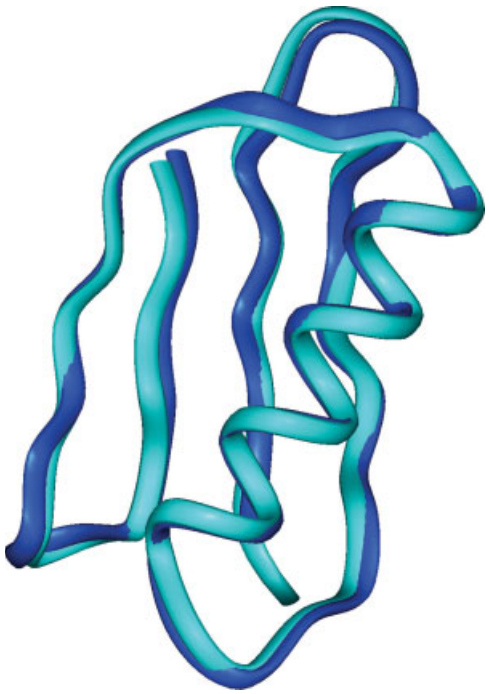

Figure 5

Comparison of the minimized average structure from the AMD trajectory (pale blue) and the structure determined using the GAF-based Dynamic-Meccano *algorithm (dark blue) applied to RDCs simulated from the trajectory (backbone rmsd 0.56 Å).*

results (Fig. 6) are very similar to those found for Scenario II.

Dependence on the alignment tensors

So far we have presented results of analysis of RDCs simulated using alignment tensors that have been determined from experimental measurements made in five different alignment media. To test the dependence of the approach on the nature of the relative independence of these alignment media, we have repeated the analysis using three, and four linearly independent alignment tensors (as measured by the normalized scalar product of the tensors[41]). The results of the application of Scenario II are shown in Figure 7. In comparison to results from Scenario II using data simulated from the experimental alignment tensors order parameters are again reproduced very closely, with a slightly smaller rms. Importantly the accuracy of the results is very similar when only three linearly independent tensors are used (Fig. 8).

DISCUSSION

This detailed analysis of the behavior of GAF based approaches to the determination of protein dynamics on the basis of RDCs indicates the robustness of the procedures to quantitatively evaluate amplitudes and modes of protein motions along the polypeptide chain. Analysis of fast librational modes represented in the classical MD simulation results in very close reproduction of the modes extracted directly from the trajectory. This is perhaps not surprising given that the 3D GAF model was originally inspired by observing common motional modes present in classical MD simulations of proteins. Possibly more surprisingly the methods reproduce larger amplitude motional modes present in the AMD ensemble quite precisely for all three tested scenarios, with slightly closer reproductions when the mean structure is already known, or when truly linearly independent tensors are available. The reproduction of the three amplitudes of the 3D GAF motions is also particularly impressive when one considers that while in general the angle between the

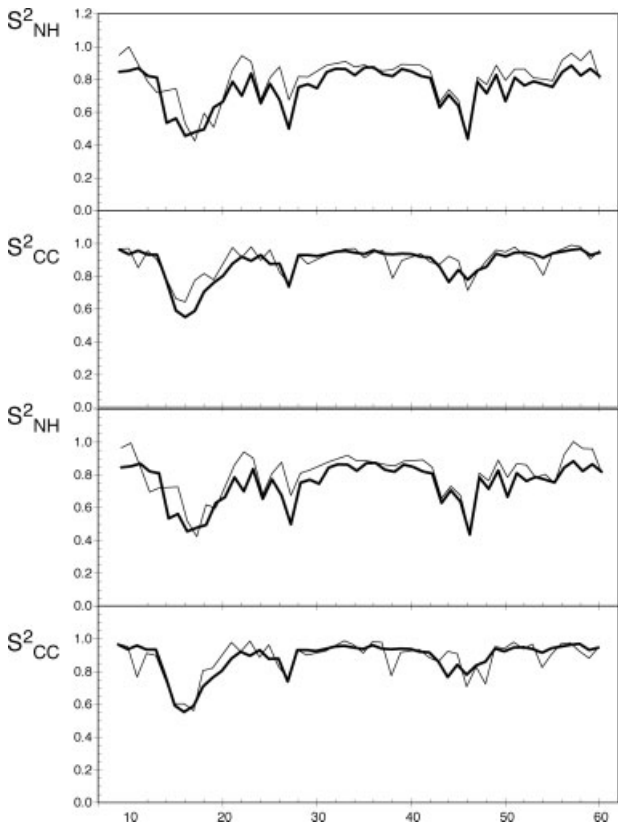

Figure 6

*Reproduction of generalized order parameters from AMD ensemble using Scenarios II and III. (**a**) Comparison of S^2_{NH} extracted from the AMD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin) using Scenario II. (**b**) Comparison of $S^2_{C'C\alpha}$ values calculated using Scenario II (same line thickness). (**c**) Comparison of S^2_{NH} extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin) using Scenario III. (**d**) Comparison of $S^2_{C'C\alpha}$ values calculated using Scenario III (same line thickness).*

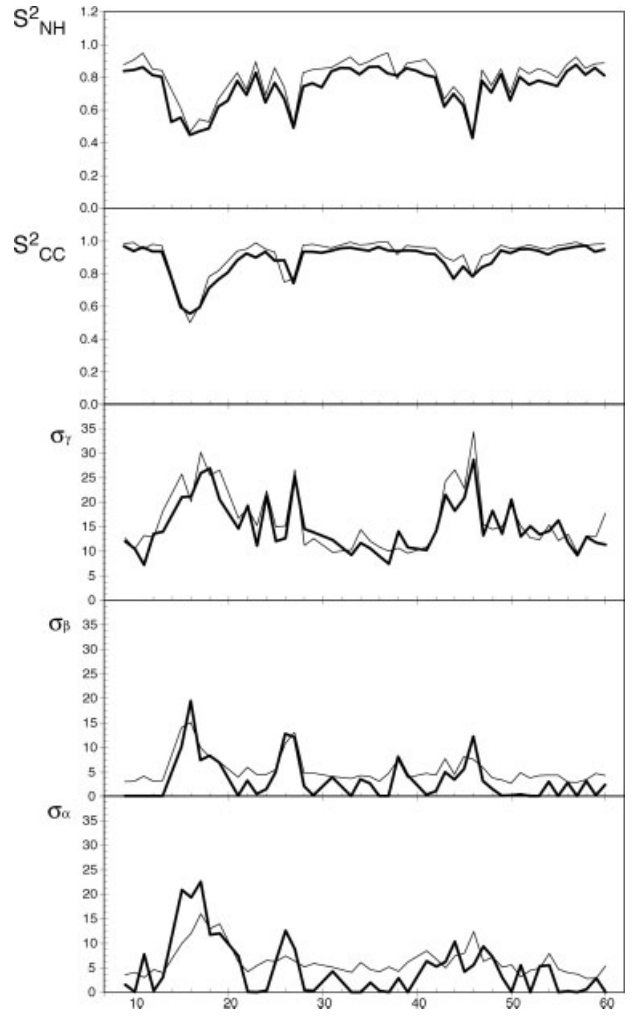

Figure 7

*Reproduction of motional parameters from AMD ensemble using Scenario II and RDCs simulated using four linearly independent alignment tensors. (**a**) Comparison of S^2_{NH} extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**b**) Comparison of $S^2_{C'C\alpha}$ values (same line thickness). (**c**) Comparison of σ_γ extracted from the AMD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**d**) Comparison of σ_β values (same line thickness). (**e**) Comparison of σ_α values (same line thickness).*

C^α—C^α vector and the γ-axis in the peptide plane varies from about 3 to 10°, in some cases, specifically in mobile loop regions, the actual axes of reorientation in the AMD trajectories can diverge by up to 20° from the peptide plane axes shown in Figure 2. We can conclude that the 3D GAF model provides a robust approximation to a number of different motional regimes and associated amplitudes.

The simulation used to produce all of the data in Figure 4 essentially reproduces, as closely as possible from simulation, the conditions encountered experimentally in our recent 3D GAF study of RDCs measured from protein GB3. The accurate determination of all of the expected parameters extracted explicitly from the AMD trajectory, including the quantitative amount of dynamics present, provides strong support for the conclusions that we drew from this study. The quantitative aspect of the study is an important aspect that should not be underestimated. Unless the true amplitude of the alignment tensors can be estimated, by taking into account the possible presence of dynamics of the sites used to estimate the tensors, some of the pervasive dynamics may be absorbed

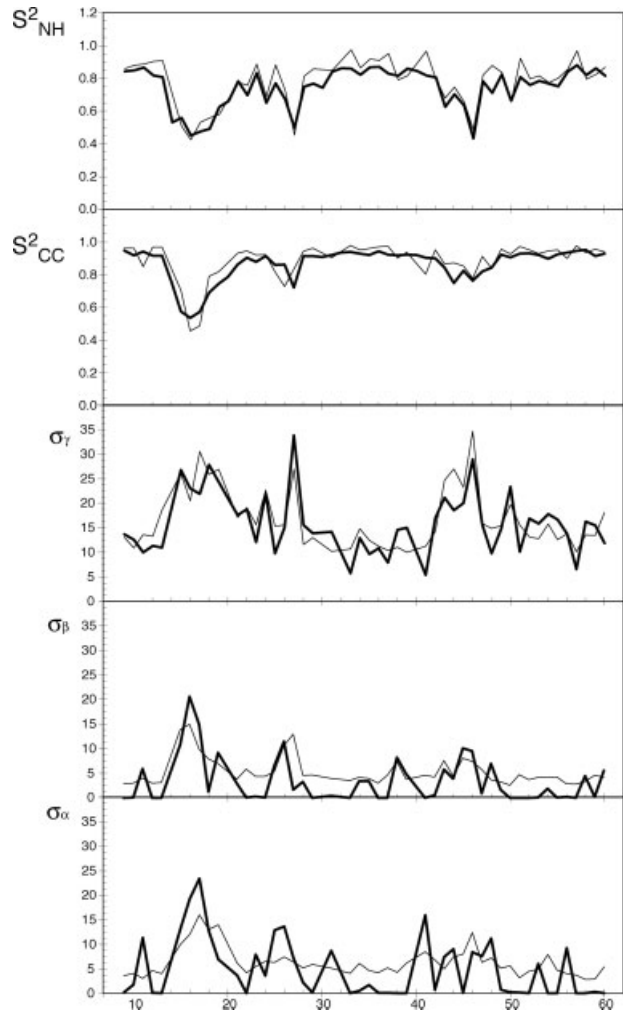

Figure 8

*Reproduction of motional parameters from AMD ensemble using Scenario II and RDCs simulated using three linearly independent alignment tensors. (**a**) Comparison of S^2_{NH} extracted from the MD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**b**) Comparison of $S^2_{C'C\alpha}$ values (same line thickness). (**c**) Comparison of σ_γ extracted from the AMD trajectory (thick) and from the RDCs simulated with noise from the same trajectory, using the 3D GAF model (thin). (**d**) Comparison of σ_β values (same line thickness). (**e**) Comparison of σ_α values (same line thickness).*

into the eigenvalues. This will be particularly true if a pervasive dynamic component is present of axial symmetry, that would simply act to reduce the effective tensor eigenvalues. The order parameters calculated from the MD trajectories using Eq. (1) would however be directly averaged by such a component. The close reproduction of both distribution and range of the generalized order parameters extracted from the MD simulation and determined using the 3D GAF model demonstrates that this source of artifact is essentially insignificant if one applies the 3D GAF approach as described here. Of course we are aware that the MD and AMD trajectories only provide a virtual representation of the real system, but we can confidently conclude that the 3D GAF approach is sufficiently robust to reproduce the motional amplitudes present in these trajectories.

A further potential source of artifact of the 3D GAF analysis is the dependence of the approach on the structural model used to describe the average conformation in solution. Although in our original study three different structural models (two X-ray structures and an NMR-refined structure) were shown to contribute essentially identical distributions of the three orthogonal amplitudes, the detection of very similar motions in a study that simultaneously determined both average structure and backbone dynamics from RDCs probably provides the strongest evidence that the dependence of earlier results on the structural model was negligible.[26] Importantly in this study the dynamically averaged structure determined using *Dynamic Meccano*[26] was also shown to provide a significantly better description of the conformational properties in solution than a bias-free static approach performed in parallel. The results from the current study strongly support these results, with essentially identical motional modes being determined from both the structure-dependent (Scenario I), structure-free (Scenario III) and simultaneous determination of structure and dynamics-based (Scenario II) analyses.

CONCLUSIONS

In conclusion we have used classical and accelerated MD simulation to characterize the ability of recently proposed approaches based on the description of peptide plane reorientational modes using 3D GAF model to reproduce true motional behavior on the backbone of polypeptide chains. It was not clear whether the 3D GAF analysis of RDCs would quantitatively reproduce local amplitudes, as this requires not only a faithful description of the motional modes, but also an accurate determination of the alignment tensors. The results demonstrate the robustness of the approaches to alignment tensor parameterisation and of the 3D GAF model to reproduce both the amplitude and direction of peptide motions. Generalized order parameters describing the amplitudes of the angular order parameters for different vectors are found to be remarkably well reproduced, even for complex or large amplitude motions, demonstrating the general accuracy of the 3D GAF model.

REFERENCES

1. Mittermaier A, Kay LE. New tools provide new insights in NMR studies of protein dynamics. Science 2006;312:224–227.
2. Kay LE, Torchia D, Bax A. Backbone dynamics of proteins as studied by N15 inverse detected heteronuclear NMR-spectroscopy—application to Staphylococcal Nuclease. Biochemistry 1989;28:8972–8979.
3. Palmer AG, III. NMR probes of molecular dynamics: overview and comparison with other techniques. Annu Rev Biophys Biomol Struct 2001;30:129–155.
4. Bruschweiler R. New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. Curr Opin Struct Biol 2003;13:175–183.
5. Eisenmesser EZ, Bosco DA, Akke M, Kern D. Enzyme dynamics during catalysis. Science 2002;295:1520–1523.
6. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D. Intrinsic dynamics of an enzyme underlies catalysis. Nature 2005;438:117–121.
7. Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM, Kay LE. Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. Nature 2004;430:586–590.
8. Tjandra N, Bax A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 1997;278:1111–1114.
9. Tolman JR, Flanaghan JM, Kennedy MA, Prestegard JH. NMR evidence for slow collective motions in cyanometmyoglobin. Nat Struct Biol 1997;4:292–297.
10. Prestegard JH, Al-Hashimi HM, Tolman JR. NMR structures of biomolecules using field oriented media and residual dipolar couplings Q Rev Biophys 2000;33:371–424.
11. Blackledge M. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. Prog NMR Spectr 2005;46:23–61.
12. Tolman JR, Al-Hashimi HM, Kay LE, Prestegard JH. Structural and dynamic analysis of residual dipolar coupling data for proteins. J Am Chem Soc 2001;123:1416.
13. Meiler J, Prompers JJ, Peti W, Griesinger C, Bruschweiler R. Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. J Am Chem Soc 2001;123:6098–6107.
14. Peti W, Meiler J, Bruschweiler R, Griesinger C. Model-free analysis of protein backbone motion from residual dipolar couplings. J Am Chem Soc 2002;124:5822–5833.
15. Briggman KB, Tolman JR. De Novo determination of bond orientations and order parameters from residual dipolar couplings with high accuracy. J Am Chem Soc 2003;125:10164–10165.
16. Ulmer TS, Ramirez BE, Delaglio F, Bax A. Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. J Am Chem Soc 2003;125:9179–9191.
17. Clore GM, Schwieters CD. 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 2004;126:2923–2938.
18. Bernado P, Blackledge M. Local dynamic amplitudes on the protein backbone from dipolar couplings: toward the elucidation of slower motions in biomolecules. J Am Chem Soc 2004;126:4907–4920.

19. Bouvignies G, Bernado P, Blackledge M. Protein backbone dynamics from N-H-N dipolar couplings in partially aligned systems: a comparison of motional models in the presence of structural noise. J Magn Reson 2005;173:328–338.
20. Bernado P, Blackledge M. Local dynamic amplitudes on the protein backbone from dipolar couplings: toward the elucidation of slower motions in biomolecules. J Am Chem Soc 2004;126:7760–7761.
21. Showalter SA, Bruschweiler R. Quantitative molecular ensemble interpretation of NMR dipolar couplings without restraints. J Am Chem Soc 2007;129:4158–4159.
22. Lakomek NA, Fares C, Becker S, Carlomagno T, Meiler J, Griesinger C. Side-chain orientation and hydrogen-bonding imprint supra-Tau(c) motion on the protein backbone of ubiquitin. Ang Chem Int Ed Engl 2005;44:7776–7778.
23. Tolman JR. A novel approach to the retrieval of structural and dynamic information from residual dipolar couplings using several oriented media in biomolecular NMR spectroscopy. J Am Chem Soc 2002;124:12020–12031.
24. Bouvignies G, Bernado P, Meier S, Cho K, Grzesiek S, Brüschweiler R, Blackledge M. Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. Proc Natl Acad Sci 2005;102:13885–13890.
25. Bremi T, Brüschweiler R. Locally anisotropic internal polypeptide backbone dynamics by NMR relaxation. J Am Chem Soc 1997; 120:6672–6673.
26. Bouvignies G, Markwick PRL, Brüschweiler R, Blackledge M. Simultaneous determination of protein backbone structure and dynamics from residual dipolar couplings. J Am Chem Soc 2006; 128:15100–15101.
27. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Biol 2002;9:646–652.
28. Karplus M, Kuriyan J. Molecular dynamics and protein function. Proc Natl Acad Sci USA 2005;9:6679–6685.
29. Levy RM, Karplus M, Wolynes PG. NMR relaxation parameters in molecules with internal motion—exact Langevin trajectory results compared with simplified relaxation models. J Am Chem Soc 1981;103:5998–6011.
30. Brüschweiler R, Roux B, Blackledge MJ, Griesinger C, Karplus M, Ernst RR. Influence of rapid intramolecular motion on NMR cross-relaxation rates—a molecular dynamics study of Antamanide in solution. J Am Chem Soc 1992;114:2289–2302.
31. Case DA. Molecular dynamics and NMR spin relaxation in proteins. Acc Chem Res 2002;35:325–331.
32. Hamelberg D, Mongan J, McCammon JA. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. J Chem Phys 2004;120:11919–11929.
33. Markwick PRL, Bouvignies G, Blackledge M. Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. J Am Chem Soc 2007;129:4724–4730.
34. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical-integration of Cartesian equations of motion of a system with constraints—molecular dynamics of N-alkanes. J Comput Phys 1977;23:327–341.
35. Cheatham TE, Miller JL, Fox T, Darden TA, Kollman PA. Molecular dynamics simulations on solvated biomolecular systems—the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. J Am Chem Soc 1995;117:4193–4194.
36. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
37. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–725.
38. Case DA, Darden TA, Cheatham TE, III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister C, Caldwell JW, Ross WS, Kollman PA. 2004. AMBER 8, University of California, San Francisco.
39. Chandrasekhar I, Clore GM, Szabo A, Gronenborn AM, Brooks BR. A 500-ps molecular dynamics simulation study of interleukin-1 beta in water—correlation with NMR spectroscopy and crystallography. J Mol Biol 1992;226:239–250.
40. Brüschweiler R, Wright PE. NMR order parameters of biomolecules—a new analytical representation and application to the Gaussian axial fluctuation model. J Am Chem Soc 1994;116:8426–8427.
41. Sass HJ, Cordier F, Hoffmann A, Rogowski M, Cousin A, Omichinski J, Lowen H, Grzesiek S. Purple membrane induced alignment of biological macromolecules in the magnetic field. J Am Chem Soc 1999;121:2047–2055.