Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 173 (2005) 328–338
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Communication

Protein backbone dynamics from N–H^N dipolar couplings in partially aligned systems: a comparison of motional models in the presence of structural noise

Guillaume Bouvignies[1], Pau Bernadó[1], Martin Blackledge*

Institut de Biologie Structurale Jean-Pierre Ebel, U.J.F.—C.N.R.S.—C.E.A., 41 rue Jules Horowitz, 38027 Grenoble Cedex, France

Received 27 October 2004; revised 22 December 2004
Available online 28 January 2005

Abstract

Residual dipolar couplings (RDCs) provide excellent probes for the exploration of dynamics in biomolecules on biologically relevant time-scales. Applying geometric motional models in combination with high-resolution structures to fit experimental RDCs allows the extraction of local dynamic amplitudes of peptide planes in proteins using only a limited number of data points. Here we compare the behaviour of three simple and intuitive dynamic modes: the Gaussian axial fluctuation model (1D-GAF), the two-site jump model, and a model supposing axially symmetric motion about a mean orientation. The requirement of a structural model makes this kind of methodology potentially very sensitive to structural imprecision. Numerical simulations of RDC dynamic averaging under different regimes show that the anisotropic motional models are more geometrically stringent than the axially symmetric model making it more difficult to alias structural noise as artificial dynamic amplitudes. Indeed, it appears that the model extracts accurate motional amplitudes even in the presence of significant structural error. We also show that a two-site jump model, also assuming the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ as rotation axis, can only be distinguished from the previously developed GAF model beyond amplitude/jumps of around 40°. The importance of appropriate estimation of the molecular alignment tensor for determination of local motional amplitudes is also illustrated here. We demonstrate a systematic scaling of extracted dynamic amplitudes if a static structure is assumed when determining the alignment tensor from dynamically averaged RDCs. As an example an artificial increase of 0.14 (0.85 compared to the expected 0.71) is observed in the extracted S^2 if a pervasive 20° GAF motion is present that is ignored in the tensor determination. Finally we apply a combined approach using the most appropriate motional model, to complete the analysis of dynamic motions from protein G.

Keywords: Residual dipolar couplings; Protein; Dynamics; Mobility; Structure

1. Introduction

Recently intense interest has focussed on the use of dilute liquid crystals to weakly align proteins and thereby induce a small degree of anisotropy into the orientational sampling experienced by the molecule of interest [1,2]. Residual dipolar couplings measured in this way provide highly precise information defining the orientation of internuclear bonds relative to a molecule-fixed frame, a characteristic that makes these parameters particularly powerful for biomolecular structure determination [3–5]. These measurements also report on averages over relatively long time-scales (up to the millisecond range) as well as over the entire population ensemble. The dynamic information contained in these averages is therefore highly complementary to motions detected from NMR spin relaxation studies and potentially of great interest for the identification of specific motional

* Corresponding author. Fax: +33 4 76885494.
E-mail address: Martin.Blackledge@ibs.fr (M. Blackledge).
[1] These authors contributed equally to this work.

1090-7807/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2005.01.001

models. This importance was recognised since the earliest work on weak alignment of proteins, and has stimulated the development of numerous methodological approaches to the extraction of dynamic information from dipolar couplings [6–11].

Two model-free interpretations of the dynamic averaging of RDCs have been recently published [12,13] and applied to the protein ubiquitin [14,15]. In these approaches, the directionality of dynamical averaging can be derived if RDCs measured in more than five alignment systems are available. We have recently investigated the possibility of assuming a simple geometric model for local anisotropic motions to interpret conformational averaging of residual dipolar couplings, thereby reducing the number of independent measurements required for the characterisation of local motions. Backbone dynamics are certainly complex processes, suggesting that no single geometric model will be universal. Nevertheless anisotropic motions, in the form of Gaussian axial fluctuations (GAF), around the average position have been shown from molecular dynamics simulations and spin relaxation measurements to be both valid and useful in characterising fundamental backbone motions in proteins [16–20]. We introduced an analytical expression incorporating the averaging of RDC under the influence of GAF motions in terms of peptide plane orientation and motional amplitude (σ). Assuming the former is known (from a high-resolution structural model for example) N–^{N}H reorientation amplitudes can be described using the single parameter σ. We have recently presented the application of this approach to the study of average anisotropic peptide plane dynamics in secondary structural elements from six different proteins from single alignment media, detecting average σ values of around 15° [21]. This analysis provided statistically significant improvement compared to static, or axially symmetric descriptions, thereby validating the approach. In the case of Lysozyme, for which a high-resolution crystal structure is available, we were also able to detect significantly different average dynamic amplitudes in loop and secondary structural regions.

In a further study we extended this approach to the modelling of local motions with the GAF model, using data from more than one alignment medium [22]. The success of this technique relies on the fact that anisotropic dynamics average differently depending on the orientation of the peptide plane, allowing a more accurate characterisation of the dynamic modes with increasing number of alignment media. The major disadvantage of the approach compared to the model-free analyses described earlier is that the model-dependent method requires high-resolution structural information in order to define the orientation of the local averaging in each peptide plane frame. It is therefore of utmost importance that care is taken to avoid aliasing of local structural noise (defined as the difference between the known structural model and the actual conformation in solution) into fictional motional amplitudes. Here we use numerical simulations performed in the presence of three different alignment tensors to test statistical methods developed to minimise this possibility. In particular we compare the GAF motional model to motion assuming cylindrical symmetry and to a simple two-site jump about the $^{\alpha}$C–$^{\alpha}$C axis. This latter model turns out to be analytically indistinguishable from the 1D-GAF motion up to a certain amplitude of motion. Interestingly we find that the GAF model exhibits more robust behaviour with respect to structural noise than the cylindrically symmetric model, implying that motional parameters extracted using this model can be treated with more confidence. Application of these models is illustrated using data from five alignment media from protein G [23].

2. Theory

2.1. Dynamic averaging of RDCs: a common treatment

The average dipolar coupling between two spins can be expressed as follows:

$$\langle D_{jk}(\theta,\phi)\rangle = b_{jk}\left[A_{\mathrm{a}}\left\langle\frac{3\cos^2\theta-1}{2}\right\rangle+\frac{3}{4}A_{\mathrm{r}}\langle\sin^2\theta\cos 2\phi\rangle\right], \tag{1}$$

where the angular brackets indicate the averaging over all sampled conformations. A_{a} and A_{r} are the axial and rhombic components of the alignment tensor, and the constant b_{jk} is given by

$$b_{jk} = -\frac{\mu_0}{4\pi}\frac{\gamma_j\gamma_k\hbar}{r_{jk}^3}.$$

Eq. (1) can be recast using averaged spherical harmonics [24],

$$\langle D_{jk}(\theta,\phi)\rangle = \sqrt{\frac{16\pi}{5}}b_{jk}\left[A_{\mathrm{a}}\langle Y_{2,0}(\theta,\phi)\rangle+\sqrt{\frac{3}{8}}A_{\mathrm{r}}(\langle Y_{2,2}(\theta,\phi)\rangle+\langle Y_{2,-2}(\theta,\phi)\rangle)\right]. \tag{2}$$

Following the logic presented by Meiler et al. [12] we can then write the average of these spherical harmonics in a frame defined by the geometry of the peptide plane, using the rotational properties of spherical harmonics. If we apply an Euler rotation $R(\alpha,\beta,\gamma)$, the averaged spherical harmonics are transformed as follows:

$$R(\alpha,\beta,\gamma)Y_{2,M}(\theta,\phi) = \sum_{M'=-2}^{+2} \mathrm{e}^{-\mathrm{i}\alpha M'} d_{M',M^{(2)}}(\beta)\,\mathrm{e}^{-\mathrm{i}\gamma M}Y_{2,M'}(\theta,\phi), \tag{3}$$

where $d^{(2)}_{M',M}$ are the Wigner rotation matrices. Eq. (2) can then be written:

$$\langle D_{jk}(\theta,\phi)\rangle = \sqrt{\frac{16\pi}{5}} b_{jk}\left[A_a\left(\sum_{M'=-2}^{+2} e^{-i\alpha M'} d^{(2)}_{M',0}(\beta)\langle Y_{2,M'}\rangle\right) + \sqrt{\frac{3}{8}}A_r\left(\sum_{M'=-2}^{+2} e^{-i\alpha M'} d^{(2)}_{M',2}(\beta)\, e^{-i2\gamma}\langle Y_{2,M'}\rangle + \sum_{M'=-2}^{+2} e^{-i\alpha M'} d^{(2)}_{M',-2}(\beta)\, e^{+i2\gamma}\langle Y_{2,M'}\rangle\right)\right]. \quad (4)$$

Notice that the Euler rotation $R(\alpha,\beta,\gamma)$ here defines the transformation from the alignment tensor frame to the peptide plane frame. As a consequence, all peptide planes will have a different set of Euler angles, allowing us to treat all sites equivalently.

Eq. (4) allows extraction of five averaged spherical harmonics if RDCs in five or more alignment media have been measured [12]. Here we describe the different motional regimes using analytical expressions, and consequently reduce the number of unknowns. Three different internal dynamic motional modes have been used to introduce dynamics on the N–H^N residual dipolar coupling definition: axially symmetric motion, one dimension GAF, and two-site jump model.

2.2. Axially symmetric motional model

The axially symmetric model is a simple and intuitive motional model broadly used in NMR. Dynamics of a vector under this motional model are characterized by a symmetric distribution of this vector around the average position. Dynamic averaging under this regime is known, its analytical description is presented here for completeness.

Euler angles definition used to describe this model place z axis of the frame as the direction of the average position of the vector under study. Analytically this can be expressed as,

$$\langle e^{\pm i\phi}\rangle = \langle e^{\pm 2i\phi}\rangle = 0, \quad (5)$$

consequently,

$$\langle Y_{2,\pm 2}(\theta,\phi)\rangle = \langle Y_{2,\pm 1}(\theta,\phi)\rangle = 0. \quad (6)$$

Eq. (4) is then simplified to,

$$\langle D_{jk}(\theta,\phi)\rangle = \sqrt{\frac{16\pi}{5}} b_{jk}\langle Y_{2,0}(\theta,\phi)\rangle\left[A_a d^{(2)}_{0,0}(\beta) + \sqrt{\frac{3}{8}}A_r\left(d^{(2)}_{0,2}(\beta)\, e^{-i2\gamma} + d^{(2)}_{0,-2}(\beta)\, e^{+i2\gamma}\right)\right], \quad (7)$$

which becomes

$$\langle D_{jk}(\theta,\phi)\rangle = b_{jk}\left\langle\frac{3\cos^2\theta-1}{2}\right\rangle\left[A_a\left(\frac{3\cos^2\beta-1}{2}\right) + \frac{3}{4}A_r\sin^2\beta\cos 2\gamma\right], \quad (8)$$

where $\beta = -\theta_{static}$ and $\gamma = -\phi_{static}$ with θ_{static} and ϕ_{static} are the polar angles that define the geometry of the symmetry axis of the movement, which is equivalent to the internuclear vector. As a consequence static RDC, $D_{jk,static}$, is linearly scaled,

$$\langle D_{jk}(\theta,\phi)\rangle = S.D_{jk,static}(\theta_{static},\phi_{static}), \quad (9)$$

where S is the scaling factor that depends on the amplitude of the motion but not on the position of the vector with respect to the alignment tensor. S can be compared to order parameters, S^2, derived from NMR relaxation studies [25,26]. In a general way S^2 can be expressed using spherical harmonics,

$$S^2 = \frac{4\pi}{5}\sum_{M=-2}^{+2}\langle Y_{2,M}(\theta,\phi)\rangle\langle Y^*_{2,M}(\theta,\phi)\rangle \quad (10)$$

under axially symmetric motion this is transformed into

$$S^2 = \frac{4\pi}{5}\langle Y_{2,0}(\theta,\phi)\rangle^2 = \left\langle\frac{3\cos^2\theta-1}{2}\right\rangle^2. \quad (11)$$

2.3. Gaussian axial fluctuation motional model—GAF

The GAF motional model describes motional averaging as a distribution of conformations centred around a known orientation, with the $(i-1,i)$ peptide plane undergoing reorientations about the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis. This model has been found to accurately describe backbone motions in molecular dynamic simulations and has been applied for the interpretation of relaxation rates, and a description of this model for interpretation of dynamically averaged RDCs that assumes a perpendicular axis of rotation with respect to the N–^{N}H vector (ortho-GAF) has recently been presented [21]. This is presented more fully here for completeness.

Under the influence of GAF motion where the y axis of the peptide plane frame is placed along the dynamically active $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ vector ($\phi = 0$), and using averaging of the trigonometric functions under GAF,

$$\langle\cos(m\theta)\rangle = \int_{-\infty}^{+\infty} p(\theta)\cos(m\theta)\,d\theta = \exp\left(\frac{-m^2\sigma^2}{2}\right), \quad (12)$$

where

$$p(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-\theta^2}{2\sigma^2}\right). \quad (13)$$

Eq. (4) is transformed into,

$$\langle D_{\mathrm{NH}}\rangle = b_{jk}\left[\frac{A_{\mathrm{a}}}{4}\{s_1(3\cos^2\beta-1)+3s_2\sin^2\beta\cos 2\alpha\} + \frac{3}{8}A_{\mathrm{r}}\left\{s_1\sin^2\beta\cos 2\gamma + 2s_2\left(\cos^4\frac{\beta}{2}\cos 2\delta_1 + \sin^4\frac{\beta}{2}\cos 2\delta_2\right)\right\}\right], \quad (14)$$

where

$$\delta_1=\alpha+\gamma;\quad \delta_2=\alpha-\gamma;\quad s_1=1+3e^{-2\sigma^2};\quad s_2=1-e^{-2\sigma^2}. \quad (15)$$

An analytical expression for S^2 has been derived under a GAF motional model with an amplitude of σ [16],

$$S^2=1-\frac{3}{4}(1-e^{-4\sigma^2}). \quad (16)$$

2.4. Two-site jump model

This motional model describes the situation where the site of interest experiences jumps between two well-defined, energetically degenerate, conformational states. In an equivalent way to the previously described ortho-GAF it is convenient to describe the jumps as rotations around axis $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ with an amplitude of 2θ. The orthogonal approximation has again been used in order to simplify the analytical development.

In the peptide plane frame the angular density under a two-site jump motional model can be written,

$$p(x)=\frac{1}{2}(\delta(x+\theta)+\delta(x-\theta)). \quad (17)$$

Spherical harmonics are then averaged,

$$\begin{aligned}\langle Y_{2,-2}\rangle &= \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{-2i\varphi},\\ \langle Y_{2,-1}\rangle &= 0,\\ \langle Y_{2,0}\rangle &= \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta-1),\\ \langle Y_{2,2}\rangle &= -\frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{+2i\varphi},\\ \langle Y_{2,1}\rangle &= 0,\end{aligned} \quad (18)$$

where θ represents the half angular amplitude of the jump. The equation of a dynamically averaged RDC under this motional regime becomes,

$$\langle D_{\mathrm{NH}}\rangle = b_{jk}\left[\frac{A_{\mathrm{a}}}{4}\{s'_1(3\cos^2\beta-1)+3s'_2\sin^2\beta\cos 2\alpha\} + \frac{3}{8}A_{\mathrm{r}}\left\{s'_1\sin^2\beta\cos 2\gamma + 2s'_2\left(\cos^4\frac{\beta}{2}\cos 2\delta_1 + \sin^4\frac{\beta}{2}\cos 2\delta_2\right)\right\}\right] \quad (19)$$

with

$$\delta_1=\alpha+\gamma;\quad \delta_2=\alpha-\gamma;$$
$$s'_1=2(3\cos^2\theta-1);\quad s'_2=2\sin^2\theta. \quad (20)$$

Notice that the form of Eq. (19) developed for two-site jump model is equivalent to that developed for ortho-GAF (14) where only terms s'_1 and s'_2 differ. The generalised order parameter of a site experiencing a two-site jump of 2θ amplitude can be expressed as:

$$S^2=\frac{3\cos^2 2\theta+1}{4}. \quad (21)$$

3. Methods

3.1. Overview of the simulation approach

In order to assess the behaviour of the different motional models in the presence of structural error we have adopted the following simulation protocol: A structural model comprising 500 amino acids has been constructed with perfect local geometry for each peptide plane, and random secondary and tertiary fold. This ensures that orientational space is sufficiently well sampled to allow reasonable conclusions to be drawn from analysis of each scenario applied to the 500 peptide planes. Before adding structural noise to the protein, data are simulated from this conformation in the presence or absence of the different dynamic modes for three different alignment tensors. This simulates the scenario where the conformation accurately represents the average conformation in solution. Thermal noise is also present in these simulations.

We then address the effects of structural noise by changing the conformation of the model in a systematic way such that *each plane* is subject to structural error of the same amplitude. The reason we have simulated data from a large number of randomly distributed peptide planes is not to simulate a real protein experiencing the same structural error and local dynamics throughout the sequence, an unlikely concept, but rather to analyse the expected effects of structural noise on extracted dynamic amplitudes for all potential orientations of a given plane with respect to the tensors. Examining 500 planes with different orientations under the same conditions should allow a clear picture of the expected behaviour for an arbitrarily oriented plane. As far as tensor determination is concerned for this case, we make the assumption that neither structural noise or dynamics are significantly affecting the rest of the molecule. We are therefore looking at differential effects here, and thus need only invoke the tensor that was used to simulate the data. This allows us to isolate the effect we are interested in: each simulation of 500 planes then characterises the behaviour of a plane of arbitrary orientation with a defined motion in the presence of a particular degree of structural noise.

This approach is relevant in the case of crystal structures of proteins where, for example the expected conformational differences between solution and crystal may be much more pronounced in loop regions than in secondary structural elements, and will therefore be heterogeneously distributed along the sequence. The fact that uncertainty in the alignment tensor determination will also have an effect on the extracted amplitudes if pervasive motion is present (a different scenario than the one described above) is also addressed in the final part of our manuscript.

3.2. Fitting of RDCs to protein structures

Parameters describing local amplitude motions have been extracted for sites using the dynamic models described above. N–H^N RDCs measured in several alignment systems are combined to extract individual amplitudes minimising a merit function:

$$\chi_r^2(\{S,\sigma,\theta\}) = \sum_m \left(\frac{D_{r,m}^{\exp} - \left\langle D_{r,m}^{\text{calc}} \right\rangle}{\delta_{r,m}} \right)^2, \tag{22}$$

where $\langle D_{r,m}^{\text{calc}} \rangle$ corresponds to the dynamically averaged RDC (in one of the previously described modes) for residue r in alignment media m. $\delta_{r,m}$ is the uncertainty of the experimental RDC. An experimental error of 1.0 Hz was used for RDCs measured in protein G and for analysis of numerically simulated RDCs. A value of 1.02 Å was used for the N–H^N distance as in previous studies [21].

3.3. Statistical criteria for dynamic model selection

Internal motional parameters were extracted fitting RDCs to structures using the three different dynamic-averaged definitions of dipolar couplings described above. Dynamic parameters were only accepted if they passed different statistical tests based on Monte-Carlo simulations. Three different tests were performed for each site independently of the number of the RDCs available. The *Uncertainty Test* consists of an estimation of the uncertainty of dynamic parameter extracted from the fit (defined as the standard deviation of the resulting set of calculated amplitude parameters). The fitting procedure was repeated 1000 times with synthetic RDCs using a normal distribution of amplitude equivalent to the experimental error of back-calculated RDC centred on the value derived from the best-fit of the experimental data. In order to accept a fit, this uncertainty should be smaller than an empirically calibrated threshold value.

The χ^2 *test* was performed by comparing the best-fit χ_r^2 value (Eq. (22)), to the distribution resulting from the Monte-Carlo simulations described above (χ_r^2 should be within the 95% best χ_r^2 from the Monte-Carlo simulations). If either the *Uncertainty Test* or the χ^2 *test* are not fulfilled, the *F test* is then applied in order to select the fit if the dynamic model satisfies the data better than a static model, using a 90% confidence level. The *F test* was applied explicitly, such that fictitious data-sets were simulated from the static model (resulting from the fit in the absence of dynamics), again with random noise added to the back-calculated RDCs, and these datasets were fit to both the dynamic and the static models. This allows an estimation of the probability that the increase in the quality of the fit when introducing an additional parameter (measured by the reduction in χ_r^2) is due to random fluctuation.

3.4. Numerical simulations

Robustness and reliability of dynamic parameters extracted using this methodology were tested using numerical simulations with respect to the nature of the averaging and the presence of different degrees of structural error. A virtual 500 amino-acid protein was created from randomly selected values for ϕ and ψ dihedral angles, using a standard geometrical definition for peptide planes. This random conformation ensures a good orientational sampling of the resulting peptide planes. For this 'structure,' synthetic dynamically averaged RDCs were calculated with different motional amplitudes and regimes with respect to three different alignment tensors, all of them sharing magnitude descriptors ($A_a = 20.0 \times 10^{-4}$ and $A_r = 13.0 \times 10^{-4}$) but oriented differently, with Euler angles $(0°, 0°, 0°)$; $(60°, 60°, 60°)$; and $(-60°, -60°, -60°)$. These angles ensure that the static RDCs from the different tensors are not strongly correlated. For 1D-GAF-averaged RDCs, σ motional amplitudes of 20°, 30°, and 40° were used in simulations as explained previously [21]. Comparable dynamic amplitudes for axially symmetric motional modes were obtained by scaling static RDCs by S values of 0.843, 0.707, and 0.597, respectively (equivalent to $\theta = 18.9°$, 26.2°, and 31.22° half-angles, respectively). Random noise selected from a Gaussian distribution centred at 0.0 and with a standard deviation of 0.5 Hz was added to the resulting RDCs in order to introduce experimental uncertainty.

For the study of structural noise effects, N–H^N averaged couplings with the same motional amplitude resulting from alignment due to the three different tensors were fitted locally using the different dynamic models described in Eqs. (9), (14), and (19). Structures with successively increasing amounts of structural noise were built from the original conformation, by rotating every N–H^N vector (0°, 3°, 5°, and 10°) with respect to a randomly selected vector perpendicular to it. Extracted parameters describing internal motion were selected according to the same statistical tests presented above and compared with the values used to generate the averaged RDCs from numerical simulation.

The simulated dynamically averaged data were also fitted to the original structure, with 0° degrees of structural noise using the incorrect motional model. This tests the ability to distinguish between axially symmetric and anisotropic motional models.

3.5. Analysis of experimental data from protein G

For an absolute estimation of amplitude motions in the case of experimental data a precise determination of the alignment tensor magnitude is mandatory. Moreover, definition of the alignment tensor is needed for the description of peptide plane position in the tensor frame (α, β, and γ) in Eqs. (14) and (19) for GAF and two-site jump models, respectively. In this case C'–$^{\alpha}C$ RDCs were used to fit alignment tensors to the 1.1 Å resolution structure of protein G (1igd [27]) using the program MODULE [28].

Model selection for the experimental data from protein G applies the above-mentioned tests, and takes the best-fitting model as the most appropriate.

4. Results and discussion

4.1. Comparison of two-site jump model with GAF

The derivation and description of RDC averaging under the two-site jump dynamic model (Eq. (19)) strongly resembles the description of the ortho-GAF averaged RDC (Eq. (14)). This similarity derives from the assumption, common to both models, that the reorientation occurs about the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis, although the dynamic behaviour of the N–H^N vector under these dynamic regimes is quite different, in the GAF case a continuum of conformations is sampled, whereas under the two-site jump, the vector samples two discrete conformations. In Fig. 1 numerical values for s_1, s_1', s_2, and s_2' terms in Eqs. (15) and (20) for different angular σ and θ amplitudes for GAF and two-site jump models are shown. The models are evidently equivalent with respect to the target function until σ and θ are above 40°, leading to equivalent derived χ^2_r for the two models. Here the GAF model is cited when the models are equivalent, although for angles below 40° the two-site jump is equally valid. This of course means that all analysis presented here for the 1D-GAF model is also valid for the two-site jump model up to an amplitude of 40°.

4.2. Differentiation between axially symmetric and anisotropic motional models

The ability to distinguish between different motional models has been explored using numerical simulation (Table 1). Not surprisingly, when dynamically averaged RDCs under a given regime are modelled using the same model, derived order parameters are in excellent agreement with those used to generate data, presenting low values of χ^2_r. Note that the numerical simulations apply a true GAF motion of the peptide plane about the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis, while the analytical model here is a simplification that supposes orthogonality of the N–^{N}H vector relative to the motional axis (in fact this vector is oriented at roughly 80° to the $^{\alpha}C_{i-1}$–$^{\alpha}C_i$ axis). This approximation has a small, but systematic effect, slightly reducing the extracted amplitude (0.73 ± 0.02 instead of 0.71 and 0.39 ± 0.02 instead of 0.36). In general the quality of the fitting is significantly worse when an incorrect motional model is used, indicating that a comparison of χ^2_r for the axially symmetric and anisotropic models should yield the most appropriate regime.

4.3. Effects of structural noise

The potential for incorrectly identifying local dynamics due to structural imprecision has also been explored

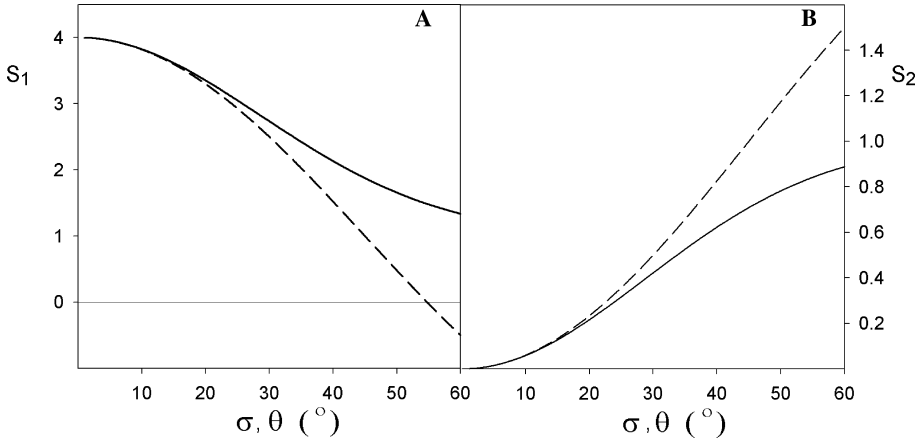

Fig. 1. s_1 (A) and s_2 (B) terms from Eqs. (14) and (19), corresponding to the GAF model, solid lines, and two-site jump model, dashed lines, with respect to the internal dynamics parameter σ or θ.

Table 1
Average order parameters derived from fitting of dynamically averaged RDCs to axially symmetric and GAF motional models

Simulated model/amplitude	Ortho-GAF model		Axially symmetric model	
	$\langle S^2\rangle$[a]	$\langle\chi^2\rangle$[b]	$\langle S^2\rangle$[a]	$\langle\chi^2\rangle$[b]
1D-GAF $\sigma = 20°$, $S^2 = 0.71$	0.73 ± 0.03	1.4	0.72 ± 0.07	6.0
1D-GAF $\sigma = 40°$, $S^2 = 0.36$	0.39 ± 0.02	4.9	0.31 ± 0.09	85.8
Ax.Sym. $S = 0.843$, $S^2 = 0.71$	0.78 ± 0.06	9.7	0.71 ± 0.03	0.5
Ax.Sym. $S = 0.597$, $S^2 = 0.36$	0.50 ± 0.08	59.8	0.36 ± 0.03	0.5

[a] Averaged order parameters over all selected sites.
[b] Averaged χ^2 function over all selected sites.

using numerical simulation. Motional parameters obtained from GAF and axially symmetric motional models in the presence of increasing degrees of structural noise are shown in Table 2.

Average GAF amplitudes (σ) derived from noisy structures are very close to the simulation values, at least for structural noise levels less than 10°. At this higher level of structural noise, small amplitude motions can be increased slightly (average of 27° compared to 20°), but we note that only 13% of the 500 sites are accepted. Of these only 2% (10/500) satisfy the χ^2 criterion, the rest being accepted via the F test, and these 10 sites have amplitudes in the range $(21 \pm 4)°$. For larger motions the extracted amplitude closely resembles the simulated amplitude, even at the highest level of structural noise, indicating that this approach is quite robust with respect to this source of error. Importantly, application of statistical testing for model selection decreases the standard deviation significantly by removing the sites that are more sensitive to aliasing in the presence of structural noise. Thus in the presence of 10° structural noise standard deviations become 4.0°, 2.5°, and 1.6° smaller for 20°, 30°, and 40° amplitude of motion after model selection. For a given degree of structural noise, when the simulated dynamic amplitude increases, the number of residues that is retained using the selection criteria increases. This is because the larger amplitude averaging dominates the effects of structural error, and the minimum of the target function is better defined.

The same kind of analysis was performed with RDCs simulated using axially symmetric motion, again summarised in Table 2. There is again a systematic decrease in the standard deviation of the derived parameter S after model selection, and the same tendency is observed in terms of percentage of residues selected. However, the underestimation of the average S values (overestimation of motional amplitude) when structural imperfection increases is more pronounced, and equally significant for large amplitude motions.

Fig. 2 shows derived dynamic parameters from anisotropic and axially symmetric models, expressed in terms of the respective order parameters, S^2 for simulated dynamics equivalent to S^2 of 0.71 and 0.36. As mentioned above, average ortho-GAF-derived order parameters, $\langle S^2_{\text{O-GAF}}\rangle$, are slightly overestimated in the absence of structural noise, due to the fact that the N–H^N vector is not actually perpendicular to the rotation axis. As structural noise increases the extracted amplitude falls, but remains closer to simulated values than the axially symmetric order parameters $\langle S^2_{\text{axial}}\rangle$, that are more rapidly affected by the presence of structural noise. Notably the standard deviation of GAF-derived order

Table 2
Effects of structural noise in derived motional amplitudes using ortho-GAF and axially symmetric motional models

	0° Struct. noise			3° Struct. noise			5° Struct. noise			10° Struct. noise		
σ_{simul}[a]	$\langle\sigma_{\text{all}}\rangle$[b]	$\langle\sigma_{\text{sel}}\rangle$[c]	%[d]	$\langle\sigma_{\text{all}}\rangle$[b]	$\langle\sigma_{\text{sel}}\rangle$[c]	%[d]	$\langle\sigma_{\text{all}}\rangle$[b]	$\langle\sigma_{\text{sel}}\rangle$[c]	%[d]	$\langle\sigma_{\text{all}}\rangle$[b]	$\langle\sigma_{\text{sel}}\rangle$[c]	%[d]
20°	18.9 ± 1.5	19.0 ± 1.2	98	19.2 ± 3.6	19.9 ± 2.9	66	19.2 ± 5.5	21.4 ± 4.7	33	20.8 ± 9.2	27.1 ± 5.7	13
30°	28.3 ± 2.1	28.3 ± 1.6	98	28.6 ± 3.3	28.7 ± 2.4	93	28.7 ± 4.3	29.2 ± 3.0	80	29.9 ± 7.6	32.5 ± 5.1	43
40°	37.5 ± 2.5	37.6 ± 2.1	99	37.8 ± 3.5	37.8 ± 2.5	96	38.0 ± 3.9	38.1 ± 3.3	95	39.1 ± 6.7	39.9 ± 5.1	79
S_{simul}[e]	$\langle S_{\text{all}}\rangle$[b]	$\langle S_{\text{sel}}\rangle$[c]	%[d]	$\langle S_{\text{all}}\rangle$[b]	$\langle S_{\text{sel}}\rangle$[c]	%[d]	$\langle S_{\text{all}}\rangle$[b]	$\langle S_{\text{sel}}\rangle$[c]	%[d]	$\langle S_{\text{all}}\rangle$[b]	$\langle S_{\text{sel}}\rangle$[c]	%[d]
0.84	0.84 ± 0.02	0.84 ± 0.02	100	0.83 ± 0.07	0.82 ± 0.06	65	0.82 ± 0.11	0.79 ± 0.11	32	0.77 ± 0.17	0.62 ± 0.17	16
0.71	0.71 ± 0.02	0.71 ± 0.02	100	0.70 ± 0.07	0.70 ± 0.06	95	0.69 ± 0.10	0.69 ± 0.08	83	0.66 ± 0.16	0.60 ± 0.14	42
0.60	0.60 ± 0.02	0.60 ± 0.02	100	0.59 ± 0.06	0.59 ± 0.06	98	0.59 ± 0.09	0.59 ± 0.08	94	0.56 ± 0.14	0.54 ± 0.11	74

[a] 1D-GAF amplitude used for numerical averaging.
[b] Average and standard deviation over 500 extracted dynamical parameters corresponding to all sites of the virtual protein.
[c] Average and standard deviation over those sites that pass the model acceptance criteria explained in Section 3.
[d] Percentage of residues that pass the model acceptance criteria.
[e] Order parameter used to create dynamically averaged RDCs under the axially symmetric regime.

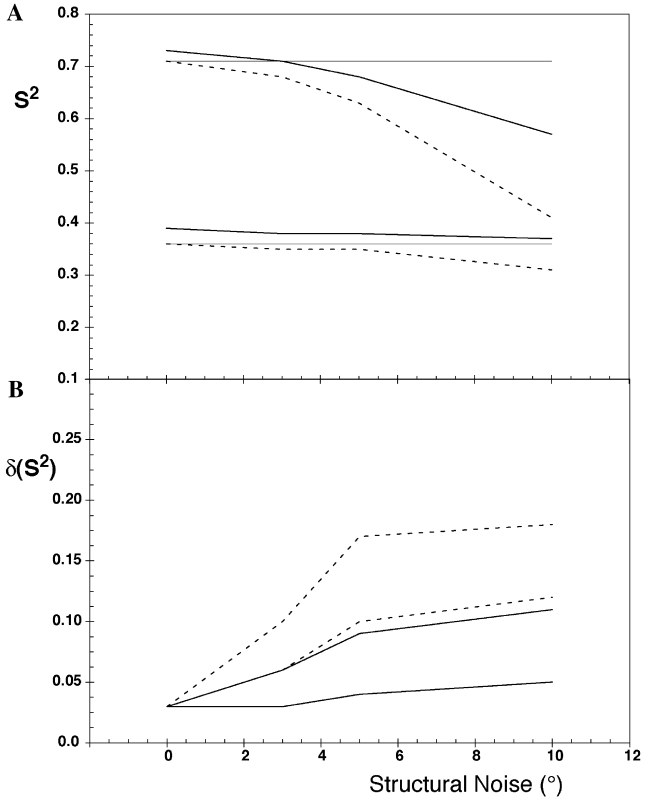

Fig. 2. Effects of structural noise in order parameters derived from 1D-GAF and axially symmetric motional models. (A) Solid lines: Averaged order parameters over all 500 sites that fulfilled filtering criteria applying the ortho-GAF motional model to model data simulated with $S^2 = 0.71$ (top line) and $S^2 = 0.36$ (bottom line) 1D-GAF motional amplitudes. Dashed lines: Averaged order parameters over all sites that fulfilled filtering criteria applying the axially symmetric motional model to model data simulated with $S^2 = 0.71$ (top line) and $S^2 = 0.36$ (bottom line) axially symmetric motional amplitudes. (B) Average of the Monte-Carlo derived standard deviations over the same sites. Solid lines: Ortho-GAF motional model with $S^2 = 0.71$ (bottom line) and $S^2 = 0.36$ (top line). Dashed lines: Axially symmetric motional model with $S^2 = 0.71$ (bottom line) and $S^2 = 0.36$ (top line).

parameters are systematically smaller, indicating that the degree of uncertainty one should expect in the presence of a high levels of structural noise is lower for this model. This suggests that amplitudes extracted using the axially symmetric motional model are less robust with respect to structural noise than those derived from the anisotropic model, due, no doubt to the directionality of the anisotropic dynamic averaging that can be less easily aliased by the presence of randomly directed local structural error.

We note that 10° of structural noise is quite high, this can be shown if we look at the quality of the fit of static data from the perfect structure against the structures with pervasive 0°, 3°, 5°, and 10° noise (data not shown). Even in the absence of systematic noise or dynamics 10° noise levels throughout the protein clearly give worse correlations than those often found between crystal structures and solution state RDC, again underlining the fact that these levels of noise cannot be present throughout the protein in the general case, while at 5°, the correlation is no better than commonly observed correlations.

We have described three quite different motional averaging models here, selected for their inherent simplicity and intuitive nature. These characteristics by no means ensure their validity, however, and it is clear that more complex modes of motion will almost certainly be present in proteins. It may be that the most appropriate analysis of the ability of these models to extract dynamic information from RDCs would be to determine motional amplitudes from a molecular dynamics simulation [12], and to compare the extracted amplitudes with those actually occurring.

4.4. Estimation of the alignment tensor

The importance of accurate estimation of the alignment tensor can be demonstrated relatively simply using this simulation system. RDCs averaged by GAF motions of pervasive amplitude 20° were numerically simulated in the presence of the three independent alignment tensors from throughout the peptide chain. Four 20 amino acid sections undergoing 1D-GAF motions of amplitude 40° were interspersed in this sequence. Not surprisingly extraction of accurate amplitudes is achieved throughout the protein if the correct, non-averaged eigenvalues are used to describe the tensors. However, if the tensors are derived by fitting to the structure assuming no motion, and amplitudes are determined from the data with respect to these scaled tensors, an average shift of 0.14 (0.85 compared to the expected 0.71) is observed in the extracted S^2 in the regions undergoing 20° motions. Interestingly in the regions where the larger amplitude is present, values that are closer to the simulated values are extracted, but in this case the quality of the fitting is generally very poor (Fig. 3).

4.5. Fitting of different models in protein G

The three different dynamic models have been applied to complete the recently published study of backbone conformational sampling in protein G where RDCs measured in five different alignment media were used in this study to fit ortho-GAF amplitudes [23]. Here, all available N–H^N RDCs were fitted using the three motional models described above. The profile of derived S^2 from the best-fitting models is shown in Fig. 4. No sites exist where the two-site jump model is required. As previously observed, an excellent correlation between S^2 derived from RDC and NMR relaxation analysis [29] is found for protein G [22], although in this case a number of additional sites could be fitted using the axially

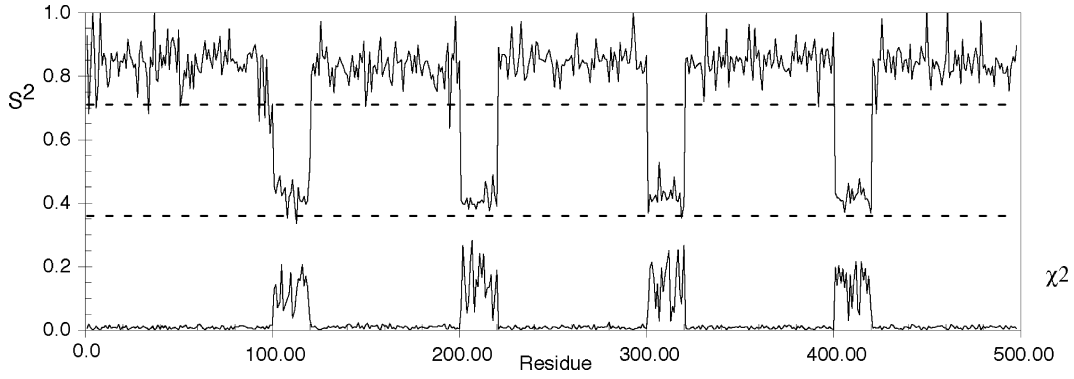

Fig. 3. The effects of extracting motional amplitudes (S^2) using incorrect alignment tensor parameters. In this case the alignment tensor parameters were optimised relative to the structure using dynamically averaged N–^{N}H RDC data. A simulated GAF amplitude of 20° was imposed throughout the molecule, with four regions of 20 amino acids containing 40° GAF motion. The alignment tensor determined using this data is described by $A_a = 18.00$, $A_r = 11.76 \times 10^{-4}$, while the perfect tensor has $A_a = 20.00$, $A_r = 13.00 \times 10^{-4}$. The extracted ortho-GAF amplitudes are therefore significantly higher than the simulated dynamics amplitudes (shown by the dashed lines). The local χ^2 function is plotted at the bottom of the figure.

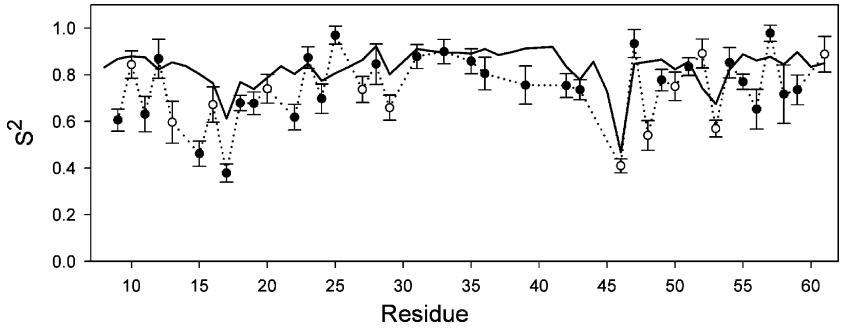

Fig. 4. Order parameter profiles for protein G derived from relaxation, solid lines, and RDCs measured in five alignment media, dashed lines. Symbols correspond to the selected motional model, open circles for axially symmetric, and filled circles for GAF model.

symmetric motional model. Again in the α-helix the RDC-derived S^2 are similar to the relaxation ones, suggesting a high integrity of this region up to the ms time-scale, although several sites present larger additional motions compared to the relaxation-derived motion, suggesting dynamic activity on the time-scale probed only by RDCs. Both anisotropic and axially symmetric models are present along the backbone. The ortho-GAF model dominates, representing 28/40 of the sites. No evident correlation is found between selected model or dynamic amplitude and topology, although the α-helix presents mainly GAF-like motions. Interestingly, in those sites where significantly different amplitudes exist, the GAF model is preferable, with lower χ^2_r values (e.g., residues K9, L17, and T22). An exception to this behaviour is residue G46 that is much better described when using an axially symmetric dynamic model. G46 is highly flexible and observes the largest dynamic amplitude in the ps-ns time-scale [29] $S^2 = 0.47 \pm 0.01$. Under these circumstances a more dynamically active axis than is supposed in the two anisotropic motional models could be expected, and an axially symmetric motion could better account for this.

5. Conclusions

Residual dipolar couplings represent a unique observable to probe dynamics up to the millisecond time-scale. In previous articles we have applied the GAF motional model for the extraction of backbone dynamics from N–^{N}H RDCs. In this article we have tested two additional motional models: the commonly used axially symmetric and a simple two-site jump model. Analytical development of this latter model under the ortho approximation (assuming the N–H^{N} vector to be perpendicular to $^{\alpha}$C–$^{\alpha}$C axis) highlights the fact that it is impossible to distinguish between these averaging processes unless the amplitude of the motion/jump is larger than 40°.

The approach discussed here requires a structural model in order to extract dynamics from RDCs. This requirement renders the approach particularly sensitive to the precision of this structural model. To minimise the possibility of aliasing structural noise to fictitious motional amplitudes we have designed a filtering protocol ensuring the robustness of the methodology based on statistical testing and the precision of the extracted

dynamic parameter. Here we have performed extensive numerical simulations to analyse the behaviour of the approach in the presence of structural noise, specifically exploring the possibility of aliasing structural noise into fictitious backbone dynamic amplitudes. These simulations clearly demonstrate that the anisotropic motional model extracts accurate motional amplitudes even in the presence of significant structural error. The model is geometrically more demanding than the axially symmetric model, and as a consequence there is much less likelihood that structural noise can be wrongly interpreted as dynamics. Indeed, systematic overestimation of dynamics and lack of precision is observed when fitted with axially symmetric model in highly noisy structures. Finally it is clear that the methodology proposed here is limited by the angular sampling available to the $N–^{N}H$ vector, and that more robust local dynamic modelling would be available, and data from fewer alignment media would be required, if multiple different couplings could be modelled simultaneously from each peptide plane. We are currently working toward this goal.

Acknowledgments

This work is supported by the Commissariat à l'Energie Atomique and the Centre Nationale de Recherche Scientifique. P.B. is a recipient of an EMBO post-doctoral fellowship.

References

[1] N. Tjandra, A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, Science 278 (1997) 1111–1114.
[2] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, J.H. Prestegard, Nuclear magnetic dipole interactions in field-oriented proteins information for structure determination in solution, Proc. Natl. Acad. Sci. USA 92 (1995) 9279–9283.
[3] J.H. Prestegard, H.M. Al-Hashimi, J.R. Tolman, NMR structures of biomolecules using field oriented media and residual dipolar couplings, Q. Rev. Biophys. 33 (2000) 371–424.
[4] A. Bax, G. Kontaxis, N. Tjandra, Dipolar couplings in macromolecular structure determination, Methods Enzymol. 339 (2001) 127–174.
[5] E. de Alba, N. Tjandra, NMR dipolar couplings for the structure determination of biopolymers in solution, Prog. Nucl. Magn. Reson. Spectrosc. 40 (2002) 175–197.
[6] J.R. Tolman, Dipolar couplings as a probe of molecular dynamics and structure in solution, Curr. Opin. Struct. Biol. 11 (2001) 532–539.
[7] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, J.H. Prestegard, NMR evidence for slow collective motions in cyanometmyoglobin, Nat. Struct. Biol. 4 (1997) 292–297.
[8] A. Bax, N. Tjandra, Are proteins even floppier than we thought? Nat. Struct. Biol. 4 (1997) 254–256.
[9] M.W.F. Fischer, J.A. Losonczi, J.L. Weaver, J.H. Prestegard, Domain orientation and dynamics in multidomain proteins from residual dipolar couplings, Biochemistry 38 (1999) 9013–9022.
[10] F.F. Tian, H.M. Al-Hashimi, J.L. Craighead, J.H. Prestegard, Conformational analysis of a flexible oligosaccharide using residual dipolar couplings, J. Am. Chem. Soc. 123 (2001) 485–492.
[11] J.R. Tolman, H.M. Al-Hashimi, L.E. Kay, J.H. Prestegard, Structural and dynamic analysis of residual dipolar coupling data, J. Am. Chem. Soc. 123 (2001) 1416–1424.
[12] J. Meiler, J.J. Prompers, W. Peti, C. Griesinger, R. Brüschweiler, Model-free approach do the dynamic interpretation of residual dipolar couplings in globular proteins, J. Am. Chem. Soc. 123 (2001) 6098–6107.
[13] J.R. Tolman, 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. 124 (2002) 12020–12030.
[14] W. Peti, J. Meiler, R. Brüschweiler, C. Griesinger, Model-free analysis of protein backbone motion from residual dipolar couplings, J. Am. Chem. Soc. 124 (2002) 5822–5833.
[15] K.B. Briggman, J.R. Tolman, De novo determination of bond orientations and order parameters from residual dipolar couplings, J. Am. Chem. Soc. 125 (2003) 10164–10165.
[16] R. Brüschweiler, P.E. Wright, New order parameters of biomolecules: a new analytical representation and application to the gaussian axial fluctuation model, J. Am. Chem. Soc. 116 (1994) 8426–8427.
[17] T. Bremi, R. Brüschweiler, Locally anisotropic internal polypeptide backbone dynamics by NMR relaxation, J. Am. Chem. Soc. 119 (1997) 6672–6673.
[18] S.F. Lienin, T. Bremi, B. Brutscher, R. Brüschweiler, R.R. Ernst, Anisotropic intramolecular dynamics of ubiquitin characterized by NMR relaxation and MD computer simulation, J. Am. Chem. Soc. 120 (1998) 9870–9879.
[19] M.W.F. Fischer, L. Zeng, Y. Pang, W. Hu, A. Majumdar, E.R.P. Zuiderweg, Experimental characterization of models for backbone picosecond dynamics in proteins. Quantification of NMR auto- and cross-correlation relaxation mechanisms involving different nuclei of the peptide plane, J. Am. Chem. Soc. 119 (1997) 12629–12642.
[20] Y. Pang, E.R.P. Zuiderweg, Determination of protein backbone ^{13}CO chemical shift anisotropy tensors in solution, J. Am. Chem. Soc. 122 (2000) 4841–4842.
[21] P. Bernadó, M. Blackledge, Anisotropic small amplitude peptide plane dynamics in proteins from residual dipolar couplings, J. Am. Chem. Soc. 126 (2004) 4907–4920.
[22] P. Bernadó, M. Blackledge, Local dynamic amplitudes on protein backbone from dipolar couplings: towards the elucidation of slower motions in biomolecules, J. Am. Chem. Soc. 126 (2004) 7760–7761.
[23] T.S. Ulmer, B.E. Ramirez, F. Delaglio, A. Bax, Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy, J. Am. Chem. Soc. 125 (2003) 9179–9191.
[24] N.R. Skrynnikov, N.K. Goto, D. Yang, W.-Y. Choy, J.R. Tolman, G.A. Mueller, L.E. Kay, Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with β-cyclodextrin, J. Mol. Biol. 295 (2000) 1265–1273.
[25] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, J. Am. Chem. Soc. 104 (1982) 4559–4570.
[26] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, J. Am. Chem. Soc. 104 (1982) 4546–4559.
[27] J.P. Derrick, D.B. Wigley, The third IgG-binding domain from Streptococcal protein G: an analysis by X-ray crystallography of

the structure alone and in a complex with Fab, J. Mol. Biol. 243 (1994) 906–918.
[28] P. Dosset, J.-C. Hus, D. Marion, M. Blackledge, A novel interactive tool for rigid-body modelling of multi-domain macromolecules using residual dipolar couplings, J. Biomol. NMR 20 (2001) 223–231.
[29] J.B. Hall, D. Fushman, Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G, J. Biomol. NMR 27 (2003) 261–275.