Anisotropic Collective Motion Contributes to Nuclear Spin Relaxation in Crystalline Proteins

Józef R. Lewandowski,[†] Julien Sein,[†] Martin Blackledge,[*,‡] and Lyndon Emsley[*,†]

Université de Lyon, CNRS/ENS-Lyon/UCB-Lyon 1, Centre de RMN à Très Hauts Champs, 69100 Villeurbanne, France, and Protein Dynamics and Flexibility, Institut de Biologie Structurale Jean Pierre Ebel, UMR 5075, CNRS/CEA/UJF, 38027 Grenoble, France

Received August 20, 2009; E-mail: martin.blackledge@ibs.fr; lyndon.emsley@ens-lyon.fr

The development of an atomic-resolution description of protein structure and dynamics is crucial for understanding numerous biophysical processes. Molecular motion is at the core of such important phenomena as enzymatic catalysis, molecular recognition, signaling, ligand binding, and protein folding.[1] NMR spectroscopy allows access to site-specific information about biomolecular motions in both solution and the solid state. For NMR analysis, one of the fundamental differences between solution and solid samples is that overall isotropic tumbling is absent in the latter. This makes solid-state NMR spectroscopy a very attractive technique for probing protein motions, implying that site-specific differences in relaxation report uniquely on local motions. Indeed, solid-state NMR studies of microcrystalline proteins to date have generally assumed that only internal random motion makes a significant contribution to relaxation rates. However, even in the absence of overall isotropic tumbling, it is possible that anisotropic collective motion (ACM) exists in microcrystalline solids, either in the form of small-amplitude overall reorientation or as other collective motions that are likely to be pertinent to function.[2] Numerous studies have inferred concerted atomic displacement on the basis of X-ray data,[3] indicating that rigid-body motions may on average account for up to 60% of the B factors.[4−6]

Here we show how ACM in solids should contribute to, and thus could be determined by, experimental NMR relaxation rates. We develop a model for ACM-induced longitudinal spin relaxation and show that ACM compatible with X-ray B-factor data can explain a large part of ^{15}N R_1 values in the model protein Crh.

It is important to note that in general, the models for describing protein dynamics in solution, such as the commonly used "model-free" approach, cannot be used in unmodified form to describe protein dynamics in solids without incurring errors. Fortunately, many models can be adapted for describing motions in spinning solids, provided that the orientation dependence of magic-angle spinning (MAS) and powder averaging are properly treated. In this spirit, Giraud and co-workers[7,8] demonstrated that after a number of modifications, the "diffusion in a cone" model could be used for describing local uncorrelated motions in crystalline solids.

Here we have adapted the three-dimensional Gaussian axial fluctuation (3D GAF)[9] model, which is used in solution for analyzing local motions, in a manner that renders it suitable for evaluation of ACM contributions to R_1 values in solids. In our 3D GAF model, the collective motion of a protein segment is described
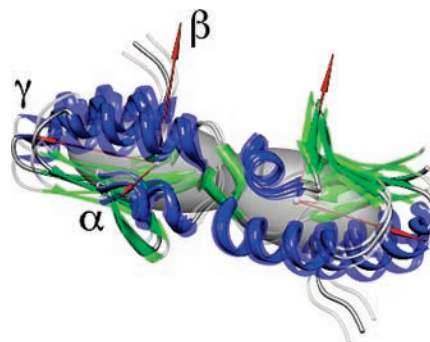
Figure 1. Representation of the 3D GAF model developed to compute the contribution of ACM to spin−lattice relaxation rates. Here, the GAF axes coincide with the axes of the inertia tensors of the monomers of the Crh dimer (ellipsoids).

by GAFs about three orthogonal axes α, β, and γ with amplitudes σ_α, σ_β, and σ_γ (see Figure 1).

The 3D GAF internal correlation function of the NH dipolar vector fluctuations in a protein segment, $C_{NH}(t)$, can be expressed as follows:

$$C_{NH}(t) = \frac{4\pi}{5} \sum_{l,k_1,k_2,m_1,m_2=-2}^{2} e^{-i\pi(k-k_1)/2}$$
$$\langle e^{-ik_2\alpha(t)+ik_1\alpha(0)} \rangle \langle e^{il\beta(t)-il\beta(0)} \rangle \langle e^{im_2\gamma(t)-im_1\gamma(0)} \rangle$$
$$d_{k_1l}^{(2)}\left(\frac{\pi}{2}\right) d_{k_2l}^{(2)}\left(\frac{\pi}{2}\right) d_{m_1k_1}^{(2)}\left(\frac{\pi}{2}\right) d_{m_2k_2}^{(2)}\left(\frac{\pi}{2}\right) Y_{2,m_1}(e_{NH}) Y_{2,m_2}^{*}(e_{NH}) \quad (1)$$

in which the polar coordinate set $e_{NH} = (\theta_{NH}, \varphi_{NH})$ represents the N−H vector in the instantaneous molecular frame (α, β, γ) of the protein segment, $Y_{2,m}$ is a second-degree spherical harmonic, and $d_{kl}^{(2)}(\phi)$ is a reduced Wigner matrix element. The derivation of eq 1 in the context of local uncorrelated motions and the forms of the dihedral correlation functions $\langle e^{im_2\gamma(t) - im_1\gamma(0)} \rangle$ can be found in ref 10 and are also detailed in the Supporting Information (SI). We use eq 1 to individually calculate $C_0(t)$, $C_1(t)$, and $C_2(t)$ (where the subscripts indicate the orders of the spherical harmonics) for any given set of fluctuation amplitudes σ_α, σ_β, and σ_γ and GAF correlation times $\tau_{GAF,\alpha}$, $\tau_{GAF,\beta}$, and $\tau_{GAF,\gamma}$ [where $\tau_{GAF} = 1/6D$, where D is a diffusion constant for the diffusion process in a harmonic potential; τ_{GAF} should not be confused with the effective correlation time $\tau_{eff,m}$, i.e., the characteristic decay time of $C_m(t)$ approximated by a single exponential]. To obtain spectral densities, and in the absence of an analytical expression, each of the normalized correlation functions is then fitted to a sum of three exponentials: $S^2 + a_1e^{-t/\tau_1} + a_2e^{-t/\tau_2} + a_3e^{-t/\tau_3}$, where $S^2 + a_1 + a_2 + a_3 = 1$ (see Figure SI6 in the SI for typical fits).

[†] Université de Lyon.
[‡] Institut de Biologie Structurale Jean Pierre Ebel.

The seven parameters (i.e., a_1, a_2, a_3, τ_1, τ_2, τ_3, and S^2) for each correlation function are used to define the spectral densities $J_2(\omega)$, $J_1(\omega)$, and $J_0(\omega)$ in the molecular frame as sums of Lorentzians. Subsequently, the spectral densities in the molecular frame are brought to the laboratory frame by a consecutive series of transformations, as shown in eq 2:

$$J_a(\omega) = \sum_{b,b'=-2}^{2} \sum_{m=-2}^{2} D_{mb}^{(2)*}(\Omega_{CM}) D_{ba}^{(2)*}(\Omega_{ML}) D_{mb}^{(2)}(\Omega_{CM}) D_{b'a}^{(2)}(\Omega_{ML}) J_m(\omega) \quad (2)$$

where $D_{ij}^{(2)}$ is an element of the Wigner matrix initiating the change of frame through the Euler angles. The effects of MAS and powder averaging are determined through the procedure given in ref 7, where the transformed spectral densities are used to calculate R_1 values for each orientation β_{CM}. The contributions from both the fluctuation of the NH dipolar vector and the ^{15}N chemical shift anisotropy are considered,[6] yielding:

$$R_1^{MAS}(\beta_{CM}) = \frac{1}{4}\left(\frac{\mu_0}{4\pi}\frac{\gamma_H\gamma_N}{\langle r_{NH}\rangle^3}\hbar\right)^2 [J_0^{MAS}(\omega_H - \omega_N) +$$
$$3J_1^{MAS}(\omega_N) + 6J_2^{MAS}(\omega_H + \omega_N)] + \frac{1}{3}\omega_N^2(\sigma_\| - \sigma_\perp)^2 J_1^{MAS}(\omega_N) \quad (3)$$

In the final step, we calculate the explicit averaged sum (EAS)[7] over the range of β_{CM} values to obtain the average R_1 in the laboratory frame. MATLAB[11] code for simulating ACM-induced ^{15}N R_1 values is available from our Web site.[12]

Using our implementation of the 3D GAF model, we can evaluate the extent to which the Crh relaxation data can be explained by the presence of small-amplitude ACM. In the first step, we explore how collective motions predicted by approaches relying on the analysis of X-ray B factors could contribute to ^{15}N R_1. It should be noted that in general it is difficult or impossible to separate the static and dynamic disorder contributions to B factors, and thus, the validity of such analyses can vary considerably.[13] Assuming here a non-negligible contribution of dynamic disorder, we analyzed the B factors from the X-ray diffraction structure of the Crh dimer[14] using several methods, including normal mode analysis (NMA),[15] the Gaussian network model (GNM),[16] the anisotropic network model (ANM)[17] and the translation libration screw (TLS) model[18−20] (see the SI for details of the analyses). All of the models yielded a picture compatible with the structure undergoing a small-amplitude scissoring motion with the largest rotation around the axis parallel to the α axis of Figure 1. Since the picture of collective motions that emerged was roughly the same from all the methods, for the sake of computational convenience we used the TLS model as implemented in the publicly available TLSMD algorithm[20] to probe the extent to which, if present, such motions could contribute to the ^{15}N R_1 values. In this approach, protein chains are divided into multiple segments that are modeled as rigid bodies undergoing TLS motions. Here we report on the two relative extremes: the one- and six-segment TLS analysis of chain A of the Crh dimer. One-segment TLS, in which the entire monomer is treated as a single rigid body (see Figure 1), is the simplest approximation. In the six-segment TLS analysis, each monomer is treated as six separate rigid bodies (residues 1−10, 11−28, 29−40, 41−56, 57−65, and 66−86, corresponding approximately to the secondary structure elements treated as rigid bodies). The six-segment TLS, having 6 times more
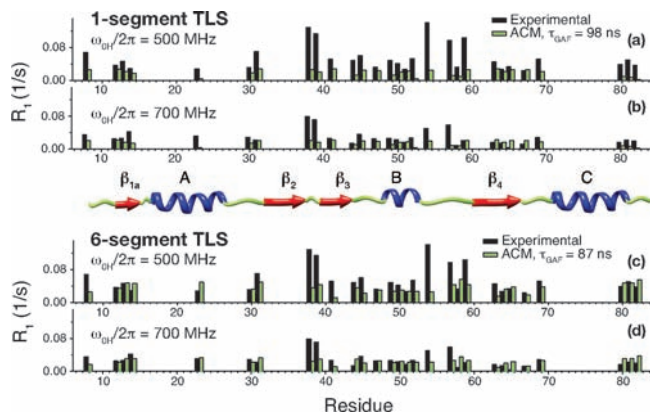
Figure 2. Comparison between experimental ^{15}N R_1 values in microcrystalline Crh [black bars; (a, c) $\omega_{0H}/2\pi = 500$ MHz, (b, d) $\omega_{0H}/2\pi = 700$ MHz] and R_1 values simulated using the 3D GAF model with segmental axial fluctuations predicted by TLS analysis of X-ray B factors (green bars). The correlation time of the motion was treated as a single fit parameter in an experimental-error-weighted fit. The best-fit correlation times were (a, b) $\tau_{GAF} = 98$ ns for one-segment TLS amplitudes and (c, d) $\tau_{GAF} = 87$ ns for six-segment TLS amplitudes. The axial fluctuations from one-segment (residues 1−86) TLS analysis were $\sigma_\alpha = 4.3°$, $\sigma_\beta = 1.6°$, $\sigma_\gamma = 0.01°$, and those from six-segment TLS analysis were: (residues 1−10) $\sigma_\alpha = 8.6°$, $\sigma_\beta = 0.01°$, $\sigma_\gamma = 2.8°$; (residues 11−28) $\sigma_\alpha = 6.7°$, $\sigma_\beta = 2.1°$, $\sigma_\gamma = 0.01°$; (residues 29−40) $\sigma_\alpha = 7°$, $\sigma_\beta = 3.2°$, $\sigma_\gamma = 0.01°$; (residues 41−56) $\sigma_\alpha = 4.4°$, $\sigma_\beta = 0.01°$, $\sigma_\gamma = 2.9°$; (residues 57−65) $\sigma_\alpha = 9.4°$, $\sigma_\beta = 2.4°$, $\sigma_\gamma = 0.01°$; (residues 66−86) $\sigma_\alpha = 6.8°$, $\sigma_\beta = 0.01°$, $\sigma_\gamma = 3.5°$. The 3D GAF axes coincide with the axes of inertia for the segments.

fitted parameters, naturally leads to better agreement between the simulated and experimental B factors (Figure SI1). It should be noted that this is not necessarily equivalent to a better description of the internal dynamics in the absence of a separation of static and dynamic disorder contributions to the B factors.

We next took the results of the TLS analysis, which reproduced the B factors, and computed ^{15}N R_1 values due to the rotational part of the motion. It should also be noted that for the simulations, the 3D GAF axes were chosen to coincide with the axes of inertia frame for each segment (as calculated using TENSOR2[21]).

We note that TLS provides only tentative amplitudes and directions of the motions but not the time scales. The question remains whether such small-amplitude motions occurring on the expected nanosecond time scale would yield appreciable contribution to R_1. To answer this question, we fit the experimental Crh R_1 values obtained at $\omega_{0H}/2\pi = 500$ MHz and $\omega_{0H}/2\pi = 700$ MHz to an ACM model assuming motional amplitudes from TLS analysis with a single GAF correlation time ($\tau_{GAF,\alpha} = \tau_{GAF,\beta} = \tau_{GAF,\gamma} = \tau_{GAF}$) as a fit parameter. Figure 2 shows results of such fits with axial fluctuations from one-segment (Figure 2a) and six-segment (Figure 2b) TLS analyses. In both cases, the simulated R_1 values due to a few degrees of ACM account for a significant fraction of the experimental rates, with larger overall contributions for the six-segment TLS amplitudes with $\tau_{GAF} = 87$ ns (which correspond to $\tau_{eff,1} < 12$ ns).

It is important to point out that the ACM contributions to R_1 values are strongly dependent on the direction of the NH bonds with respect to the motional axes. This means that in the presence of collective motion of a segment, the motion contributions can vary significantly (or even be undetectable for some residues) across the segment when the analysis is performed locally.

Since B factors in principle may contain a substantial contribution from static disorder, it is instructive to compare motions predicted
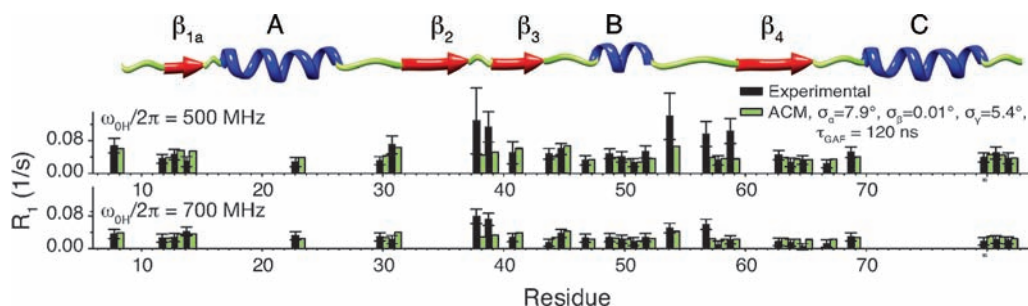
Figure 3. Comparison between measured ^{15}N R_1 values in Crh [black bars; (top) 11.74 T, (bottom) 16.45 T] and R_1 values simulated using the 3D GAF model with Crh treated as a single rigid body (green bars). Both the GAF correlation time and fluctuation amplitudes were fitted simultaneously to spin−lattice rates measured at 11.7 and 16.5 T. The best-fit correlation time was $\tau_{GAF} = 120$ ns, and the axial fluctuations were $\sigma_\alpha = 7.9°$, $\sigma_\beta = 0.01°$, $\sigma_\gamma = 5.4°$. The GAF axes coincide with the axes of inertia for the Crh dimer. Error bars give the experimental error.[7]

on their basis to the motions fitted directly to NMR data without input from X-ray data. Accordingly, in the work discussed in this paragraph, we disregarded the B factor analyses and added the fluctuation amplitudes as fit parameters in our model. For simplicity, we approximated the ACM with one-segment fluctuations with the 3D GAF axes coinciding with the inertia axes of the Crh dimer. Figure 3 shows the results of such a fit. Overall, the ACM of the dimer with $\sigma_\alpha = 7.9°$, $\sigma_\beta = 0.01°$, $\sigma_\gamma = 5.4°$, and $\tau_{GAF} = 120$ ns was sufficient to account (within the experimental error) for most of the experimental rates. We note that since at this point we explicitly neglected the short-range motion (SRM; i.e., the fluctuations of adjacent dihedral angles)[22] and accounted only for a single overall motion, it is likely that the fluctuations were overestimated. The omission of SRM can also explain why the largest deviations between the simulated and experimental rates were observed for the residues in loop regions (e.g., D38, G39, and G58): loop residues are expected to contain a larger contribution to R_1 due to the local fluctuations. We also note that $\tau_{eff,1}$ [$C_1(t)$ yields the largest contribution to R_1] here is $\sim$12 ns, which is in the same range as the correlation time for rotational diffusion of Crh in solution ($\tau_c \approx 11$ ns at room temperature). This similarity of time scale may result from similar solvent−protein interactions in crystalline and solution environments. In solution, these interactions lead to complete rotational diffusion, while in the crystal they result only in small-amplitude rocking of the protein within the lattice.

Our study suggests that ACM should be detectable using NMR spectroscopy. Conversely, neglecting ACM could lead to substantial errors in interpretation. Unambiguous determination of ACM requires more experiments, but the perspective is particularly exciting since it provides a probe of the larger-amplitude domain motions that are often involved in processes such as enzymatic catalysis, substrate binding, regulation and allosteric behavior, and motor functions.

In conclusion, we have introduced a 3D GAF model for treating the influence of anisotropic collective motions on site-specific relaxation rates in crystalline proteins. We have demonstrated that small-amplitude (<10°) collective motions compatible with the dynamic picture provided by techniques other than solid-state NMR can lead to a large contribution to the ^{15}N R_1 values if they occur on a relaxation-active time scale. ACM should thus be systematically included in the future analysis of NMR data for dynamics in microcrystalline proteins along with uncorrelated local motions.

With this perspective, more than simplifying the analysis, the absence of isotropic reorientation in solids may provide improved access to the elusive nanosecond collective motions in proteins that are often linked to protein function and folding.

Supporting Information Available: Details of NMA, GNM, ANM, and TLS analyses of the B factors in Crh dimer, including animations of the predicted motions; additional details of the 3D GAF formalism; larger versions of Figures 2 and 3; and fits of correlation functions. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Boehr, D. D.; Dyson, H. J.; Wright, P. E. *Chem. Rev.* **2006**, *106*, 3055.
(2) Gerstein, M.; Lesk, A. M.; Chothia, C. *Biochemistry* **1994**, *33*, 6739.
(3) Kuriyan, J.; Weis, W. I. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2773.
(4) Song, G.; Jernigan, R. L. *J. Mol. Biol.* **2007**, *369*, 880.
(5) Stec, B.; Zhou, R. S.; Teeter, M. M. *Acta Crystallogr.* **1995**, *D51*, 663.
(6) Soheilifard, R.; Makarov, D.; Rodin, G. *Phys. Biol.* **2008**, *5*, 026008.
(7) Giraud, N.; Blackledge, M.; Goldman, M.; Bockmann, A.; Lesage, A.; Penin, F.; Emsley, L. *J. Am. Chem. Soc.* **2005**, *127*, 18190.
(8) Sein, J.; Giraud, N.; Blackledge, M.; Emsley, L. *J. Magn. Reson.* **2007**, *186*, 26.
(9) Bremi, T.; Bruschweiler, R. *J. Am. Chem. Soc.* **1997**, *119*, 6672.
(10) Lienin, S. F. Ph.D. Thesis, Eidgenossische Technische Hochschule, 1998.
(11) *MATLAB*, version 7.6.0.324(R2008a); The Math Works: Natick, MA, 2008.
(12) http://www.ens-lyon.fr/crmn/index.html (accessed Nov 6, 2009); see the Research/Software section.
(13) Moore, P. B. *Structure* **2009**, *17*, 1307.
(14) Juy, M.; Penin, F.; Favier, A.; Galinier, A.; Montserret, R.; Haser, R.; Deutscher, J.; Bockmann, A. *J. Mol. Biol.* **2003**, *332*, 767.
(15) Alexandrov, V.; Lehnert, U.; Echols, N.; Milburn, D.; Engelman, D.; Gerstein, M. *Protein Sci.* **2005**, *14*, 633.
(16) Bahar, I.; Atilgan, A. R.; Erman, B. *Folding Des.* **1997**, *2*, 173.
(17) Eyal, E.; Yang, L.; Bahar, I. *Bioinformatics* **2006**, *22*, 2619.
(18) Schomaker, V.; Trueblood, K. N. *Acta Crystallogr.* **1968**, *B24*, 63.
(19) Painter, J.; Merritt, E. A. *Acta Crystallogr.* **2006**, *D62*, 439.
(20) Painter, J.; Merritt, E. A. *J. Appl. Crystallogr.* **2006**, *39*, 109.
(21) Dosset, P.; Hus, J. C.; Blackledge, M.; Marion, D. *J. Biomol. NMR* **2000**, *16*, 23.
(22) Brüschweiler, R. *J. Chem. Phys.* **1995**, *102*, 3396.

JA907067J