Journal of Biomolecular NMR, **24:** 123–132, 2002.
KLUWER/ESCOM

Principal component method for assessing structural heterogeneity across multiple alignment media

Jean-Christophe Hus & Rafael Brüschweiler*
Carlson School of Chemistry and Biochemistry, Clark University, Worcester, MA 01610, U.S.A.

Received 11 June 2002; Accepted 16 August 2002

Key words: heterogeneous conformational ensemble, multiple alignment media, NMR, principal component analysis, protein structure, residual dipolar couplings

Abstract

The recent availability of residual dipolar coupling measurements in a variety of different alignment media raises the question to what extent biomolecular structure and dynamics are differentially affected by their presence. A computational method is presented that allows the sensitive assessment of such changes using dipolar couplings measured in six or more alignment media. The method is based on a principal component analysis of the covariance matrix of the dipolar couplings. It does not require a priori structural or dynamic information nor knowledge of the alignment tensors and their orientations. In the absence of experimental errors, the covariance matrix has at most five nonzero eigenvalues if the structure and dynamics of the biomolecule is the same in all media. In contrast, differential structural and dynamic changes lead to additional nonzero eigenvalues. Characteristic features of the eigenvalue distribution in the absence and presence of noise are discussed using dipolar coupling data calculated from conformational ensembles taken from a molecular dynamics trajectory of native ubiquitin.

Introduction

The plethora of alignment media that has become available in recent years for the measurement of residual dipolar couplings (RDCs) of biomolecules (for recent reviews see Prestegard et al., 1999; Bax et al., 2001; de Alba and Tjandra, 2002; Al-Hashimi and Patel, 2002) raises the question to what extent biomolecular structure and dynamics are influenced by their presence. Some alignment media, such as phages, correspond to suspensions of oriented particles that are separated by hundreds of Å while others, such as gels, constitute preferentially elongated cavities. For most of the time the macromolecular solutes reside far away from the alignment medium, but when they diffuse within a close distance of the medium they interact via non-bonding interactions that disrupt the isotropy of the orientational distribution leading to the nonzero dipolar couplings. In case these interactions are comparable in strength to intramolecular interactions they may change the biomolecule's structure and its intramolecular dynamics. The magnitude of such effects can be studied by comparison of residual dipolar couplings measured in different alignment media. For example, from dipolar coupling data of T4 lysozyme recorded in two different liquid-crystalline media it was found that the relative domain orientation remains essentially unchanged (Goto et al., 2001).

The recently developed model-free analysis of RDCs (Meiler et al., 2001) provides detailed information on averaging effects in RDC data collected in multiple alignment media. Its application to experimental N-H RDCs of ubiquitin for 11 alignment media indicates a significant amount of intramolecular flexibility (Peti et al., 2002). The method does not discriminate between homogeneous effects, which are protein fluctuations that are identical in all media, and heterogeneous effects, which stem from structural or dynamic changes between different media.

Recently, backbone N-H RDCs of calbindin D9k in the presence (1) of paramagnetic ions, (2) of a non-

*To whom correspondence should be addressed. E-mail: bruschweiler@nmr.clarku.edu

ionic liquid crystalline phase, and (3) of both at the same time were reported (Barbieri et al., 2002). From the good correlation between the RDCs measured for (3) and the sum of the RDCs of (1) and (2) it was concluded, without explicitly using a 3D structural model of the protein, that the orienting medium is innocent with respect to the structure of the protein in solution. However, this conclusion may not hold for all conceivable microscopic models of the alignment process. Let us assume that the picture for the dilute liquid-crystalline alignment medium described above applies, i.e. at a given instant the alignment medium affects only the small fraction of protein molecules in the sample that are close to the liquid crystal molecules. On the other hand, the paramagnetic alignment dominates at a given instant the alignment behavior of the large majority of protein molecules in the sample, which are all molecules except for the small fraction of molecules interacting with the liquid crystal. The RDCs measured in situation (3) therefore correspond in very good approximation to the sum of the RDCs of the paramagnetically aligned protein molecules and the RDCs of the proteins interacting with the liquid crystal molecules. This is the same as when the RDCs are independently measured (situations (1) and (2)) and the corresponding couplings are added a posteriori. This is true whether the interacting and non-interacting protein molecules have the same 3D structure or not. Therefore, for this particular alignment model the above experimental analysis is inconclusive with respect to the effect of the alignment medium on the protein structure.

An analysis method is presented here to assess the effect of the alignment media on protein structure that does not depend on the microscopic details of the alignment mechanism. It is demonstrated that a principal component analysis applied to RDC sets measured in at least six media with different alignment tensors is highly sensitive to structural heterogeneity effects induced by the media. After a presentation of the method, it is applied to synthetic RDC sets calculated from ubiquitin conformations obtained from a molecular dynamics (MD) computer simulation with and without the addition of random Gaussian noise.

Method

Principal component analysis of dipolar couplings

We consider residual dipolar couplings $D_i^{(k)}$ belonging to internuclear vectors $i = 1,..., N$ of a protein measured in M different alignment media $k = 1, ..., M$. From the $M \cdot N$ couplings a $N \times N$ weighted covariance matrix C is calculated according to

$$C_{ij} = \frac{1}{M-1} \sum_{k=1}^{M} w_k (D_i^{(k)} - \langle D_i \rangle)(D_j^{(k)} - \langle D_j \rangle) \quad (1)$$
$$i, j = 1, ..., N,$$

where $\langle D_i \rangle = \frac{1}{M} \sum_{k=1}^{m} D_i^{(k)}$ is the average of coupling D_i over the M alignments. The weights w_k in Equation (1) allow one to balance the importance of the RDC sets belonging to different alignments k. In the following, these weights are set to $w_k = 1/\sigma_k^2$, where $\sigma_k^2 = \frac{1}{N-1} \sum_{i=1}^{N} (D_i^{(k)} - \langle D_i \rangle)^2$ is the variance of the dipolar couplings collected in medium k.

Next, a principal component analysis is applied to the weighted covariance matrix C by diagonalization

$$C|q\rangle = \lambda_q |q\rangle, \quad (2)$$

where $|q\rangle (q = 1, ..., N)$ are the N normalized eigenvectors and λ_q are the corresponding eigenvalues. All eigenvalues fulfill $\lambda_q \geq 0$. The eigenvalues are sorted with the smallest one being λ_1. In the absence of experimental noise, *no more than five eigenvalues differ from zero* for the following two cases:
(a) If the protein adopts a static structure that remains unchanged in all media or
(b) if the protein is present as a multiple conformational ensemble that is identical in all media.

In case (a) the covariance matrix is related to the metric matrix of rank $L = 2$, which exhibits $2L + 1$ nonzero eigenvalues in analogy to the properties of the metric matrix in distance geometry (Prompers and Brüschweiler, 2002; Hus and Brüschweiler, 2002). Case (b) follows from (a) because of the linear averaging behavior of RDCs (cf. Equation 9 of Meiler et al., 2001). The appearance of additional nonzero eigenvalues is therefore direct evidence of structural or dynamic heterogeneity or the presence of experimental errors as is discussed below. The proposed method is termed 'Self-Consistency of Dipolar Couplings Analysis' (SECONDA).

Collectivity of eigenvectors

The collectivity κ, which is a scalar measure for each of the N eigenvectors $|q\rangle$ of C, is defined as (Brüschweiler, 1995; Prompers et al., 2001a):

$$\kappa_q = \frac{1}{N} \exp\left\{ -\sum_{j=1}^{N} \left|\langle q|_j\right|^2 \log \left|\langle q|_j\right|^2 \right\} \cdot 100\%, \quad (3)$$

where $|q\rangle_j$ is the jth component of the normalized eigenvector $|q\rangle$. κ ranges between $100/N$ % and 100%, which is the percentage of the RDCs that significantly contribute to eigenvector $|q\rangle$. The κ_q values directly indicate whether variations of the RDCs between different media occur in a collective or more localized fashion. Each eigenvalue and eigenvector pair can then be conveniently represented by the two numbers λ and κ.

Calculation of dipolar couplings

In the Results section different models of protein behavior and alignment tensors are tested with respect to their effect on the eigenvalue distribution. For this purpose, residual dipolar couplings are calculated for a given protein structure with vector orientations $(\theta_i^{(k)}, \varphi_i^{(k)})$ $(i = 1, ..., N)$ defined in the eigenframe of alignment tensor $\mathbf{D}^{(k)}$ according to

$$\begin{aligned} D_i^{(k)} \;=\; & D_a^{(k)} \left\{ 3\cos^2\theta_i^{(k)} - 1 + \right. \\ & \left. \tfrac{3}{2} R^{(k)} \sin^2\theta_i^{(k)} \cos 2\varphi_i^{(k)} \right\}, \end{aligned} \quad (4)$$

where $D_a^{(k)} = D_{zz}^{(k)}/2$ is the axial component and $R^{(k)} = 2/3 \cdot (D_{xx}^{(k)} - D_{yy}^{(k)})/D_{zz}^{(k)}$ is the rhombicity of the alignment tensor $\mathbf{D}^{(k)}$.

If instead of a single structure a conformational ensemble consisting of P conformers is present, the $D_i^{(k)}$ values of Equation 4 are first averaged over all conformers

$$\overline{D_i^{(k)}} = \frac{1}{P} \sum_{p=1}^{P} D_{i,p}^{(k)}, \quad (5)$$

where $D_{i,p}^{(k)}$ is the RDC belonging to vector i of conformation p in alignment medium k.

Molecular dynamics simulation

Ensembles of conformations of ubiquitin were generated by a MD simulation using the program CHARMM (Brooks et al., 1983; MacKerell et al., 1998). An all-atom representation of the X-ray structure of ubiquitin (Vijay-Kumar et al., 1987) was embedded in a cubic box including 2909 explicit water molecules. The simulation was performed at a temperature of 300 K for 11 ns under periodic boundary conditions. More details on the trajectory are described elsewhere (Lienin et al., 1998; Prompers et al., 2001b). The first 1 ns of the MD simulation was used for equilibration and a total of 1000 snapshots with a time increment of 10 ps were analyzed from the final 10 ns of the trajectory. Two additional ensembles were generated with a finer time resolution (exhibiting smaller average pairwise root-mean-square deviations (RMSD)) by storing 600 snapshots every 1 fs between 10 ns and 10.0006 ns and by storing 600 snapshots every 100 fs between 10 ns and 10.06 ns, respectively. The structures of each ensemble were orientationally aligned with respect to a reference structure.

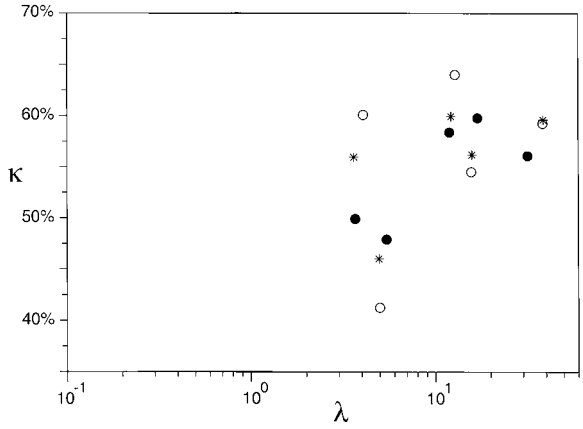

Figure 1. Collectivity κ vs. eigenvalue λ distribution of RDC covariance matrix of 68 backbone N-H vectors of ubiquitin for 11 different alignments calculated from the MD snapshot at 10 ns (filled circles) and a homogeneous protein ensemble represented by 100 MD conformations sampled with a time increment of 100 ps (open circles). The * symbols indicate κ, λ values for an ensemble consisting only of the 20 first MD conformations of the above ensemble.

Results and discussion

Homogeneous protein behavior

Residual dipolar couplings were calculated for 68 backbone N-H vectors of ubiquitin for 11 alignment tensors using the same MD snapshot at 10 ns. The 68 N-H vectors include all backbone N-H vectors of ubiquitin except for the N-terminal group and the N-H groups of the 4 C-terminal residues that exhibit sig-

nificantly increased mobilities. The alignment tensors, each specified by five parameters (3 Euler angles, D_a, and R), were chosen randomly. The allowed values for R lie between 0 and 2/3, whereas the range of D_a was not restricted (due to the division by σ_k^2 in Equation (1), D_a has no influence on the weighted covariance matrix).

The weighted covariance matrix **C** was determined according to Equations 1 and 4 and a principal component analysis was performed by diagonalization of **C**. For each eigenvector $|q\rangle$ the collectivity κ was calculated according to Equation 3. The collectivities κ_j are plotted vs. the eigenvalues λ_j in Figure 1 (κ, λ plot) for the reference structure at 10 ns (filled circles). Only five eigenvalues are finite while the other six eigenvalues are smaller by at least a factor of 10^6. The latter can be considered as zero for all practical purposes. The five finite eigenvalues vary between 3.7 and 31.9 whereas the collectivities range between 47.9% and 59.8%.

The high degree of singularity of the covariance matrix is unchanged when the dipolar couplings are averaged over an ensemble consisting of 100 structures with 100 ps increment using Equation 5 (open circles). The average pairwise backbone RMSD of this ensemble is 0.96 ± 0.25 Å and the average computed N-H S^2 order parameter (Lipari and Szabo, 1982) is 0.83 ± 0.11. The nonzero eigenvalues change only little and the corresponding collectivities vary by less than 10%. If a subset of the ensemble is used instead consisting of the first 20 structures, the λ, κ values lie typically between the two cases (Figure 1). Thus, the overall sensitivity of the covariance matrix to homogeneous protein behavior is relatively low.

Heterogeneous protein behavior

Additional nonzero eigenvalues emerge when the protein structure varies among media, i.e., when the protein behavior is heterogeneous across the media. This is demonstrated by assigning 11 MD conformations to the 11 alignment tensors. To assess the sensitivity of the eigenvalues on structural changes, 9 sets of conformations are defined each consisting of 11 MD conformations sampled with a constant time increment of 1 fs, 2 fs, 10 fs, 50 fs, 200 fs, 1 ps, 10 ps, 100 ps, and 900 ps, respectively. The larger the time increment the larger are on average the structural differences between snapshots. The eigenvalue and collectivity distributions of these ensembles are depicted in Figure 2. There are 10 nonzero eigenvalues λ_2, ..., λ_{11} corresponding to the number of degrees of freedom (number of alignment tensors minus one). The eigenvalues form two distinct clusters containing 5 eigenvalues each. The 5 largest eigenvalues, which reflect structural properties (Hus and Brüschweiler, 2002), stay close to the original values, whereas the other five eigenvalues are sensitive to the time increment. For short time increments a clear gap is visible between the two eigenvalue clusters. A quantitative measure for the size of the gap is the ratio of the fifth largest eigenvalue λ_7 and sixth largest eigenvalue λ_6,

$$\rho = \lambda_7/\lambda_6. \quad (6)$$

ρ gradually decreases for larger time steps taking a maximal value at $\rho = 28.9$ for $\Delta t = 1$ fs and a minimal value of $\rho = 1.8$ for $\Delta t = 100$ ps (Figure 3a). Generally, the smaller the gap the less consistent are the dipolar couplings of the structural ensemble with a unique protein structure.

The ρ values can be correlated with respect to the average pairwise backbone RMSD of each ensemble (Figure 3b) as well as the average angular S^2 order parameter of the N-H vectors (Figure 3c). The pairwise average RMSD varies between 0.02 Å and 0.98 Å for the different ensembles and it is strongly correlated with $1 - \langle S^2 \rangle$ (correlation coefficient $r = 0.992$, Figure 3d). With increasing $\langle$RMSD$\rangle$, ρ decreases rapidly and fluctuates about a mean value. The two quantities obey in good approximation an analytical relationship $\rho = 0.58/\langle \text{RMSD} \rangle + 1.58$, which is indicated as solid line in Figure 3b. An analogous relationship exists between ρ and $1 - \langle S^2 \rangle$ (see caption of Figure 3).

Repetition of the above calculation for 1000 sets of randomly chosen alignment tensors yields information about the dependence of the gap on the alignment tensors. In Figure 4, the mean value and standard deviation of ρ are displayed for the different ensembles with increasing time increments. While ρ clearly depends on the alignment tensor set as can be seen from the standard deviations, the trends of Figure 3a are reproduced, i.e., ρ decreases with an increasing ensemble diversity and approaches an average value of $\rho = 2.04$.

Eigenvector properties for simplified models of structural change

The eigenvectors of the covariance matrix provide direct information on collective structural changes induced by the presence of the alignment media. This is illustrated by using two simplified models of structural changes. In the first case, the α-helix including

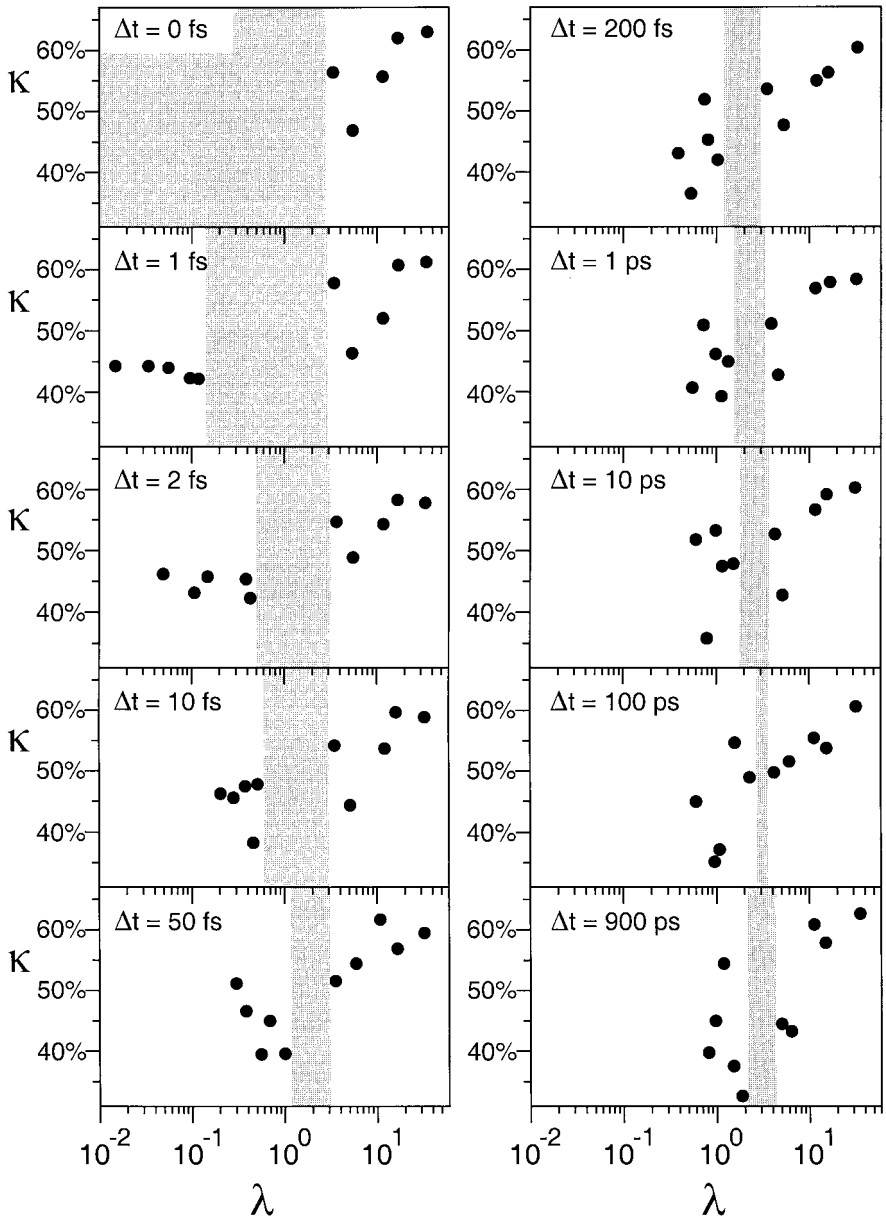

Figure 2. Eigenvalue and collectivity distributions of RDC covariance matrices calculated for heterogeneous protein systems by assigning the 11 alignments different MD snapshots with increasing time increments of 1 fs, 2 fs, 10 fs, 50 fs, 200 fs, 1 ps, 10 ps, 100 ps, and 900 ps.

residues 23 to 31 is rotated about an axis orthogonal to its symmetry axis (defined as the mean direction of all backbone N-H vectors) and the N-H vector of residue Ile23 while the rest of the protein was kept fixed. In the second case, the N-terminal β-sheet including residues 1 to 17 is rotated along an axis orthogonal to the sheet. The rotation angles were varied from $-20°$ to $20°$ in constant increments of $4°$. The κ, λ plots for both cases are given in Figure 5, panels a and b. The gap ρ is 17.6 for the helix motion and 10.7 for the sheet motion.

Due to the reduced number of degrees of freedom for this type of motion only 9 instead of 10 eigenvalues are effectively nonzero. The eigenvectors are displayed in Figures 5c, d. While the five largest eigenvectors $|7\rangle, \ldots |11\rangle$ reflect average structural properties and are virtually identical for the two models, the following four eigenvectors selectively emphasize

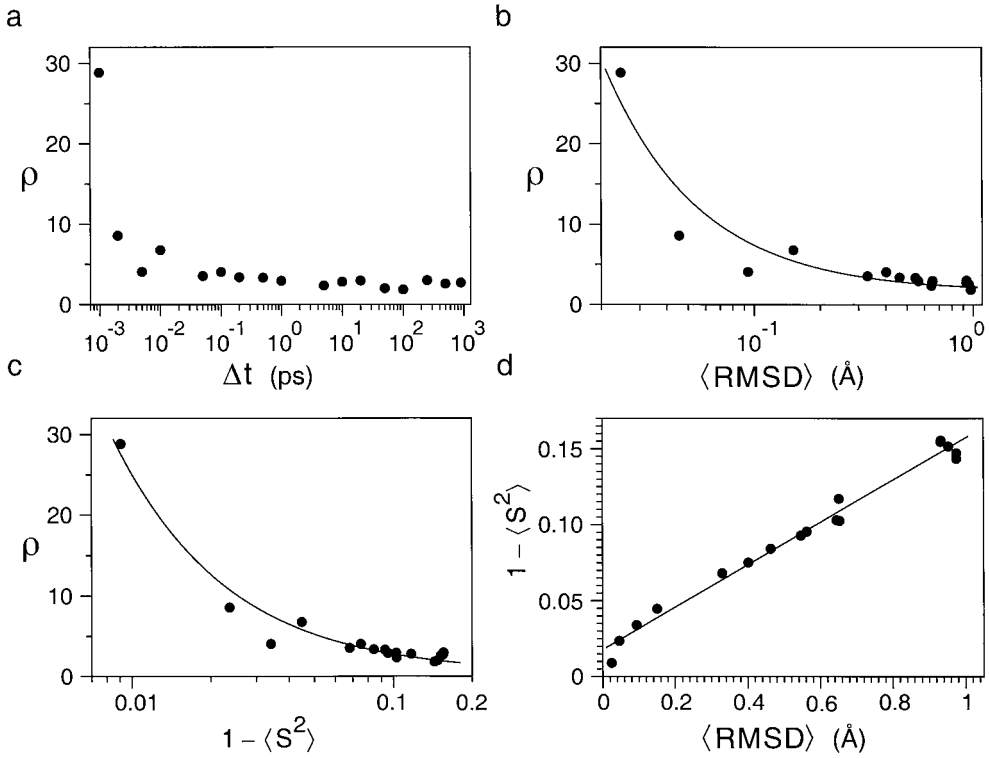

Figure 3. Correlation plot of gap $\rho = \lambda_7/\lambda_6$ and MD time step Δt (panel a), ρ and average pairwise backbone RMSD (panel b), and ρ and average S^2 N-H order parameter (panel c). In panel d, $1 - \langle S^2 \rangle$ is plotted versus $\langle$RMSD$\rangle$ with the straight line corresponding to the relationship $1 - \langle S^2 \rangle = 0.139 \cdot \langle \text{RMSD} \rangle + 0.018$ with a correlation coefficient of 0.992. Sets of MD snapshots were used with the following time increments: 1 fs, 2 fs, 5 fs, 10 fs, 50 fs, 100 fs, 200 fs, 500 fs, 1 ps, 5 ps, 10 ps, 20 ps, 50 ps, 100 ps, 250 ps, 500 ps and 900 ps.

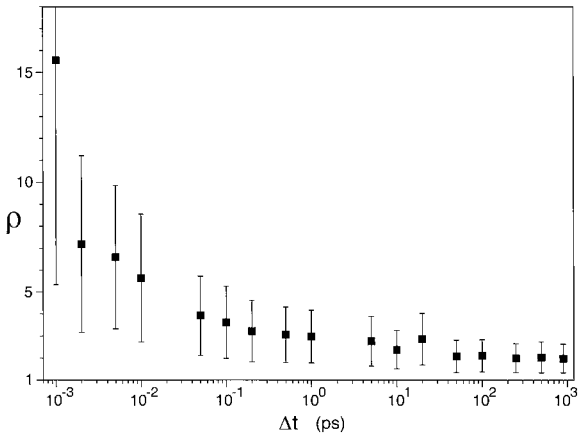

Figure 4. Mean value and standard deviation of ρ for increasing time increments Δt for 1000 calculations using the same MD snapshots but different alignments tensors.

the secondary structural elements that undergo reorientational modulation. Thus, the eigenvectors provide a detailed picture of the dominant reorientational structural changes that are induced by the different alignment media.

Effect of experimental noise

Nonzero eigenvalues do also emerge due to experimental uncertainties (noise) present in the RDC data. To quantitatively assess this effect on the eigenvalue distribution, random Gaussian noise was added to the RDCs for the reference structure with standard deviations corresponding to 0%, 2%, 5%, 10%, 20%, and 30% of D_a. The eigenvalue and collectivity distributions of the resulting principal component analysis are displayed in Figure 6. The same random number generator seed was used for all panels. As in the heterogeneous ensemble case, there are a total of 10 nonzero eigenvalues. The five largest eigenvalues, λ_7, ..., λ_{11}, and their collectivities κ_7, ..., κ_{11} remain largely unaffected by the amount of noise added. In contrast, eigenvalues λ_2, ..., λ_6 grow steadily as the noise level is increased. This behavior is also reflected in the gap ρ, which monotonously decreases with increasing

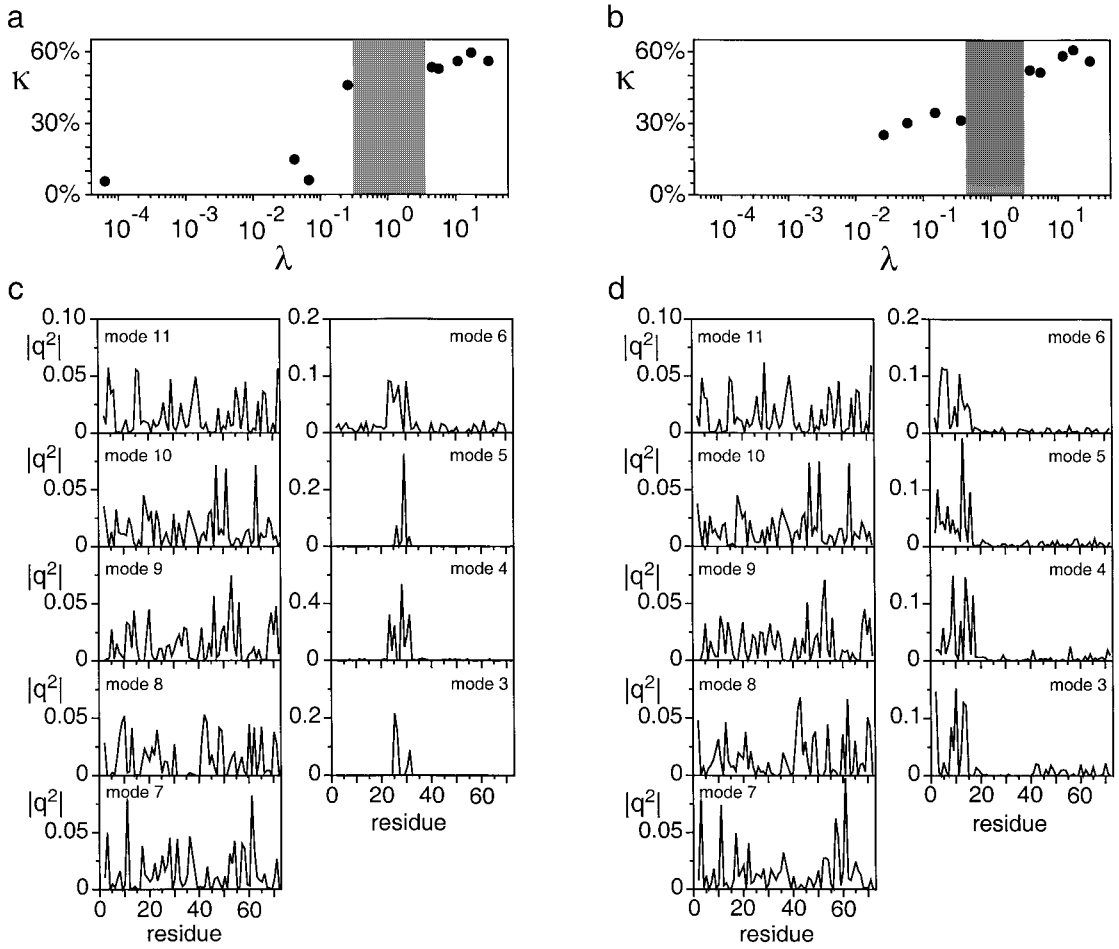

Figure 5. κ, λ plots of reorientational rigid-body motions of (a) the central α-helix and (b) the N-terminal β-sheet with respect to the rest of ubiquitin. Panels c and d: Squared elements of eigenvectors plotted vs. residue number. Panels a and c correspond to reorientations of the helix by ±20° about an axis perpendicular to its symmetry axis (see text). Panels b and d correspond to reorientations of the N-terminal β-sheet about the axis that is perpendicular to the sheet.

noise taking values between 517.5 and 3.3. The dependence of ρ on the noise level is well parametrized by the relationship $\rho = 2064.8/(\text{noise }\%)^2 + 1.42$. Very similar dependencies are obtained for other random number generator seeds and other alignment tensors suggesting that this type of relationship may hold more generally.

The gap ρ was determined for 7 to 50 alignment media with variable amounts of noise using the reference snapshot at 10 ns (Figure 8). Each point corresponds to the average over 50 randomly chosen alignment tensor sets. The gap increases with the number of alignments reflecting an effective reduction of noise effects.

The simultaneous presence of noise and structural heterogeneity was analyzed by adding random Gaussian noise to the ensembles of Figure 2. For example, for noise levels of 5% and 10% the gap ρ belonging to the ensemble with a 1 fs time step is reduced by a factor of 1.33 and 2.21, respectively. If the character and amount of experimental noise is known, it is possible in principle to determine a noise-corrected ρ value, which then solely reflects the amount of structural heterogeneity. If however the nature of the experimental errors is unknown, the (uncorrected) ρ value yields an upper bound for structural heterogeneity based on the relationships plotted in Figure 3. Conversely, if a certain amount of structural heterogeneity is assumed, SECONDA allows one to obtain an estimate of the experimental errors.

The manifestation of heterogeneous reorientational rigid-body movements described in Figure 5 with re-

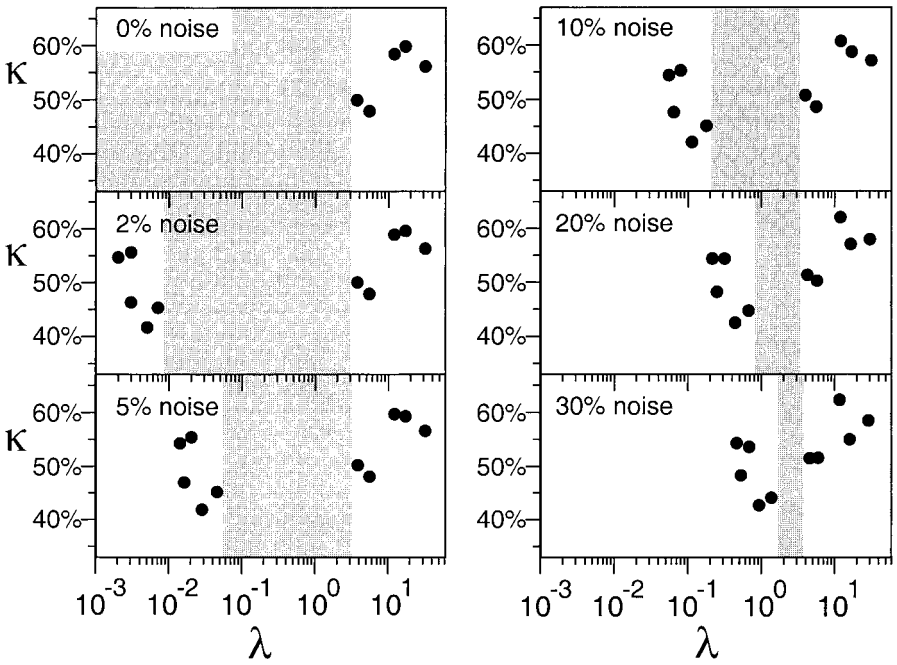

Figure 6. Eigenvalue and collectivity distributions of RDC covariance matrices calculated for the MD reference snapshot of ubiquitin for 11 different alignments with Gaussian noise added corresponding to 0%, 2%, 5%, 10%, 20%, and 30% of D_a.

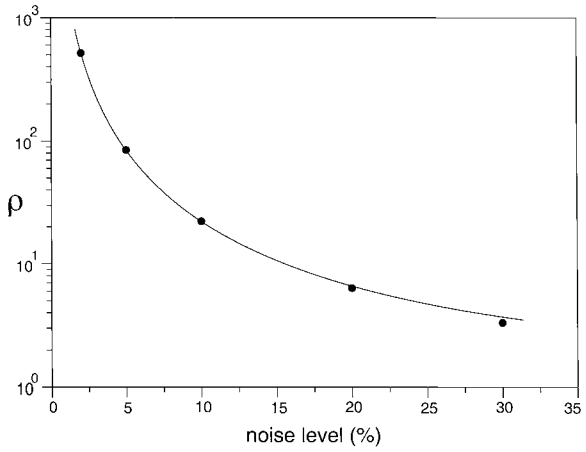

Figure 7. Gap $\rho = \lambda_7/\lambda_6$ of Figure 6 as a function of the noise level. The solid line corresponds to the hyperbolic function defined in the text.

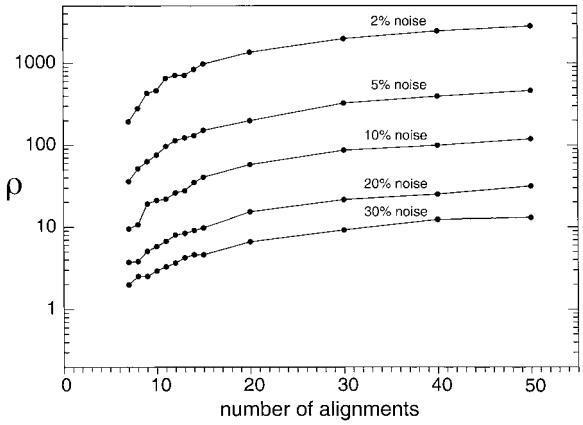

Figure 8. Dependence of gap ρ on the number of alignments in the presence of variable amounts of noise. Each point was determined by averaging over 50 randomly chosen alignment tensor sets.

orientational amplitudes of ±20° was analyzed in the presence of variable amounts of noise with the results displayed in Figure 9. The cumulative amplitudes were calculated as the sum of the squared elements of the four dynamical eigenvectors weighted with the square-root of their eigenvalues. For noise levels up to 10%, the dynamic regions can be easily identified, whereas for noise levels above 20% the dynamic features can no longer be unambiguously discerned from noise effects.

Conclusion

A new analysis method has been described that allows one to test the self-consistency of RDC data collected in multiple alignments with respect to differential structural changes induced by the media. The

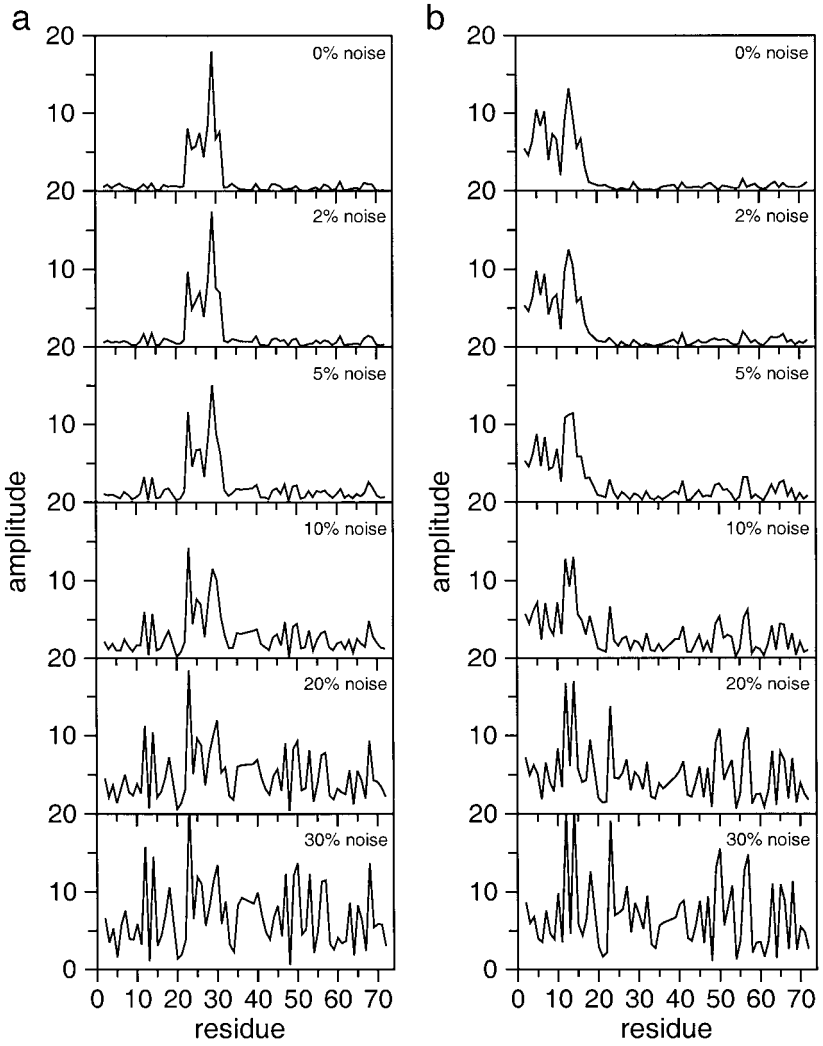

Figure 9. Manifestation of heterogeneous reorientational rigid-body changes induced by different alignment media (see Figure 5) in the presence of variable amounts of noise. (a) Fluctuations of central helix, (b) fluctuations of N-terminal β-sheet. The same random number generator seed was used for each graph.

SECONDA method is based on a principal component analysis of the covariance matrix of the RDC data. The gap between the five largest eigenvalues and the remaining eigenvalues is a measure for the degree of structural heterogeneity. The eigenvectors depict collective structural changes induced by the alignment media. The presence of multiple conformations has no effect on the gap as long as the same internal conformers occur in all media with unaltered populations, while the presence of noise in the dipolar couplings narrows the gap. The method is most robust if the experimental noise level is low and if the alignment tensors probe distinct directions. Application of SECONDA to experimental RDC sets of ubiquitin is currently under way. Because SECONDA does neither require a 3D structural model nor explicit knowledge of the alignment tensors, it should be applicable to RDC data of folded, partially folded, and non-folded biomolecular systems.

Acknowledgement

This work was supported by NSF Grant MCB-0211512.

References

Al-Hashimi, H.M. and Patel, D.J. (2002) *J. Biomol. NMR*, **22**, 1–8.

Barbieri, R., Bertini, I., Lee, Y.-M., Luchinat, C. and Velders, A.H. (2002) *J. Biomol. NMR*, **22**, 365–368.

Bax, A., Kontaxis, G. and Tjandra, N. (2001) *Meth. Enzymol.*, **339**, 127–174.

Brooks, R.B., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.

Brüschweiler, R. (1995) *J. Chem. Phys.*, **102**, 3396–3403.

de Alba, E. and Tjandra, N. (2002) *Progr. NMR Spectrosc.*, **40**, 175–197.

Goto, N.K., Skrynnikov, N.R., Dahlquist, F.W. and Kay, L.E. (2001) *J. Mol. Biol.* **308**, 745–764.

Hus, J.-C. and Brüschweiler, R. (2002) *J. Chem. Phys.*, **117**, 1166–1172.

Lienin, S. F., Bremi, T., Brutscher, B., Brüschweiler, R. and Ernst, R. R. (1998) *J. Am. Chem. Soc.*, **120**, 9870–9879.

Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104**, 4546–4559.

MacKerell, Jr., A.D., Bashford, D., Bellott, M., Dunbrack, Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, III, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D. and Karplus, M. (1998) *J. Phys. Chem.*, **B102**, 3586–3616.

Meiler, J., Prompers, J.J., Peti, W., Griesinger, C. and Brüschweiler, R. (2001) *J. Am. Chem. Soc.*, **123**, 6098–6107.

Peti, W., Meiler, J., Brüschweiler, R. and Griesinger, C. (2002) *J. Am. Chem. Soc.*, **124**, 5822–5833.

Prestegard, J.H., Tolman, J.R., Al-Hashimi, H.M. and Andrec, M. (1999) Protein structure and dynamics from field-induced residual dipolar couplings, In *Biological Magnetic Resonance*, Vol. 17: *Structure Computation and Dynamics in Protein NMR*, Plenum Publishers, New York, pp. 311–355.

Prompers, J.J. and Brüschweiler, R. (2002) *J. Am. Chem. Soc.*, **124**, 4522–4534.

Prompers, J.J., Lienin, S.F. and Brüschweiler, R. (2001a) In *Biocomputing: Proceedings of the 2001 Pacific Symposium*, Altman, R.B., Dunker, A.K., Hunter, L., Lauderdale, K. and Klein, T.E. (Eds.), World Scientific, Singapore, pp. 79–88.

Prompers, J.J., Scheurer, C. and Brüschweiler, R. (2001b) *J. Mol. Biol.*, **305**, 1085–1097.

Vijay-Kumar, S., Bugg, C.E. and Cook, W.J. (1987) *J. Mol. Biol.*, **194**, 531–544.