Journal of Biomolecular NMR, **23:** 127–137, 2002.
KLUWER/ESCOM

Evaluation of uncertainty in alignment tensors obtained from dipolar couplings

Markus Zweckstetter[a,b,*] & Ad Bax[a,*]
[a]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, U.S.A.;* [b]*Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany*

Received 25 February 2002; Accepted 22 April 2002

Key words: alignment tensor, backbone dynamics, dipolar coupling, error analysis, liquid crystal, order matrix, protein structure determination, ubiquitin

Abstract

Residual dipolar couplings and their corresponding alignment tensors are useful for structural analysis of macromolecules. The error in an alignment tensor, derived from residual dipolar couplings on the basis of a known structure, is determined not only by the accuracy of the measured couplings but also by the uncertainty in the structure (structural noise). This dependence is evaluated quantitatively on the basis of simulated structures using Monte-Carlo type analyses. When large numbers of dipolar couplings are available, structural noise is found to result in a systematic underestimate of the magnitude of the alignment tensor. Particularly in cases where only few dipolar couplings are available, structural noise can cause significant errors in best-fitted alignment tensor values, making determination of the relative orientation of small fragments and evaluation of local backbone mobility from dipolar couplings difficult. An example for the protein ubiquitin demonstrates the inherent limitations in characterizing motions on the basis of local alignment tensor magnitudes.

Introduction

Residual dipolar couplings (rDCs) can be observed in solution when a molecule is aligned with the magnetic field, either as a result of its own magnetic susceptibility anisotropy (Gayathri et al., 1982; Bothner-By, 1996), or caused by an anisotropic environment such as an oriented liquid crystalline phase (Saupe and Englert, 1963; Emsley, 1996) or an anisotropically compressed gel (Tycko et al., 2000; Sass et al., 2000). When alignment can be kept sufficiently weak, the NMR spectra retain the simplicity normally observed in regular isotropic solution, while allowing quantitative measurement of a wide variety of rDCs, even in macromolecules (Tolman et al., 1995; Tjandra and Bax, 1997). These couplings can either be used for further refinement of structures determined by conventional methods (Tjandra et al., 1997) or for determining a structure directly (Delaglio et al., 2000; Hus et al., 2000). They can also be used for structure validation (Cornilescu et al., 1998; Clore and Garrett, 1999), for determination of relative orientations of molecular fragments or domains (Tjandra et al., 1997; Losonczi et al., 1999; Bewley and Clore, 2000; Clore, 2000; Goto et al., 2001), or for studying dynamic effects (Fischer et al., 1999; Meiler et al., 2001; Tolman et al., 2001). Analysis of the relation between a molecule's 3D shape and its alignment tensor values can yield important insights into all of these processes (Zweckstetter and Bax, 2000; Bewley and Clore, 2000).

If well-defined structures of either complete macromolecules, their domains, or even smaller fragments are available, an alignment tensor that describes the average orientation of this (sub)structure with respect to the magnetic field can be calculated from the observed dipolar couplings (Tjandra et al., 1996;

*To whom correspondence should be addressed. E-mails: bax@nih.gov; mzwecks@gwdg.de

Losonczi et al., 1999). This is increasingly used to determine relative orientations of protein domains and RNA fragments, and also for identification of mobile protein segments (Vermeulen et al., 2000; Fischer et al., 1999; Tolman et al., 2001; Almond and Duus, 2001; Wang et al., 2001). In order to avoid overinterpretation of the data in such applications, it is critical to have accurate estimates of the uncertainties in the magnitude and orientation of the alignment tensor. These uncertainties can be particularly large when only a small number of residual dipolar couplings is available for the individual substructures.

This report focuses primarily on the effect of uncertainty in the atomic coordinates of the structure used, so-called structural noise, on alignment tensors that are calculated from rDCs on the basis of such a structure. Results and conclusions in this study are based mostly on simulated data, where the 'true' structure and dipolar couplings are known exactly. The effect of structural noise is simulated by adding random errors to the bond vector orientations, whereas simply adding noise to the dipolar couplings simulates the effect of measurement noise. The effect of structural noise is shown to vary greatly, depending on the orientations of the internuclear vectors relative to the alignment tensor. Modules to determine the uncertainty in the alignment parameters as a function of measurement or structural noise are implemented in the software package PALES (Zweckstetter et al., unpublished).

Methods

The alignment tensor can conveniently be derived from experimental dipolar couplings and a known structure, either by singular value decomposition (SVD) (Losonczi et al., 1999) or by an iterative multidimensional least-squares minimization of the difference between back-calculated rDCs and experimental ones (Tjandra et al., 1996). SVD is more stable than iterative least-squares minimization but does not permit certain parameters, such as the alignment tensor magnitude or rhombicity, to be held constant. All fits carried out in this study are based on SVD.

Simulated dipolar vector orientations are generated as follows: First, 10,000 vectors are created that are uniformly distributed on a sphere, using a double-cubic-lattice method (Eisenhaber et al., 1995). Then, a random subset of N_{DC} vectors is selected from this ensemble, and corresponding 'true' dipolar couplings are calculated using

$$d^{PQ}(\theta_{PQ}, \phi_{PQ}) = D_a[(3\cos^2\theta_{PQ} - 1) + \tfrac{3}{2}R\sin^2\theta_{PQ}\cos(2\phi_{PQ})], \quad (1)$$

where (θ_{PQ} ,ϕ_{PQ}) are the polar angles of vector P-Q. Most of the graphs shown are derived for a magnitude of the alignment tensor, D_a, of 10 Hz, and rhombicity, $R = 0.3$, but the effect of varying R is also considered.

The effect of structural uncertainty is simulated by slightly reorienting the selected vector orientations in a random manner, such that the deviations between the original and final vectors are described by a Gaussian cone-shaped distribution, with a standard deviation σ^{cone}, and a relative probability of $\sin(\beta)\exp(-\beta^2/\sigma^2_{cone})$ for an angle β between the original and modified orientation.

The angle ϑ between the orientation of two alignment tensors is calculated according to Sass et al. in a five-dimensional linear vector space, and expresses the collinearity of alignment tensors, independent of their magnitude (Sass et al., 1999). Also following Sass, a generalized alignment tensor magnitude is defined, which includes the effect of rhombicity:

$$GMag = (2D_a/D^{PQ}_{max})[\pi(4 + 3R^2)/5]^{1/2}, \quad (2)$$

where $D^{PQ}_{max} = -\mu_0(h/2\pi)\gamma_P\gamma_Q/(4\pi^2 r^3_{PQ})$, with μ_0 the permittivity of vacuum, h Planck's constant, γ the magnetogyric ratio, and r_{PQ} the PQ internuclear distance. However, because much of the recent literature reports D_a, and visualization of the meaning of the parameters is frequently carried out using diagonalized alignment tensors, most of the results reported here refer to D_a. For the same reason, the deviation of the principal z and transverse axes of the alignment tensor from their true orientations is also considered.

Results and discussion

The effect of noise on best-fitted alignment tensors is evaluated in two different ways. First, we analyze for a large number of simulated structures how the errors in derived alignment tensors correlate, on average, with parameters such as the number of experimentally available dipolar couplings and the uncertainty in the structure coordinates. Next, we evaluate for a large set of individual structures how the error estimate obtained by Monte-Carlo procedures correlates with the true error in the derived alignment tensor. Proper parameterization of such Monte-Carlo procedures is

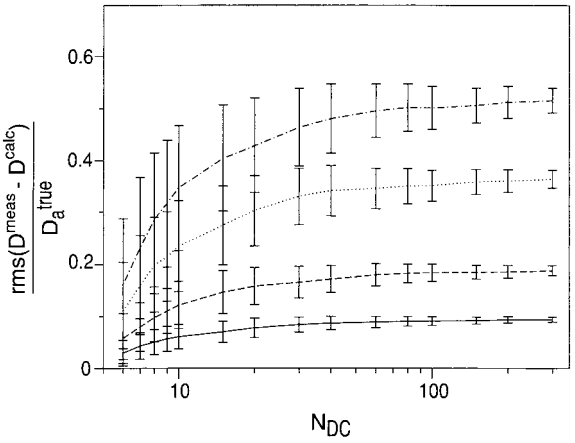

Figure 1. Ratio of the RMSD between measured and back-calculated dipolar couplings, $D^{meas,i}$ and $D^{calc,i}$, $rms(D^{meas} - D^{calc}) = [\sum_{i=1}^{NDC} (D^{meas,i} - D^{calc,i})^2/N_{DC}]^{1/2}$, and the true magnitude of the alignment tensor (D_a^{true}). Noise-free D^{meas} values are generated for a randomly selected initial structure, using $D_a^{true} = 10$ Hz, and R = 0.3, according to Equation 1. D^{calc} values are dipolar couplings that are back-calculated after an SVD fit of D^{meas} values is made to a structure in which structural noise has been added to the original structure. Results are shown as a function of the number of dipolar couplings, N_{DC}, and are averages over 1000 different initial structures. Calculations are carried out for four different levels of structural noise: 2.5° (—), 5° (– –), 10° (....) and 15° (·---). Assuming an isotropic distribution of internuclear vectors, rmsd values vary only weakly with rhombicity of the alignment tensor and scale with $(4 + 3\ R^2)^{1/2}$ (Clore and Garrett, 1999). Error bars correspond to one standard deviation.

found to be critical when evaluating the outcome of such evaluations.

Structural quality and dipolar coupling RMSD

Figure 1 shows the relation between structural noise and the average RMSD of the input and back-calculated dipolar couplings. Average RMSD values are obtained using 1000 randomly selected starting structures, each consisting of N_{DC} (from 6 to 300) vectors. The calculation is repeated for structural noise amplitudes, σ^{cone}, of 2.5, 5, 10 and 15°. As expected for an SVD fit with five adjustable parameters, Figure 1 shows that if only six dipolar couplings are included in the fit, the RMSD is invariably quite small. The RMSD rapidly reaches a plateau value when more than *ca* 25 couplings are included. For example, for structural noise with Gaussian amplitude $\sigma^{cone} = 5°$, the RMSD becomes about 18% of D_a, which is comparable to what is seen when comparing experimental dipolar couplings with high-resolution crystal structures. For example, 17% is obtained when comparing experimental ubiquitin dipolar couplings with its 1.8-Å X-ray structure (Ottiger and Bax, 1998) and 11% is found when comparing dipolar couplings for the third IgG-binding domain of streptococcal protein G with its 1.1-Å X-ray structure (B. Ramirez, unpublished). Much lower agreement, corresponding to >10° structural noise is found when comparing dipolar couplings measured for a range of other proteins to either lower resolution (> 2 Å) crystal structures or to NMR structures solved in the absence of dipolar couplings, including structures with small (< 0.7 Å) backbone RMSD (data not shown). The fact that for the wide variety of proteins studied to date the agreement between dipolar couplings and X-ray crystal structure invariably improves with higher resolution suggests that the residual in the dipolar coupling fit is often dominated by structural noise, particularly when considering globular domains, and not by local dynamics.

For a given amount of structural noise, the rms error in D_a depends only very weakly on the rhombicity of the alignment tensor (see legend to Figure 1). Note, that these structural noise estimates are based on the assumption that errors in dipolar coupling measurements are small compared to structural uncertainties (as is the case with current measurement techniques and overall alignment strengths discussed in this paper) and highly flexible protein parts, such as the C-terminus of ubiquitin (residues 73–76), are excluded.

Error in derived alignment tensor

Figure 2 shows the rms error in the derived alignment tensor parameters for different amounts of structural noise. As expected, for large numbers of dipolar couplings ($N_{DC} \geq 20$) the rms error in the orientation of the alignment tensor z-axis, $\Delta\psi_z$, decreases approximately with $N_{DC}^{1/2}$, for all four levels of structural noise (Figure 2B). Remarkably, the fractional error in the magnitude, D_a, of the SVD-derived alignment tensor shows a very different behavior (Figure 2A): At low levels of structural noise, the error decreases with increasing N_{DC}, whereas at $\sigma^{cone} = 5°$ the fractional error reaches a plateau value of *ca*. 0.02 and the plateau value increases to *ca*. 0.18 for $\sigma^{cone} = 15°$. This plateau value results from a systematic underestimate of the alignment tensor magnitude in the case of structural noise, which would not occur if the error were only in the measurement instead of in the structure (see below, and Figures 5C,E). The magnitude of this systematic error increases rapidly with in-

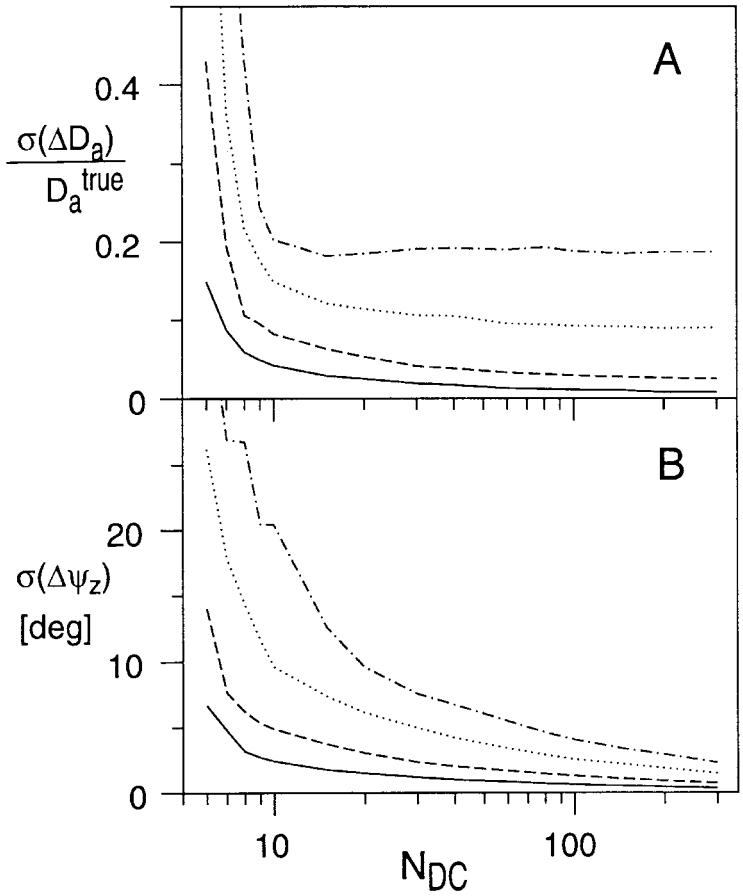

Figure 2. Uncertainty in magnitude (A) and orientation (B) of alignment tensors as a function of the number of dipolar couplings, N_{DC}, for four different levels of structural noise: 2.5° (-), 5° (– –), 10° (....) and 15° (·---). $\Delta D_a = D_a^{true} - D_a^{svd}$ indicates the difference between the SVD-derived alignment magnitude and its true value, D_a^{true}, whereas $\Delta\psi_z$ is the deviation from the true orientation of the z-axis of the alignment tensor; $\sigma(\Delta D_a) = [\sum_{i=1}^{1000} (\Delta D_a^i)^2/1000]^{1/2}$ and $\sigma(\Delta\psi_z) = [\sum_{i=1}^{1000} (\Delta\psi_z^i)^2 / 1000]^{1/2}$ are the rms values of these errors. Values shown are averages obtained from 1000 randomly generated structures; errors for individual structures can be significantly higher (see text).

creasing magnitude of the structural noise (Figure 3), particularly for high N_{DC} values.

The error in the SVD-derived rhombicity, ΔR, correlates directly with the fractional error in D_a (Figure 4). For $N_{DC} = 15$ (solid lines in Figure 4), the ratio of the error in the magnitude of the rhombicity and the fractional error in D_a changes relatively little with increasing structural noise. However, at larger N_{DC} values the error in R decreases relative to the error in D_a. This results from the above-mentioned systematic underestimate in the SVD-derived D_a value for the case of structural noise and large N_{DC}. Because R is the ratio of the z and the transverse component of the alignment tensor, both of which are systematically underestimated in the case of structural noise, there is no such systematic bias in R. As a consequence, the error in R decreases relative to the error in D_a with increasing structural noise. Numerical calculations show that the rms error in the transverse orientation of the alignment tensor, $\sigma(\Delta\psi_{xy})$, is directly related to the rms error in z axis orientation, $\sigma(\Delta\psi_z)$:

$$\sigma(\Delta\psi_{xy}) \approx 2\sigma(\Delta\psi_z)/(3R). \quad (3)$$

Clearly, Equation 3 only applies to cases where $2\sigma(\Delta\psi_z)/(3R) << \pi/2$; for larger values the transverse orientation will be undefined.

A numerical example

A simple practical example serves to illustrate how the above numbers can be used to estimate the expected uncertainty in the SVD-derived alignment parameters. The example discussed below assumes that for a given structure the SVD analysis yields $D_a = 7.5$ Hz,

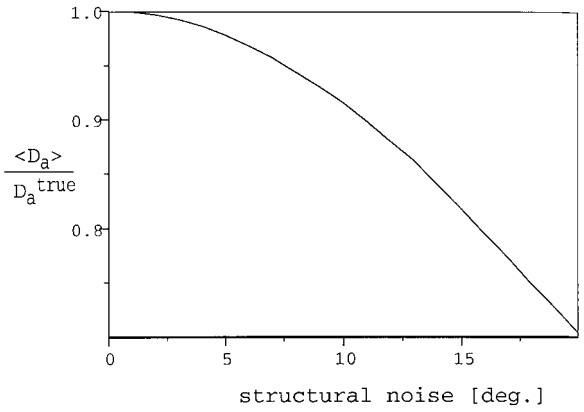

Figure 3. Systematic underestimation of the average magnitude of alignment as a function of structural noise for the case of large numbers of dipolar couplings. In the simulation 200 couplings were used, and values shown are averages obtained from 1000 different structures, i.e. $<D_a> = [\sum_{i=1}^{1000} D_{a,i}]/1000$, where $D_{a,i}$ is the alignment magnitude obtained for structure i by a single SVD.

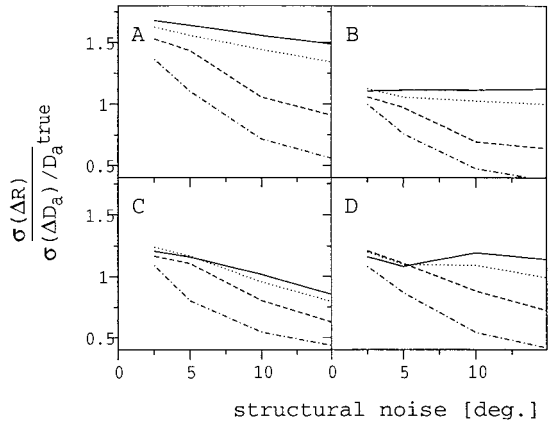

Figure 4. Uncertainty in alignment tensor rhombicity as a function of structural noise for four different numbers of dipolar couplings: $N_{DC} = 15$ (–); $N_{DC} = 20$ (....); $N_{DC} = 40$ (– –); $N_{DC} = 100$ (·---·). Panels A–D differ in the rhombicity, R, of the alignment tensor used for calculating the exact dipolar couplings: R = 0.0 (A); R = 0.1 (B); R = 0.2 (C), R = 0.6 (D). Symbols σ (ΔD_a) and σ (ΔR) are the root-mean-square values of the differences between the SVD-derived D_a and R and their true values, D_a^{true} and R^{true}, for an individual structure. Rms values are obtained by repeating SVD calculations for 1000 different, randomly selected structures.

R = 0.25, with a 3 Hz rmsd of the SVD fit, using $N_{DC} = 50$. In Figure 1 an RMSD/D_a = 3/7.5 = 0.4 ratio for $N_{DC} = 50$ falls between the 10° and 15° structural noise curves, and interpolation suggests an approximate structural noise level of about 12°. Figure 3 indicates that 12° structural noise yields a D_a underestimate by about 11% relative to the true value, which therefore is estimated at 8.3 Hz. In a second iteration, the estimate for RMSD/D_a = 3/8.3 is slightly smaller, pointing to a structural noise of 10°. Figure 2B then indicates an estimated error of 4° for the *z*-axis orientation. The estimated uncertainty in the transverse orientation then follows from Equation 3 to be about 10°. The uncertainty in R is derived from Figure 4C, using σ (Δ D_a)/$D_a^{true} \approx 0.1$ (Figure 2A). For $N_{DC} = 50$, interpolation between the $N_{DC} = 40$ and $N_{DC} = 100$ curves in Figure 4C yields an uncertainty in R of 0.075.

It is important to realize that these estimated uncertainties are averages, derived from a large number of simulated structures. For any individual structure, errors may be considerably higher or lower, and analysis of uncertainties for individual structures can only be derived with Monte-Carlo methods, discussed below.

Estimation of errors: Losonczi Monte-Carlo method

In order to obtain a better representation of the uncertainty in SVD-derived alignment tensor values, obtained for a given structure, Losonczi et al. (1999) proposed to repeat the SVD calculation many times (1000 times in the present study), but each time with different Gaussian noise added to the experimental rDCs. In this so-called Monte-Carlo approach, only those solutions are accepted for which all back-calculated rDCs are within a given margin of the original experimental dipolar couplings. Originally, it was suggested to set the amplitude of the added noise two to three times higher than the measurement uncertainty, in order to account indirectly for uncertainties in the structure. However, the results obtained depend both on the amplitude of this added noise and on the acceptance margin used. Instead, we use a fixed recipe, described below, for selection of these parameters that yields close agreement between the derived uncertainties and the standard errors, obtained from the simulated results. In the present study, the PALES program (M. Zweckstetter, unpublished) is used, which iteratively adjusts the amplitude of the noise added to the dipolar couplings, such that an adjustable fraction of the solutions (80% in our study) are accepted when using an acceptance margin that is two-fold larger than the rms amplitude of the added noise. In practice, this 80% acceptance criterion corresponds approximately to a rms noise that is comparable to two to three times the rms obtained in the initial SVD fit.

Estimation of errors: Structural noise Monte-Carlo method

A second method for estimating the uncertainty in the derived alignment tensor has also been implemented in the program PALES. In this so-called 'structural noise Monte-Carlo method', structural noise is added to the original structure (see Methods) with an amplitude that is derived from Figure 1 to match the RMSD between the experimental and back-calculated dipolar couplings. The spread in the alignment parameters obtained for these noise-corrupted structures, when using the coupling constants back-calculated for the original structure (i.e., yielding a perfect fit if no structural noise were added), then provides another unbiased measure for the spread in the alignment parameters. On average, the Losonczi Monte-Carlo method, when implemented in the way described above, and structural noise Monte-Carlo method yield uncertainties that are quite similar to one another. However, as will be shown for the case of ubiquitin, some differences can occur when considering small fragments.

Losonczi Monte-Carlo evaluation of individual structures

Below, ensembles of N_{DC} internuclear vectors are randomly selected from a set of 10 000 uniformly distributed orientations. In the Losonczi Monte-Carlo method, Gaussian noise is then added to the exact dipolar couplings, calculated from Equation 1, and the standard deviation of this noise is adjusted to correspond to 5° structural noise (i.e., 1.8 Hz; see Figure 1). For a given starting structure, the Losonczi Monte-Carlo procedure is repeated 1000 times using different sets of noise-containing dipolar couplings, and the average alignment tensor is calculated from the 'accepted' SVD solutions. In order to simulate variations in the non-uniformity of the distribution of internuclear vectors, the whole procedure is repeated 100 times, randomly selecting different starting ensembles (with the same number of internuclear vectors N_{DC}) from the set of 10,000 uniformly distributed orientations. For each of these 100 starting structures, average D_a values and errors in the orientation of the Monte-Carlo-averaged tensor are shown in Figure 5. Although averaging D_a values is not strictly valid due to the non-linearity of alignment tensors, average values are shown as recent work focusing on quantitative analysis of motion from dipolar couplings were based on those (Tolman et al., 2001). Analysis of the results obtained for simulated data confirms that adding noise to the data as part of the Losonczi Monte-Carlo procedure followed by averaging of the resulting tensors generally yields magnitudes quite similar to those obtained from a single SVD with no additional noise added to the data (data not shown).

The performance of the Losonczi Monte-Carlo method is evaluated for three different cases: (i) absence of measurement and structural noise, (ii) presence of measurement noise but no structural noise, and (iii) presence of structural noise but no measurement noise. These three different cases are simulated by addition of noise to exact (noise-free) dipolar couplings/structures a single time prior to a Monte-Carlo evaluation. Thus, this addition of noise is performed for simulating situations practically encountered when working with dipolar couplings and not related to the Monte-Carlo procedure itself. During the Losonczi Monte-Carlo procedure noise is then added many times to dipolar couplings to assess potential variations in alignment tensor values.

Figures 5A and 5B apply to the case where, as part of the Monte-Carlo procedure, ~1.8 Hz random noise is added to the couplings calculated for the exact structure (i.e., in the absence of measurement and structural noise). Except for the smallest ensembles ($N_{DC} = 6$), the averaged SVD results all cluster tightly around the correct value. For $N_{DC} = 6$, the spread in calculated alignment tensor values is very high for the few structures that yield large outliers, identifying these Monte-Carlo SVD results as unreliable.

Figures 5C and 5D correspond to the case where there is no structural noise, but 1.8 Hz rms noise has been added to the simulated dipolar couplings, prior to the Monte-Carlo procedure. This corresponds to the case where the structure is perfectly known, but the measurement has a rms error of 1.8 Hz. Depending on the ensemble of starting vectors, substantial errors in both magnitude and orientation of the alignment tensor can be present, which decrease with increasing N_{DC}.

Figures 5E and 5F correspond to the case where 5° structural noise is added to the bond vector orientations, but the measurement is assumed to be perfect. This reflects the common situation where one-bond ^{15}N-^{1}H and or ^{13}C-^{1}H dipolar couplings are measured for a protein or nucleic acid whose atom positions are not known exactly, whereas measurement error in the dipolar couplings is negligible. Comparison of Figures 5E and 5F with 5C and 5D reveals that for small numbers of dipolar couplings ($N_{DC} \leq 10$), structural noise results in an up to ca. 50% larger Monte-Carlo

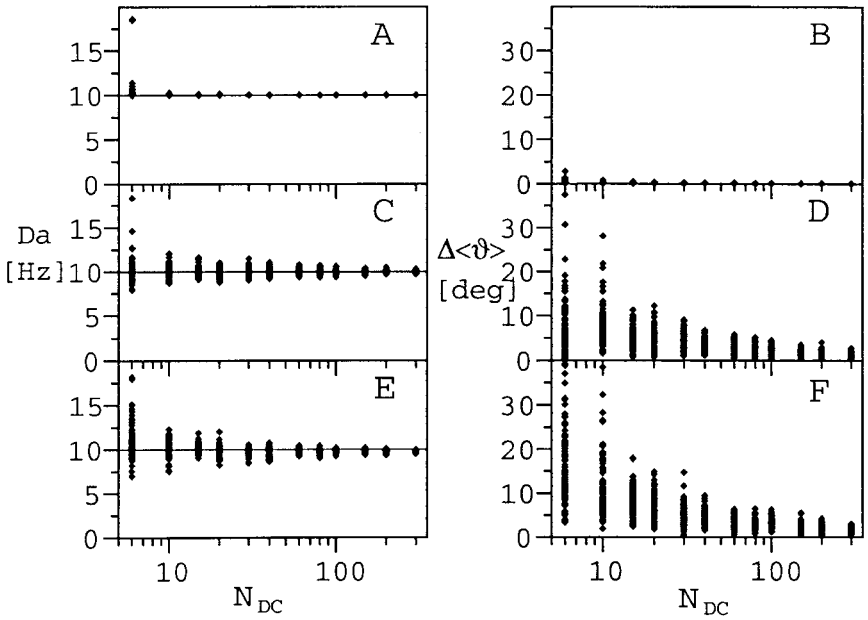

Figure 5. Average alignment tensor values as a function of the number, N_{DC}, of dipolar couplings when using the Losonczi Monte-Carlo method, with 100 different ensembles of internuclear vectors (test structures) for each N_{DC}. $D_a = (\sum_{i=1}^{1000} D_a^i)/1000$ is the alignment tensor magnitude, averaged over 1000 Monte-Carlo calculations; $\Delta<\vartheta> = (\sum_{i=1}^{1000} |\vartheta^i|)/1000$ denotes the corresponding average error in the generalized orientation, where ϑ^i is the angle between the true tensor orientation and the SVD-derived orientation for the i-th Monte-Carlo calculation. A and B show variations for ensembles without measurement or structural noise; C and D are for structures with measurement noise ($\sim$ 1.8 Hz) but without structural noise; E and F are for the absence of measurement noise but with random errors in the vector orientations following a Gaussian distribution of 5° standard deviation. In A, C and E the distributions in alignment tensor magnitude is shown, whereas B, D and F depict the angles between the orientation of the Monte-Carlo averaged alignment tensor and its true orientation. The magnitude used for calculating true dipolar couplings ($D_a = 10$ Hz) is indicated by straight lines in A, C and E (the rhombicity was 0.3).

spread than for the case where noise is added to the experimental data, whereas for large N_{DC} the spread in SVD results is only ca. 25% higher.

Importantly, the data shown in Figure 5E confirm that for $N_{DC} \geq 20$ the SVD results tend to underestimate the true magnitude of the alignment tensor, whereas the opposite is seen for very low N_{DC} values. This systematic underestimate increases with higher structural noise, from 2% to 9%, and 18% for Gaussian structural noise with standard deviations of 5, 10 and 15°, respectively (Figure 3). This underestimate results from the fact that structural noise unevenly influences experimental dipolar couplings. Couplings for vectors that are at an orientation where the predicted value varies steeply with orientation effectively contain more structural noise than those for orientations nearly parallel to any of the principal axes of the alignment tensor. One possible way of addressing this problem is to weight the dipolar couplings during SVD according to their dependence on structural noise, after an initial, unweighted SVD has defined the alignment tensor orientation and approximate magnitude. In practice, however, we find the improvement in the accuracy of D_a obtained with such a procedure to be modest and more than offset by an increase in the uncertainty of the alignment tensor orientation (data not shown).

Clearly, the results shown in Figure 5 indicate that substantial errors in the magnitude and orientation of the alignment tensor can result from either measurement error or uncertainty in the input structure. A key question is, however, whether the uncertainty in this alignment tensor can be accurately estimated from the spread in the SVD results when adding noise to the experimental dipolar couplings. Figures 6A, 6B show the ratio of the error in the average alignment tensor and the root-mean-square spread of the SVD results for the case of noise-contaminated dipolar couplings, when conducting the Losonczi Monte-Carlo analysis. As can be seen from this figure, for most input structures the true errors in D_a and tensor orientation are smaller than the Monte-Carlo spreads ($\Delta<D_a>/\sigma_{MC}D_a < 1$), but exceptions occur. In particular, for $N_{DC} = 6$, the true error can be more than three times larger than the

standard deviation observed in the Monte-Carlo simulations. This demonstrates that in particular for small numbers of dipolar couplings the accuracy of SVD-results and Monte-Carlo evaluations is strongly related to the linear independence of involved internuclear vectors.

For the situation where the error in the original SVD fit is dominated by structural noise, the Losonczi Monte-Carlo analysis tends to underestimate the true error in the alignment tensor parameters (Figures 6C and 6D). For example, even for $N_{DC} = 300$ the true error can be more than two-fold larger than the standard deviation observed in the Monte-Carlo analysis, and the ratio can be considerably larger for $N_{DC} = 6$. This applies to both the magnitude and orientation of the alignment tensor, and similar results are obtained when evaluating uncertainties in the alignment tensor by the structural noise Monte-Carlo method (data not shown).

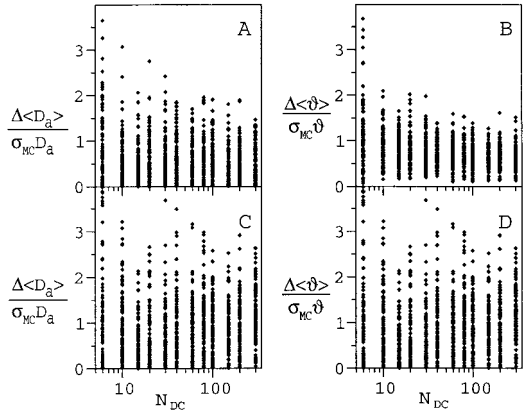

Figure 6. Distribution of the ratio of the true errors ($\Delta<D_a> = [(\sum_{i=1}^{1000} D_a^i)/1000 - D_a^{True}]$ and $\Delta<\vartheta>$) in the averaged Losonczi Monte-Carlo results, and the corresponding rmsd uncertainties, $\sigma_{MC}D_a = \{(\sum_{i=1}^{1000} D_a^i D_a^i)/1000 - [(\sum_{i=1}^{1000} D_a^i)/1000]^2\}^{1/2}$ and $\sigma_{MC}\vartheta = \{(\sum_{i=1}^{1000} \vartheta^i\vartheta^i)/1000 - [(\sum_{i=1}^{1000} \vartheta^i)/1000]^2\}^{1/2}$, where ϑ^i indicates the angle between a single alignment tensor during the i-th Losonczi Monte-Carlo procedure and its average orientation. (A,B) With 1.8 Hz measurement noise but no structural noise; (C,D) with 5° structural noise but no measurement noise. These figures are derived from the same simulations used to generate Figures 5 C/D and E/F. The variations shown result from different distributions of internuclear vectors. In total 100 different ensembles of internuclear vectors were tested for each N_{DC}. Input magnitude and rhombicity were 10 Hz and 0.3, respectively. For $N_{DC} = 6$, Δ/σ-ratios fall outside the window shown and can reach up to 5.

Evaluation of fragments in ubiquitin

As is clear from the above, the spread in alignment magnitudes and orientations obtained with the Losonczi Monte-Carlo procedure provides a reasonable estimate for the uncertainty in the molecular alignment parameters. These uncertainties need to be taken into account when evaluating whether alignment tensors for different fragments in the same protein have different magnitudes (requiring dynamic effects). Figure 7 addresses this question for 3-residue fragments of ubiquitin, with on average about 12 couplings per fragment, as a function of position along the polypeptide backbone. In this analysis we consider two different structural models available for ubiquitin, namely the high-resolution NMR structure (Cornilescu et al., 1998), refined with residual dipolar couplings (Figure 7 A/C/E/G), and the 1.8-Å crystal structure (Vijay-Kumar et al., 1987) (Figure 7 B/D/F/H).

In Figures 7A-D, black curves mark the generalized magnitude of the local alignment tensors as obtained from an initial SVD fit, i.e. without adding additional noise to the dipolar couplings. Colored lines correspond to the uncertainty estimates (one standard deviation in Figures 7A,B; extreme values in 7C,D) when using either the Losonczi Monte-Carlo method (green) or the structural noise Monte-Carlo method (red). As described above, in the Losonczi Monte-Carlo method the amplitude of the noise added to the dipolar couplings is adjusted automatically such that an adjustable fraction of the solutions (80% in our study) are accepted when using an acceptance margin that is two-fold larger than the rms amplitude of the added noise. In the structural noise Monte-Carlo method, on the other hand, structural noise is added to the original structure (see Methods) with an amplitude that is derived from Figure 1 to match the RMSD between the experimental and back-calculated dipolar couplings. Because addition of noise to the structure or to the dipolar couplings typically leads to an asymmetric distribution of SVD results, the average of the Monte-Carlo positive and negative standard deviations (obtained by averaging the values of either the two red or the two green curves) generally does not coincide exactly with the results from the initial SVD (black line).

As can be seen from Figure 7A, for the fragment centered at Leu^8 the generalized magnitude falls more than one standard deviation below what is observed for the full protein, both when considering the Losonczi Monte-Carlo and the structural noise Monte-Carlo procedures. However, when considering the entire ensemble of Monte-Carlo results (Figure 7C), several of these exceed the value obtained for the full

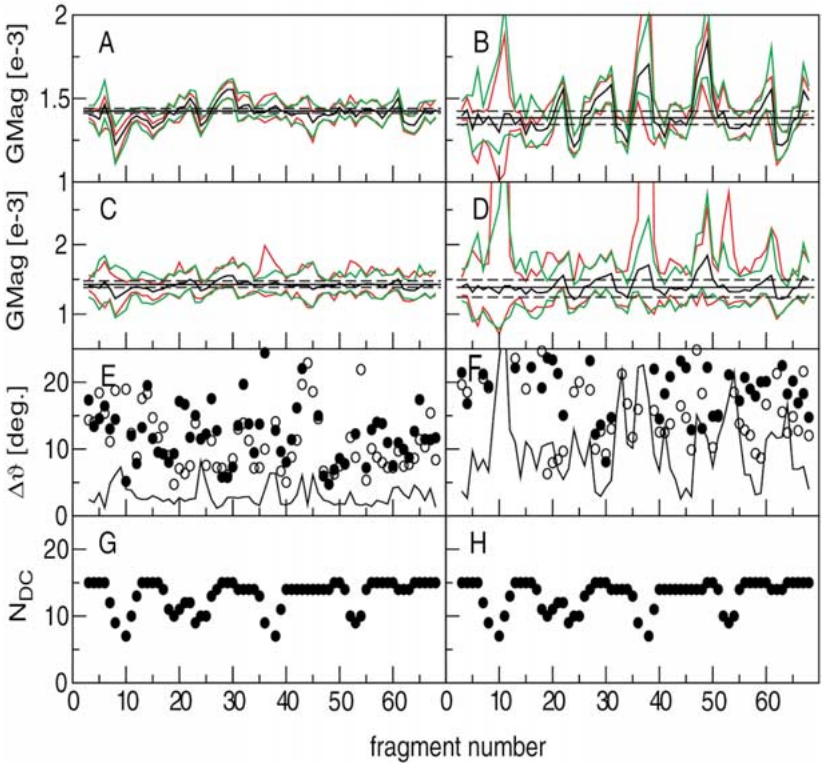

Figure 7. Variation in SVD-derived local alignment tensor values for 3-residue fragments of ubiquitin as a function of position along the polypeptide backbone. Experimental dipolar couplings comprise five different types of backbone couplings (N-H; C'-N; C'-H^N; C^α-H^α; C'-C^α; in total 295 couplings) observed in 5% w/v bicelles (Ottiger and Bax, 1998). SVD results are shown for the NMR structure (A/C/E/G) (Cornilescu et al., 1998) (PDB code 1D3Z) and the 1.8 Å X-ray structure (B/D/F/H) (PDB code 1UBQ) (Vijay-Kumar et al., 1987). In A-D, the straight solid lines correspond to the generalized magnitude (Sass et al., 1999) of the alignment tensor (GMag) as obtained from an initial SVD fit, i.e. without recourse to the Monte-Carlo procedure, and using all 295 dipolar couplings. The rms uncertainty in this value, as obtained from the distribution of magnitudes when using the Losonczi Monte-Carlo procedure on the full protein is marked by dashed lines (A,B). Dashed lines in (C,D) mark the extreme values, obtained when using this Monte-Carlo procedure. Black curves in A-D show the GMag obtained by a single SVD on individual fragments. Green lines in A and B are error estimates obtained from the Losonczi Monte-Carlo RMSD, whereas C and D show the range of solutions (using 1000 noise sets). The corresponding results for the structural noise Monte-Carlo evaluation are shown in red. (E,F) Difference in the local alignment tensor orientation, $\Delta\vartheta$, (as obtained by a single SVD on the corresponding fragment) relative to that determined from a SVD on the full structure are plotted as solid lines. Maximum angular deviations obtained from the Losonczi Monte-Carlo (○) and the structural noise Monte-Carlo (●) methods are also shown. (G,H) Number of available dipolar couplings for each fragment. Fragments are labeled with the number of the central residue.

protein. Therefore, the data of Figure 7A cannot be interpreted as proof of internal motion at Leu[8]. Interestingly, when considering the same dipolar couplings but using the X-ray structure (which differs from the solution structure by less than 0.4 Å rmsd for the backbone atoms) this fragment shows alignment parameters that are very close to the value obtained for the full protein (Figure 7B). This confirms that very small differences in local structure can have substantial effects on the apparent magnitude of observed alignment parameters. For several other fragments the local alignment tensor also falls more than one Monte-Carlo standard deviation above or below the alignment tensor value obtained for the full protein. However, for none of these does the entire Monte-Carlo ensemble fall outside the values for the full protein. We therefore conclude that the dipolar couplings used in this evaluation do not provide evidence for motion of the ubiquitin polypeptide backbone. When comparing the Losonczi Monte-Carlo method (green lines in Figure 7) with the structural noise Monte-Carlo method (red), similar patterns are obtained. Differences between the two methods tend to be largest for small N_{DC} values.

Figures 7E and 7F show the difference in the local alignment tensor orientation from the global alignment tensor. Clearly, substantial differences are observed, particularly for the X-ray structure. However, for nei-

ther structure does the difference for any 3-residue fragment exceed the maximum deviation obtained with the structural noise Monte-Carlo method. The spread in local alignment tensor orientation obtained with the Losonczi Monte-Carlo method tends to be smaller than that of the structural noise Monte-Carlo method. However, in both cases the high degree of uncertainty indicates that it is very difficult to spot rigid body structural rearrangements for small fragments from such a local dipolar coupling analysis.

As expected, the degree in uncertainty of the local alignment tensor correlates inversely with the number of dipolar couplings available for a given fragment. For example, some of the largest uncertainties in the X-ray structure derived alignment parameters are observed around residues Gly^{10} and Pro^{38}, two fragments for which only seven dipolar couplings are available. The smaller increase in uncertainty when evaluating the NMR structure is believed to be artificial: the smaller number of dipolar couplings available for the NMR structure calculation resulted essentially in exact fits between the structure and the dipolar couplings during the structure determination. In contrast, when the number of dipolar couplings available for a given fragment significantly exceeds the degrees of freedom it generally is impossible to obtain a perfect fit between structure and dipolar couplings because the assumptions regarding bond angles, peptide bond planarity, and uniform local dynamics made during NMR structure calculation are not strictly valid.

The better agreement between the local alignment tensor magnitude and orientation observed for the NMR structure over the X-ray structure does not imply that the NMR structure describes ubiquitin better than the X-ray structure. It merely serves to illustrate that the spread obtained in the Monte-Carlo evaluation is lower for structures that show good agreement with the dipolar couplings because lower noise values are used in their Monte-Carlo evaluations.

Based on the results of Figure 7, we conclude that the ubiquitin dipolar couplings provide no conclusive evidence for extensive backbone motions, even though such motions can be detected from ^{15}N relaxation studies for several regions in ubiquitin, e.g., residues T9-T12, and Q49-D52 (Tjandra et al., 1995). However, results plotted in Figure 7 neither confirm nor exclude the presence of large amplitude motions on a time scale slower than the overall rotational diffusion rates. The large uncertainty in the magnitude of alignment when analyzing small fragments merely poses upper limits for the amplitude of such motions.

Concluding remarks

When measuring one-bond dipolar couplings in weakly aligned proteins, the error in the measurement frequently is more than an order of magnitude smaller than the range of dipolar couplings measured. In this case, the residual when fitting the dipolar couplings to a reference structure generally will be dominated by the uncertainty in the coordinates of the reference structure. Both the Losonczi and structural noise Monte-Carlo methods provide reasonable estimates for the uncertainty in the derived alignment parameters. However, in the case of structural noise and a large number of dipolar couplings ($>\sim$20) the magnitude of the alignment tensor is systematically underestimated. When fitting dipolar couplings to very high resolution X-ray crystal structures (determined at a resolution $\leq$1.5 Å), the effect of structural noise is expected to be most severe for one-bond ^{15}N-^{1}H dipolar couplings, where the ^{1}H positions are typically added to the original coordinates by model building. In our experience, this frequently results in a less good fit of these couplings to the structure than for those between $^{13}C^{\alpha}$-$^{13}C'$ and $^{13}C'$-^{15}N.

The values for the alignment magnitude plotted in Figure 7 are, on average, derived from more than 12 dipolar couplings each. Considering that the structural noise in ubiquitin is rather small (ca. 5°), this corresponds to a favorable system for detecting the presence of motions from dipolar couplings measured in a single medium. However, as can be seen from Figure 7, even in this favorable case the maximum uncertainties in the magnitude of alignment for many of the three-residue fragments exceed 25%, which corresponds to uncertainties in the squared generalized order parameter, S^2, of nearly 50%. We therefore conclude that evaluation of local backbone mobility from variations in local alignment magnitude using a single medium is very difficult.

Only an approximate estimate for the uncertainties of the derived alignment parameters can be derived from Figures 1–4. A more precise estimate for a given structure and a set of experimental data requires use of either the Losonczi or structural noise Monte-Carlo methods. Uncertainty estimates obtained with the Losonczi Monte-Carlo procedure are quite sensitive to the assumed error and the acceptance cut-off criterion used. Therefore, it is recommended to follow the guidelines outlined in this paper when using this method.

Analogous to the evaluation shown in Figure 7, estimation of the uncertainty in the alignment tensor magnitudes for different domains in a given protein is needed in order to evaluate whether an observed difference in alignment tensor magnitude (requiring domain motion) or orientation (requiring a different relative domain orientation from the reference structure) is statistically significant.

Convenient tools for conducting the analyses described above, including the standard Monte-Carlo procedure and the structural noise Monte-Carlo method have been incorporated into the software package PALES (M. Zweckstetter, unpublished), which can be obtained from the authors.

Acknowledgements

We thank Tobias Ulmer for useful suggestions during the preparation of this manuscript. M.Z. is the recipient of a DFG Emmy Noether-Fellowship. This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the NIH.

References

Almond, A. and Duus, J.O. (2001) *J. Biomol. NMR*, **20**, 351–363.
Bewley, C.A. and Clore, G.M. (2000) *J. Am. Chem. Soc.*, **122**, 6009–6016.
Bothner-By, A.A. (1996) in *Encyclopedia of Nuclear Magnetic Resonance*, Grant, D.M. and Harris, R.K. (Eds.), Wiley, Chichester, pp. 2932–2938.
Clore, G.M. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 9021–9025.
Clore, G.M. and Garrett, D.S. (1999) *J. Am. Chem. Soc.*, **121**, 9008–9012.
Cornilescu, G., Marquardt, J.L., Ottiger, M. and Bax, A. (1998) *J. Am. Chem. Soc.*, **120**, 6836–6837.
Delaglio, F., Kontaxis, G. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 2142–2143.
Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. and Scharf, M. (1995) *J. Comput. Chem.*, **16**, 273–284.
Emsley, J.W. (1996) in *Encyclopedia of Nuclear Magnetic Resonance*, Grant, D.M. and Harris, R.K. (Eds.), Wiley, Chichester, pp. 2788–2799.
Fischer, M.W.F., Losonczi, J.A., Weaver, J.L. and Prestegard, J.H. (1999) *Biochemistry*, **38**, 9013–9022.
Gayathri, C., Bothnerby, A.A., Vanzijl, P.C.M. and Maclean, C. (1982) *Chem. Phys. Lett.*, **87**, 192–196.
Goto, N.K., Skrynnikov, N.R., Dahlquist, F.W. and Kay, L.E. (2001) *J. Mol. Biol.*, **308**, 745–764.
Hus, J.C., Marion, D. and Blackledge, M. (2000) *J. Mol. Biol.*, **298**, 927–936.
Losonczi, J.A., Andrec, M., Fischer, M.W.F. and Prestegard, J.H. (1999) *J. Magn. Reson.*, **138**, 334–342.
Meiler, J., Prompers, J.J., Peti, W., Griesinger, C. and Bruschweiler, R. (2001) *J. Am. Chem. Soc.*, **123**, 6098–6107.
Ottiger, M. and Bax, A. (1998) *J. Am. Chem. Soc.*, **120**, 12334–12341.
Sass, H.J., Musco, G., Stahl, S.J., Wingfield, P.T. and Grzesiek, S. (2000) *J. Biomol. NMR*, **18**, 303–309.
Sass, J., Cordier, F., Hoffmann, A., Cousin, A., Omichinski, J.G., Lowen, H. and Grzesiek, S. (1999) *J. Am. Chem. Soc.*, **121**, 2047–2055.
Saupe, A. and Englert, G. (1963) *Phys. Rev. Lett.*, **11**, 462–464.
Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.
Tjandra, N., Feller, S.E., Pastor, R.W. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 12562–12566.
Tjandra, N., Grzesiek, S. and Bax, A. (1996) *J. Am. Chem. Soc.*, **118**, 6264–6272.
Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M. and Bax, A. (1997) *Nat. Struct. Biol.*, **4**, 732–738.
Tolman, J.R., Al-Hashimi, H.M., Kay, L.E. and Prestegard, J.H. (2001) *J. Am. Chem. Soc.*, **123**, 1416–1424.
Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 9279–9283.
Tycko, R., Blanco, F.J. and Ishii, Y. (2000) *J. Am. Chem. Soc.*, **122**, 9340–9341.
Vermeulen, A., Zhou, H.J. and Pardi, A. (2000) *J. Am. Chem. Soc.*, **122**, 9638–9647.
Vijay-Kumar, S., Bugg, C.E. and Cook, W.J. (1987) *J. Mol. Biol.*, **194**, 531–544.
Wang, L.C., Pang, Y.X., Holder, T., Brender, J.R., Kurochkin, A.V. and Zuiderweg, E.R.P. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 7684–7689.
Zweckstetter, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 3791–3792.