ChemPubSoc Europe

DOI: 10.1002/cphc.201800626

CHEMPHYSCHEM Articles

Special Issue

Minute Additions of DMSO Affect Protein Dynamics Measurements by NMR Relaxation Experiments through Significant Changes in Solvent Viscosity

Johan Wallerstein[a] and Mikael Akke*[a]

Studies of protein–ligand binding often rely on dissolving the ligand in dimethyl sulfoxide (DMSO) to achieve sufficient solubility, and then titrating the ligand solution into the protein solution. As a result, the final protein–ligand solution contains small amounts of DMSO in the buffer. Here we report how the addition of DMSO impacts studies of protein conformational dynamics. We used ^{15}N NMR relaxation to compare the rotational diffusion correlation time (τ_C) of proteins in aqueous buffer with and without DMSO. We found that τ_C scales with the viscosity of the water–DMSO mixture, which depends sensitively on the amount of DMSO and varies by a factor of 2 across the relevant concentration range. NMR relaxation studies of side chains dynamics are commonly interpreted using τ_C as a fixed parameter, obtained from backbone ^{15}N relaxation data acquired on a separate sample. Model-free calculations show that errors in τ_C, arising from mismatched DMSO concentration between samples, lead to significant errors in order parameters. Our results highlight the importance of determining τ_C for each sample or carefully matching the DMSO concentrations between samples.

1. Introduction

Studies of ligand binding to proteins are fundamental to our understanding of biological processes and the development of efficient drugs for treatment of diseases. Synthetic organic compounds designed to bind to proteins with high affinity are often poorly soluble in water, because this property promotes binding to hydrophobic pockets and generates a favorable free energy of desolvation.[1,2] Mixed solvents are therefore often used to solubilize such ligands in order to characterize binding in vitro. Dimethyl sulfoxide (DMSO) is an effective dipolar solvent that is completely miscible with water and many organic liquids and has a low chemical reactivity. For this reason, DMSO is a preferred solvent, because ligands can be solubilized at high concentration in DMSO and the resulting solution can then be titrated into an aqueous protein buffer to yield a homogeneous solution containing the ligand–protein complex and a small amount of DMSO (typically a few percent and usually less than 10%). While high concentrations of DMSO typically denature proteins, low concentrations (< 10%) are not expected to affect the structure or stability of the folded protein appreciably.[3–5] However, progressive destabilization has been observed prior to the onset of global unfolding,[6,7] and effects on stability, aggregation, and ligand binding might be observed for some proteins already at low DMSO concentrations.[8]

Naturally, it is important to establish that DMSO does not interfere with ligand binding, lest characterization of the binding thermodynamics becomes corrupted. Previous studies have concluded that hydrophobic or hydrogen bond interactions on the surface of a protein are not sufficient to produce DMSO–protein complexes with life times greater than a few nanoseconds; instead a suitable DMSO-binding pocket or cleft appears to be required.[9] Here we address the effects of DMSO on the dynamical properties of proteins, with particular focus on the overall rotational correlation time, τ_C, as well as internal conformational fluctuations on the pico- to nanosecond timescale as manifested in the form of order parameters determined by the model-free approach.

In this context, it is important to recognize that the viscosity of water–DMSO mixtures depends sensitively on the amount of DMSO. This is particularly so in the dilute regime with DMSO mole fractions of 0–0.06 (0–20% v/v), across which the viscosity increases linearly by more than a factor of 3.[10–13] Thus, even small variations in the residual DMSO concentration can have significant effects on the solvent viscosity, which in turn affects the rotational correlation time of the solutes and the interpretation of NMR relaxation data in terms of internal dynamics.

We investigated the effects of varying DMSO concentrations on the rotational diffusion constant, τ_C, using as model systems two small proteins that have been studied extensively before: the B1 domain of Streptococcal protein G (56 residues, M_r 6.24 kDa, denoted PGB1) and the carbohydrate-binding domain of human galectin-3 (138 residues, M_r 15.71 kDa, denoted galectin-3C). Galectin-3 is an actively pursued drug target in many laboratories,[14] and we have previously investigated the

[a] J. Wallerstein, Prof. M. Akke
Biophysical Chemistry, Center for Molecular Protein Science
Department of Chemistry, Lund University
Box 124, SE-221 00 Lund, Sweden
E-mail: mikael.akke@bpc.lu.se

Special Issue An invited contribution to a Special Issue on BioNMR Spectroscopy

role of conformational entropy and hydration in ligand binding to galectin-3C.[15,16] Thus, it is important to investigate whether DMSO has any effects on the structure and dynamics of galectin-3C.

2. Results and Discussion

We successively added aliquots of DMSO to aqueous samples of galectin-3C such that samples with DMSO concentrations of $C_{DMSO} = 1.0$, 2.0, 3.9, 7.5, and 10.9% (v/v) were generated. At each concentration, we acquired a ^{1}H-^{15}N HSQC spectrum and monitored the ^{1}H and ^{15}N chemical shifts of the backbone amide groups. In the case of PGB1 we did not titrate in DMSO, but added it in a single step to yield a final concentration of $C_{DMSO} = 10.7\%$. For both proteins, we measured ^{15}N laboratory frame relaxation rates R_1 and R_2 on the sample with the highest concentration ($C_{DMSO} \approx 11\%$).

2.1. Transient and Nonspecific Binding of DMSO to the Protein Surface

We monitored chemical shift changes of the protein backbone amide ^{1}H and ^{15}N resonances as a function of DMSO concentration, see Equation (1). For the two proteins studied here, the addition of DMSO generates only minor chemical shift changes in the protein NMR spectrum. Figure 1 shows the chemical shift changes induced by addition of DMSO for galectin-3C (Figure 1A) and PGB1 (Figure 1B). The maximum chemical shift change in galectin-3C is 0.077 ppm, observed for the C-terminal residue I250 (Figure 1A), and 28 residues (25%) have $\Delta\delta_{HN} > 0.03$ ppm. For PGB1 the maximum chemical shift change is 0.052 ppm (loop residue V21) (Figure 1B) and 11 residues (21%) have $\Delta\delta_{HN} > 0.03$ ppm.

These chemical shift perturbations are considerably smaller than those typically expected for ligand binding, where in the case of galectin-3C there are 5–6 residues showing perturbations between 0.12-0.20 ppm upon binding of ligand.[16] A rule of thumb indicates that ligand binding to a well-defined site is expected to elicit a shift change of at least 0.2 ppm,[17] which is a factor of 2–3 greater than the highest values measured here. The modest shift changes with increasing DMSO concentration indicate that the binding affinities are very weak, with dissociation constants on the order of at least several hundred mM, and most likely considerably greater than that. This estimate is in line with results for specific DMSO binding to the active site in lysozyme, where $K_d = 0.4$ M, which is expected to be significantly higher in affinity than the non-specific interactions observed here. Thus, the DMSO affinity is much too weak to interfere with ligand binding.

In the case of galectin-3C, interactions with DMSO are observed primarily for solvent-exposed sites in loops or hairpins connecting β-strands (Figure 1D). Interestingly, there are no larger chemical shift changes within the binding site for natural and designed ligands of pharmaceutical interest, which is located on the backside of the protein in the view of Figure 1D.

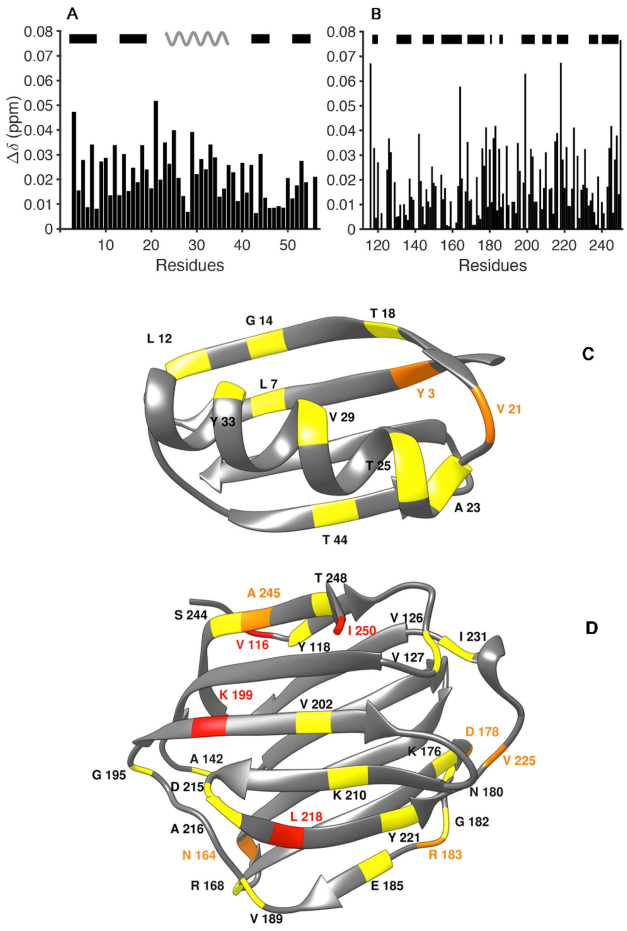

Figure 1. Changes in protein ^{1}H and ^{15}N chemical shifts induced by DMSO. (A, B) $\Delta\delta$ is plotted versus residue number for PGB1 (A) and galectin-3C (B). The secondary structure is indicated above each plot: black bar, β-strand; grey wave, α-helix. C, D) $\Delta\delta$ color coded onto the structure of PGB1 (C; PDB id 1PGB[18]) and ligand-free galectin-3C (D; PDB id 3ZSL[19]). The carbohydrate binding site in galectin-3C is located on the back side in the present view. The color scheme ranges from yellow (0.03 ppm $< \Delta\delta <$ 0.04 ppm) via orange to red ($\Delta\delta > 0.06$ ppm).

Reassuringly, these observations reinforce the conclusion that DMSO does not compete with designed ligands for the same site. For PGB1, the shift changes are evenly distributed over the smaller protein (Figure 1C). In either protein, no particular type of residue (hydrophobic, polar, etc.) is more influenced than others. For example, the five residues in galectin-3C with the largest chemical shift changes include charged residues (K and D), hydrophobic (V and L), and polar (N); similar results are observed for PGB1 (Figures 1C and D). Taken together, these results suggest that DMSO interacts mainly with solvent exposed regions of proteins.

2.2. Water-DMSO Mixture: Rotational Diffusion Correlation Time Scales with Viscosity

We determined ^{15}N R_1 and R_2 relaxation rate constants for 53 well-resolved backbone amides in PGB1 and 113 in galectin-3C (Figure 2). There are very significant differences between the

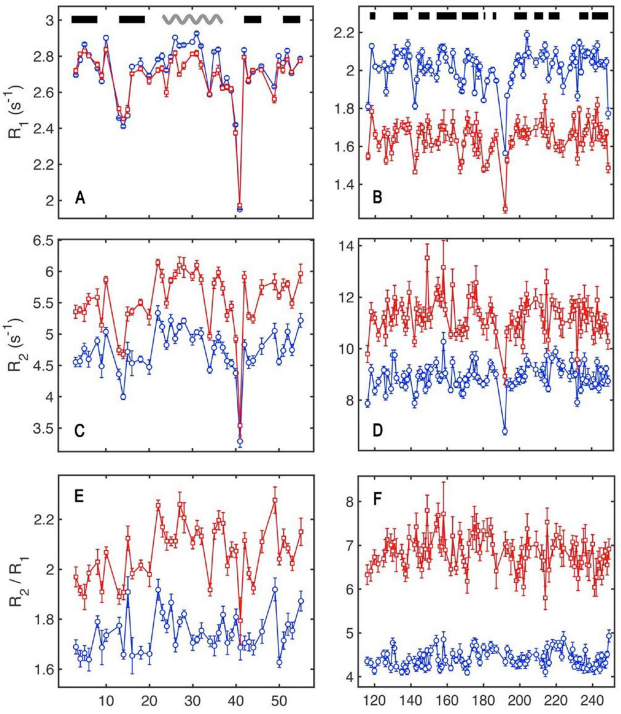

Figure 2. Comparison of protein backbone ^{15}N relaxation rates acquired on samples dissolved in aqueous buffer or a water–DMSO mixture of 11% (v/v) DMSO. R_1 relaxation rates are for PGB1 (A) and galectin-3C (B). R_2 relaxation rates for PGB1 (C) and galectin-3C (D). R_2/R_1 ratio for PGB1 (E) and galectin-3C (F). Blue circles indicate data sets acquired in aqueous buffer without DMSO; red squares indicate data sets acquired with a mole fraction of 0.03 or 11% (v/v) DMSO. Data were measured at a magnetic field strength of 11.7 T. The secondary structure is indicated at the top of panels (A) and (B): black bar, β-strand; wave, α-helix.

relaxation rate constants measured on samples with 11% DMSO and without DMSO, indicating significant changes in the rotational diffusion correlation times. The R_2 values are roughly 20% greater for both proteins in the sample containing 11% DMSO (Figures 2C and D). The R_1 rate constants for PGB1 appear nearly independent of the change in correlation time due to addition of DMSO (Figure 2A), while for galectin-3C (Figure 2B) there is a marked decrease in R_1 with addition of DMSO. This result is explained by the fact that R_1 is dominated by the spectral density term $J(\omega_N)$, which has its maximum at $\omega_N\tau_C \approx 1$ [cf. Equations (3) and (6) without the factors of 1/9]. Near its maximum, $J(\omega_N)$ depends weakly on minor changes in τ_C. It turns out that $\omega_N\tau_C \approx 1$ for PGB1 at the static magnetic field strength used here, 11.74 T (500 MHz); calculations yield $\omega_N\tau_C = 0.95$ (control) and 1.18 (DMSO) for PGB1, while the corresponding values for galectin-3C are $\omega_N\tau_C = 2.38$ (control) and 2.99 (DMSO).

We determined τ_C from the R_2/R_1 ratios using ROTDIF[20] on trimmed data sets that excluded those residues with R_2/R_1 values outside of 1 standard deviation from the mean in the samples without DMSO (denoted 'control'); the same residues were excluded from the data sets resulting from samples with DMSO. For PGB1 the trimmed data set (41 residues) had $<R_2/R_1> = 1.77 \pm 0.12$ (control) and $<R_2/R_1> = 2.08 \pm 0.13$ (DMSO; Figure 2E). For galectin-3C the corresponding numbers (95 residues) are $<R_2/R_1> = 4.45 \pm 0.20$ (control) and $<R_2/R_1> = 6.90 \pm 0.39$ (DMSO; Figure 2F). These R_2/R_1 ratios can be compared with numbers for the full data sets, which are in the case of PGB1 (53 residues) $<R_2/R_1> = 1.79 \pm 0.18$ (control) and $<R_2/R_1> = 2.10 \pm 0.15$ (DMSO). For galectin-3C (113 residues) we have $<R_2/R_1> = 4.61 \pm 0.62$ (control) and $<R_2/R_1> = 7.06 \pm 0.79$ (DMSO).

The resulting correlation times determined by ROTDIF are 7.07 ± 0.15 ns (control) and 9.34 ± 0.20 ns (DMSO) for galectin-3C, and 2.99 ± 0.31 ns (control) and 3.71 ± 0.22 ns (DMSO) for PGB1. The best-fit diffusion tensor model is fully anisotropic in the case of galectin-3C, with p-values of 0.01 (control) and 0.002 (DMSO), comparing the best-fit model with an axially symmetric diffusion tensor. The anisotropy of galectin-3C is $2D_{ZZ}/(D_{XX}+D_{YY}) = 1.12 \pm 0.05$ (control) and 1.11 ± 0.04 (DMSO). For PGB1, an axially symmetric model is preferred over the isotropic one, with p-values of 0.005 (control) and 0.006 (DMSO). PGB1 shows an anisotropy of 1.33 ± 0.25 (control) and 1.22 ± 0.13 (DMSO).

We compared the observed change in τ_C upon addition of 11% DMSO with that expected from the change in viscosity, as predicted by the Stokes-Einstein relationship, Equation (2). The ratio of the τ_C values measured with or without DMSO is $\tau_{C,DMSO}/\tau_{C,0} = 1.32 \pm 0.04$ for galectin-3C and 1.24 ± 0.15 for PGB1.

The viscosity of binary water-DMSO mixtures have been reported for different mole fractions,[10–13] albeit not at the exact same conditions as those used here. In the region of low DMSO concentration, the viscosity of the solvent mixture shows a nearly linear dependence on DMSO content. Thus, to obtain the viscosity η_{DMSO} of a 10.8% mixture (corresponding to a water mole fraction of $x = 0.970$) at 28 °C, we first interpolate between published data at 25 °C,[10] and then correct for the 3 degrees temperature difference using an Arrhenius-like expression, $\eta(T) = A \times \exp[B/T]$.[11] We determined the parameters A and B from temperature-dependent (25–65 °C) data measured at the water mole fractions $x = 0.945$ and $x = 0.975$,[10] which yielded $\eta(0.945)$ and $\eta(0.975)$ at 28 °C. Next, we performed a linear interpolation between these two temperature-corrected viscosities to determine $\eta(0.970)$ at 28 °C, which is the required value of η_{DMSO}. The calculations give $\eta_{DMSO} = 1.079$ mPa s. The viscosity of pure water at 28 °C is $\eta_0 = 0.832$ mPa s. Thus, the viscosity ratio is $\eta_{DMSO}/\eta_0 = 1.079/0.832 = 1.30$. This value is identical, within errors, to the values of $\tau_{C,DMSO}/\tau_{C,0}$ determined by NMR for galectin-3C and PGB1. Hence, the correlation time for rotational diffusion scales with the viscosity as expected from the Stokes-Einstein relationship [Eq. (2)], indicating that the hydration layer of the proteins in the presence of DMSO behaves just like the bulk solvent mixture. This observation is in line with the interpretation that DMSO does not accumulate at the protein surface to such an extent that it perturbs the hydration layer beyond the effects it exerts in the bulk solvent mixture.

The concentration of DMSO used here, 11% (v/v), is arguably higher than those typically encountered when titrating protein samples with poorly soluble ligands. Thus, the present results are likely to indicate a worst-case scenario. Having ascertained that the effect on τ_C of DMSO can be

predicted by the Stokes-Einstein relationship, we estimated the expected effect of lower DMSO concentrations, covering the range 1–5%. We calculated η for this concentration range by interpolating across the interval $x = 1.000–0.975$. The resulting values of $\eta_{DMSO}/\eta_0 = \tau_{C,DMSO}/\tau_{C,0}$ are 1.13, 1.05 and 1.025 for 5%, 2% and 1% (v/v) DMSO. We note that as a rule of thumb the percent increase in τ_C, compared to a sample without DMSO, equals the v/v percentage of DMSO multiplied by a factor 2.5. Thus, τ_C is significantly affected already at quite modest additions of DMSO that commonly result from titration with poorly water-soluble ligands. At concentrations higher than 5% DMSO, the rule-of-thumb scaling factor already begins to deviate from linearity, as gauged from our results reported above for 10% DMSO.

Many polar, organic solvents produce binary mixtures with water that exhibit a similar dependence of viscosity on molality as that observed for water–DMSO, i.e., a rapid increase in viscosity upon small additions of organic solvent and with a maximum around $x = 0.7–0.8$.[13] Thus, effects similar to those described here for water–DMSO are expected also for other organic solvents, such acetonitrile or dimethylformamide, which might also be used as ligand-carriers or conceivably occur as a residual from protein purification procedures.

2.3. Potential Effect of Between-Sample Variations in DMSO Concentration on Model-Free Interpretation of Relaxation Data

Protein dynamic studies based on NMR relaxation methods often target the backbone amides, but side-chain dynamics typically show greater response to ligand binding. Specifically, order parameters for methyl bearing side-chains have been highlighted as very useful probes for monitoring changes in conformational entropy upon ligand binding.[21,22] Model-free analysis of methyl ^{2}H relaxation experiments, typically takes τ_C, determined by ^{15}N relaxation data, as a fixed parameter.[23,24] Thus, accurate estimates of changes in order parameters and conformational entropy are critically dependent on correct measurement of τ_C, which depends sensitively on the concentration of DMSO (as shown above), which in turn might differ between the samples used for backbone and side-chain relaxation experiments.

We investigated the effects of using an incorrect τ_C value in model-free analysis of methyl dynamics. An initial analysis based on the model-free expression [Eq. (6)] reveals that the systematic error in order parameter due to an incorrect τ_C value is $\Delta S^2 = (\partial S^2/\partial \tau_C)\Delta\tau_C$, and the leading term of the relative error is approximately $\Delta S^2/S^2 = \Delta\tau_C/\tau_C$. Thus, invoking the rule of thumb described above, $\Delta\tau_C = 2.5\tau_C\Delta C_{DMSO}$, we expect that the relative error in S^2 can be described approximately as $\Delta S^2/S^2 = 2.5\Delta C_{DMSO}$.

To obtain further insight, we calculated ^{2}H relaxation rates using Equations (6) and (7a–d), as a function of the model-free parameters S^2, τ_e, and τ_C. Figure 3 shows how ^{2}H relaxation rate constants vary with τ_C for the case of a relatively rigid methyl axis, $S^2 = 0.8$. The shaded regions correspond to the change in τ_C resulting from a change in C_{DMSO} from 0 to 11% for PGB1 (grey) and galectin-3C (cyan). As observed, significant changes in relaxation rate constants arise as a consequence of relatively small changes in C_{DMSO}, demonstrating that a more in-depth analysis is warranted to outline the systematic errors in model-free parameters resulting from incorrect τ_C values.

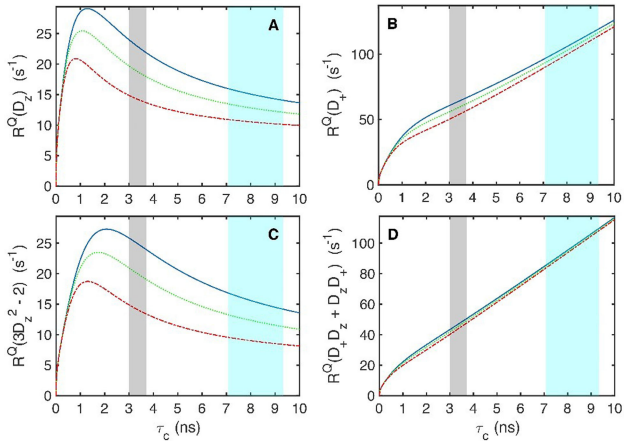

Figure 3. Calculated ^{2}H-relaxation rate constants plotted as function of the global correlation time τ_C. A) longitudinal magnetization R(D_Z), B) transverse magnetization R(D_+), C) quadrupolar order magnetization R($3D_Z^2-2$), and D) transverse anti-phase magnetization R($D_+D_Z+D_ZD_+$). Results are shown for three static magnetic field strengths: 11.7 T (blue, solid), 14.1 T (green, dot) and 18.8 T (red, dash) assuming $S^2 = 0.8$. The shaded areas highlight regions of τ_C corresponding to DMSO concentrations in the range from 0 (left-hand edge) to 11% (right-hand edge) for PGB1 (grey) and galectin-3C (cyan).

To investigate in more detail how fitted order parameters are affected by discrepancies between the actual τ_C of the protein in the sample used for ^{2}H methyl relaxation studies and the τ_C value used in the model-free fits, we first generated synthetic relaxation data sets comprising the 4 relaxation rate constants of Equations (7a–d), and then fitted model-free parameters to the resulting data sets, while keeping τ_C fixed to a value, denoted τ_C(fit), different from that used to generate the rate constants, denoted τ_C(input). We generated sets of relaxation rate constants using Equations (6) and (7) by varying S^2 from 0 to 1 in steps of 0.02, and τ_C over three ranges of values. The three ranges of τ_C values correspond in each case to a range of C_{DMSO} from 0 to 11%, starting at either 3.0 ns, 7.1 ns, or 14.2 ns, which equal τ_C at $C_{DMSO} = 0$ for PGB1, galectin-3C, and a hypothetical protein twice the size of galectin-3C. Each τ_C range was covered by 25 grid points. The model-free parameter τ_e was varied linearly from 20 to 200 ps in step with the change of S^2 from 1 to 0, as a first-level representation of physically reasonable models; however, the actual value of τ_e has only minor effects on the fitted parameters. In fitting model-free parameters to the synthetic relaxation data, we set τ_C(fit) to the value corresponding to $C_{DMSO} = 4\%$.

Figure 4 shows how the fitted S^2 is affected by a mismatch of τ_C. The heat maps represent the error in the fitted order parameter ΔS^2 as a function of the input values of τ_C and S^2. As expected from Equation (6), ΔS^2 increases as τ_C(fit) underestimates the real τ_C, and the effect is weakly dependent on protein size (cf. Figures 4A–C). Notably, for a given error in τ_C,

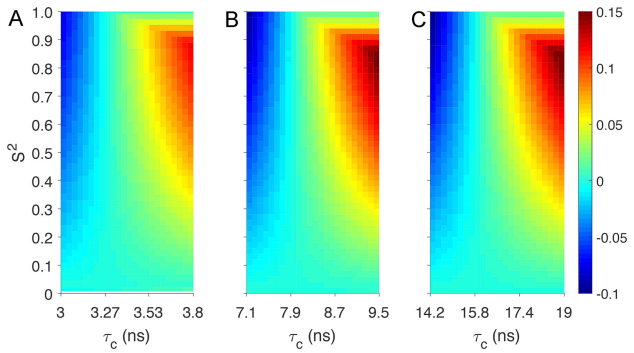

Figure 4. Error in the fitted methyl-axis order parameter due to error in τ_C. The heat map encodes $\Delta S^2 = S^2(\text{fit}) - S^2(\text{input})$ as a function of τ_C(input) and S^2(input). The input values of S^2 and τ_C were used to generate ^{2}H relaxation rate constants, which were subsequently taken as input for model-free fits using a fixed value of τ_C(fit) to yield S^2(fit) and τ_e(fit); see the text for details. The panels show results for three different ranges of τ_C, where the fixed value used in the fit is τ_C(fit) = 3.27 ns (A), τ_C(fit) = 7.9 ns (B), or τ_C(fit) = 14.2 ns (C), which corresponds to the case C_{DMSO} = 4% and coincides with the second tick mark.

the value of ΔS^2 depends also on S^2. For a small protein like PGB1, underestimating the DMSO concentration by 5% results in ΔS^2 as high as 0.1 (Figure 4A). For medium-sized and larger proteins, the same level of mismatch leads to ΔS^2 approaching 0.15 for rigid side chains with $S^2 \approx 0.8$ (Figures 4B and C). Figure 4 also confirms the expected decrease in S^2 (blue color) for the reverse case where τ_C is overestimated in the model-free fits. These results bear out the analytical estimate presented above, and also provide a detailed picture of how ΔS^2 varies with both the error in τ_C and the inherent motional amplitudes modeled by S^2. We find that quite modest levels of mismatch can lead to errors in order parameters that are significantly greater than the standard error of the estimates expected for methyl ^{2}H relaxation experiments.

3. Conclusions

The present results show that addition of small amounts of DMSO to aqueous protein solutions results in an increase in the global rotational correlation time as a function of solvent viscosity, in full agreement with the Stokes–Einstein equation. This result suggests that DMSO does not selectively perturb the hydration shell of the protein, which appears to behave just like the bulk solvent mixture.

Because the viscosity of water–DMSO mixtures depends sensitively on the amount of DMSO, τ_C varies significantly between samples that differ by only a few percent in DMSO concentration. Frequently, interpretation of NMR relaxation data for protein side-chains make use of τ_C values determined separately from ^{15}N backbone relaxation data on a near-identical sample. The present results indicate that very minor (on the order of 1–2%) differences in DMSO concentration between two such samples lead to incorrect side-chain order parameters, which could be devastating for comparative studies of different ligand–protein complexes. For this reason it is critical that τ_C is determined on the actual sample in question, or corrected for in a rigorous way. The same conclusion applies to any other method that depends on the overall correlation time, such as ligand-binding assays based on fluorescence depolarization of fluorophore-tagged competitive inhibitors.[25]

Minor chemical shift perturbations are pervasive across the protein surface, indicating that DMSO interacts transiently and non-specifically with the proteins, as might be expected since DMSO contributes up to 10% of the solvent volume. In the specific case of galectin-3C, DMSO does not compete with designed or natural ligands for the binding site.

Overall, these results demonstrate that reliable results can be attained when monitoring the effects of ligand binding on fast time-scale conformational fluctuations in proteins, also in the presence of low amounts of DMSO in the buffer, provided that the effect on τ_C is taken into account explicitly.

Experimental Section

NMR Sample Preparation

PGB1 and galectin-3C were expressed and purified as described.[15,16,26] The PGB1 construct includes three mutations introduced to avoid post-translational modifications, namely T2Q, N8D and N37D.[26] Uniformly ^{15}N/^{13}C-labeled samples at concentrations of 0.4–0.5 mM were prepared with 5% D_2O at pH 7.40, using 5 mM HEPES buffer for galectin-3C and a buffer-free solution for PGB1. Two samples, with or without added DMSO, were prepared for each protein. In the case of PGB1 DMSO was added in a single aliquot to yield $C_{DMSO} = 10.7 \pm 0.2\%$ (v/v) DMSO, whereas the galectin-3C sample was prepared by adding $5 + 5 + 10 + 20 + 20\ \mu$L aliquots of DMSO to the NMR tube to yield a final $C_{DMSO} = 10.9 \pm 0.5\%$ DMSO. Since DMSO was added in several steps to galectin-3C, the uncertainty in the final DMSO concentration is somewhat greater for this sample.

NMR Relaxation Experiments and Data Processing

NMR experiments were performed at a temperature of $28.0 \pm 0.2\,^{\circ}$C and a static magnetic field strength of 11.7 T. Temperature calibration was done with a type-T copper-constantan thermocouple element with one electrode placed in an ice-water bath and the other electrode placed in an NMR tube in water and positioned at the sample location inside the magnet. Longitudinal (R_1) and transverse (R_2) backbone ^{15}N relaxation experiments[27,28] were performed using relaxation delays of 0–1 s for R_1 and 0–0.2 s for R_2, each sampled by 12 data points. R_2 was measured with a 1.2 ms delay between refocusing 180° pulses in the CPMG train. Experiments were acquired with spectral widths of 8012.8 Hz (^{1}H) and 1519.5 Hz (^{15}N) in the R_1 experiment, and 6009.6 Hz and 1519.5 Hz in the R_2 experiment; in both experiments the number of points in the ^{1}H and ^{15}N dimensions were 1024 and 128. The ^{1}H carrier was placed on the water frequency and the ^{15}N carrier was placed in the center of the backbone amide region at 120 ppm. Spectra were processed using NMRPipe.[29] The processing protocol involved cosine apodization functions, zero filling to twice the number of increments in all dimensions, and baseline correction in the ^{1}H dimension. Peak intensities were extracted using CcpNmr Analysis and fitted to a single exponential decay using the boot strap error method.[30]

Chemical Shift Differences

Chemical shift differences were evaluated on a per-residue basis as the weighted Euclidean distance from the chemical shift observed in the absence of DMSO [Eq. (1)]:

$$\Delta\delta_{HN} = [(\Delta\delta_H)^2 + 0.1(\Delta\delta_N)^2]^{1/2} \quad (1)$$

where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift differences for the ^{1}H and ^{15}N nuclei, respectively. The weighting factor of 0.1 for ^{15}N shifts is based on the ratio $\gamma(^{15}N)/\gamma(^1H)$, where γ is the gyromagnetic ratio.[17]

Analysis of ^{15}N Relaxation Data

The Stokes-Einstein relationship gives the isotropic rotational correlation time for globular proteins with approximately spherical shape [Eq. (2)]:[31]

$$\tau_C = 4\pi\eta r_H^3/(3k_BT) \quad (2)$$

where η is the viscosity of the solvent (mixture), r_H is the hydrodynamic radius of the protein, k_B is Boltzmann's constant, and T is the temperature.

$$R_1 = d^2/4[J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H+\omega_N)] + c^2J(\omega_N) \quad (3)$$

$$R_2 = d^2/8[4J(0) + J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H+\omega_N)] + c^2/6[4J(0) + 3J(\omega_N)] + R_{ex} \quad (4)$$

$$NOE = 1 + \{\gamma_H/\gamma_N[6J(\omega_H+\omega_N) - J(\omega_H-\omega_N)]\}/R_1 \quad (5)$$

where ω_i is the Larmor frequencies of nuclide i, $J(\omega)$ is the spectral density, $d = \mu_0 h\ \gamma_H\gamma_N < r_{HN}^{-3} > /(8\pi^2)$, μ_0 is the permeability of free space, h is Planck's constant, γ_H and γ_N are the gyromagnetic ratios of ^{1}H and ^{15}N, respectively, r_{HN} is the distance between the two nuclei, $c = \gamma_N B_0 \Delta\sigma/3^{1/2}$, B_0 is the static magnetic field strength, $\Delta\sigma$ is the chemical shielding anisotropy of ^{15}N, and R_{ex} is the exchange contribution to R_2.

We determined τ_C by fitting the rotational diffusion tensor to the ^{15}N relaxation R_2 and R_1 data, using the MATLAB version of ROTDIF (v. 7)[20] with protein structures 1PGB[18] (PGB1) and 3ZSL[19] (galectin-3C) to extract the ^{1}H–^{15}N bond vector orientations in the molecular frame. Hydrogens were added to the PDB structures with the *addh* function in Chimera.[32] Specifically, ROTDIF determines the best-fit diffusion tensor based on the ratio $(2R'_2/R'_1-1)^{-1} = 3J(\omega_N)/4J(0)$, where the prime indicates that the rates are modified by subtracting the contributions from high-frequency components of the spectral density in Equations (3) and (4).[33] Normally, these components are determined from the NOE [Eq. (5)]. We used a fixed NOE value 0.8 for all residues; this approach leads to minor deviations in τ_C of 1–5%, which is within experimental errors. The standard errors of the fitted parameters were estimated using Monte Carlo simulations covering 300 samples.[34]

Model-free analysis of simulated ^{2}H relaxation data for methyl side-chains. We used the model-free formalism[24,35,36] to calculate ^{2}H relaxation rate constants as a function of model-free parameters. The spectral density function is modelled as [Eq. (6)]:[24]

$$J(\omega) = (1/9)S^2\tau_C/(1 + (\omega\tau_C)^2] + (1-(1/9)S^2)\tau/(1 + (\omega\tau)^2] \quad (6)$$

where S^2 is the order parameter of the methyl axis, $\tau = (1/\tau_C + 1/\tau_e)^{-1}$, and τ_e is the effective correlation time for internal motions. We considered 4 different ^{2}H relaxation rate constants [Eqs. (7a–d)]:[23]

$$R(D_Z) = (3/40)\{e^2qQ/h\}^2[J(\omega_D) + 4J(2\omega_D)] \quad (7a)$$

$$R(3D_Z^2-2) = (3/40)\{e^2qQ/h\}^2[3J(\omega_D)] \quad (7b)$$

$$R(D_+) = (1/80)\{e^2qQ/h\}^2[9J(0) + 15J(\omega_D) + 6J(2\omega_D)] \quad (7c)$$

$$R(D_+D_Z + D_ZD_+) = (1/80)\{e^2qQ/h\}^2[9J(0) + 3J(\omega_D) + 6J(2\omega_D)] \quad (7d)$$

where e^2qQ/h is the quadrupolar coupling constant, e is the elementary charge, eq is the principal component of the electric field gradient tensor, Q is the nuclear quadrupole moment, and h is Planck's constant. e^2qQ/h was set to 167 kHz.[23] The simulated data sets were subsequently fitted using the same model-free expressions as those used to generate the data, but with a fixed value of τ_C, so as to simulate the effect of mismatched DMSO concentrations between samples. The model-free fits were carried out using in-house MATLAB scripts.

Acknowledgements

This research was supported by the Knut and Alice Wallenberg Foundation (KAW 2013.022). Protein production was carried out by the Lund Protein Production Platform (LP3) at Lund University. We thank David Fushman, Kristofer Modig, Olof Stenström, and Ulrich Weininger for helpful discussions.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: order parameter · ligand binding · drug design · side-chain dynamics · rotational diffusion

[1] A. J. Ruben, Y. Kiso, E. Freire, *Chem. Biol. Drug Des.* **2006**, *67*, 2–4.
[2] E. Freire, *Drug Discovery Dev.* **2008**, *13*, 869–874.
[3] T. Arakawa, Y. Kita, S. N. Timasheff, *Biophys. Chem.* **2007**, *131*, 62–70.
[4] I. K. Voets, W. A. Cruz, C. Moitzi, P. Lindner, E. P. G. Areas, P. Schurtenberger, *J. Phys. Chem. B* **2010**, *114*, 11875–11883.
[5] A. N. L. Batista, J. M. Batista, V. S. Bolzani, M. Furlan, E. W. Blanch, *Phys. Chem. Chem. Phys.* **2013**, *15*, 20147–20152.
[6] A. L. Jacobson, C. L. Turner, *Biochemistry* **1980**, *19*, 4534–4538.
[7] M. Jackson, H. H. Mantsch, *Biochim. Biophys. Acta* **1991**, *1078*, 231–235.
[8] A. Tjernberg, N. Markova, W. J. Griffiths, D. Hallen, *J. Biomol. Screening* **2006**, *11*, 131–137.
[9] H. Johannesson, V. P. Denisov, B. Halle, *Protein Sci.* **1997**, *6*, 1756–1763.
[10] J. M. G. Cowie, P. M. Toporowski, *Can. J. Chem.* **1961**, *39*, 2240–2243.
[11] S. A. Schichman, R. L. Amey, *J. Phys. Chem.* **1971**, *75*, 98–102.
[12] J. Catalan, C. Diaz, F. Garcia-Blanco, *J. Org. Chem.* **2001**, *66*, 5846–5852.
[13] T. M. Aminabhavi, B. Gopalakrishna, *J. Chem. Eng. Data* **1995**, *40*, 856–861.
[14] H. Blanchard, X. Yu, P. M. Collins, K. Bum-Erdene, *Expert Opin. Ther. Pat.* **2014**, *24*, DOI 10.1517/13543776.2014.947961.
[15] C. Diehl, S. Genheden, K. Modig, U. Ryde, M. Akke, *J. Biomol. NMR* **2009**, *45*, 157–169.
[16] C. Diehl, O. Engström, T. Delaine, M. Håkansson, S. Genheden, K. Modig, H. Leffler, U. Ryde, U. J. Nilsson, M. Akke, *J. Am. Chem. Soc.* **2010**, *132*, 14577–14589.
[17] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.
[18] T. Gallagher, P. Alexander, P. Bryan, G. L. Gilliland, *Biochemistry* **1994**, *33*, 4721–4729.

[19] K. Saraboji, M. Håkansson, S. Genheden, C. Diehl, J. Qvist, U. Weininger, U. J. Nilsson, H. Leffler, U. Ryde, M. Akke, et al., *Biochemistry* **2012**, *51*, 296–306.
[20] O. Walker, R. Varadan, D. Fushman, *J. Magn. Reson.* **2004**, *168*, 336–345.
[21] K. K. Frederick, M. S. Marlow, K. G. Valentine, A. J. Wand, *Nature* **2007**, *448*, 325–330.
[22] J. A. Caro, K. W. Harpole, V. Kasinath, J. Lim, J. Granja, K. G. Valentine, K. A. Sharp, A. J. Wand, *Proc. Natl. Acad. Sci.* **2017**, *114*, 6563–6568.
[23] O. Millet, D. R. Muhandiram, N. R. Skrynnikov, L. E. Kay, *J. Am. Chem. Soc.* **2002**, *124*, 6439–6448.
[24] N. R. Skrynnikov, O. Millet, L. E. Kay, *J. Am. Chem. Soc.* **2002**, *124*, 6449–6460.
[25] P. Sorme, B. Kahl-Knutsson, M. Huflejt, U. J. Nilsson, H. Leffler, *Anal. Biochem.* **2004**, *334*, 36–47.
[26] S. Lindman, W. F. Xue, O. Szczepankiewicz, M. C. Bauer, H. Nilsson, S. Linse, *Biophys. J.* **2006**, *90*, 2911–2921.
[27] J. Kördel, N. J. Skelton, M. Akke, A. G. Palmer, W. J. Chazin, *Biochemistry* **1992**, *31*, 4856–4866.
[28] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, L. E. Kay, *Biochemistry* **1994**, *33*, 5984–6003.
[29] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–293.
[30] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, P. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins Struct. Funct. Bioinf.* **2005**, *59*, 687–696.
[31] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, M. Rance, N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice*, Elsevier Academic Press, San Diego, **2007**.
[32] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[33] D. Fushman, N. Tjandra, D. Cowburn, *J. Am. Chem. Soc.* **1998**, *120*, 10947–10952.
[34] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, Cambridge, **1986**.
[35] B. Halle, H. Wennerström, *J. Chem. Phys.* **1981**, *75*, 1928–1943.
[36] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.

Manuscript received: July 2, 2018
Accepted manuscript online: August 13, 2018
Version of record online: September 3, 2018