J|A|C|S
COMMUNICATIONS

Probing Supramolecular Structure from Measurement of Methyl ^{1}H−^{13}C Residual Dipolar Couplings

Remco Sprangers and Lewis E. Kay*

*Departments of Medical Genetics, Biochemistry, and Chemistry, The University of Toronto,
Toronto, Ontario M5S 1A8, Canada*

Received August 3, 2007; E-mail: kay@pound.med.utoronto.ca

Among the most significant advances in solution-based NMR spectroscopy studies of biological molecules in the past decade are methods for the introduction and subsequent measurement of residual dipolar couplings (RDCs) in weakly aligned systems.[1,2] In the case of protein applications, most of these measurements focus on the backbone, and ^{1}H−^{15}N RDCs, in particular, are now routinely used in structure determination. Applications have, however, been limited to systems smaller than 100 kDa[3,4] as the spectral quality of proton−nitrogen correlation maps decreases rapidly with increasing molecular weight. For larger systems, proton−carbon methyl spectra that are recorded using the methyl-TROSY principle can be of very high quality.[5] We have recently established that detailed information on supramolecular machines can be obtained using methyl-TROSY-based experiments and highly deuterated protein samples that contain selectively protonated methyl groups in isoleucine, leucine, and valine residues.[6−8] Most of our efforts have focused on obtaining insight into protein dynamics and molecular interactions; here we show that methyl-TROSY spectroscopy can also be used to probe the structure of protein complexes with molecular masses in the hundreds of kDa in a direct manner via measurement of ^{1}H−^{13}C RDCs.

Several methods have been introduced for the measurement of RDCs in methyl groups;[9−12] however, these applications suffer from the fact that they mix fast and slowly relaxing magnetization components, leading to significantly reduced signal-to-noise in spectra of slowly tumbling macromolecules. The TROSY−HMQC[5]-based scheme of Figure 1 insures that such mixing does not occur. It departs from HSQC-based approaches that quantify methyl ^{1}H−^{13}C RDCs from splittings in the ^{13}C dimension since the long-lived signal of interest in the methyl-TROSY experiment used here derives from a component that is not split by ^{1}H−^{13}C coupling in F_1.[5] Instead, RDCs are obtained from splittings in F_2, exploiting the IPAP principle,[13,14] whereby in-phase (IP) and anti-phase (AP) doublet spectra (in F_2) are combined to produce a pair of data sets, each containing only one of the two doublet signals.[14] The experiment is best performed on samples with high levels of deuteration and where only one of the two isopropyl methyl groups of Leu and Val is protonated to minimize line broadening from spin relaxation contributions[15] and long-range proton−proton dipolar couplings that in some cases can lead to artifacts.[16] An additional advantage of the methyl-TROSY dipolar coupling experiment is that the slowly relaxing proton magnetization is not split by ^{1}H−^{1}H dipolar couplings in F_2 since the sum of all local ^{1}H dipolar fields is zero for the transitions of interest.

The pulse scheme of Figure 1 was applied to a sample of a U-[^{2}H], Ile-δ_1-[^{13}CH$_3$], Leu, Val-[^{13}CH$_3$,^{12}CD$_3$]-labeled 360 kDa protein complex comprising 14 copies of the α-subunit of the 20S proteasome[7] that we have referred to as $\alpha_7\alpha_7$ previously. We observed only one of the F_2 doublet components in each subspectrum, both in the absence and in the presence of alignment media (Figure 2), leading to significant spectral simplification in comparison to data sets recorded with the TROSY-HMQC scheme without carbon decoupling during acquisition and without the IPAP approach. Very high quality spectra were obtained, even for the fractionally aligned sample where the protein concentration was 200 μM (in monomer). From multiple measurements, we estimate the precision of the extracted couplings to be about 1 Hz, which is also reflected in the high correlation between couplings extracted from measurements on samples that have different residual alignment (Figure 3a).

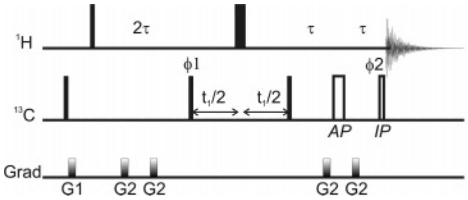

Figure 1. Methyl-TROSY-based pulse scheme for the measurement of methyl ^{1}H−^{13}C RDCs in high molecular weight protein complexes. The pulses labeled AP (IP) are only applied in the experiment that yields the anti-phase (in-phase) spectrum. The two separately recorded experiments (IP and AP) are subsequently added/subtracted to give spectra that contain only one of the two F_2 doublet components. See SI for details.

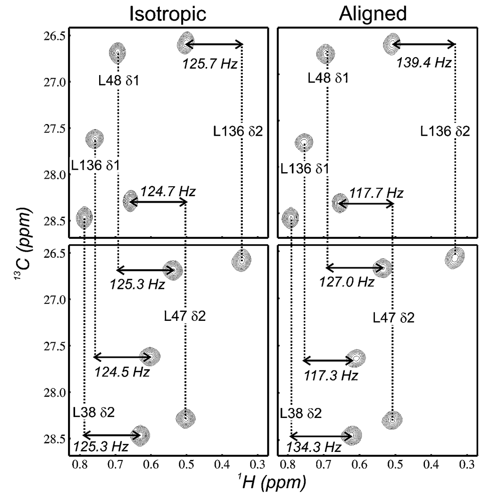

Figure 2. Spectrum of the 360 kDa “half-proteasome”[7] in the isotropic (left) and aligned (right) phases. The top (bottom) spectra contain only the downfield (upfield) component of the carbon-coupled spectrum. See SI for details.

Extracted residual dipolar couplings for $\alpha_7\alpha_7$ are between −10 and 17 Hz in the presence of ∼40 mg/mL Pf1 phage. As has been recognized early, RDCs contain information on both protein structure and dynamics,[1,17] and in the limit of axially symmetric internal motion about an averaging axis, the measured RDC can be expressed as RDC$_{\text{measured}}$ = RDC$_{\text{static}}$ × S, where RDC$_{\text{static}}$ is the

10.1021/ja075846i CCC: $37.00

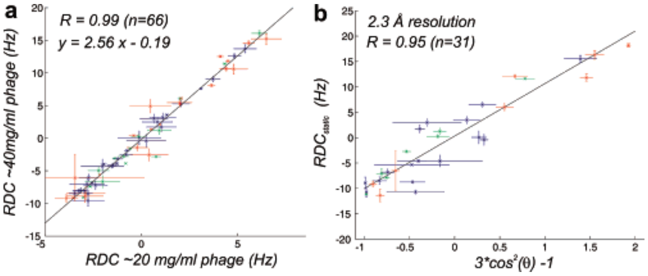

Figure 3. (a) Correlation between residual dipolar couplings measured in the presence of ∼20 and ∼40 mg/mL Pf1 phage (ASLA Biotech Ltd.). Green, blue, and red points indicate data from isoleucine, leucine, and valine residues, respectively. RDC values are those measured directly from splittings in spectra, Figure 2, and are not "corrected" for the effects of internal dynamics. (b) Correlation between $3\cos^2\theta - 1$ from the 2.3 Å resolution crystal structure of the full proteasome and RDC_{static} ($RDC_{measured}/S$). The error bars in $3\cos^2\theta - 1$ are obtained from the seven unique monomers in the crystal. See SI for details.

coupling that would be obtained for a bond along the time-averaged orientation in the absence of internal protein dynamics and S is the reduction in the size of the observed coupling due to internal motions. For structured parts of the protein backbone, these motions are usually small, and the dynamic component of the $^1H-^{15}N$ RDC is often assumed to be uniform.[18] Side chains, on the other hand, are often differentially mobile, with dynamics that can be asymmetric, leading to complications in the interpretation of RDCs in terms of structure.[11] Here values of RDC_{static} were obtained from $RDC_{measured}$ by dividing measured values with the square root of the previously determined order parameter squared, S^2 (ref 7), that is reasonable for side chains with fast time scale, low amplitude motions. It should be noted that slower time scale dynamics also affect $RDC_{measured}$; however, we have not made an attempt to correct for these motions here.

For the axially symmetric 20S proteasome, the measured RDCs are a function of the angle θ between the $H_3C_{methyl}-C$ bond and the symmetry axis of the protein according to

$$RDC = A \times \langle (3\cos^2\theta - 1) \rangle$$

where A is a constant that depends, among other things, on the degree of alignment and the angular brackets $\langle\ \rangle$ denote internal averaging. Figure 3b shows the correlation between RDC_{static} and $3\cos^2\theta - 1$ for values of θ obtained from a 2.3 Å resolution crystal structure of the 20S proteasome in complex with the 11S activator[19] for residues with limited internal dynamics ($S^2 > 0.6$, $S > 0.77$). A high correlation ($R = 0.95$) is obtained (i) despite the intermediate resolution of the crystal structure, where the high degree of structural variability is reflected in the large error bars for $3\cos^2\theta - 1$, and (ii) the fact that RDCs have been "corrected" for fast time scale motions only in an approximate manner. A more rigorous correction would have to take into account the exact nature of the motion; this level of detail is not available presently. Notably, the correlation is still good ($R = 0.92$) when all residues are included, independent of the amplitude of motion, as illustrated in SI, Figure 1. Small changes in structure can have a significant effect on calculated values of $3\cos^2\theta - 1$. Of interest, a very poor correlation is obtained when a 3.4 Å resolution crystal structure[20] is used (see SI, Figure 2) where the side-chain conformations are likely biased by default database rotamer conformations and not solely by diffraction data.

In summary, we have shown that accurate methyl $^1H-^{13}C$ residual dipolar couplings can be measured for very high molecular weight complexes. Such couplings can be used as upper bounds in structure calculations[11] or in programs that generate ensembles based on constraints that include the effects of dynamics, such as methyl group order parameters.[21] In combination with chemical shift mapping and molecular docking,[22] methyl group RDCs will be able to provide direct structural information for weakly interacting or transient biomolecular complexes that can be hard to obtain by X-ray crystallography or electron microscopy.

Acknowledgment. R.S. acknowledges the Canadian Institutes of Health Research (CIHR) Training Program in Protein Folding and Disease for a postdoctoral fellowship. This work was supported by a grant from the CIHR. L.E.K. is the recipient of a Canada Research Chair in Biochemistry.

Supporting Information Available: Details for Figures 1−3 and figures of the correlation between all RDC_{static} and $3\cos^2\theta - 1$ for Ile, Leu, and Val residues based on the 2.3 and 3.4 Å resolution crystal structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.
(2) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(3) Bax, A.; Grishaev, A. *Curr. Opin. Struct. Biol.* **2005**, *15*, 563−570.
(4) Tugarinov, V.; Kay, L. E. *J. Mol. Biol.* **2003**, *327*, 1121−1133.
(5) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420−10428.
(6) Sprangers, R.; Gribun, A.; Hwang, P. M.; Houry, W. A.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16678−16683.
(7) Sprangers, R.; Kay, L. E. *Nature* **2007**, *445*, 618−622.
(8) Velyvis, A.; Yang, Y. R.; Schachman, H. K.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8815−8820.
(9) Kaikkonen, A.; Otting, G. *J. Am. Chem. Soc.* **2001**, *123*, 1770−1771.
(10) Kontaxis, G.; Bax, A. *J. Biomol. NMR* **2001**, *20*, 77−82.
(11) Ottiger, M.; Delaglio, F.; Marquardt, J. L.; Tjandra, N.; Bax, A. *J. Magn. Reson.* **1998**, *134*, 365−369.
(12) Vogeli, B.; Kovacs, H.; Pervushin, K. *J. Am. Chem. Soc.* **2004**, *126*, 2414−2420.
(13) Yang, D. W.; Nagayama, K. *J. Magn. Reson. A* **1996**, *118*, 117−121.
(14) Ottiger, M.; Delaglio, F.; Bax, A. *J. Magn. Reson.* **1998**, *131*, 373−378.
(15) Tugarinov, V.; Kay, L. E. *J. Biomol. NMR* **2004**, *28*, 165−172.
(16) Tjandra, N.; Bax, A. *J. Magn. Reson.* **1997**, *124*, 512−515.
(17) Peti, W.; Meiler, J.; Bruschweiler, R.; Griesinger, C. *J. Am. Chem. Soc.* **2002**, *124*, 5822−5833.
(18) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732−738.
(19) Forster, A.; Masters, E. I.; Whitby, F. G.; Robinson, H.; Hill, C. P. *Mol. Cell* **2005**, *18*, 589−599.
(20) Lowe, J.; Stock, D.; Jap, B.; Zwickl, P.; Baumeister, W.; Huber, R. *Science* **1995**, *268*, 533−539.
(21) Lindorff-Larsen, K.; Best, R. B.; Depristo, M. A.; Dobson, C. M.; Vendruscolo, M. *Nature* **2005**, *433*, 128−132.
(22) Dominguez, C.; Boelens, R.; Bonvin, A. M. *J. Am. Chem. Soc.* **2003**, *125*, 1731−1737.

JA075846I