Communication

De Novo Determination of Bond Orientations and Order Parameters from Residual Dipolar Couplings with High Accuracy

Kathryn B. Briggman, and Joel R. Tolman

J. Am. Chem. Soc., **2003**, 125 (34), 10164-10165• DOI: 10.1021/ja035904+ • Publication Date (Web): 30 July 2003

J|A|C|S
COMMUNICATIONS

Published on Web 07/30/2003

De Novo Determination of Bond Orientations and Order Parameters from Residual Dipolar Couplings with High Accuracy

Kathryn B. Briggman and Joel R. Tolman*

Department of Chemistry, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218

Received May 1, 2003; E-mail: tolman@jhu.edu

The ability to measure residual dipolar couplings (RDCs) has opened up new prospects for the characterization of the structural and dynamic properties of biomolecules in solution.[1] Notable benefits, such as the ability to accurately establish order between remote portions of a molecule, have been realized almost immediately. Yet other anticipated benefits have been slower to materialize. This in large part arises from the unfortunate circumstance that the conversion of RDCs into useful structural constraints, not to mention information concerning dynamics, is not straightforward unless prior structural knowledge is available. While there has been steady progress in the development of alternative strategies for the interpretation of RDC data,[2,3] it is still generally necessary to make some prior assumptions about structure or dynamics to proceed. For cases in which RDC data can be collected in many ($\geq$5) different alignment media, we have recently proposed a method, called DIDC (direct interpretation of dipolar couplings),[4] which dispenses entirely with the need for prior information (including resonance assignments). Moreover, it can determine bond orientations to high accuracy alongside a description of their respective amplitudes of motion expressed as generalized order parameters. We present here an experimental demonstration of the method for the amide N−H bonds of the protein ubiquitin.

The DIDC approach is somewhat analogous to the reconstruction of a 3D object (the structure and dynamics of the macromolecule) based on a series of snapshots taken from different perspectives, in this case, RDC datasets collected utilizing different alignment media (Figure 1). If it can be assumed that the properties of the macromolecule do not change between different alignment media, DIDC simply provides the recipe for proceeding directly from raw measurements to the desired bond orientations and order parameters. The method is capable of delivering high accuracies, due to both its built-in signal averaging properties and that the only uncertainties introduced by the method itself (as opposed to uncertainties of measurement) are any deviations arising from selection, among a limited range of possible solutions, of the one which exhibits the least internal flexibility on average. Instead of assuming a rigid model for purposes of data interpretation, DIDC utilizes a model of minimum flexibility which remains consistent with the data.

In the present application to the protein ubiquitin, a total of 11 RDC datasets were employed, of which six correspond to physically different media as opposed to differences in concentrations or ionic strengths. With the exception of two bicelle datasets taken from the literature,[5] data were collected redundantly using both frequency-based (IPAP-HSQC)[6] and intensity-based (HSQC-PEC^2)[7] experimental approaches to ensure the absence of measurable systematic errors. The most challenging aspect of the DIDC approach is the requirement that RDC datasets be collected in at least five different alignment media. Therefore, the first step is the assessment of whether sufficient data have been collected to meet this requirement.

This assessment proceeds via singular value decomposition (SVD) of the data matrix, D, which consists of the raw coupling measurements row-indexed by residue and column-indexed according to the alignment medium employed. Shown in Figure 2 are the singular values of the data matrix D formed from the 57 ^{15}N−^{1}H RDC measurements obtained for ubiquitin in each of the 11 datasets acquired. The presence of five singular values with significant magnitude indicates that, barring the persistence of serious errors in the data, the data collectively represent the required five independent datasets. Potential complications can arise if a change in alignment media were to perturb the structure or dynamics of the protein. This possibility can be evaluated on the basis of the consideration of the six (and higher) singular values, provided that at least six different media have been employed, as is the case here. The magnitude of the sixth singular value is not significantly distinguishable from the noise, supporting a conclusion that the structure and dynamics of ubiquitin remain consistent between alignment medium, in overall agreement with the conclusions of Hus et al.[9]

Once assured that the RDC data represent the required five independent media, the determination of bond orientations and order parameters proceeds according to our previous description.[4] Shown in Figure 3 are the residue-by-residue generalized order parameters and the angular deviations from the X-ray structure, along with the associated uncertainties. Strikingly, the RDC-derived order parameters exhibit some correlation ($r = 0.63$) with those obtained

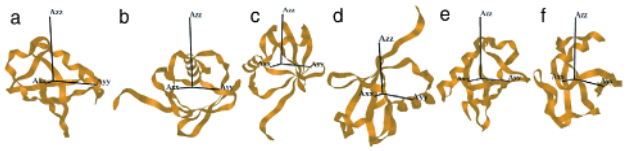

Figure 1. RDC "snapshots" of the protein ubiquitin obtained used six different alignment media:[1a] (a) DHPC/DMPC/CTAB, (b) Bacteriophage Pf1, (c) DHPC/DMPC, (d) *n*-dodecylpenta(ethylene glycol)/*n*-hexanol, (e) purple membranes, and (f) Helfrich phase. The graphics were created using the program Module.[8]

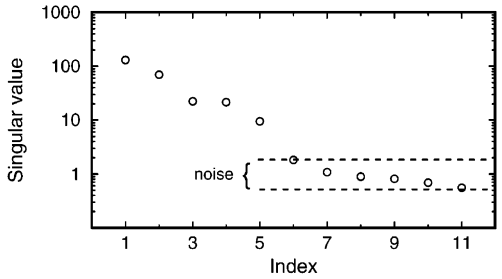

Figure 2. A plot of the singular values of the data matrix can reveal how strongly each of the required five independent alignments are present in the data. If the molecular structure and dynamics do not change between media, the acquisition of additional data can only enhance the magnitude of the first five singular values. On the other hand, the incoherence of random errors will ensure that they are distributed (more or less) evenly across all singular values. The eigenvectors corresponding to, and identified by, the largest five singular values form the basis of DIDC.

10.1021/ja035904+ CCC: $25.00

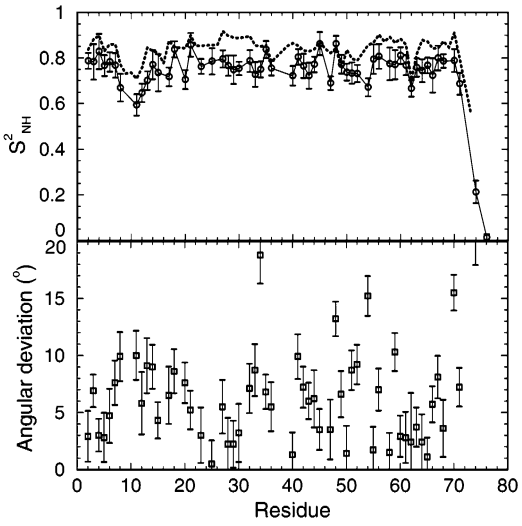

Figure 3. Residue-by-residue comparison of the DIDC results with the amide N−H orientations derived from an X-ray structure (1UBQ)[11] and picosecond time scale order parameters resulting from analysis of ^{15}N relaxation rates. Error bars are the result of the analysis of 10 000 synthetic datasets generated by addition of normally distributed random noise to the data with $\sigma = 0.2$ Hz. This corresponds to the largest experimental uncertainty among all the datasets. Top: Comparison of RDC-derived order parameters squared (○) and those from a previous ^{15}N spin relaxation study[10] (· · ·). As the overall scaling for RDC-derived order parameters is arbitrary,[3,4] the above were scaled to be equal to or less than the corresponding ^{15}N spin relaxation order parameters. Bottom: Angular deviation between (i.e., angles subtended by) DIDC N−H bond vectors and those derived from the X-ray coordinates (1UBQ), for which ^{1}H positions were calculated using the program MOLMOL.[11] Note that for purposes of comparison the proper N−H labeling was chosen, but in general RDCs are invariant to interchange of nuclei.

from ^{15}N spin relaxation rates. This is despite that RDCs have a much broader sensitivity to different motional time scales (picoseconds to milliseconds) than ^{15}N spin relaxation derived order parameters, which reflect amplitudes of motion occurring on just the picosecond time scale. Notably, the resulting dipolar order parameters exhibit considerably less variation between residues than was reported in previous studies of ubiquitin, which all proceeded on the basis of prior structural information.[3b,4] Although the effects of structural accuracy have not been thoroughly investigated, we believe that the removal of the need for prior knowledge of alignment tensors (and hence the dependence on structural accuracy) delivers gains in precision and accuracy for DIDC relative to these previous approaches. This conclusion is further supported by the correlation obtained with spin relaxation order parameters, an occurrence which would seem improbable by chance.

The structural precision of DIDC is remarkably high (Figure 3), with uncertainties in determined N−H bond orientations of ca. 2°. The resulting N−H bond orientations are also highly accurate, as shown in Figure 4. The de novo determination of N−H vector orientations by DIDC reproduces, essentially within experimental uncertainty, the results obtained upon refinement of N−H orientations using the same RDC data and starting from either solid or solution state coordinates. These results are encouraging for the prospect of applying DIDC to systems considerably less favorable than ubiquitin, while maintaining high resolution.

The prospects for the routine application of DIDC as the basis for NMR structure determination hinge, in large part, on whether

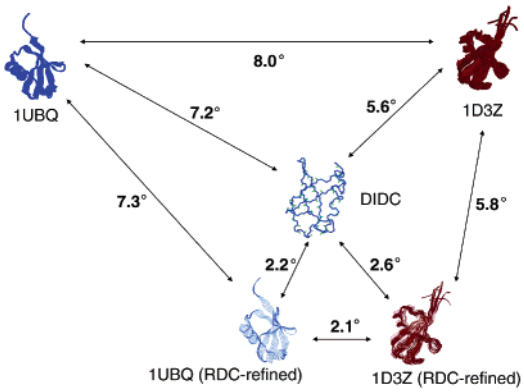

Figure 4. Angular RMSDs between N−H bond orientations derived from X-ray (1UBQ)[12] and NMR (1D3Z)[13] coordinates, the orientations resulting after refinement using RDCs and the N−H bond orientations obtained using DIDC.

the ability to modulate alignment demonstrated for ubiquitin proves to be typical or an exception. If not an exception, then potentially significant new applications become accessible to NMR spectroscopy. Studies of intermediate time scale (10^{-9}−10^{-5} s) dynamics are clearly among these. However, the ability to determine bond orientations with this level of precision and accuracy is also promising for structure determination of larger proteins, for which the availability of a sufficient number of structural constraints is often a limiting factor.

Acknowledgment. This work was supported by start-up funds provided by Johns Hopkins University.

Supporting Information Available: Details of alignment media and corresponding alignment tensors. Table of RDC data. RDC order parameters resulting from the refinement-based approach (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Prestegard, J. H.; Kishore, A. I. *Curr. Opin. Chem. Biol.* **2001**, *5*, 584−590. (b) Tolman, J. R. *Curr. Opin. Struct. Biol.* **2001**, *11*, 532−539. (c) MacDonald, D.; Lu, P. *Curr. Opin. Struct. Biol.* **2002**, *12*, 337−343.
(2) (a) Hus, J. C.; Marion, D.; Blackledge, M. *J. Am. Chem. Soc.* **2001**, *123*, 1541−1542. (b) Tian, F.; Valafar, H.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 11791−11796. (c) Al-Hashimi, H. M.; Gorin, A.; Majumdar, A.; Gosser, Y.; Patel, D. J. *J. Mol. Biol.* **2002**, *318*, 637−649. (d) Tolman, J. R.; Al-Hashimi, H. M.; Kay, L. E.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 1416−1424. (e) Giesen, A. W.; Homans, S. W.; Brown, J. M. *J. Biomol. NMR* **2003**, *25*, 63−71.
(3) (a) Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Bruschweiler, R. *J. Am. Chem. Soc.* **2001**, *123*, 6098−6107. (b) Peti, W.; Meiler, J.; Bruschweiler, R.; Griesinger, C. *J. Am. Chem. Soc.* **2002**, *124*, 5822−5833. (c) Hus, J. C.; Bruschweiler, R. *J. Chem. Phys.* **2002**, *117*, 1166−1172.
(4) Tolman, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 12020−12030.
(5) Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 12334−12341.
(6) Ottiger, M.; Delaglio, F.; Bax, A. *J. Magn. Reson.* **1998**, *131*, 373−378.
(7) (a) Tolman, J. R.; Prestegard, J. H. *J. Magn. Reson., Ser. B* **1996**, *112*, 245−252. (b) Cutting, B.; Tolman, J. R.; Nanchen, S.; Bodenhausen, G. *J. Biomol. NMR* **2002**, *23*, 195−200.
(8) Dosset, P.; Hus, J. C.; Marion, D.; Blackledge, M. *J. Biomol. NMR* **2001**, *20*, 223−231.
(9) Hus, J. C.; Bruschweiler, R. *J. Biomol. NMR* **2002**, *24*, 123−132.
(10) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.
(11) Koradi, R.; Billeter, M.; Wuethrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.
(12) Vijaykumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.
(13) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836−6837.

JA035904+