J|A|C|S COMMUNICATIONS

Published on Web 09/18/2009

One-Sample Approach to Determine the Relative Orientations of Proteins in Ternary and Binary Complexes from Residual Dipolar Coupling Measurements

Marco Tonelli,[†] Larry R. Masterson,[‡] Gabriel Cornilescu,[†] John L. Markley,[†] and Gianluigi Veglia*[,‡,§]

National Magnetic Resonance Facility at Madison, Department of Biochemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706-1544, and Departments of Biochemistry, Molecular Biology, and Biophysics, and Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431

Received June 17, 2009; E-mail: vegli001@umn.edu

NMR spectroscopy is a powerful method for analyzing changes in the structure and dynamics of macromolecular complexes at the atomic level.[1−4] NMR distance restraints used to define the conformational space of protein−protein assemblies are commonly derived from intermolecular nuclear Overhauser effects,[5−7] paramagnetic relaxation enhancements,[8,9] cross-saturation,[10] and chemical shift perturbations.[11] However, because full analysis requires spectral information from each partner in the complex, multiple samples are used for such measurements. Thus the full atomic details of protein assemblies, which only NMR can offer *in solution*, can be time and cost restrictive. Other strategies have exploited amino acid selective[12] or asymmetric labeling patterns[13,14] to facilitate measurements of intermolecular distances. Unfortunately, only the interface is characterized and the full backbone conformation and relative orientation of protein−protein complexes is not defined. In some cases, the approach requires pulse sequences which are insensitive for large assemblies with inherently short transverse relaxation times (T_2).

The conformational space of macromolecular complexes can also be determined from orientational restraints derived from residual dipolar couplings (RDCs) obtained in anisotropic alignment media.[15,16] Again, a common strategy for avoiding spectral overlap and assigning NMR signals to a specific subunit is to prepare multiple samples with differential labeling patterns. Particular care is necessary in the case of RDC measurements because slight changes in the experimental conditions can alter the alignment tensor, such that orientational constraints from different samples cannot be correlated directly.

The implementation of a spin−echo difference during a constant time period of an HSQC was originally introduced by Bax and co-workers[17] for the measurement of side chain dihedral angles. We recently applied a similar constant time spin−echo filter element to replace a two-dimensional version of the triple resonance HNCO pulse sequence to identify sequential pairs of amino acids in large proteins and enzymes[18] and extended this idea to enable chemical shift perturbation mapping of samples containing three isotopically labeled species in solution.[19]

Here we report an approach that, combined with an asymmetric isotopic labeling scheme, enables simultaneous measurement of RDCs from subunits of binary and ternary complexes with high sensitivity. Unlike recently reported schemes,[20,21] the pulse sequence used in this approach is shorter and provides higher sensitivity, particularly for large systems (>40 kDa). Also, the approach supports the unambiguous identification of spectral information from each subunit of a binary or ternary complex.

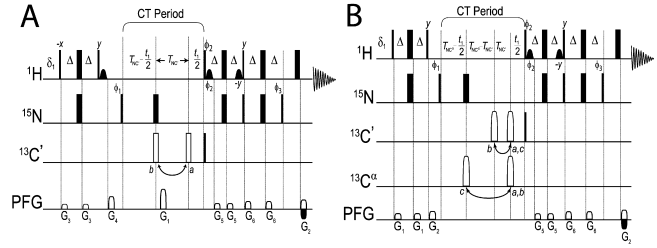

Figure 1. Schematic for the gradient-selected TROSY-based pulse sequences for binary (A) or ternary (B) protein mixtures. A reference spectrum is obtained by applying the 180° ^{13}C pulses (open pulses) at position *a*, while ^{13}C′ or ^{13}C$^\alpha$ suppression is obtained with pulses position *b* or *c*, respectively. Delay durations: $\Delta = 2.4$ ms; $\delta_1 = 1.5$ s; $T_{NC'} = 16.5$ ms; $T_{NC\alpha} = 23.5$ ms. Further details, including spectral editing to obtain each subunit, are provided in the Supporting Information.

The complete details for the detection and deconvolution of a spectrum containing three isotopically labeled species were reported in a previous publication,[19] and only a brief description is provided here. The subunits of the ternary complex are labeled differentially as follows: uniformly ^{15}N (U−^{15}N) labeled (species *A*), uniformly ^{15}N and ^{13}C (U−^{15}N,^{13}C) labeled (species *B*), and U−^{15}N and selectively ^{13}C′ or ^{13}C$^\alpha$ labeled (species *C*). Selective ^{13}C′ or ^{13}C$^\alpha$ labeling can be achieved for recombinant proteins expressed in *Escherichia coli* by utilizing 1-^{13}C or 2-^{13}C glucose, respectively, as the sole carbon source.[22] Signals from either or both ^{13}C labeled species (species *B* or *C*) can be suppressed in a *selective manner* by using a [^{1}H/^{15}N]-HSQC sequence containing a modified constant-time period which leads to $J_{NC'}$ or $J_{NC\alpha}$ modulation (Figure 1). Thus resonances from either species *A* or species *A* and *C* can be produced in subspectra. Linear combinations of these subspectra and a reference spectrum with all three species, lead to the observation of the three individual subunits.[19]

To measure RDCs, we adapted these sequences to utilize sensitivity-enhanced TROSY or anti-TROSY spin-state selection and allowed J_{HN} coupling to become active during chemical shift evolution so that ^{1}H−^{15}N splittings could be measured accurately (Figure 1). Alternatively, RDCs can be extracted by measuring the ^{1}H−^{15}N half-splitting between a pair of TROSY and decoupled HSQC spectra.[23] Because the introduction of $J_{NC'}$ or $J_{NC\alpha}$ modulation is used only for the purpose of suppressing the detection of resonances, this modulation has no effect on the accuracy of the ^{1}H−^{15}N splittings measured. As shown in Figure 1, the sequence is simplified to increase sensitivity in cases of a binary complex (Figure 1A), or the full pulse sequence (Figure 1B) is used in the

[†] University of Wisconsin—Madison.
[‡] Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota.
[§] Department of Chemistry, University of Minnesota.

10.1021/ja904766g CCC: $40.75 © 2009 American Chemical Society

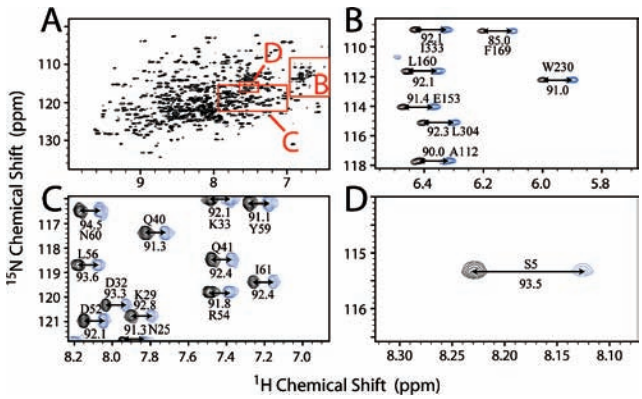

Figure 2. Spectral analysis of the convoluted spectrum containing [U−^{2}H, U−^{15}N]-MBP, [U−^{15}N,^{13}C]-ubiquitin, and [^{15}N-Ser^{5}, ^{13}C′-Ala^{4}]- Kemptide. TROSY-detection was used to obtain all ^{15}N resonances in the sample (A). Isotropic *J*-coupling values were obtained from spectral editing of TROSY and anti-TROSY data sets as shown for MBP (B), ubiquitin (C), and Kemptide (D).

case of a ternary complex. The total constant time period is ∼33 ms ($^{1}/_{2}$ $J_{NC'}$) or ∼49 ms ($^{1}/_{2}$ $J_{NC\alpha}$), respectively.[19]

As a proof of concept, we applied this approach to a ternary mixture of noninteracting proteins: maltose binding protein (MBP, 44 kDa), ubiquitin (8 kDa), and Kemptide (0.7 kDa). Proteins were either expressed in *Escherichia coli* BL21(DE3) to obtain uniform labeling ([U−^{2}H, U−^{15}N]-MBP and [U−^{15}N,^{13}C]-ubiquitin) or synthesized using solid phase synthesis with standard Fmoc chemistry ([^{15}N-Ser^{5},^{13}C′-Ala^{4}]-Kemptide). The NMR sample contained 1.5 mM MBP, 1 mM ubiquitin, and 0.5 mM Kemptide.

The pulse sequence from Figure 1B was used to obtain the isotropic couplings for the ternary mixture shown in Figure 2. All ^{1}H−^{15}N correlations were observed in the reference spectrum (Figure 2A), and the subspectra for each component were obtained by using spectral editing of the suppressed spectra (Figure 2B−D). Weak alignment was then introduced using pf1 phage (12 mg/mL), and the couplings were remeasured to obtain backbone RDC values for the well-folded proteins ubiquitin and MBP. Back-calculated RDC values based on the crystal structures 1UBQ (ubiquitin)[24] and 1OMP (MBP)[25] matched well with experimentally measured values obtained in this study (Figure S1, Supporting Information). Outliers from the MBP structure may arise from a small difference in pH (∼7.3 used here vs 7.0 used previously) and possibly from the slightly different construct used in this study which has an N-terminal His$_6$ tag and an artifact of a TEV-cleavage site that left four extra residues at the C-terminus.[26] Thus, coupling values from this single sample were sufficient to define the backbone conformational space of these structured proteins, along with their relative alignment tensors.

This new approach for measuring backbone RDC values in a ternary complex eliminates the need for multiple samples, removes errors from sample inconsistencies, and ultimately reduces costs related to the preparation of multiple samples. Although shown only for a tertiary mixture of proteins, this approach will work for soluble and membrane proteins, so long as the reference HSQC experiment can be obtained. In the case of membrane proteins, lanthanides or polyacrylamide gels may be used to introduce weak alignment so that the detergent or bicelle are not perturbed.[16,27]

When a binary mixture of proteins is studied, the signal-to-noise (*S*/*N*) resulting from the pulse sequence (Figure 1A) used to detect the ^{15}N-labeled species is $\sqrt{2}$ higher than that of approaches that rely on a HNCO-based sequence,[18] as has been proposed recently.[20,21] In addition, the time period in which magnetization is transverse in HNCO-based experiments is ∼68 ms, whereas the sequences presented here only require half this time period (∼33 ms) for a binary mixture or ∼49 ms for a ternary mixture. Our laboratories have shown that, for large proteins, the constant-time spin−echo filter used in this approach has superior *S*/*N* and works well when HNCO-based sequences provide insufficient *S*/*N*, particularly in cases in which T_2 values are less than ∼50 ms.[18]

In summary, we present a new approach for the acquisition of backbone amide RDCs for binary or ternary complexes using a single sample. Together with the applicability of these sequences to monitor chemical shift perturbations in titration experiments,[19] this method provides useful orientational restraints for high-resolution studies of protein complexes.[11] Since numerous biological processes rely on protein assemblies or transient interactions, this approach should be well suited for a wide range of applications.

Acknowledgment. Financial support provided by NIH Grants GM64742, GM072701 and HL080081 (G.V.), RR02301 and GM66326 (J.L.M.) and AHA support from 0615546Z (L.R.M.). NMR data were collected at NMRFAM (NIH: P41RR02301, P41GM66326, RR02781, and RR08438; NSF: DMB-8415048, OIA-9977486, and BIR-9214394). U. of Minnesota NMR Facility is supported by NSF (BIR-961477).

Supporting Information Available: Pulse sequence and spectral editing details, and experimental RDC fitting to available structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Teague, S. J. *Nat. Rev. Drug Discovery* **2003**, *2*, 527–541.
(2) Clarkson, J.; Campbell, I. D. *Biochem. Soc. Trans.* **2003**, *31*, 1006–1009.
(3) Kay, L. E. *J. Magn. Reson.* **2005**, *173*, 193–207.
(4) Betz, M.; Saxena, K.; Schwalbe, H. *Curr. Opin. Chem. Biol.* **2006**, *10*, 219–225.
(5) Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1997**, *4*, 849–853.
(6) Zwahlen, C.; Legault, P.; Vincent, S. J. F.; Greenblatt, J.; Konrat, R.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 6711–6721.
(7) Post, C. B. *Curr. Opin. Struct. Biol.* **2003**, *13*, 581–588.
(8) Clore, G. M.; Tang, C.; Iwahara, J. *Curr. Opin. Struct. Biol.* **2007**, *17*, 603–616.
(9) Arora, A.; Tamm, L. K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 540–547.
(10) Shimada, I. *Methods Enzymol.* **2005**, *394*, 483–506.
(11) Zuiderweg, E. R. *Biochemistry* **2002**, *41*, 1–7.
(12) Igarashi, S.; Osawa, M.; Takeuchi, K.; Ozawa, S.; Shimada, I. *J. Am. Chem. Soc.* **2008**, *130*, 12168–12176.
(13) Walters, K. J.; Ferentz, A. E.; Hare, B. J.; Hidalgo, P.; Jasanoff, A.; Matsuo, H.; Wagner, G. *Methods Enzymol.* **2001**, *339*, 238–58.
(14) Traaseth, N. J.; Ha, K. N.; Verardi, R.; Shi, L.; Buffy, J. J.; Masterson, L. R.; Veglia, G. *Biochemistry* **2008**, *47*, 3–13.
(15) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. *Chem. Rev.* **2004**, *104*, 3519–3540.
(16) Bax, A.; Grishaev, A. *Curr. Opin. Struct. Biol.* **2005**, *15*, 563–570.
(17) Vuister, G. W.; Wang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 5334–5335.
(18) Tonelli, M.; Masterson, L. R.; Hallenga, K.; Veglia, G.; Markley, J. L. *J. Biomol. NMR* **2007**, *39*, 177–185.
(19) Masterson, L. R.; Tonelli, M.; Markley, J. L.; Veglia, G. *J. Am. Chem. Soc.* **2008**, *130*, 7818–7819.
(20) Ortega-Roldan, J. L.; Jensen, M. R.; Brutscher, B.; Azuaga, A. I.; Blackledge, M.; van Nuland, N. A. *Nucleic Acids Res.* **2009**, *37*, e70.
(21) Bermel, W.; Tkach, E. N.; Sobol, A. G.; Golovanov, A. P. *J. Am. Chem. Soc.* **2009**, *131*, 8564–8570.
(22) Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D. F.; Religa, T. L.; Akke, M.; Kay, L. E. *J. Biomol. NMR* **2007**, *38*, 199–212.
(23) Kontaxis, G.; Clore, G. M.; Bax, A. *J. Magn. Reson.* **2000**, *143*, 184–196.
(24) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531–544.
(25) Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A. *Biochemistry* **1992**, *31*, 10657–10663.
(26) Masterson, L. R.; Bortone, N.; Yu, T.; Ha, K. N.; Gaffarogullari, E. C.; Nguyen, O.; Veglia, G. *Protein Expr. Purif.* **2009**, *64*, 231–236.
(27) Veglia, G; Opella, S. J. *J. Am. Chem. Soc.* **2000**, *122*, 11733–11734.

JA904766G