J Biomol NMR (2013) 56:331–335
DOI 10.1007/s10858-013-9751-9

ARTICLE

Divide and conquer is always best: sensitivity of methyl correlation experiments

Kaustubh Sinha · Linda Jen-Jacobson · Gordon S. Rule

Received: 13 May 2013 / Accepted: 8 June 2013 / Published online: 15 June 2013

Abstract The HMCM [CG]CBCA experiment (Tugarinov and Kay in J Am Chem Soc 125:13868–13878, 2003) correlates methyl carbon and proton shifts to C_γ, C_β, and C_α resonances for the purpose of resonance assignments. The relative sensitivity of the HMCM[CG]CBCA sequence experiment is compared to a divide-and-conquer approach to assess whether it is best to collect all of the methyl correlations at once, or to perform separate experiments for each correlation. A straightforward analysis shows that the divide-and-conquer approach is intrinsically more sensitive, and should always be used to obtain methyl-C_γ, C_β, and C_α correlations. The improvement in signal-to-noise associated with separate experiments is illustrated by the detection of methyl-aliphatic correlations in a 65 kDa protein-DNA complex.

Keywords Methyl resonance assignment · Protein-DNA complex · Pulse sequence

Introduction

NMR has made a significant contribution to our understanding of the role of dynamics in the function of proteins, enzymes, and nucleic acids (Mittermaier and Kay 2006; Markwick et al. 2008; Göbl and Tjandra 2012; Kleckner and Foster 2011; Shajani and Varani 2007). Historically, NMR studies on proteins and nucleic acids have been restricted to relatively small molecules, on the order of 25 kDa in the case of proteins. The reduction in proton–proton dipolar relaxation provided by protein deuteration (Sattler and Fesik 1996), combined with destructive relaxation interference (TROSY, Tzakos et al. 2006; Fernandez and Wider 2003; Salzmann et al. 1998), has allowed investigators to obtain near complete assignment of mainchain and C_β atoms in large proteins (Gardner et al. 1998; McCallum et al. 1999; Salzmann et al. 2000). Deuteration restricts the observation of proton resonances to mainchain H_N atoms, limiting NOEs to those protons. This limitation causes challenges in the determination of the tertiary structure of larger proteins because of the reduced number of inter-residue distance constraints. This problem has been circumvented by selective reprotonation of the protein, either randomly at all aliphatic sites (LeMaster and Richards 1988), or focusing solely on methyl groups from Ile, Leu, and Val (Goto and Kay 2000), Thr (Sinha et al. 2011; Velyvis et al. 2012), or Ala (Ayala et al. 2009).

There are a number of advantages associated with labeling methyl groups (Ruschak and Kay 2010). The dense network of dipolar couplings between the methyl groups within the core of the protein provides distance constraints for structure determination of larger proteins (Mueller et al. 2000), and provides convenient sites for the characterization of dynamics by ^{13}C and ^{2}H relaxation (Tugarinov and Kay 2005), as well as relaxation dispersion (Baldwin et al. 2010). Thr residues are frequently found at protein-nucleic acid interfaces (Biswas et al. 2009), thus can report on the structure and dynamics of critical regions in these complexes.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9751-9) contains supplementary material, which is available to authorized users.

K. Sinha · G. S. Rule (✉)
Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
e-mail: rule@andrew.cmu.edu

L. Jen-Jacobson
Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15213, USA

The assignment of methyl resonances is prerequisite for the utilization of these groups for protein structure determination or for the interpretation of relaxation measurements. Methyl resonances can be assigned by predicting methyl–methyl NOES, either alone (Xu and Matthews 2013), or in combination with paramagnetic relaxation enhancement (Venditti et al. 2011). If the protein contains a lanthanide ion binding site, pseudocontact shifts can also be used to assign methyl resonances (John et al. 2007). In cases where the N, H_N, CO, C_α, and C_β assignments are known, methyl resonances can be readily assigned by correlation of the methyl resonances to the already assigned atoms. This correlation can be accomplished using COSY type transfers from the methyl to other atoms. A significant advance in the use of COSY type correlation experiments was attained by labeling Leu and Val such that only one methyl was labeled, thus producing a linear arrangement of coupled carbons for these residues (see Tugarinov and Kay 2003). This advance in labeling permitted the development of a suite of NMR experiments that correlated methyl resonances to assigned N, H_N, CO, C_α, and C_β atoms, leading to near complete methyl assignments in a 723 residue protein. One of the key experiments developed in that work is the HMCM[CG]CBCA pulse sequence, which simultaneously correlates the methyl carbon and proton shifts to the rest of the carbons on the sidechain in a single spectrum. The HMCM[CG]CBCA experiment has been successfully used to assign methyl resonances in a number of smaller proteins (Wang et al. 2012; Krejcirikova and Tugarinov 2012; Zhuravleva et al. 2012; Chan et al. 2012). In the case of larger systems, this experiment has been replaced by a "divide-and-conquer" strategy, whereby the shifts of the individual carbons, e.g. C_α, are collected in separate experiments (Sprangers and Kay 2007). To the best of our knowledge there has not been a clear analysis in the literature regarding the relative sensitivities of each approach, which would provide guidance to users regarding which experiments to perform. A straightforward analysis shows that the divide-and-conquer approach should always be used, regardless of the size of the protein.

Materials and methods

Spectra were acquired on a Bruker AMX spectrometer, operating at 600 MHz (^{1}H) using a standard room temperature probe. The pulse sequences were tested on the 130 residue RNA binding domain of *E. coli* rho protein (Rho130) and the 245 residue homodimeric EcoRV-DNA complex. Both proteins were perdeuterated and methyl labeled as described by Goto et al. (1999). The proteins were expressed from a standard T7 expression system using C3013 cells (New England Biolabs) as the host. Studier's PG medium (Studier 2005) was used with deuterated uniformly ^{13}C labeled glucose, ^{15}N labeled ammonium sulfate, and ~100 % D_2O. The cells were grown to an A_{550} of 0.6 at 37 °C. At this point, methyl-protonated, ^{13}C-labeled (uniform, except one methyl is ^{12}C) α-ketoisovalerate (100 mg/L) and methyl-protonated uniformly deuterated and ^{13}C ketobutyrate (50 mg/L), was added to the media, as described by Goto et al. (1999). Isotopically labeled compounds were obtained from either Cambridge Isotopes or Sigma Aldrich. The cells were allowed to grow for 60 min and isopropylthiogalactoside (IPTG) (1 mM) was added to induce protein expression. The cells were harvested 3 h after induction. Rho130 was purified as described previously (Briercheck et al., 1998) and the purification scheme for EcoRV will be presented elsewhere. The EcoRV-DNA complex was generated by adding the DNA duplex (5′-GCAAAGATATCTTTCG-3′; IDT) to the protein in 1:1 stoichiometric ratio.

Results and discussion

The HMCM[CG]CBCA sequence was used to acquire methyl correlation spectra on a small 130 residue protein (Rho130) and a 65 kDa EcoRV-DNA complex (see Fig. 1). Although most (but not all) of the expected correlations were observed for Rho130, the EcoRV-DNA complex gave unexpectedly weak signal in the HMCM[CG]CBCA experiment due to enhanced relaxation of the protein methyls from protons on the non-deuterated DNA. Although it was only possible to obtain a small number of aliphatic-methyl correlations using the HMCM[CG]CBCA experiment, most of the expected correlations were obtained using the divide-and-conquer strategy.

The divide-and-conquer experiments are more sensitive than the HMCM[CG]CBCA experiment for a number of reasons. Because the HMCM[CG]CBCA experiment simultaneously correlates the methyl resonances to all of the other aliphatic carbons (e.g. Ile C_γ, C_β, C_α) it is necessary to produce four different product operator terms ($4I_z C_x^\gamma C_z^\beta$, $2I_z C_y^\beta$, $4I_z C_z^\beta C_x^\alpha$, $8I_z C_x^\gamma C_y^\beta C_x^\alpha$) prior to recording the aliphatic carbon shift. Only the first three lead to single-quantum frequencies, hence one-fourth of the signal is lost. Second, since all three carbon peaks are acquired in the same spectrum, the original magnetization from the methyl is divided among the three peaks, thus each peak represents one-fourth of the original magnetization. Third, because the C_γ and C_β frequencies are detected after the magnetization has been transferred to the C_α, all three of the signals suffer the large relaxation loss associated with the long delay required to relay the magnetization from the methyl to the C_α carbon. By acquiring the shift of each carbon in separate

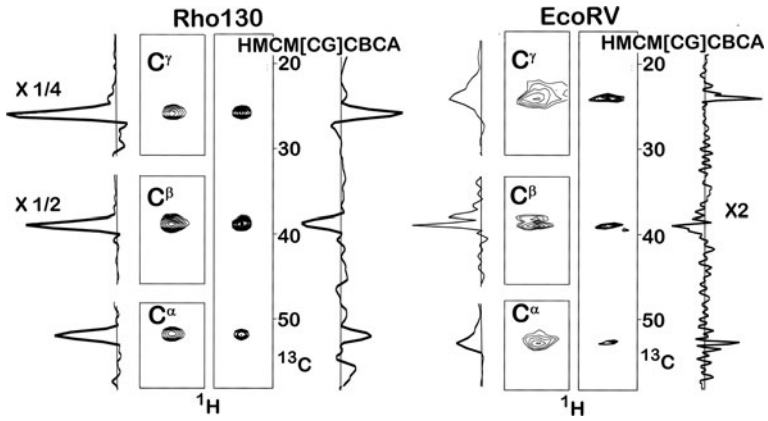

Fig. 1 Leucine Methyl Correlation Spectra of Rho130 and the EcoRV-DNA Complex. The methyl carbon region was sampled using 21 complex points in all experiments. The C_α resonances are broadened by unresolved coupling to the ^{15}N amide nitrogen. The concentration of Rho130 was 1.0 mM and spectra were acquired at 30 °C. In the case of the Rho130 sample, all four spectra were acquired under identical condition; 8 scans with the same spectral width (50 ppm) and number of points (40) in the non-methyl carbon dimension. The total acquisition time/experiment was 9 h. In the rho130 spectra the one dimensional trace for the C_β peaks are divided by two and the trace for the C_γ peak was divided by four. The concentration of the EcoRV-DNA complex was 0.8 mM and the spectra were acquired at 35 °C. In the case of EcoRV-DNA sample the spectral widths for the aliphatic carbons in the C_α C_β, C_γ, and the HMCM[CG]CBCA experiments were 20, 60, 20, and 50 ppm, respectively. These were sampled using 32, 40, 32, and 40 complex points, respectively. A total of 40, 16, 8, and 40 scans were acquired for the C_α, C_β, C_γ, and HMCM[CG]CBCA experiments, respectively, giving total acquisition times of 35.8, 17.92, 7.17, and 44.8 h, respectively. The entire one dimensional trace for the HMCM[CG]CBCA experiment is multiplied by two. The increase in signal intensity for one residue from the EcoRV-DNA sample, after scaling for the total experimental time, is 2.0, 8.8, and 6.0 for the C_α, C_β, C_γ, experiments. These ratios are the lower limit because many correlations could not be observed in the HMCM[CG]CBCA experiment. The C_β experiment could have been run in ½ the time (8 scans), which would have given a total acquisition time for all three experiments of approximately 52 h, i.e. not much longer than the entire HMCM[CG]CBCA experiment. Pulse sequences can be found in figure S1, along with additional guidelines for defining the acquisition time for the detection of non-methyl carbons

experiments only the intensity of the C_α is affected by the long transfer time, effectively increasing the signal of the C_β and C_γ resonances. Since only one signal is being detected at a time, the full magnetization from the methyl is also detected, increasing the signal by four-fold for the C_γ and C_β resonances. The increase in signal intensity is smaller for the C_α resonances because it is necessary to lengthen the last polarization transfer step from 7 to 14 ms to allow for complete conversion of the anti-phase term $2C_z^\beta C_x^\alpha$ to C_y^α prior to the evolution period that records the C_α frequency (see supplementary material figure S1 for pulse sequences). Regardless of the increase in polarization transfer time, the sensitivity of the C_α experiment still exceeds that of the HMCM[CG]CBCA experiment, except for very short carbon T_2 times, on the order of 10 ms (see figure S3). Although it is necessary to run all three experiments to obtain the same chemical shift information as obtained in the HMCM[CG]CBCA experiment, the additional acquisition time is significantly reduced due to the increase in sensitivity of the C_β and C_γ experiments. For example, at a carbon T_2 of 40 ms, the sensitivities of the C_α, C_β, and C_γ experiments are approximately 2.8, 5.7, and 11.4 times that of the HMCM[CG]CBCA experiment. In addition to the gains in sensitivity, the C_α and C_γ experiments can be acquired with a much narrower sweepwidth (see figure S4), reducing the acquisition time of the experiment. The increase in sensitivity allows the C_α, C_β, and C_γ experiments to be acquired in approximately 24/30, 5/30 and 1/30, respectively, of the time required for the HMCM[CG]CBCA experiment and still yield a 2–4 fold increase in intensity.

The predicted gains in sensitivity are essentially realized in practice. In Fig. 1 we show spectra of Rho130 and the EcoRV-DNA complex that were obtained using the C_α, C_β, and C_γ and the HMCM[CG]CBCA experiments. Signals from 15 Leu residues in Rho130 gave relative intensities of 1.9 ± 0.5, 4.4 ± 0.9, and 8.4 ± 3.6 for the C_α, C_β, and C_γ peaks, respectively, compared to the HMCM[CG]CBCA experiment. Note that the sensitivity gain for the C_α experiment is a lower limit since a number of C_α peaks were missing in the HMCM[CG]CBCA experiment, but present in the C_α experiment. In the case of the larger EcoRV-DNA complex, only a few correlations were observed in the HMCM[CG]CBCA experiment hence it is difficult to compare sensitivities. When C_α resonances could be observed, they were approximately two fold less intense in the HMCM[CG]CBCA spectra than in the experiment that only detected the C_α shifts.

In conclusion, because the sensitivity of detecting methyl-aliphatic correlations is significantly increased by recording each carbon shift in a separate experiment,

separate experiments should be run, even for smaller proteins. A similar philosophy has also been historically applied when recording correlations between NH resonances and C_β resonances. In this case, the standard HNCACB experiment, which detects both C_α and C_β shifts simultaneously, is often run with the delay for magnetization transfer between the C_α and C_β spins set to $1/2J_{cc}$, causing complete conversion of the C_x^α product operator to anti-phase $2C_z^\beta C_x^\alpha$, prior to recording of the carbon chemical shift. This approach transfers all of the magnetization from the NH group to the C_β peak, instead of splitting it between the C_α and C_β peaks.

Acknowledgments We thank Virgil Simplaceanu for maintenance of the NMR spectrometers and Dr. C. Sanders (Vanderbilt University) for providing the Bruker code for the HMCM[CG]CBCA experiment. This work was supported by the Mellon College of Science at CMU and an NIH Merit grant 5R37-GM029207 to L.J–J.

References

Ayala I, Sounier R, Use N, Gans P, Boisbouvier J (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. J Biomol NMR 43:111–119

Baldwin AJ, Religa TL, Hansen DF, Bouvignies G, Kay LE (2010) $^{13}CHD_2$ methyl group probes of millisecond time scale exchange in proteins by ^1H relaxation dispersion: an application to proteasome gating residue dynamics. J Am Chem Soc 132(32):10992–10995. doi:10.1021/ja104578n

Biswas S, Guharoy M, Chakrabarti P (2009) Dissection, residue conservation, and structural classification of protein-DNA interfaces. Proteins 74:643–654

Briercheck DM, Wood TC, Allison TJ, Richardson JP, Rule GS (1998) The NMR structure of the RNA binding domain of E. coli rho factor suggests possible RNA-protein interactions. Nat Struct Biol 5:393–399

Chan PH, Weissbach S, Okon M, Withers SG, McIntosh LP (2012) Nuclear magnetic resonance spectral assignments of α-1,4-galactosyltransferase LgtC from neisseria meningitidis: substrate binding and multiple conformational states. Biochemistry 51:8278–8292

Fernandez C, Wider G (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. Curr Opin Struct Biol 13:570–580

Gardner KH, Zhang X, Gehring K, Kay LE (1998) Solution NMR Studies of a 42 KDa Escherichia Coli Maltose Binding Protein/β-Cyclodextrin Complex: chemical Shift Assignments and Analysis. J Am Chem Soc 120(45):11738–11748. doi:10.1021/ja982019w

Göbl C, Tjandra N (2012) Application of solution NMR spectroscopy to study protein dynamics. Entropy 14(3):581–598. doi:10.3390/e14030581

Goto NK, Kay LE (2000) New developments in isotope labeling strategies for protein solution NMR spectroscopy. Curr Opin Struct Biol 10:585–592

Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated ^{15}N-, ^{13}C-, ^2H-labeled proteins. J Biomol NMR 13:369–374

John M, Schmitz C, Park AY, Dixon NE, Huber T, Otting G (2007) Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. J Am Chem Soc 7:13749–13757

Kleckner IR, Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. Biochim Biophys Acta. doi:10.1016/j.bbapap.2010.10.012

Krejcirikova A, Tugarinov V (2012) 3D-TROSY-based backbone and ILV-methyl resonance assignments of a 319-residue homodimer from a single protein sample. J Biomol NMR 54:135–143. doi:10.1007/s10858-012-9667-9

LeMaster DM, Richards FM (1988) NMR sequential assignment of Escherichia coli thioredoxin utilizing random fractional deuteriation. Biochemistry 27:142–150

Markwick PR, Malliavin T, Nilges M (2008) Structural biology by NMR: structure, dynamics, and interactions. PLoS Comput Biol 4(9):e1000168. doi:10.1371/journal.pcbi.1000168

McCallum SA, Hitchens TK, Rule GS (1999) Solution structure of the carboxyl terminus of a human class mu glutathione s-transferase: NMR assignment strategies in large proteins. J Mol Biol 285:2119–2132

Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. Science 312:224–228

Mueller GA, Choy WY, Yang D, Forman-Kay JD, Venters RA, Kay LE (2000) Global folds of proteins with low densities of noes using residual dipolar couplings: application to the 370-residue maltodextrin-binding protein. J Mol Biol 300:197–212

Ruschak AM, Kay LE (2010) Methyl groups as probes of supramolecular structure, dynamics and function. J Biomol NMR 46:75–87. doi:10.1007/s10858-009-9376-1

Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc Natl Acad Sci U S A. 95(23):13585–13590

Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K (2000) NMR Assignment and Secondary Structure Determination of an Octameric 110 kDa Protein Using TROSY in Triple Resonance Experiments. J Am Chem Soc 122:7543–7548. doi:10.1021/ja0003268

Sattler M, Fesik SW (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. Structure 4:1245–1249

Shajani Z, Varani G (2007) NMR studies of dynamics in RNA and DNA by ^{13}C relaxation. Biopolymers 86:348–359. doi:10.1002/bip.20650

Sinha K, Jen-Jacobson L, Rule GS (2011) Specific labeling of threonine methyl groups for NMR studies of protein-nucleic acid complexes. Biochemistry 50:10189–10191. doi:10.1021/bi201496d

Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445:618–622. doi:10.1038/nature05512

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207–234

Tugarinov V, Kay LE (2003) Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. J Am Chem Soc 125:13868–13878. doi:10.1021/ja030345s

Tugarinov V, Kay LE (2005) Quantitative ^{13}C and ^2H NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface. Biochemistry 44:15970–15977. doi:10.1021/bi0519809

Tzakos AG, Grace CRR, Lukavsky PJ, Riek R (2006) NMR Techniques for Very Large Proteins and RNAs in Solution. Annu Rev Biophys Biomol Struct 35:319–342

Velyvis A, Ruschak AM, Kay LE (2012) An economical method for production of (2)H, (13)CH3-threonine for solution NMR studies of large protein complexes: application to the 670 kDa

proteasome. PLoS ONE 7(9):e43725. doi:10.1371/journal.pone.0043725

Venditti V, Fawzi NL, Clore GM (2011) Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl–methyl nuclear Overhauser enhancement spectroscopy. J Biomol NMR 51:319–328

Wang X, Vu A, Lee K, Dahlquist FW (2012) CheA–Receptor Interaction Sites in Bacterial Chemotaxis. J Mol Biol 422:282–290. doi:10.1016/j.jmb.2012.05.023

Xu Y, Matthews S (2013) MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins. J Biomol NMR 55:179–187. doi:10.1007/s10858-012-9700-z

Zhuravleva A, Clerico EM, Gierasch LM (2012) An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. Cell 151:1296–1307