J Biomol NMR (2014) 59:51–56
DOI 10.1007/s10858-014-9834-2

COMMUNICATION

Sparsely-sampled, high-resolution 4-D omit spectra for detection and assignment of intermolecular NOEs of protein complexes

Su Wang · Pei Zhou

Received: 10 March 2014 / Accepted: 22 April 2014 / Published online: 1 May 2014

Abstract Unambiguous detection and assignment of intermolecular NOEs are essential for structure determination of protein complexes by NMR. Such information has traditionally been obtained with 3-D half-filtered experiments, where scalar coupling-based purging of intramolecular signals allows for selective detection of intermolecular NOEs. However, due to the large variation of $^{1}J_{HC}$ scalar couplings and limited chemical shift dispersion in the indirect proton dimension, it is difficult to obtain reliable and complete assignments of interfacial NOEs. Here, we demonstrate a strategy that combines selective labeling and high-resolution 4-D NOE spectroscopy with sparse sampling for reliable identification and assignment of intermolecular NOEs. Spectral subtraction of component-labeled complexes from a uniformly-labeled protein complex yields an "omit" spectrum containing positive intermolecular NOEs with little signal degeneracy. Such a strategy can be broadly applied to unbiased detection, assignment and presentation of intermolecular NOEs of protein complexes.

Keywords Fast NMR · Sparse sampling · Protein complex · Intermolecular NOE

The advancement of NMR technology has greatly enhanced the sensitivity of NMR experiments and has made it increasingly feasible to determine high-resolution structures of protein complexes that play important roles in various cellular processes by solution NMR. In comparison with the well-established procedure to determine the high-resolution structure of a single-chain protein, structure determination of a protein complex presents unique challenges. In particular, the limited number of interfacial NOEs that define macromolecular complexes require reliable detection and unambiguous assignment to ensure the proper assembly of individual components of the protein complex.

Detection of intermolecular NOEs is traditionally achieved by half-filtered experiments (Breeze 2000; Otting and Wuthrich 1990). The strategy involves isotopic labeling of a component of the protein complex and spectroscopic suppression of protons attached to ^{13}C or ^{15}N nuclei prior to NOE transfer, followed by selective detection of ^{13}C or ^{15}N-attached protons. Theoretically, this approach selectively detects dipolar interactions from protons in the unlabeled component to protons in the isotopically-labeled component within the same protein complex, thus ensuring all of the observed NOE crosspeaks are intermolecular. However, the large variation of $^{1}J_{HC}$ couplings (125–200 Hz) makes it difficult to achieve perfect suppression of isotope-attached proton signals, even with specially-designed adiabatic or composite pulses to compensate for the variation of $^{1}J_{HC}$ couplings (Stuart et al. 1999; Zwahlen et al. 1997). Additionally, since the protein-protein interface often consists of the same type of functional groups (e.g., methyl-methyl packing) that share similar chemical shifts, intermolecular NOEs are often located very close to the diagonal positions. Thus, incompletely suppressed diagonal signals can cause significant interference with intermolecular NOE signals. Finally, a lack of chemical shift dispersion in the indirect proton dimension often causes assignment ambiguities, precluding reliable assignment of interfacial NOEs.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-014-9834-2) contains supplementary material, which is available to authorized users.

S. Wang · P. Zhou (✉)
Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA
e-mail: peizhou@biochem.duke.edu

The recent development of fast acquisition of high-resolution 4-D NOESY experiments based on sparse time-domain sampling and high-performance spectral reconstruction in the frequency domain offers a convenient solution to these challenges (Coggins et al. 2012; Coggins and Zhou 2008; Hiller et al. 2009; Hoch and Stern 2001; Hyberts et al. 2009; Tugarinov et al. 2005; Wen et al. 2012; Werner-Allen et al. 2010). Compared with the 3-D ^{13}C half-filtered NOESY–HSQC experiment, there should be little loss of sensitivity in a high-resolution 4-D ^{13}C HMQC–NOESY–HSQC experiment, as both experiments would require similar durations of active J-coupling periods for building up H-C correlations, either for isotope signal suppression (as in the half-filtered experiment) or for signal selection (as in the 4-D NOE experiment). On the other hand, 4-D NOE spectroscopy offers much better signal separation than a 3-D half-filtered NOESY–HSQC experiment. Importantly, in a subunit-labeled protein complex, a 4-D ^{13}C HMQC–NOESY–HSQC experiment naturally selects *intramolecular* NOEs within the labeled component. Since clean selection of isotope-attached proton signals is much easier to achieve than perfect suppression of the same signals, 4-D NOESY is much less affected by the variation of $^{1}J_{HC}$ couplings, thus ensuring the quality of the intramolecular NOE spectra. Hence detecting *intermolecular* NOEs is as simple as comparing the 4-D NOE spectrum of uniformly-labeled protein complex samples with that of subunit-selectively labeled samples. To facilitate data analysis, NOE signals of subunit-selectively labeled protein complexes can be subtracted from the uniformly-labeled protein complex to generate an "*omit*" spectrum—a concept analogous to the omit electron density map in crystallography (Bhat 1988) and the 2-D double difference NOE spectroscopy employing selective protein deuteration (Nudelman et al. 2011)—that only contains *intermolecular* NOE crosspeaks for unbiased data interpretation. Although such an approach is conceptually simple, it can only be achieved with a high-resolution 4-D NOESY experiment, but not with a 3-D NOESY–HSQC experiment, as the latter does not allow for selective detection of NOEs within isotopically-enriched subunits. Therefore, its application in biomolecular NMR has been prohibited by the lengthy data acquisition times of conventional high-resolution 4-D experiments (typically many weeks). Recent development of sparse sampling techniques (Coggins et al. 2010; Kazimierczuk et al. 2010) overcomes this obstacle by making it practical to collect high-resolution 4-D NOESY spectra using only a fraction of the measurement time compared to the conventional approach. Although sparse sampling introduces aliasing artifacts proportional to the intensity and square root of the number of the NMR signals, these artifacts can be suppressed using iterative reconstruction algorithms such as SCRUB (Coggins et al. 2012). As an added benefit, by subtracting NOE signals of individually labeled components from the uniformly labeled protein complex in the time domain, one not only removes the very strong diagonal signals, but also drastically reduces the number of NOE signals for reconstruction, both of which further enhance the quality of the reconstructed omit spectrum containing intermolecular NOEs.

We first demonstrate such an approach for unambiguous detection and assignment of intermolecular NOEs of the human Y-family polymerase (Pol) ι UBM1-ubiquitin complex (Fig. 1) using the sparsely-sampled 4-D ^{13}C HMQC–NOESY–HSQC experiment (pulse sequence shown in Figure S1). Three samples of gb1UBM1-ubiquitin complexes were prepared, including (1) a uniformly ^{13}C/^{15}N-labeled complex, (2) a ubiquitin selectively labeled complex, and (3) a gb1UBM1 selectively labeled complex, all at 3 mM concentrations. Such a high concentration is necessary due to the weak binding affinity of the UBM1-ubiquitin interaction (K_d = 130 μM) to ensure that the majority of the protein remains in the complex form. Three identical 4-D ^{13}C HMQC–NOESY–HSQC experiments, each taking 43 h, were recorded for each sample, using cosine-weighted random concentric shell sampling with 3,190 sampling points digitized on a 64 × 96 × 64 grid for the c(t1)–h(t2)–C(t3) indirect dimensions (Coggins and Zhou 2008), corresponding to 0.8 % of sampling points required for Nyquist sampling. The omit spectrum was generated by subtracting scaled FIDs of component-labeled complexes from those of the uniformly-labeled complex and reconstructed by SCRUB (Coggins et al. 2012). To account for the minor difference of NOE crosspeaks intensities in the three samples, it is preferable to slightly over-subtract the component-specific NOE signals so that all of the diagonal signals are negative. As a result, scaling factors of 1.07 and 1.05 are used for the gb1UBM1-labeled sample and the ubiquitin-labeled sample, respectively. This ensures that the remaining positive signals represent genuine intermolecular NOEs, whereas intramolecular crosspeaks are either completely removed or appear as weak negative signals (Fig. 1).

Structural analysis of the human Pol ι UBM1-ubiquitin complex reveals a helix-turn-helix motif of UBM1 that recognizes ubiquitin (Figure S2, Table S1). Accordingly, a large number of intermolecular NOEs can be observed in the 4-D omit spectrum, providing ample constraints to define the UBM1-ubiquitin binding interface. In particular, numerous intermolecular NOEs between UBM1 residues (V502 from the N-terminal tail, V506 and F507 from the α1, L510 from the turn between the two helices, I514 and I518 from α2 and F527 from the C-terminal tail) and the L8 methyl groups of ubiquitin are detected in the 4-D omit spectrum (Fig. 1a, b), consistent with the notion that

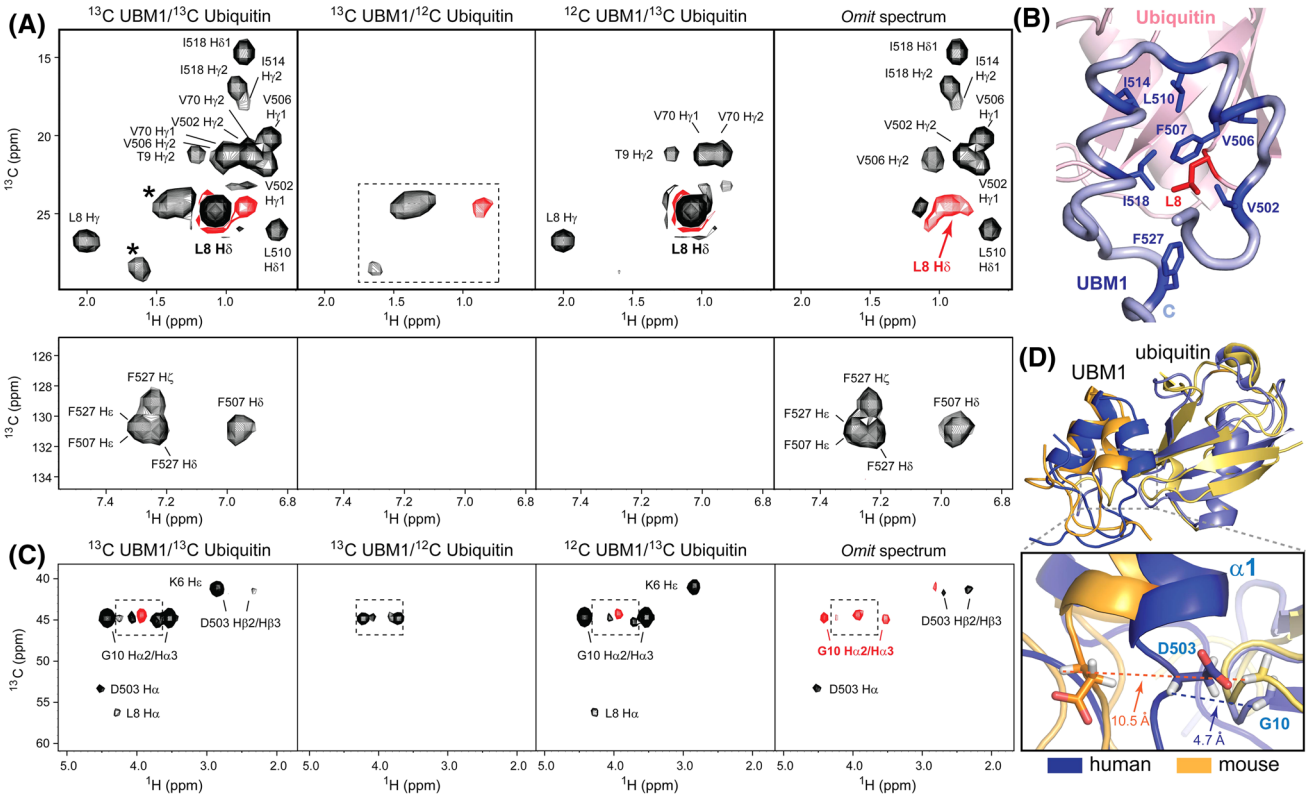

Fig. 1 Omit spectrum of the human Pol ι UBM1-ubiquitin complex showing intermolecular NOEs. 4-D ^{13}C HMQC–NOESY–HSQC spectra are collected for the gb1UBM1-ubiquitin complex with both components or with individual components ^{13}C-labeled. Reconstruction of the difference time domain signals of the uniformly labeled protein complex from component-labeled samples generates an omit spectrum containing only intermolecular NOEs. Slight over-subtraction of time domain data from individual components generates negative diagonal signals (*red*) in the omit spectrum and ensures all of the positive cross-peaks originate from intermolecular NOEs. **a** (upper, aliphatic regions; lower, aromatic regions) shows sections of F1-F2 slices of the corresponding 4-D spectra centered at 24.53 ppm in F3 and 1.05 ppm in F4, displaying NOEs to the ubiquitin L8 methyl groups. Boxed peaks or peaks labeled with asterisks are off-plane signals. **b** Interface of the human Pol ι UBM1-ubiquitin complex, showing an interaction network centered at L8 of ubiquitin. **c** 2-D slices of the 4-D ^{13}C HMQC–NOESY–HSQC spectra centered at 45.07 ppm in F3 and 3.51 ppm in F4, illustrating NOEs to the G10 Hα3 proton of ubiquitin. Boxed peaks are off-plane signals. **d** Overlay of the human Pol ι UBM1-ubiquitin complex (*blue*; this work) with the mouse Pol ι UBM1-ubiquitin complex (*yellow*; PDB ID 2KWV) by superimposing the ubiquitin part. The zoomed-in panel shows the distances between the ubiquitin G10 Hα3 proton and the Hα proton of a conserved UBM1 Asp residue (D503 in human Pol ι UBM1 and D501 in mouse Pol ι UBM1)

UBM1 interacts with ubiquitin through a binding surface centered at L8 of ubiquitin in a manner similar to that of UBM2 (Bienko et al. 2005; Bomar et al. 2010; Burschowsky et al. 2011; Cui et al. 2010).

Interestingly, structural comparison of the human Pol ι UBM1-ubiquitin complex determined by 4-D NOE with the previously reported mouse Pol ι UBM1-ubiquitin complex (PDB 2KWV) (Burschowsky et al. 2011) has revealed larger-than-expected structural deviation at the N-terminus of α1 of UBM1 (Fig. 1d). In the human Pol ι UBM1-ubiquitin complex, the N-terminus of α1, including the highly conserved N-cap residue of α1, D503, is located in the vicinity of the ubiquitin surface and is involved in the interaction with ubiquitin. Such a configuration is well supported by intermolecular NOEs between D503 Hα and Hβ protons of human UBM1 and the Hα3 protons of G10 of ubiquitin (Fig. 1c). In contrast, in the mouse Pol ι UBM1-ubiquitin complex structure (Burschowsky et al. 2011), the N-terminus of α1 is tilting substantially away from ubiquitin, in which the shortest distances between ubiquitin G10 Hα protons (i.e. Hα3) and the Hβ protons of the corresponding Asp in UBM1 (D501 in mouse Pol ι UBM1) are over 7.4 Å. The ubiquitin G10 Hα3 and mouse UBM1 D501 Hα protons are located further apart, with a distance of ~10.5 Å, reflecting a lack of intermolecular NOEs between these protons.

The strategy described above can be conveniently extended to homooligomeric complexes, such as the homotrimeric foldon complex. Foldon is the C-terminal domain of bacteriophage T4 fibritin, which forms a trimeric β-hairpin propeller, with the molecular weight of about 9 kDa for the trimer (Boudko et al. 2004; Guthe et al. 2004;

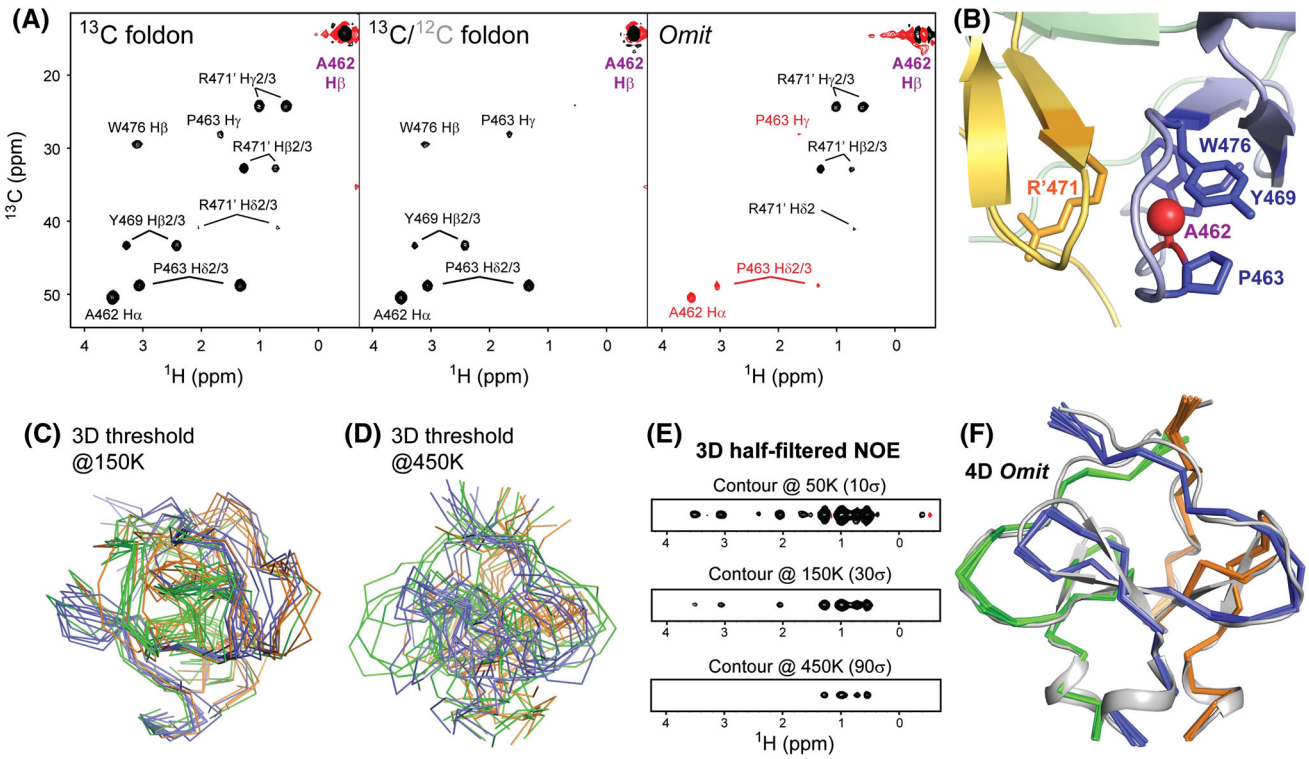

Fig. 2 Omit spectrum of the foldon trimer. **a** F1–F2 slices of the corresponding 4-D spectra centered at 14.12 ppm in F3 and −0.49 ppm in F4, displaying NOEs to the A462 methyl group. 4-D ^{13}C HMQC–NOESY–HSQC spectra are collected for foldon with uniformly (left) or 25% (middle) ^{13}C-labeled samples. Subtraction of these two spectra generates an omit spectrum (right) containing positive intermolecular NOE and negative (*red*) diagonal signals and intramolecular NOEs. **b** Interface of two subunits of the foldon trimer, showing an intersubunit interaction between R471 and A462 and intrasubunit interactions between W476, Y469, P463 and A462. **c**, **d** Show the results of structure calculation using the 3-D half-filtered NOE peak list with intensity thresholds of 150 and 450 K, respectively. **e** A slice of the 3-D half-filtered NOE spectrum centered at the A462 methyl group, plotted at contour levels of 50, 150 and 450 K, respectively. **f** The structural ensemble of trimeric foldon (subunits colored in *blue*, *green* and *orange*) calculated from the CYANA-assigned 4-D intermolecular NOE peak list from the omit spectrum superimposed with the crystal structure (*grey*; PDB ID: 1OX3)

Tao et al. 1997). In this case, two sets of sparsely sampled 4-D ^{13}C HMQC–NOESY–HSQC spectra using cosine-weighted random concentric shell sampling were collected for a 0.33 mM uniformly ^{13}C-labeled foldon sample and for a 1.32 mM sample containing 25% ^{13}C-labeled protein and 75% unlabeled (^{12}C) protein, each taking 86 h. Since intermolecular NOEs are *statistically diluted* in the second sample, signal subtraction of the second 4-D NOE dataset from the first dataset generates an omit spectrum containing intermolecular NOEs. A scaling factor of 1.7 was chosen so that all diagonal NOEs are slightly negative in the omit spectrum (e.g., the negative diagonal signal of the A462 methyl group in Fig. 2a), ensuring that the remaining positive NOE crosspeaks in the omit spectrum arise from genuine intermolecular NOEs (e.g., black colored crosspeaks between R471' and A462 in Fig. 2a) and distinguishing these intermolecular signals from the slightly negative intramolecular NOEs (e.g., intramolecular interactions between A462 and its neighbors; Fig. 2a, b).

It should be noted that in contrast to the heterodimeric complex, it is not possible to obtain *pure* intramolecular NOEs using a homooligomeric sample containing both isotopically-labeled and unlabeled components. Instead, one obtains a spectrum containing intramolecular NOEs and *significantly weakened* intermolecular NOEs due to *statistical dilution* of intermolecular dipolar interactions. Therefore, the 4-D omit spectrum in this case is less sensitive than the corresponding 3-D half-filtered NOE experiment in detecting intermolecular NOEs. However, as demonstrated with the foldon trimer, despite the superior sensitivity of the 3-D half-filtered NOE experiment (compare the omit spectrum in Fig. 2a and the half-filtered spectrum plotted at different contour levels in Fig. 2e), due to assignment ambiguities, these crosspeaks cannot be analyzed by automated procedures in CYANA (Güntert 2004) to generate a properly converged fold, even when NOE crosspeaks are filtered at 30- and 90-fold above the spectral noise, respectively (Fig. 2c, d). In contrast, the inclusion of

an additional dimension in the 4-D omit spectrum largely eliminated the assignment ambiguity for the automated NOE analysis protocol. Since the identity of the subunits contributing to intermolecular NOEs for a homooligomeric complex with the C3 symmetry cannot be differentiated (e.g., interaction of A–B vs. A–C), these intermolecular NOEs are symmetrized in the order of A–B, B–C, and C–A, with the counter rotating pairs of A–C, B–A, C–B implemented as ambiguous constraints. When these constraints are used in the CYANA calculation, they readily generate a highly-converged foldon structure with 0.6 Å backbone RMSD deviation from the corresponding crystal structure (PDB 1OX3), clearly demonstrating the advantage of 4-D omit spectrum.

Reliable detection and unambiguous assignment of intermolecular NOEs is essential to NMR-based structure determination of protein complexes. Such a task is traditionally carried out by using 3-D half-filtered NOE experiments. Although the half-filtered experiments are easy to set up and have resulted in successful structural resolution of many protein complexes, they offer limited separation for interfacial signals with similar proton chemical shifts, and the incomplete suppression of diagonal signals may further complicate data analysis. The approach demonstrated in this work overcomes the limitations of 3-D half-filtered experiments through 4-D omit spectra enabled by fast NMR techniques to allow for reliable detection and unambiguous assignment of intermolecular NOEs. Superficially, such a strategy requires the collection of multiple 4-D NOE spectra and seems less efficient than the 3-D half-filtered experiment. However, one should bear in mind that these 4-D NOE spectra not only allow for detection of intermolecular NOEs, but also provide complete distance information for defining the subunit and complex structures in full, whereas the traditional half-filtered NOE experiment alone is insufficient to define the structure of a protein complex. Although the heterodimeric UBM1-ubiquitin complex demonstrated here utilized a 3 mM sample due to the weak binding affinity ($K_d = 130$ μM), the same approach was also applied to a 0.33 mM foldon sample and more recently to a tight binding heterodimeric complex with only 0.8 mM sample concentration (data not shown). Hence, the 4-D omit spectrum approach presented in this study is widely applicable to heterodimeric complexes and homooligomeric complexes to achieve unambiguous assignment of intermolecular NOEs and to enable reliable structure calculation using automated protocols. These advantages render such a strategy particularly attractive to structural studies of protein complexes by biomolecular NMR.

Acknowledgments This work was supported in part by the National Institutes of Health/National Institute of Allergy and Infectious Diseases (AI055588) to P. Z.

References

Bhat TN (1988) Calculation of an OMIT map. J Appl Cryst 21:279–281

Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR et al (2005) Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science 310:1821–1824. doi:10.1126/science.1120615

Bomar MG, D'Souza S, Bienko M, Dikic I, Walker GC, Zhou P (2010) Unconventional ubiquitin recognition by the ubiquitin-binding motif within the Y family DNA polymerases iota and Rev1. Mol Cell 37:408–417. doi:10.1016/j.molcel.2009.12.038

Boudko SP, Strelkov SV, Engel J, Stetefeld J (2004) Design and crystal structure of bacteriophage T4 mini-fibritin NCCF. J Mol Biol 339:927–935. doi:10.1016/j.jmb.2004.04.001

Breeze AL (2000) Isotope-filtered NMR methods for the study of biomolecular structure and interactions. Prog Nucl Magn Reson Spectrosc 36:323–372

Burschowsky D, Rudolf F, Rabut G, Herrmann T, Peter M, Wider G (2011) Structural analysis of the conserved ubiquitin-binding motifs (UBMs) of the translesion polymerase iota in complex with ubiquitin. J Biol Chem 286:1364–1373. doi:10.1074/jbc.M110.135038

Coggins BE, Zhou P (2008) High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN. J Biomol NMR 42:225–239. doi:10.1007/s10858-008-9275-x

Coggins BE, Venters RA, Zhou P (2010) Radial sampling for fast NMR: concepts and practices over three decades. Prog Nucl Magn Reson Spectrosc 57:381–419. doi:10.1016/j.pnmrs.2010.07.001

Coggins BE, Werner-Allen JW, Yan A, Zhou P (2012) Rapid protein global fold determination using ultrasparse sampling, high-dynamic range artifact suppression, and time-shared NOESY. J Am Chem Soc 134:18619–18630. doi:10.1021/ja307445y

Cui G, Benirschke RC, Tuan HF, Juranic N, Macura S, Botuyan MV, Mer G (2010) Structural basis of ubiquitin recognition by translesion synthesis DNA polymerase iota. Biochemistry 49:10198–10207. doi:10.1021/bi101303t

Güntert P (2004) Automated NMR structure calculation with CYANA. Methods Mol Biol 278:353–378

Guthe S, Kapinos L, Moglich A, Meier S, Grzesiek S, Kiefhaber T (2004) Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. J Mol Biol 337:905–915. doi:10.1016/j.jmb.2004.02.020

Hiller S, Ibraghimov I, Wagner G, Orekhov VY (2009) Coupled decomposition of four-dimensional NOESY spectra. J Am Chem Soc 131:12970–12978. doi:10.1021/ja902012x

Hoch JC, Stern AS (2001) Maximum entropy reconstruction, spectrum analysis and deconvolution in multidimensional nuclear magnetic resonance. Methods Enzymol 338:159–178

Hyberts SG, Frueh DP, Arthanari H, Wagner G (2009) FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation. J Biomol NMR 45:283–294. doi:10.1007/s10858-009-9368-1

Kazimierczuk K, Stanek J, Zawadzka-Kazimierczuk A, Kozminski W (2010) Random sampling in multidimensional NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 57:420–434. doi:10.1016/j.pnmrs.2010.07.002

Nudelman I, Akabayov SR, Scherf T, Anglister J (2011) Observation of intermolecular interactions in large protein complexes by 2D-double difference nuclear overhauser enhancement spectroscopy: application to the 44 kDa interferon-receptor complex. J Am Chem Soc 133:14755–14764. doi:10.1021/ja205480v

Otting G, Wuthrich K (1990) Heteronuclear filters in two-dimensional [1H,1H]-NMR spectroscopy: combined use with isotope labelling for studies of macromolecular conformation and intermolecular interactions. Q Rev Biophys 23:39–96

Stuart AC, Borzilleri KA, Withka JM, Palmer AG 3rd (1999) Compensating for variations in $^{1}H–^{13}C$ scalar coupling constants in isotope-filtered NMR experiments. J Am Chem Soc 121:5346–5347

Tao Y, Strelkov SV, Mesyanzhinov VV, Rossmann MG (1997) Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. Structure 5:789–798

Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) High-resolution four-dimensional $^{1}H–^{13}C$ NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. J Am Chem Soc 127:2767–2775

Wen J, Zhou P, Wu J (2012) Efficient acquisition of high-resolution 4-D diagonal-suppressed methyl-methyl NOESY for large proteins. J Magn Reson 218:128–132. doi:10.1016/j.jmr.2012.02.021

Werner-Allen JW, Coggins BE, Zhou P (2010) Fast acquisition of high resolution 4-D amide-amide NOESY with diagonal suppression, sparse sampling and FFT-CLEAN. J Magn Reson 204:173–178. doi:10.1016/j.jmr.2010.02.017

Zwahlen C, Legault P, Vincent SJF, Greenblatt J, Konrat R, Kay LE (1997) Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage N-peptide/boxB RNA complex. J Am Chem Soc 119:6711–6721