Protein NMR

Deutsche Ausgabe: DOI: 10.1002/ange.201504070
Internationale Ausgabe: DOI: 10.1002/anie.201504070

Probing Invisible, Excited Protein States by Non-Uniformly Sampled Pseudo-4D CEST Spectroscopy**

Dong Long, Frank Delaglio, Ashok Sekhar, and Lewis E. Kay**

Abstract: *Chemical exchange saturation transfer (CEST) NMR spectroscopy is a powerful tool for studies of slow timescale protein dynamics. Typical experiments are based on recording a large number of 2D data sets and quantifying peak intensities in each of the resulting planes. A weakness of the method is that peaks must be resolved in 2D spectra, limiting applications to relatively small proteins. Resolution is significantly improved in 3D spectra but recording uniformly sampled data is time-prohibitive. Here we describe non-uniformly sampled HNCO-based pseudo-4D CEST that provides excellent resolution in reasonable measurement times. Data analysis is done through fitting in the time domain, without the need of reconstructing the frequency dimensions, exploiting previously measured accurate peak positions in reference spectra. The methodology is demonstrated on several protein systems, including a nascent form of superoxide dismutase that is implicated in neurodegenerative disease.*

NMR spectroscopy has emerged as an extremely powerful technique for characterizing protein dynamics over a broad spectrum of timescales,[1–3] including millisecond motions that are often important functionally.[4] Central to the emergence of NMR as a tool for probing slow motions in these macromolecules has been the development of protein-based Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion (RD) experiments[5] and more recently CEST spectroscopy.[6,7] In both cases the accurate quantification of NMR signal intensities forms the basis for extracting exchange parameters from pseudo-3D RD and CEST data sets in which a series of 2D ^{1}H–^{15}N or ^{1}H–^{13}C correlation planes are recorded as a function of different CPMG field strengths or B_1 field offsets, respectively. Of the two classes of experiment, CEST is particularly demanding because a series of up to ca. 100 2D planes are often needed to generate accurate profiles from which exchange parameters can be quantified (Figure 1). For most applications involving small proteins ($\approx$10 kDa) resolution in 2D data sets is not restricting and the majority of peaks can be analyzed. In contrast, resonance overlap is often limiting in applications to larger proteins and an obvious solution is to record a pseudo-4D CEST experiment in which a series of 3D spectra is obtained. Unfortunately, the increased resolution of cross-peaks in such data sets does not decrease the required number of B_1 field positions that must be recorded in the pseudo-4th dimension since CEST profiles must be obtained that are well sampled (Figure 1). Clearly, the measurement time for recording ca. 50–100 uniformly sampled 3D CEST data sets for each pseudo-4D experiment would be prohibitive.

Herein we present HNCO-based pseudo-4D CEST spectroscopy implemented with non-uniform sampling (NUS) in both ^{15}N and ^{13}CO dimensions, that drastically reduces acquisition times relative to the standard uniformly sampled 4D approach and improves resolution. A large number of

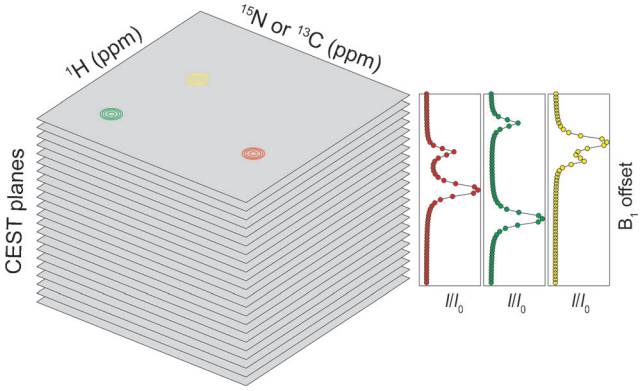

Figure 1. Schematic illustration of a "conventional" pseudo-3D CEST data set (left) and the resulting CEST profiles[6,7] (right). Experiments are recorded as a series of 2D spectra where for each spectrum a weak B_1 radio frequency field (rf, typically 10–50 Hz) is positioned at a different frequency (pseudo 3rd dimension). When the rf field is applied proximal to an excited state peak the resulting perturbation is transferred to the ground state correlation (shown here as yellow, green or red peaks), decreasing its intensity. When the rf is applied on resonance with the ground state peak its intensity disappears. The resultant profile of the intensity of the ground state peak (I) relative to the corresponding intensity in the absence of rf perturbation (I_0) gives rise to a pair of dips at the resonance positions of the major and minor conformers. Recording a series of 3D data sets in place of the 2D planes would be impractical in the absence of NUS.

[*] Dr. D. Long, Dr. A. Sekhar, Prof. L. E. Kay
Departments of Molecular Genetics
Biochemistry & Chemistry, University of Toronto
Toronto, Ontario M5S 1A8 (Canada)
E-mail: dlong@pound.med.utoronto.ca
kay@pound.med.utoronto.ca
Dr. F. Delaglio
NMR Science, Campbell CA 95008 (USA)
Prof. L. E. Kay
Hospital for Sick Children
Program in Molecular Structure and Function
555 University Avenue, Toronto, Ontario M5G1X8 (Canada)

[**] D.L. is the recipient of a Canadian Institutes of Health Research (CIHR) postdoctoral fellowship. This work was funded through a CIHR research grant to L.E.K. L.E.K. holds a Canada Research Chair in Biochemistry. The authors thank Prof. Liz Meiering, University of Waterloo, for the gift of a sample of apoSOD1^{2SH}.
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201504070.

NUS methods have been proposed with the goal in the majority of applications to reconstruct the full frequency domain data set.[8–14] However, in relaxation-based experiments the positions of cross-peaks and their line-shapes remain invariant with delays or radio-frequency fields and can be obtained a priori from reference spectra.[15] This significantly simplifies the data fitting problem to one involving reproducing peak intensities, without the need to reconstruct frequency dimensions. Indeed the process of generating frequency data that typically involves "filling in" the missing time domain points using non-linear reconstruction methods followed by Fourier transformation might potentially lead to distortions of peak intensities that are critical for the quantitative analysis described here.

In what follows we explain the approach taken in the analysis of the time-domain CEST data by focusing on one of the many 3D HNCO data sets from the pseudo-4D HNCO CEST experiment. The time domain signal for any of these data sets can be written concisely as:[11]

$$\mathbf{S} = \sum_{m=1}^{M} \alpha_m \mathbf{V}_m^{13CO} \otimes \mathbf{V}_m^{15N} \otimes \mathbf{V}_m^{1H} \quad (1)$$

where α_m is the intensity of the m-th component (cross-peak), $\mathbf{V}_m^q$ is the normalized time-domain shape vector that describes the evolution of $q \in \{13CO, 15N, 1H\}$ magnetization contributing to the m-th component (see below) and $\otimes$ denotes tensor product. Note that three shape vectors are required to describe the time domain response associated with a peak in the HNCO spectrum. The directly detected dimension can always be densely sampled and subsequently Fourier transformed. Thus, it suffices to consider the 2D interferogram involving the indirectly detected ^{15}N and ^{13}CO dimensions for each ω_H value of interest. This reduces the number of peaks that must be considered from M to N ($N < M$),

$$\mathbf{S} = \sum_{n=1}^{N} \alpha_n \mathbf{V}_n^{13CO} \otimes \mathbf{V}_n^{15N} \quad (2)$$

where N is the number of (^{13}CO,^{15}N) cross-peaks that contribute to the signal of the 2D interferogram at ω_H (e.g., within ± 0.05 ppm of the chosen ^1H frequency). In Eq. (2) $\mathbf{V}_n^q$ is given by

$$V_n^q = \exp\left(i\omega_n^q t_i - R_n^q t_i\right) \quad (3)$$

where ω_n^q is the q resonance frequency (rad/s) of a given spin, R_n^q is its transverse relaxation rate, $q \in (13CO, 15N)$, and t_i is evolution time given by the sampling schedule (Poisson-gap sampling[16] in the present applications, see Supporting Information (SI) for details). In the HNCO-based experiment (Figure S1, SI), the ^{15}N chemical shift evolution is recorded in a constant-time manner so that R_n^{15N} can effectively be set to 0 and each point of the 2D interferogram given by Eq. (2) can be written as

$$S_{i,k} = \sum_n \alpha_n \cdot e^{-R_n^{13CO} t_i^{13CO}} \cos\left(\omega_n^{13CO} t_i^{13CO}\right) \cos\left(\omega_n^{15N} t_k^{15N}\right) \quad (4)$$

where we have considered only the cosine modulation terms of the signal. The remaining three quadrature components are obtained by replacing one or both cosine terms with sine in Eq. (4). To extract CEST profiles (i.e. α_n), the 2D (^{13}C–^{15}N) non-uniformly sampled time domain data are modeled by finding $\min||\mathbf{S}^{expt} - \mathbf{S}||^2$, where $\mathbf{S}^{expt}$ is the corresponding experimental NUS data and the fitting parameters include α, ω^{13CO}, ω^{15N} and R^{13CO}. Values of ω^{13CO} and ω^{15N} can be predetermined accurately from reference spectra and need only be fine-adjusted (within ± 5 Hz of the starting values in our analyses) in the optimization.

As a first demonstration of the methodology we used a non-TROSY version of the pseudo-4D HNCO-based CEST experiment (Figure S1 A, SI) to study conformational exchange in a G48A mutant Fyn SH3 domain (7 kDa) that has previously been shown to exchange between a ground-state native conformer and a conformationally excited-state (referred to in what follows as excited-state) that corresponds to an unfolded ensemble (25 °C),[17–19] $G \underset{k_{EG}}{\overset{k_{GE}}{\rightleftharpoons}} E$.

Figure 2 A shows a comparison of representative ^{15}N CEST profiles obtained via "standard" HSQC-based pseudo-3D CEST[7] (blue) and from pseudo-4D CEST (red) that exploits the NUS approach described above. Values of (k_{ex},p_E) of ($125 \pm 3\ s^{-1}$, $5.6 \pm 0.1\%$) and ($133 \pm 5\ s^{-1}$, $5.7 \pm 0.1\%$) were obtained from pseudo 3D and 4D approaches, respectively, based on simultaneous fits of profiles as described in the SI, where $k_{ex} = k_{GE} + k_{EG}$ and p_E is the population of the excited (unfolded) state. An excellent correlation is also obtained for extracted chemical shift differences between exchanging states ($\Delta\tilde{\omega}$), Figure 3 A.

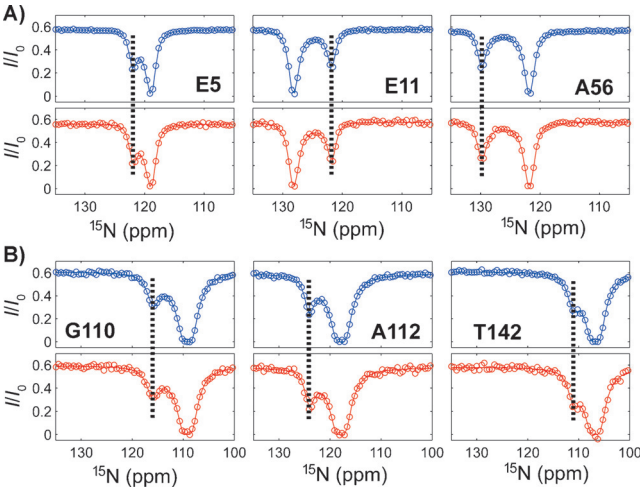

Figure 2. ^{15}N CEST profiles of selected residues of A) G48A Fyn SH3 and B) L99A T4L derived from the standard pseudo-3D HSQC-based CEST experiment[7] (blue open circles) and the pseudo-4D HNCO-based CEST approach (red open circles) described here. Solid lines are best fits to a two-state exchange model (see SI for details on experimental measurements and data analysis). The vertical dashed lines indicate the positions of minor dips from the excited state. Samples of 2 mM SH3 and 0.57 mM T4 L were used, along with recording times of 1.7 and 5.3 days, respectively. An additional study was performed using T4L data obtained in 2.6 days with similar exchange parameters and $\Delta\tilde{\omega}$ values obtained (see SI, Figure S3).

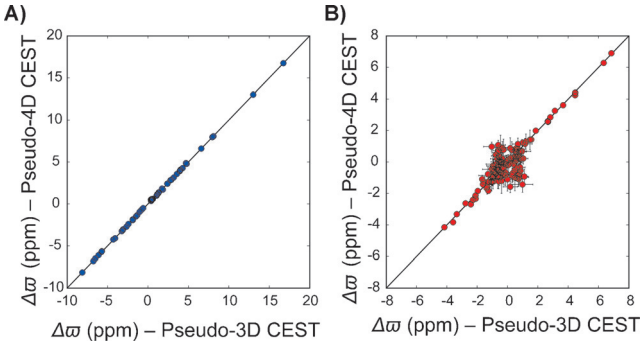

Figure 3. Linear correlation plots of $\Delta\tilde{\omega}$ values extracted from uniformly sampled, Fourier transformed pseudo-3D CEST (*X*-axis) and from non-uniformly sampled pseudo 4D HNCO-based CEST (*Y*-axis) experiments for A) G48A Fyn SH3 and B) L99A T4L. Error bars smaller than the size of the symbols are not displayed.

As a second example we considered an L99A cavity mutant of T4 lysozyme (19 kDa, L99A T4L)[20] that interconverts between ligand accessible and inaccessible conformers.[21] The increased size of this protein and its correlation time under the conditions studied ($\approx$19 ns at 10 °C) challenges CEST applications and we have therefore used a TROSY[22,23] version of the pseudo-4D experiment (Figure S1 B, SI). Figure 2 B compares a number of CEST traces generated from pseudo-3D and pseudo-4D data sets and, as with applications to smaller proteins, very similar profiles are obtained. Further, excellent agreement is noted between fitted exchange parameters from the pseudo-3D ($286 \pm 16\ s^{-1}$, $1.5 \pm 0.1\,\%$) and pseudo-4D ($292 \pm 30\ s^{-1}$, $1.6 \pm 0.1\,\%$) experiments with a good correlation between extracted $\Delta\tilde{\omega}$ values for $|\Delta\tilde{\omega}| > \approx 1$ ppm, Figure 3 B. For smaller values minor state dips are not resolved from the major dip, leading to some scattering. This is not an artefact of the NUS methodology but, rather, represents the inherent limitations of CEST given the signal-to-noise and the linewidths of the dips in the present application. Indeed, the pattern of scattering can be reproduced from fits of a pair of in silico synthetic L99A T4L CEST data sets that were calculated with extracted exchange parameters and shift differences obtained from analysis of the experimental data and to which random noise was added to recapitulate the experimental signal-to-noise (Figure S2, SI).

The advantage of the NUS pseudo-4D CEST approach becomes clear in the case of overlapping resonances in 2D ^{1}H–^{15}N spectra that prevents analysis of such residues in standard CEST applications. Figure 4 A,D show examples of such cases from both G48A Fyn SH3 (A6, W37) and L99A T4L (L13, V103). Because the ^{13}CO chemical shifts of these residues are distinct (Figure 4 B,E), their CEST profiles can be readily obtained from pseudo-4D HNCO-based experiments (Figure 4 C,F).

As a final example we applied the pseudo-4D experiment to study conformational dynamics in human superoxide dismutase (SOD1), a 153 residue 8-stranded β-barrel antioxidant enzyme that is dimeric in its functional, mature form ($Cu_2Zn_2SOD1^{S\text{–}S}$).[24] Each monomer of the symmetric dimer binds one Cu and one Zn ion, and contains a conserved disulfide bond. Misfolded conformations of SOD1 originating from nascent proteins that lack the disulfide bond and/or bound metal ions have been hypothesized to be important in the pathology of amyotrophic lateral sclerosis (ALS), a progressive neurodegenerative disease characterized by muscle atrophy and eventual death.[25] In a recent study we have established that the most immature form of the protein, apoSOD1^{2SH}, that is populated prior to metallation and disulfide bond formation, is in equilibrium with at least four other transiently populated states.[26] Two of these resemble conformations adopted by $Cu_2Zn_2SOD1^{S\text{–}S}$, corresponding to a dimeric form of the protein (Figure 5 A, top) and a conformation in which a helix present in the so-called electro-

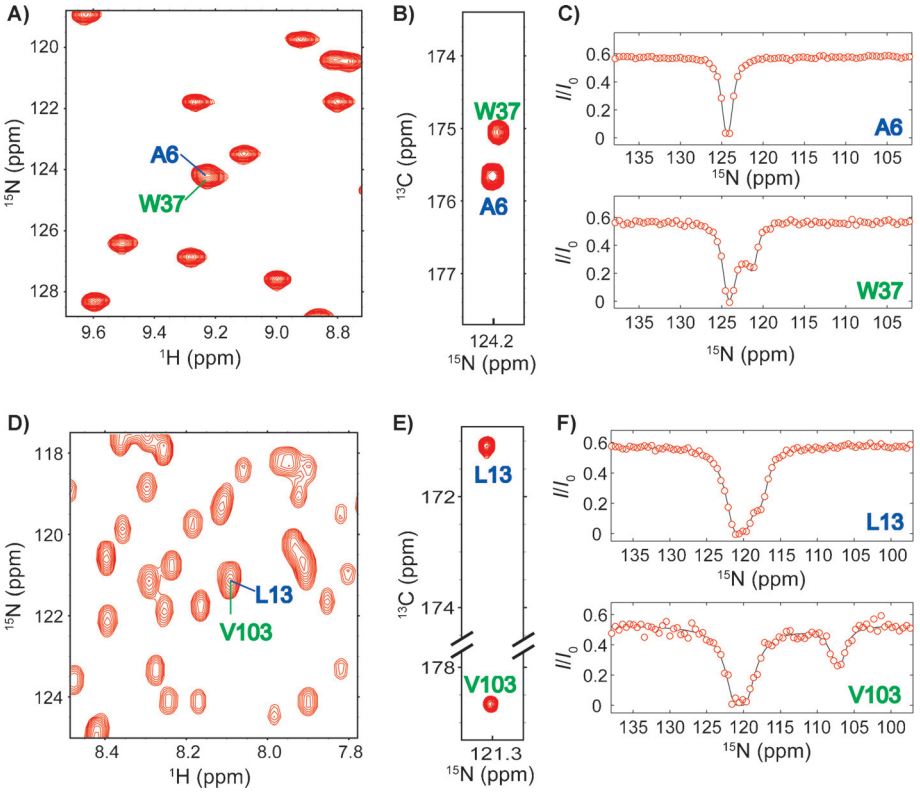

Figure 4. A,D) 2D ^{1}H–^{15}N correlation spectra of A) G48A Fyn SH3 and D) L99A T4L highlighting pairs of overlapping resonances. B,E) 3D HNCO strips for B) G48A Fyn SH3 ($^{1}H = 9.22$ ppm) and E) L99A T4L ($^{1}H = 8.09$ ppm) illustrating that the correlations can be separated along the ^{13}CO dimension. C,F) ^{15}N CEST profiles of these residues obtained from HNCO-based pseudo-4D CEST data sets.

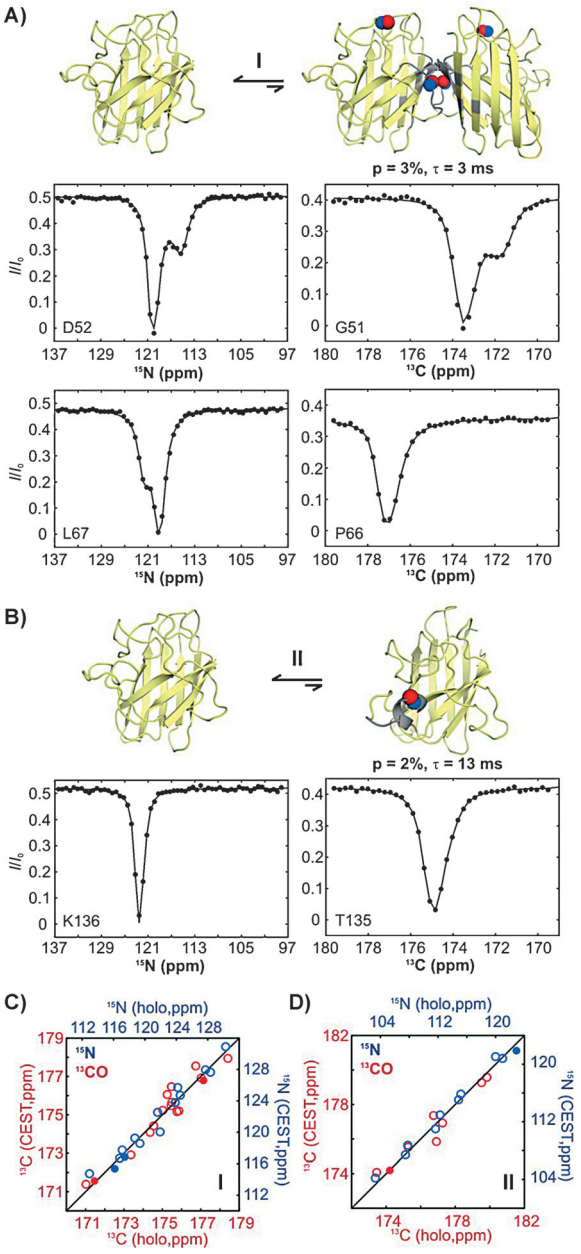

Figure 5. ^{15}N and ^{13}CO CEST profiles of apoSOD1^{2SH} residues belonging to the transient native dimerization process (A, process I) and the local helix folding process (B, process II). Regions of apoSOD1^{2SH} involved in dimerization and local folding are indicated in gray on the cartoon representations of the excited state structural models,[26] while ^{15}N and ^{13}CO nuclei giving rise to the CEST profiles illustrated here are highlighted by blue and red spheres, respectively. Note that these profiles are derived from residues that are severely overlapped in 2D spectra (Figure S5). C,D) Correlations between experimentally measured ^{15}N (blue) and ^{13}CO (red) chemical shifts for probes reporting on the processes giving rise to dimeric (C) and locally folded (D) excited state conformers (*Y*-axis) and the corresponding chemical shifts[27] of $Cu_2Zn_2SOD1^{S-S}$ (*X*-axis). The data points obtained in this work are indicated by filled circles, with the remainder obtained in a separate study.[26]

static loop folds locally (Figure 5B, top). These excited states are populated at 3% and 2%, respectively, and have lifetimes of 3 ms and 13 ms, Figure 5. This pair of transitions has been characterized using pseudo-3D CEST experiments and an excellent agreement between chemical shifts of the excited state with those of $Cu_2Zn_2SOD1^{S-S}$ is obtained (Figure 5C,D, open circles).[26] However, the excited state chemical shifts of key residues (Asp52, Leu67 and Lys136) involved in these two exchange processes could not be measured because they were present in overlapped regions of the ^{1}H–^{15}N HSQC spectrum of apoSOD1^{2SH} (Figure S5). By recording NUS based pseudo-4D CEST experiments the ^{15}N and ^{13}CO CEST profiles for these "missing" residues could be obtained and the extracted chemical shifts are in excellent agreement with the measured values from $Cu_2Zn_2SOD1^{S-S}$ (Figure 5C,D, closed circles), further validating the dimeric and locally folded structural models for these excited states.

In summary, we have presented HNCO-based pseudo-4D CEST experiments for characterizing slow conformational exchange in ^{15}N, ^{13}C labeled proteins that provide superior resolution relative to conventional pseudo-3D experiments. The non-uniform sampling approach employed reduces the total experimental time from months to days and importantly data processing is achieved in a straightforward manner by analysis in the ^{15}N–^{13}C time domains without the need for Fourier transformation. As such the method exploits chemical shift information that is already available from reference spectra, producing higher quality fits than would otherwise be obtained. The utility of the NUS pseudo-4D CEST experiments was established with applications to a number of different exchanging systems, enabling measurement of excited state ^{15}N and ^{13}CO chemical shifts of residues that are overlapped in 2D ^{1}H–^{15}N HSQC spectra. We expect that the approach of analyzing time-domain NUS data, demonstrated in the context of the CEST measurements here, will find wide applications in a range of pseudo-*n*D NMR studies that monitor systematic changes of peak intensities, including a large body of relaxation and real-time NMR experiments.

Keywords: CEST · conformational dynamics · non-uniform sampling · proteins · pseudo-4D NMR spectroscopy

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 10507–10511
Angew. Chem. **2015**, *127*, 10653–10657

[1] A. G. Palmer, J. Williams, A. McDermott, *J. Phys. Chem.* **1996**, *100*, 13293–13310.
[2] R. Ishima, D. A. Torchia, *Nat. Struct. Biol.* **2000**, *7*, 740–743.
[3] A. Mittermaier, L. E. Kay, *Science* **2006**, *312*, 224–228.
[4] A. Sekhar, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12867–12874.
[5] A. G. Palmer, C. D. Kroenke, J. P. Loria, *Methods Enzymol.* **2001**, *339*, 204–238.
[6] N. L. Fawzi, J. Ying, R. Ghirlando, D. A. Torchia, G. M. Clore, *Nature* **2011**, *480*, 268–272.
[7] P. Vallurupalli, G. Bouvignies, L. E. Kay, *J. Am. Chem. Soc.* **2012**, *134*, 8148–8161.
[8] S. G. Hyberts, H. Arthanari, G. Wagner, *Top. Curr. Chem.* **2012**, *316*, 125–148.
[9] S. G. Hyberts, G. J. Heffron, N. G. Tarragona, K. Solanky, K. A. Edmonds, H. Luithardt, J. Fejzo, M. Chorev, H. Aktas, K. Colson, et al., *J. Am. Chem. Soc.* **2007**, *129*, 5108–5116.

[10] V. Jaravine, I. Ibraghimov, V. Y. Orekhov, *Nat. Methods* **2006**, *3*, 605 – 607.
[11] V. Y. Orekhov, V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 271 – 292.
[12] J. Stanek, W. Koźmiński, *J. Biomol. NMR* **2010**, *47*, 65 – 77.
[13] Y. Matsuki, M. T. Eddy, R. G. Griffin, J. Herzfeld, *Angew. Chem. Int. Ed.* **2010**, *49*, 9215 – 9218; *Angew. Chem.* **2010**, *122*, 9401 – 9404.
[14] Y. Matsuki, M. T. Eddy, J. Herzfeld, *J. Am. Chem. Soc.* **2009**, *131*, 4648 – 4656.
[15] Y. Matsuki, T. Konuma, T. Fujiwara, K. Sugase, *J. Phys. Chem. B* **2011**, *115*, 13740 – 13745.
[16] S. G. Hyberts, K. Takeuchi, G. Wagner, *J. Am. Chem. Soc.* **2010**, *132*, 2145 – 2147.
[17] G. Bouvignies, P. Vallurupalli, L. E. Kay, *J. Mol. Biol.* **2014**, *426*, 763 – 774.
[18] D. Long, A. Sekhar, L. E. Kay, *J. Biomol. NMR* **2014**, *60*, 203 – 208.
[19] D. Long, G. Bouvignies, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8820 – 8825.
[20] A. E. Eriksson, W. A. Baase, J. A. Wozniak, B. W. Matthews, *Nature* **1992**, *355*, 371 – 373.
[21] G. Bouvignies, P. Vallurupalli, D. F. Hansen, B. E. Correia, O. Lange, A. Bah, R. M. Vernon, F. W. Dahlquist, D. Baker, L. E. Kay, *Nature* **2011**, *477*, 111 – 114.
[22] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366 – 12371.
[23] D. Nietlispach, *J. Biomol. NMR* **2005**, *31*, 161 – 166.
[24] J. S. Valentine, P. A. Doucette, S. Zittin Potter, *Annu. Rev. Biochem.* **2005**, *74*, 563 – 593.
[25] P. Zetterström, K. S. Graffmo, P. M. Andersen, T. Brännström, S. L. Marklund, *NeuroMol. Med.* **2013**, *15*, 147 – 158.
[26] A. Sekhar, J. Rumfeldt, H. R. Broom, C. M. Doyle, G. Bouvignies, E. M. Meiering, L. E. Kay, *Elife* **2015**, DOI: 10.7554/eLife.07296.
[27] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, M. S. Viezzoli, *Eur. J. Biochem.* **2002**, *269*, 1905 – 1915.

Received: May 5, 2015
Published online: July 14, 2015

