NMR Spectroscopy

International Edition: DOI: 10.1002/anie.201810868
German Edition: DOI: 10.1002/ange.201810868

Measuring Diffusion Constants of Invisible Protein Conformers by Triple-Quantum ^{1}H CPMG Relaxation Dispersion

*Tairan Yuwen, Ashok Sekhar, Andrew J. Baldwin, Pramodh Vallurupalli, and Lewis E. Kay**

Abstract: *Proteins are not locked in a single structure but often interconvert with other conformers that are critical for function. When such conformers are sparsely populated and transiently formed they become invisible to routine biophysical methods, however they can be studied in detail by NMR spin-relaxation experiments. Few experiments are available in the NMR toolkit, however, for characterizing the hydrodynamic properties of invisible states. Herein we describe a CPMG-based experiment for measuring translational diffusion constants of invisible states using a pulsed-field gradient approach that exploits methyl ^{1}H triple-quantum coherences. An example, involving diffusion of a sparsely populated and hence invisible unfolded protein ensemble is presented, without the need for the addition of denaturants that tend to destroy weak interactions that can be involved in stabilizing residual structure in the unfolded state.*

A complete description of biomolecular structure must extend beyond the most populated ground conformer to include higher energy (excited) states that are often only transiently formed. Although their low populations and short lifetimes challenge most biophysical approaches, a growing number of solution NMR methods have emerged to study these recalcitrant and most often "invisible" conformers,[1] leading to the discovery of novel mechanisms of biomolecular function.[2] The experiments that are most often exploited include Carr–Purcell–Meiboom–Gill (CPMG) and off-resonance $R_{1\rho}$ relaxation dispersion, Chemical/Dark-state Exchange Saturation Transfer (CEST/DEST), and Paramagnetic Relaxation Enhancement (PRE) approaches[1b,3] with each optimally sensitive to exchange events on different timescales. A major goal of our laboratory for over the past decade has been the development of CPMG and CEST experiments so that, ultimately, descriptions of rare (excited state) conformers will become possible at a level of detail close to what is currently attainable for highly populated ground state conformations. The available toolkit of NMR experiments for studies of excited state conformers is constantly growing, enabling measurement of chemical shifts, residual dipolar couplings, residual CSAs, pseudo-contact shifts, order parameters, hydrogen exchange rates, PREs and pKa values.[4] These experiments largely provide structural or dynamics information; much less common are methods to measure hydrodynamic properties that inform on molecular size. Herein we present a CPMG-based pulse scheme for quantifying translational diffusion constants of excited states, providing valuable hydrodynamic information.

By means of introduction consider the classic gradient spin-echo pulse scheme used to measure diffusion constants for systems with long transverse relaxation times,[5] Figure 1 A. In what follows we initially neglect relaxation and chemical exchange. The attenuation of transverse magnetization, M_+, due to diffusion is given by Equation (1)[5]

$$M_+(T,K) = M_+(0,0)\exp(-2/3K^2D\delta)\exp(-K^2DT) \quad (1)$$

where $K=\gamma g\delta$, γ is the gyromagnetic ratio of a probe-spin attached to the diffusing molecule, g is the strength of the encode/decode gradients, each of duration δ, T is the time between these gradients and D is the diffusion constant. When $T \gg \delta$, as is most often the case, the first exponential in Equation (1) can be neglected and therefore it follows that the effect of diffusion during T is to evolve M_+ according to Equation (2).[6]

$$\frac{dM_+}{dT} = -K^2DM_+ \quad (2)$$

Thus, including evolution from chemical shift, relaxation, diffusion during an interval T that separates encoding/decoding gradients and two-state chemical exchange, the Bloch–McConnell equations[7] become Equation (3) (see Supporting Information)

$$\begin{aligned}\frac{dM_+^G}{dT} &= -i\omega_G M_+^G - K^2D^GM_+^G - R_2^GM_+^G - k_{GE}M_+^G + k_{EG}M_+^E,\\ \frac{dM_+^E}{dT} &= -i\omega_E M_+^E - K^2D^EM_+^E - R_2^EM_+^E + k_{GE}M_+^G - k_{EG}M_+^E\end{aligned} \quad (3)$$

where superscripts G and E denote the interconverting

[*] Dr. T. Yuwen, Prof. L. E. Kay
Departments of Molecular Genetics, Biochemistry and Chemistry, University of Toronto
Toronto, Ontario M5S 1A8 (Canada)
E-mail: kay@pound.med.utoronto.ca
Prof. L. E. Kay
Program in Molecular Medicine, Hospital for Sick Children
555 University Avenue, Toronto, Ontario M5G 1X8 (Canada)
Dr. A. Sekhar
Molecular Biophysics Unit, Indian Institute of Science
Bangalore, Karnataka 560012 (India)
Dr. A. J. Baldwin
Physical and Theoretical Chemistry Laboratory, University of Oxford
Oxford OX1 3QZ (UK)
Dr. P. Vallurupalli
TIFR Centre for Interdisciplinary Sciences, Tata Institute of Fundamental Research
Hyderabad, Telangana 500107 (India)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201810868.

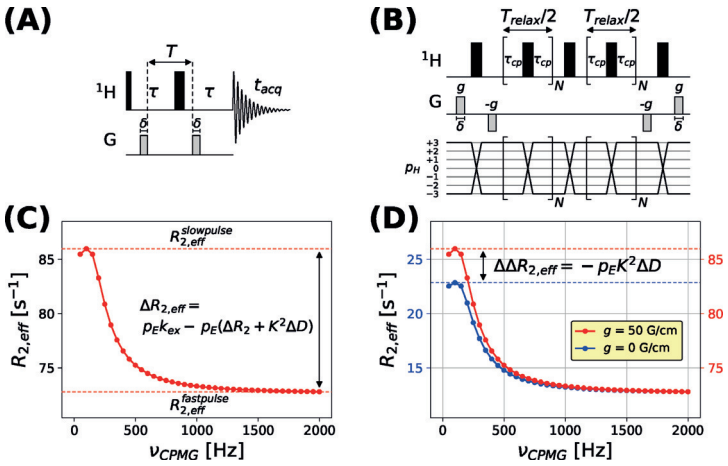

Figure 1. A) Pulsed-field gradient spin-echo based approach for measurement of diffusion constants. Only populated states that give rise to observable NMR signals can be studied. B) Schematic representation of 3Q ^{1}H CPMG-element (top) along with a coherence transfer diagram (bottom) that illustrates how coherence orders change throughout the element (see Figure 2). C) Simulated CPMG dispersion profile with expressions for $R_{2,eff}$ values in the slow ($\nu_{CPMG}\to 0$) and fast ($\nu_{CPMG}\to\infty$) pulsing regimes, $R_{2,eff}^{slowpulse}$ and $R_{2,eff}^{fastpulse}$, given in the text. D) Diffusion changes the relative sizes of dispersion profiles; as ΔD becomes increasingly negative the relative size of the profile recorded with the largest encode/decode gradients becomes larger. In (C) and (D) simulations were performed using $R_1^G = R_1^E = 2\ s^{-1}$, $R_2^G = R_2^E = 10\ s^{-1}$, $p_E = 10\%$, $k_{ex} = 100\ s^{-1}$, $\Delta\nu_{GE} = 300$ Hz (where $\Delta\nu_{GE}$ is 3 times the difference in ^{1}H chemical shifts between ground and excited state spins taking into account that it is 3Q ^{1}H magnetization that is interrogated during the CPMG period), $T_{relax} = 40$ ms, $D^G = 1.0\times10^{-6}\ cm^2\,s^{-1}$, $D^E = 0.5\times10^{-6}\ cm^2\,s^{-1}$, $\delta = 1$ ms and a pulse scheme as in (B) or Figure 2. In (C) the simulation was carried out with $g = 50\ G\,cm^{-1}$.

ground and excited states, ($k_{ex} = k_{GE} + k_{EG}$), ω_A, D^A, and R_2^A are the chemical shift (rad s^{-1}), diffusion constant, and intrinsic transverse relaxation rate of spin $A \in \{G, E\}$. The effect of diffusion during T thus modifies the Bloch–McConnell equations by additional terms proportional to K^2D that are easily subsumed into expressions for transverse relaxation via $R_{2,K}^A = R_2^A + K^2D^A$.[6] More complicated equations are necessary to include the effects of chemical exchange during the encoding/decoding gradients, but as $T \gg \delta$ this is not an important consideration. Thus, the standard Bloch–McConnell equations for chemical exchange can be used to analyze diffusion-based experiments so long as $R_{2,K}^A$ is substituted for R_2^A.[6] This realization forms the basis for understanding how the CPMG approach can be used to measure diffusion constants for invisible states, as described below.

Consider a CPMG experiment where the encode/decode gradients are placed on opposite sides of the CPMG interval (Figure 1B) and assume, further, a slowly exchanging spin system, interconverting between a pair of states. In the limit where ν_{CPMG} ($= 1/(2\zeta)$, where ζ is the time between successive 180° refocusing pulses) is small the effective transverse relaxation rate of spin G is given by $R_{2,eff}^{slowpulse} = R_{2,K}^G + p_E k_{ex}$, that includes a contribution from life-time broadening accounting for exchange in the slow limit,[8] and p_E is the fractional population of the excited state ($p_E + p_G = 1$). In contrast, when ν_{CPMG} is large such that the effective chemical shift difference between exchanging spins is "pulsed out" (fast exchange on the chemical shift timescale), the transverse relaxation rate is given by the population weighted average of relaxation rates in each state (so long as $k_{ex} \gg |\Delta R_{2,K}|$), Equation (4)[9]

$$R_{2,eff}^{fastpulse} = p_G R_{2,K}^G + p_E R_{2,K}^E = R_2^G + K^2 D^G + p_E(\Delta R_2 + K^2\Delta D) \quad (4)$$

where $\Delta R_2 = R_2^E - R_2^G$, $\Delta D = D^E - D^G$. The net size of the dispersion profile is thus given by Equation (5)

$$\Delta R_{2,eff} = R_{2,eff}^{slowpulse} - R_{2,eff}^{fastpulse} = p_E k_{ex} - p_E(\Delta R_2 + K^2\Delta D) \quad (5)$$

so that $\Delta R_{2,eff}$ depends on ΔD, with $R_{2,eff}$ a function of both D^G and ΔD (Figure 1C). By recording the experiment with a pair of different encode/decode gradient strengths corresponding to $g \neq 0$ and $g = 0$ dispersion profiles that differ in size by $\Delta\Delta R_{2,eff} = -p_E K^2 \Delta D$ are obtained, Figure 1D. Combined fits of profiles measured with $g \neq 0$ and $g = 0$ (see Supporting Information) thus provide estimates of both D^G and D^E.

Figure 2 and Figure S1 in the Supporting Information illustrate the pulse schemes that have been developed to measure D^E. Central to the experiment is the gradient encode/decode pair that sandwich the constant-time CPMG interval of duration T_{relax}. In principle, a number of different CPMG schemes can be chosen, however as the size of the diffusion effect scales as K^2 (see above), which in turn depends on g^2, we have chosen the triple-quantum (3Q) methyl ^{1}H CPMG based sequence that is optimally sensitive to the encode/decode gradients.[10] Note that the effect of the encode/decode gradients on 3Q coherences is three-fold larger than for signal quantum magnetization so that a factor of 9 increase in K^2 is obtained (i.e., g must be multiplied by 3 in the analysis of the 3Q-CPMG data). Note also that $K^2 \propto \gamma^2$ so that the effect is largest for nuclei with large gyromagnetic ratios and it is thus advisable to apply the encode/decode gradients when the magnetization of interest is ^{1}H. Finally, as described pre-

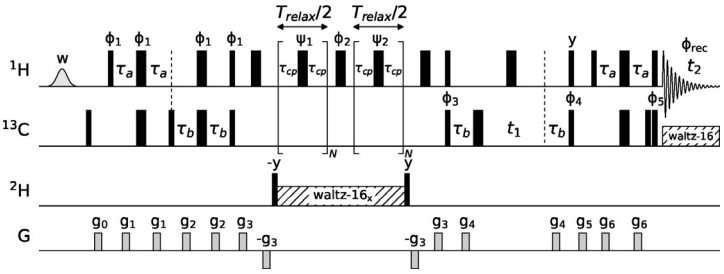

Figure 2. Methyl ^{1}H 3Q-CPMG pulse scheme for measuring excited state diffusion constants. Bipolar gradients, g_3, flanking the CPMG element of duration T_{relax} are applied to encode and decode magnetization. Note that for this pulse scheme $K = 6\gamma_H g\delta$ that includes a factor of 2 (bipolar gradient pair) and a factor of 3 (3Q). Details are provided in Ref. [11] and succinctly in the legend to Figure S1, and the source code is given in the Supporting Information.

viously,[10c,11] the penalty for ^{1}H 3Q selection is relatively small, only 1/4 of the intrinsic sensitivity of a SQ experiment is lost, neglecting relaxation effects.

An important consideration in optimizing the performance of the experiment is the size of K that should be chosen for highest sensitivity for a given T_{relax} time, since $\Delta\Delta R_{2,eff}$ increases with K^2, while the signal intensity $I(\nu_{CPMG},K)$ is attenuated by the factor $\exp(-K^2 D^G T_{relax})$. As a value of K should be chosen that optimizes $\Delta\Delta R_{2,eff}$ relative to the uncertainty in the $R_{2,eff}$ values, and noting that the uncertainty in each $R_{2,eff}$ value of the dispersion profile is proportional to $1/I(\nu_{CPMG},K)$, it can be shown that the product, Equation (6)

$$p_E K^2 \Delta D T_{relax} \exp(-K^2 D^G T_{relax}) \quad (6)$$

should be maximized and this occurs for $K^2 D^G T_{relax} = 1$. Note that this result is only approximate as $I(\nu_{CPMG},K)$ is a function of both D^G and ΔD so that optimal K values depend on both D^G and D^E. Estimates for D^E are, however, in general not a priori available.

The utility of the proposed methodology is illustrated with an application to a G48A mutant Fyn SH3 domain that has been shown previously to exchange between a folded conformer and an unfolded ensemble.[12] Figure 3 and Figure S2 show ^{1}H 3Q dispersion profiles for several methyl groups of the SH3 domain that are sufficiently large so as to permit accurate parameters to be extracted (25 °C); these clearly increase when encode/decode gradients are included (Figure 3 and Figure S2), implying that $\Delta D = D^E - D^G$ is negative, Equation (5). That the diffusion constant of the excited state is smaller than the ground state is in keeping with the fact that unfolded conformers are more extended than their folded counterparts. From simultaneous fits of dispersions profiles recorded at ^{1}H Larmor frequencies of 600 MHz ($g = 0$ G cm^{-1}) and 800 MHz ($g = 0$ G cm^{-1}, 30.6 G cm^{-1}) global exchange parameters (p_E, k_{ex}) = (7.6 ± 0.5 %, 124.6 ± 9.0 s^{-1}) are obtained, along with ^{1}H chemical shift differences between exchanging states indicated in the panels in Figure 3. In total 10 methyl groups are included in the data analysis, with D^G values varying between 1.54–1.59 × 10^{-6} cm^2 s^{-1} from individual fits, while D^E rates range from 1.01–1.41 × 10^{-6} cm^2 s^{-1}. Notably, $D^G = (1.57 \pm 0.01) \times 10^{-6}$ cm^2 s^{-1} and $D^E = (1.22 \pm 0.04) \times 10^{-6}$ cm^2 s^{-1} are obtained from a global fit, so that $D^G/D^E = 1.29 \pm 0.04$. The predicted hydrodynamic radius of the folded G48A Fyn SH3 domain based on an experimentally derived relation from Smith, Dobson and co-workers[13] (15.5 Å) is in good agreement with that calculated using the Stokes–Einstein equation (15.6 Å), the D^G value is similar to that predicted based on HYDROPRO NMR[14] (1.55 × 10^{-6} cm^2 s^{-1}) and there is excellent agreement between D^G values extracted from fits of CPMG profiles and those measured by recording a series of 2D planes using the scheme in Figure 2 with different g_3 values and with N set to 1 (Figure S3). Additionally, the ratio of diffusion values agrees well with the ratio obtained for a related SH3 domain from drk (1.31 ± 0.01), under conditions where both states are equally populated so that diffusion measurements can be made on each simultaneously.[15]

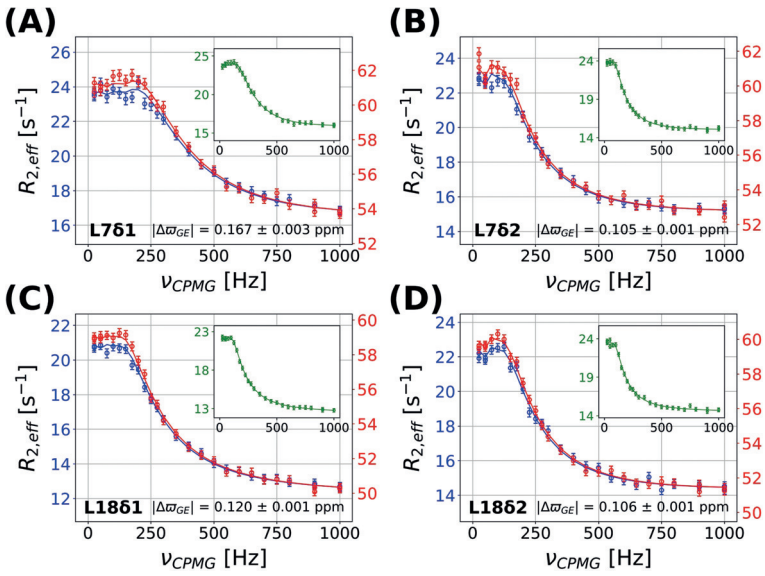

Figure 3. Experimental methyl ^{1}H 3Q dispersion profiles (circles) of selected residues from the G48A Fyn SH3 domain, 25 °C, 18.8 T, $g = 30.6$ G cm^{-1} (red) and 0 G cm^{-1} (blue) recorded with the scheme shown in Figure 2. Corresponding profiles measured at 14.1 T, $g = 0$ G cm^{-1} are shown in insets. Profiles have been scaled so that the fast pulsing regions superimpose to highlight the differences in sizes with and without encode/decode gradients. The four methyl groups selected are representatives with relatively large dispersions and clear differences in dispersion sizes measured with different encode/decode gradient strengths; profiles are simultaneously fit as described in Supporting Information (solid lines). Including more profiles in the fit does not change the extracted exchange parameters nor the obtained diffusion constants. CPMG dispersion profiles for the other 6 methyl groups included in the global fit are illustrated in Figure S2. $|\Delta\nu_{GE}|$ values can be calculated directly from $|\Delta\varpi_{GE}|$ as $|\Delta\nu_{GE}| = 3 \times |\Delta\varpi_{GE}| \times B_0$, where B_0 is the static magnetic field strength in Hz. Experimental details are included in Supporting Information, along with simulations illustrating the sensitivity of the diffusion CPMG experiment to $\Delta\nu_{GE}$ values (Figure S4). Fits of the data assumed $\Delta R_2 = 0$ s^{-1}, but this has little effect on the fitted D^G or D^E values (Figures S5 and S6).

In summary, we have presented an experiment for the measurement of self-diffusion constants of invisible, excited-states in exchange with their ground-state conformers. The experiment takes advantage of the sensitivity of 3Q coherences to pulsed field gradients, with differences in diffusion constants between ground and excited states encoded in the sizes of CPMG relaxation dispersion profiles. Optimization of the experiment is critical, Equation (6), because although $\Delta\Delta R_{2,eff}$ increases with K^2 and hence g^2, the sensitivity of the dispersion profiles decrease with K^2. It is anticipated that an important area of application will involve studies of unfolded protein conformers, as illustrated here. Such studies are typically carried out using denaturants or extremes of pH that shift the folding equilibrium towards the unfolded ensemble, increasing its population to the point where it can be readily observed in standard NMR experiments. However, this population shift often comes at a cost, whereby weak

interactions in the unfolded state are disturbed by the extreme conditions, leading to changes in hydrodynamic properties. The CPMG approach described here avoids this problem, so that accurate diffusion constants of the unfolded state can be obtained under near physiological conditions, so long as the unfolded ensemble is populated at a level of several percent. A second area of application could involve studies of oligomerization reactions, such as protein dimerization or trimerization, where $\Delta D \neq 0$, so long as exchange rates fall in the CPMG regime (millisecond lifetimes) with populations of a few percent for the excited state(s). The CPMG-diffusion scheme thus serves as a valuable addition to the expanding repertoire of experiments for measuring the molecular properties of NMR-invisible protein states.

Acknowledgements

This work was supported by grants from the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (L.E.K). T.Y. acknowledges post-doctoral support from the CIHR. A.S. acknowledges the Department of Science and Technology, Government of India for a Ramanujan Fellowship. L.E.K. holds a Canada Research Chair in Biochemistry.

Conflict of interest

The authors declare no conflict of interest.

Keywords: CPMG · invisible excited states · NMR spectroscopy · protein folding · triple quantum

How to cite: *Angew. Chem. Int. Ed.* **2018**, *57*, 16777–16780
Angew. Chem. **2018**, *130*, 17019–17022

[1] a) A. Sekhar, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12867–12874; b) N. J. Anthis, G. M. Clore, *Q. Rev. Biophys.* **2015**, *48*, 35–116.

[2] a) D. D. Boehr, D. McElheny, H. J. Dyson, P. E. Wright, *Science* **2006**, *313*, 1638–1642; b) K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972; c) B. Zhao, S. L. Guffy, B. Williams, Q. Zhang, *Nat. Chem. Biol.* **2017**, *13*, 968–974; d) I. J. Kimsey, E. S. Szymanski, W. J. Zahurancik, A. Shakya, Y. Xue, C. C. Chu, B. Sathyamoorthy, Z. C. Suo, H. M. Al-Hashimi, *Nature* **2018**, *554*, 195–201.

[3] a) A. G. Palmer, F. Massi, *Chem. Rev.* **2006**, *106*, 1700–1719; b) P. Vallurupalli, G. Bouvignies, L. E. Kay, *J. Am. Chem. Soc.* **2012**, *134*, 8148–8161; c) A. G. Palmer, *J. Magn. Reson.* **2014**, *241*, 3–17.

[4] a) P. Vallurupalli, D. F. Hansen, E. Stollar, E. Meirovitch, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 18473–18477; b) D. F. Hansen, P. Vallurupalli, L. E. Kay, *J. Biomol. NMR* **2008**, *41*, 113–120; c) P. Vallurupalli, D. F. Hansen, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11766–11771; d) D. F. Hansen, P. Vallurupalli, L. E. Kay, *J. Am. Chem. Soc.* **2009**, *131*, 12745–12754; e) A. L. Hansen, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1705–E1712; f) D. Long, G. Bouvignies, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8820–8825; g) B. Zhao, Q. Zhang, *J. Am. Chem. Soc.* **2015**, *137*, 13480–13483; h) R. S. Ma, Q. F. Li, A. D. Wang, J. H. Zhang, Z. J. Liu, J. H. Wu, X. C. Su, K. Ruan, *Phys. Chem. Chem. Phys.* **2016**, *18*, 13794–13798; i) A. Sekhar, R. Rosenzweig, G. Bouvignies, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2794–E2801; j) A. Sekhar, J. A. O. Rumfeldt, H. R. Broom, C. M. Doyle, R. E. Sobering, E. M. Meiering, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E6939–E6945.

[5] E. O. Stejskal, J. E. Tanner, *J. Chem. Phys.* **1965**, *42*, 288–292.

[6] C. S. Johnson, *J. Magn. Reson.* **1993**, *102*, 214–218.

[7] H. M. McConnell, *J. Chem. Phys.* **1958**, *28*, 430–431.

[8] A. G. Palmer, C. D. Kroenke, J. P. Loria, *Methods Enzymol.* **2001**, *339*, 204–238.

[9] a) J. Cavanagh, W. J. Fairbrother, A. G. Palmer, M. Rance, N. J. Skelton, *Protein NMR Spectroscopy*, 2nd ed., Academic Press, London, **2007**; b) A. J. Baldwin, *J. Magn. Reson.* **2014**, *244*, 114–124.

[10] a) D. Zax, A. Pines, *J. Chem. Phys.* **1983**, *78*, 6333–6334; b) G. Zheng, A. M. Torres, W. S. Price, *J. Magn. Reson.* **2009**, *198*, 271–274; c) R. Huang, J. P. Brady, A. Sekhar, T. Yuwen, L. E. Kay, *J. Biomol. NMR* **2017**, *68*, 249–255.

[11] T. Yuwen, P. Vallurupalli, L. E. Kay, *Angew. Chem. Int. Ed.* **2016**, *55*, 11490–11494; *Angew. Chem.* **2016**, *128*, 11662–11666.

[12] G. Bouvignies, P. Vallurupalli, L. E. Kay, *J. Mol. Biol.* **2014**, *426*, 763–774.

[13] D. K. Wilkins, S. B. Grimshaw, V. Receveur, C. M. Dobson, J. A. Jones, L. J. Smith, *Biochemistry* **1999**, *38*, 16424–16431.

[14] A. Ortega, D. Amorós, J. García de la Torre, *Biophys. J.* **2011**, *101*, 892–898.

[15] W. Y. Choy, F. A. A. Mulder, K. A. Crowhurst, D. R. Muhandiram, I. S. Millett, S. Doniach, J. D. Forman-Kay, L. E. Kay, *J. Mol. Biol.* **2002**, *316*, 101–112.

Manuscript received: September 20, 2018
Accepted manuscript online: October 29, 2018
Version of record online: November 21, 2018

