Protein-Structure Elucidation

International Edition: DOI: 10.1002/anie.201500640
German Edition: DOI: 10.1002/ange.201500640

Dependence of Distance Distributions Derived from Double Electron–Electron Resonance Pulsed EPR Spectroscopy on Pulse-Sequence Time**

James L. Baber, John M. Louis, and G. Marius Clore**

Abstract: *Pulsed double electron–electron resonance (DEER) provides pairwise P(r) distance distributions in doubly spin labeled proteins. We report that in protonated proteins, P(r) is dependent on the length of the second echo period T owing to local environmental effects on the spin-label phase memory relaxation time T_m. For the protein ABD, this effect results in a 1.4 Å increase in the P(r) maximum from T = 6 to 20 μs. Protein A has a bimodal P(r) distribution, and the relative height of the shorter distance peak at T = 10 μs, the shortest value required to obtain a reliable P(r), is reduced by 40% relative to that found by extrapolation to T = 0. Our results indicate that data at a series of T values are essential for quantitative interpretation of DEER to determine the extent of the T dependence and to extrapolate the results to T = 0. Complete deuteration (99%) of the protein was accompanied by a significant increase in T_m and effectively abolished the P(r) dependence on T.*

Double electron–electron resonance (DEER) is a powerful EPR method for measuring the distance between two unpaired electron spins.[1] Four-pulse DEER (Figure 1),[1d] coupled with site-directed spin labeling,[2a] has been increasingly exploited in biophysics to gain quantitative insight into the structure and function of macromolecules and their assemblies, to probe conformational transitions and determine relative populations of multiple conformational states.[2b–f] DEER also offers much promise in supplementing existing techniques of macromolecular structure determination, such as X-ray crystallography, NMR spectroscopy, and solution X-ray scattering. The commercial availability of high-powered (150 W) Q-band pulsed EPR spectrometers has greatly enhanced the prospects for more routine implementation of DEER. Sensitivity at Q-band can be up to an order of magnitude higher than that of traditional X-band,[3] resulting in increased data quality and shorter data-acquisition times. Further, although a temperature of about 50 K is optimal for sensitivity purposes, most DEER measurements have been reported at 80 K owing to the practical disadvantage of liquid-helium cooling. The commercial availability, however, of a cold head-cooled, liquid helium cryostat makes working at optimal DEER temperatures for extended periods of time practical, thus yielding further increases in sensitivity.

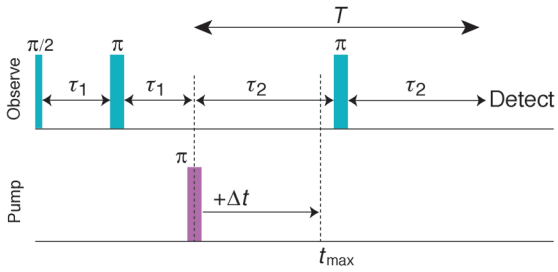

Figure 1. Four-pulse DEER.[1d] The total duration of the second echo period is T ($=2\tau_2$).

The reliability of DEER data is dependent on both the signal-to-noise ratio and the total dipolar-evolution time (t_{max} in Figure 1).[2d] Measurement of longer distances requires longer dipolar-evolution times and, therefore, longer second echo periods ($T = 2\tau_2$ in Figure 1) for the observe pulses.

The measured DEER curve is a convolution of two signals. The first is the desired cosine-modulated signal from dipolar coupling of intramolecular spin pairs. The second is an exponentially decaying background arising from dipolar coupling to other spin-labeled proteins randomly distributed in the solvent matrix. To obtain reliable distance information, the intermolecular background must be reliably deconvoluted from the desired intramolecular signal. The longer the dipolar evolution is sampled, the more reliable the deconvolution of the two signals. Thus, there are two good reasons for collecting data for as long a total dipolar-evolution time as practical. Unfortunately, the total time of the pulse sequence is limited by the relaxation properties of the electron spins. For any given sample, a compromise must be chosen between the signal-to-noise ratio and the length of the dipolar-evolution time. As compared to traditional X-band, high-power Q-band offers both long (> 10 μs for a deuterated solvent matrix at 50 K) dipolar-evolution times and a good signal-to-noise ratio. Here, we report that taking advantage of longer dipolar-evolution times afforded at Q-band reveals a significant issue of practical importance that must be taken into account when making quantitative distance or population measurements on protonated proteins by DEER: specifically,

[*] Dr. J. L. Baber, Dr. J. M. Louis, Dr. G. M. Clore
National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health
Bethesda, MD 20892-0520 (USA)
E-mail: jamesbaber@niddk.nih.gov
mariusc@mail.nih.gov

[**] We thank Alexey Potapov, Ad Bax, and Lalit Deshmukh for useful discussions, and Annie Aniana for technical assistance. This work was supported by the Intramural Program of NIDDK, NIH and by the AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201500640.

the P(r) distance distribution can depend significantly on the length of T, and therefore appropriate strategies need to be employed to take these effects into account.

DEER data were collected over a range of T values for two proteins of known structure, the albumin-binding GA domain (ABD)[4] and the immunoglobulin-binding B domain of protein A (Protein A).[5] Both proteins were nitroxide-labeled with (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl)methanethiosulfonate (MTSL) at two engineered surface cysteine sites located close to the N- and C-termini of the structured domains (see Figure S1 in the Supporting Information), and dissolved in a medium comprising a 30:70 mixture of deuterated [D_8]glycerol and 99.9% D_2O (see the Supporting Information for details of sample preparation).

Analysis of the DEER curves by Tikhonov regularization[2d,6] yielded a single broad distance P(r) distribution for the ABD protein (Figure 2a–c). However, as T increased, the short side of the P(r) distribution was progressively attenuated, resulting in a measurable shift in the P(r) maximum to longer distance (Figure 2b). In this example, the P(r) maximum was shifted by 1.4 Å to longer distance on going from $T = 6$ to 20 μs. The dependence on T is evident in the raw DEER data (Figure 2a). Zooming in to the $t = 0$ μs region, one can see that the slope of the raw data decreases as T increases. The plot of the P(r) maximum versus T yields a straight line (Figure 2c) with a y-intercept (at $T = 0$) of (33.6 ± 0.2) Å and a slope of (0.085 ± 0.015) Å μs^{-1}.

Protein A offers a more dramatic dependence of the DEER-derived P(r) distribution on T, where P(r) (obtained by Tikhonov regularization) is clearly bimodal and the relative height of the two peaks depends critically on the length of T (Figure 2d–f). The first maximum in P(r), centered at 33 Å, is severely reduced in intensity at the longest T values that were practical to measure (up to 20 μs) for the protonated protein. The bimodal nature of the distribution is readily evident in the raw DEER data acquired at $T = 6$ μs (black curve in inset of Figure 2d) and becomes increasingly less obvious as T increases. The second maximum in P(r) shows no obvious dependence on T; instead, its position varies randomly over a small range of about 0.4 Å. The results of Tikhonov regularization were normalized to the maximum of the longer distance (ca. 38 Å) of the bimodal P(r) distribution (Figure 2e), and a plot of the ratio of peak heights versus T yields an approximately straight line with a y-intercept of 1.25 ± 0.02 (at $T = 0$) and slope of (-0.047 ± 0.002) μs^{-1} (Figure 2f). A single-exponential fit to this curve is worse than the linear fit. As no physical model predicts a linear dependence, we also fit the curve to the ratio of two stretched exponentials, $a[exp(-T/T_{m1})^{x_1}/exp(-T/T_{m2})^{x_2}]$, by analogy to the stretched-exponential nature of electron spin-echo decay curves,[7] in which a is a pre-exponential factor, x_i are exponential stretch factors, and T_{mi} are apparent phase memory relaxation times. As this function has many local minima, we chose starting values for minimization of $a = 1.25$, equal to the y-intercept from the linear fit; 9 μs and 1.6 for T_{m1} and x_1, respectively, corresponding to a partially solvent accessible spin label of intermediate mobility;[7b] and 12.6 μs and 1.3 for T_{m2} and x_2, respectively, corresponding to a fully solvent accessible, highly mobile spin label.[7b] The resulting fit is more consistent with the data (red curve, Figure 2f) than the linear fit (black dashes, Figure 2f), and the physically reasonable values for the two spin-label phase memory relaxation times ($T_m = 9$ and 13 μs, close to the values obtained from a two-pulse Hahn spin-echo experiment; see Figure S5a) and stretch factors (ca. 1.5) provide an established physical model[7b] for understanding the experimental dependence of the DEER curves and P(r) distributions on the length of T.

To gain further insight into this phenomenon, Protein A was perdeuterated to levels of approximately 80% or 99% (by growing the bacteria in minimal medium in either D_2O alone[8] or D_2O supplemented with fully deuterated glucose plus deuterated Isogro growth medium, respectively; see the Supporting Information) and spin-labeled with perdeuterated MTSL. The results of these DEER measurements are reported in Figure 2g–i and j–l, respectively. The dependence of the peak ratio in the P(r) distribution on T was significantly reduced (Figure 2h) in the sample perdeuterated to a level of approximately 80% as compared to that observed with the protonated protein (Figure 2e): a plot of relative peak intensity versus T is approximately linear and can be fit to a straight line (black dashes in Figure 2i) with a y-intercept of 1.34 ± 0.01 (at $T = 0$) and slope of (-0.0097 ± 0.0003) μs^{-1}. Thus, the rate of decay in intensity of the first peak in the P(r) distribution was reduced by a factor of approximately 5 upon perdeuteration to a level of about 80%. Further, 99% perdeuteration completely abolished the T dependence of the peak ratio in the P(r) distribution (Figure 2k and l). The latter finding on 99% perdeuterated protein A also holds when the spin-label is protonated (data not shown) owing to the large hyperfine couplings of the protons in the nitroxide spin-label.[9]

The most obvious explanation for the dependence of the P(r) distributions on T is that the spin-label phase memory relaxation time T_m is affected by the local environment of the spin label, which can be closer to or further from agents, in particular protons, that will cause transverse spin relaxation, depending on the rotameric state of the spin label and/or protein side chains. Thus, as T increases, the representation of rotamer populations with shorter T_m values in the P(r) distribution will be attenuated relative to those with longer T_m values. Indeed, the experimental P(r) distributions are consistent with those predicted from a spin-label rotamer library using the program MMM[10] (see Figure S6). These findings are presaged by an earlier report that T_m is correlated to spin-label proximity to the protein surface:[7b] spin labels that are fully accessible to a deuterated solvent matrix were found to have significantly longer T_m values than those partially buried in a protein pocket or in close proximity to the protein surface. The T_m values reported by Huber et al.[7b] for a protonated protein in a deuterated glycerol/D_2O solvent mixture varied from 1.4 to 12.6 μs, depending on the position of the nitroxide spin label, and it was proposed that the T_m value could be used as a measure of the proximity of a label to the protein.

For the ABD protein, the difference in distance between the P(r) maximum at $T = 0$ (by extrapolation) and 10 μs (the minimum time necessary to obtain a reliable P(r) distribu-

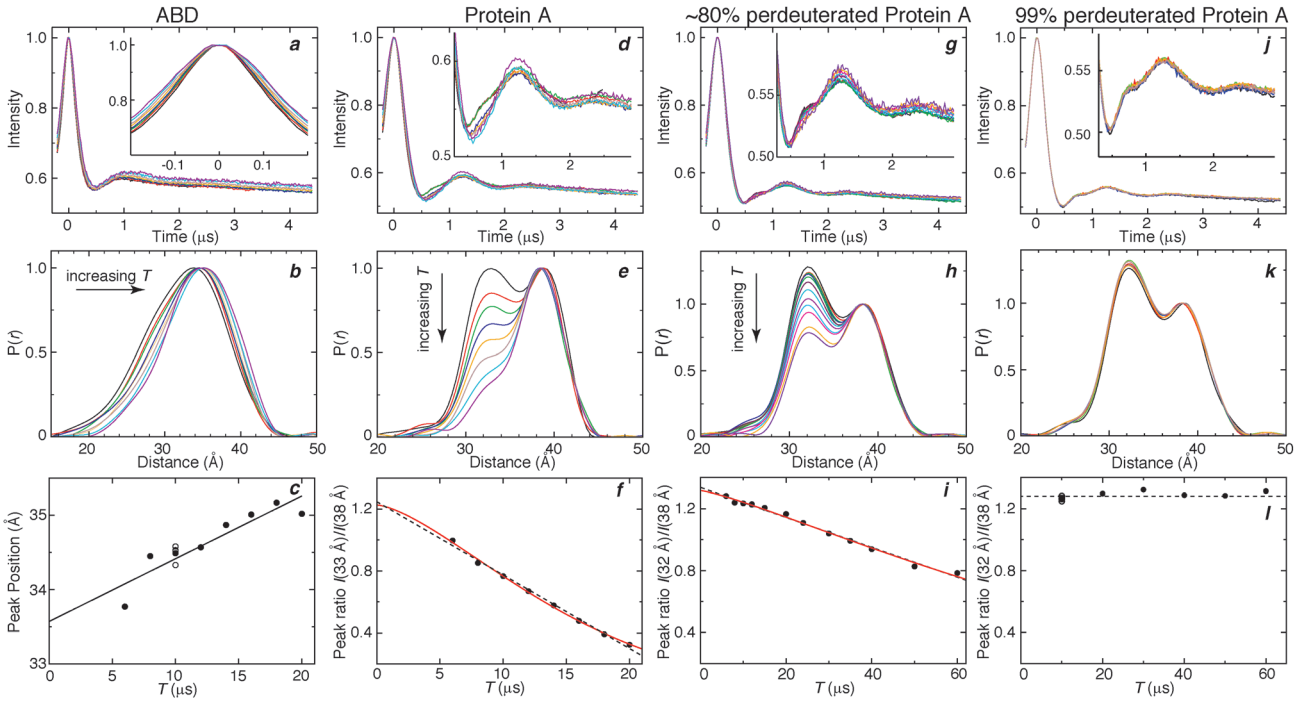

Figure 2. Effect of the total time T of the second spin-echo period on Q-band DEER measurements for ABD and Protein A. The raw DEER echo curves are shown in the top row (a, d, g, j; see Figure S2 for background-subtracted curves), P(r) distributions generated by Tikhonov regularization (TR) with DeerAnalysis 2013[6] in the middle row (b, e, h, k), and plots of the dependence of either the maximum of the P(r) distribution for protonated ABD (c) or the ratio of peaks heights for the two maxima of the bimodal P(r) distribution for Protein A (f, i, l) on T in the bottom row. a–c) Protonated ABD domain: $T=6$ (black), 8 (red), 10 (green), 12 (blue), 14 (orange), 16 (brown), 18 (turquoise), and 20 μs (violet). The original DeerAnalysis[6] results were normalized to yield a peak maximum of 1 and are presented in (b). The position of the P(r) maximum was determined by fitting the seven points around the TR-derived distribution maximum to a quadratic equation and setting the first derivative of this function to zero. The four open circles at $T=10$ μs in (c) and (l) represent the results of four different experiments and were included to provide a measure of data reproducibility (see Figure S4). d–f) Protonated Protein A: $T=6$ (black), 8 (red), 10 (green), 12 (blue), 14 (orange), 16 (brown), 18 (turquoise), and 20 μs (violet). g–i) Approximately 80% perdeuterated Protein A labeled with [D_{15}]MTSL: $T=6$ (black), 8 (red), 10 (green), 12 (blue), 15 (dark green), 20 (maroon), 24 (turquoise), 30 (violet), 35 (cyan), 40 (magenta), 50 (orange), and 60 μs (indigo). k–l) 99% perdeuterated Protein A labeled with [D_{15}]MTSL: $T=10$ (black), 20 (red), 30 (green), 40 (blue), 50 (orange), and 60 μs (brown). The original DeerAnalysis TR results for Protein A were normalized to the longer-distance peak. The shot repetition times are 3.1 ms for ABD, 3.8 ms for protonated Protein A, and 4.6 ms for perdeuterated (ca. 80% and 99%) Protein A. The T dependence of the ratio of peak heights in the P(r) distribution for Protein A (f, i, l) was fit to either a straight line (black dashes) or in the case of protonated and ca. 80% perdeuterated Protein A to the ratio of two stretched exponentials (red curves; see main text). The parameters for the latter are $a=1.2$, $T_{m1}=9$ μs, $T_{m2}=13$ μs, and $x_1=x_2=1.5$ for protonated protein A; those for approximately 80% perdeuterated Protein A are $a=1.3$, $T_{m1}=25$ μs, $T_{m2}=29$ μs, and $x_1=x_2=1.25$. These T_m values are close to those obtained from a two-pulse Hahn spin-echo experiment (see Figure S5). DEER data were collected at Q-band (33.8 GHz) on a Bruker E-580 spectrometer equipped with a 150 W traveling-wave tube amplifier and a model ER5107D2 probe. Experiments were carried out with 8 ns pump (ELDOR) π pulses, and 12 ns π/2 and 24 ns π observe pulses, with a 95 MHz frequency difference between pump and observe pulses. The pump frequency was centered at the field spectrum maximum. The τ_1 value for the first-echo-period time (see Figure 1) of 400 ns was incremented eight times in 16 ns increments to average ^{2}H modulation. The pump pulse was incremented in 16 ns steps. The sample temperature was 48 K. The bandwidth of the overcoupled resonator was approximately 120 MHz. All samples were placed in quartz tubes with a 1.1 mm internal diameter (Wilmad WG-221T-RB) and flash frozen in liquid N_2. DEER curves were sampled to approximately 4.4 μs for the various τ_2 delays, as collecting more data was not deemed to be a good use of instrument time. A cutoff of 4.3 μs was used to process all data. Cutoffs of 2.3 and 3.3 μs were used for the τ_2 (T/2) = 3 and 4 μs data sets, respectively. Acquisition was not carried out over the full τ_2 range in the latter cases because of a persistent "2 + 1" echo perturbation of the DEER curve at a time of about τ_1 from the final observe π pulse. Total data collection times varied from about 1 to about 22 h, with the goal of achieving comparable signal/noise ratios for all spectra. The pulse gate time used for echo integration was 30–34 ns. The TR parameter α (100 for protein ABD and 10 for Protein A) was determined by examination of the relevant L-curves[6] (see Figure S3). The DeerAnalysis[6] homogeneous model with a dimension of 3 (i.e. exponential background) was used to fit and automatically subtract the background; the resulting Pake patterns indicate good separation of inter- and intramolecular contributions.

tion) is 0.8 Å (Figure 2c). Although a systematic error of less than 1 Å is not particularly large and in most instances is unlikely to affect most structural conclusions, it could have a profound impact on the interpretation of any structural analyses involving small or subtle conformational changes.

For protonated Protein A, the ratio [I(33 Å)/I(38 Å)] of the peak heights at the two maxima of the P(r) distribution is reduced from an extrapolated value of 1.25 at $T=0$ to an experimental value of 0.77 at $T=10$ μs (the shortest value required to obtain a reliable P(r) distribution from the DEER data), which corresponds to an approximately 40% reduction in the apparent population of the state(s) with the shorter distance (Figure 2 f). This difference is clearly of considerable practical significance when attempting to draw conclusions on

the populations of various states from DEER data. Upon perdeuteration of the protein to approximately 80%, the effect is much smaller (ca. 7%). In this case, the ratio is reduced from an extrapolated value of 1.34 at $T=0$ to an experimental value of 1.23 at $T=10\,\mu s$ (Figure 2i). Although small, this difference could still be significant for modeling purposes; however, the relative peak intensities are clearly much more accurate than for the protonated protein. Definitive results are obtained with 99% perdeuteration of the protein, which yields an average peak ratio over all T values of 1.28 ± 0.02 (Figure 2l).

Our results yield several important conclusions of practical significance for the application of DEER in structural biology and biophysics. DEER-derived P(r) distributions of intramolecular distances between pairs of spin labels in protonated proteins are likely to be distorted owing to the dependence of the phase memory relaxation time T_m of a spin label on its local environment. A distribution of rotamers (from the spin label and/or neighboring protein side chains) all but guarantees differences in the local environment at different spin-labeling sites. Preliminary data from three other systems studied in our laboratory displayed T dependencies similar in magnitude to that reported for ABD here. The magnitude of this dependence can only be gauged by performing DEER experiments over a range of T values which becomes especially important when attempting to derive conclusions regarding relative populations of different states. Such an approach applies not only to the observation of resolved rotameric states in the P(r) distribution, as in the case of Protein A, but also to the study of global conformational transitions in proteins, such as the dissection of relative populations of open and closed states where the local environments of the spin labels are likely to be different in the various conformers. Our results also suggest two simple practical solutions for obtaining accurate P(r) distance distributions from DEER data: First, wherever possible the protein and solvent matrix should be deuterated to attenuate any distortion due to differential phase memory relaxation times (in addition to the usual advantages of longer T_m values, which yield better signal-to-noise ratios and allow longer distances to be probed).[11] At 99% deuteration, the effect is essentially abolished. Second, DEER data on protonated or partially deuterated proteins should be recorded at multiple second-echo-period times T to permit extrapolation of the results to $T=0$. Finally, when using DEER to assess small structural changes for any given perturbation, especially on protonated proteins, one must be careful to consider whether the observed variations in distance between spin-label pairs are due to an actual change in distance or a change in relative relaxation parameters that might ultimately distort the measured P(r) distance distribution.

Keywords: deuteration · distance distributions · EPR spectroscopy · phase memory time · proteins

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 5336–5339
Angew. Chem. **2015**, *127*, 5426–5429

[1] a) A. D. Milov, K. M. Salikhov, M. D. Shirov, *Fiz. Tverd. Tela* **1981**, *23*, 975–982; b) A. D. Milov, A. B. Ponomarev, Y. D. Tsvetkov, *Chem. Phys. Lett.* **1984**, *110*, 67–72; c) R. G. Larsen, D. J. Singel, *J. Chem. Phys.* **1993**, *98*, 5134–5146; d) M. Pannier, S. Veit, A. Godt, G. Jeschke, H. W. Spiess, *J. Magn. Reson.* **2000**, *142*, 331–340.

[2] a) Y. K. Shin, C. Levinthal, F. Levinthal, W. L. Hubbell, *Science* **1993**, *259*, 960–963; b) W. L. Hubbell, C. Altenbach, *Curr. Opin. Struct. Biol.* **1994**, *4*, 566–573; c) G. E. Fanucci, D. S. Cafiso, *Curr. Opin. Struct. Biol.* **2006**, *16*, 644–653; d) G. Jeschke, Y. Polyhach, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895–1910; e) C. Altenbach, A. K. Kusnetzow, O. P. Ernst, K. P. Hofman, W. L. Hubbell, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7439–7444; f) G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419–446.

[3] Y. Polyhach, E. Bordignon, R. Tschaggelar, S. Gandra, A. Godt, G. Jeschke, *Phys. Chem. Chem. Phys.* **2012**, *14*, 10762–10773.

[4] M. U. Johansson, M. de Château, M. Wikström, S. Forsén, T. Drakenberg, L. Björck, *J. Mol. Biol.* **1997**, *266*, 859–865.

[5] H. Gouda, H. Torigoe, A. Saito, M. Sato, Y. Arata, I. Shimada, *Biochemistry* **1992**, *31*, 9665–9672.

[6] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Appl. Magn. Reson.* **2006**, *30*, 473–498.

[7] a) M. Lindgren, G. R. Eaton, S. S. Eaton, B. H. Jonsson, P. Hammarström, M. Svensson, U. Carlsson, *J. Chem. Soc. Perkin Trans. 2* **1997**, 2549–2554; b) M. Huber, M. Lindgren, P. Hammarstrom, L. G. Martensson, U. Carlsson, G. R. Eaton, S. S. Eaton, *Biophys. Chem.* **2001**, *94*, 245–256.

[8] a) A. Shekhtman, R. Ghose, M. Goger, D. Cowburn, *FEBS Lett.* **2002**, *524*, 177–182; b) A. L. Hansen, P. Lundström, A. Velyvis, L. E. Kay, *J. Am. Chem. Soc.* **2012**, *134*, 3178–3189.

[9] G. Jeschke, Y. Polyhach, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895–1910.

[10] Y. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356–2366.

[11] R. Ward, A. Bowman, E. Sozudogru, H. El-Mkami, T. Owen-Hughes, D. Norman, *J. Magn. Reson.* **2010**, *207*, 164–167.

Received: January 22, 2015
Revised: February 10, 2015
Published online: March 10, 2015

