Protein Dynamics

Multi-Timescale Conformational Dynamics of the SH3 Domain of CD2-Associated Protein using NMR Spectroscopy and Accelerated Molecular Dynamics**

*Loïc Salmon, Levi Pierce, Alexander Grimm, Jose-Luis Ortega Roldan, Luca Mollica, Malene Ringkjøbing Jensen, Nico van Nuland, Phineus R. L. Markwick, J. Andrew McCammon, and Martin Blackledge**

The determination of three-dimensional atomic resolution protein structure underpins our understanding of many biological processes, providing descriptions of the molecular basis of physiologically important interactions between biochemically active molecules. However, a complete understanding of the relationship between biological activity and molecular conformation also requires a description of the thermally accessible potential energy landscape intrinsic to a protein at its physiological temperature.[1,2]

NMR ^{15}N and ^{13}C spin relaxation experiments are routinely applied for the characterization of rapid motions occurring in proteins on the pico- to nanosecond timescale.[3] However dynamics occurring on longer timescales, in the nano- to millisecond range, are potentially of greater functional interest, because many biologically important processes, such as enzyme catalysis, signal transduction, ligand binding and allosteric regulation are expected to occur on these timescales.[4,5] Although longer timescale molecular dynamics (MD) simulations are becoming more accessible,[6] there is a notable lack of experimental data against which the accuracy of such predictions can be gauged. The precise elucidation of the nature, amplitude and timescale of intrinsic motions occurring in proteins in solution therefore remains a fundamentally important challenge for structural and molecular biologists.

In response to this challenge, there has been considerable activity over the last decade, exploiting the exquisite sensitivity of residual dipolar couplings (RDCs), measured in weakly aligned proteins, to determine the dynamic averaging properties of internuclear bond vectors.[7–10] It has been demonstrated that the measurement of a sufficient number of RDCs in differently aligning media allows the accurate determination of the motional properties of the protein backbone. Two generic approaches to the interpretation of the experimental data can be distinguished—either direct analysis of the RDCs to extract averaged spherical harmonic terms describing the angular averaging of the internuclear bonds,[11–14] or exploiting MD simulation (with or without restraints) to reproduce the motional amplitudes and modes in terms of an explicit conformational ensemble.[15–21]

Such RDC-based studies allowed the identification of nano- to millisecond motions that were localized predominantly in the molecular recognition sites of the small proteins ubiquitin (Ub) and GB3.[18,22,25,26] These discoveries have prompted speculation about the role that intrinsic motions play in molecular recognition, in particular concerning selection of distinct conformers from an existing equilibrium. We note that consideration of the importance of conformational selection for the promiscuity of Ub binding has until now largely neglected the role played by intrinsic dynamics of the partner proteins.

Application of the three dimensional Gaussian Axial Fluctuation (3DGAF) model to Ub[22] resulted in dynamic modes and amplitudes that were in good agreement with those determined by comparison to restraint-free accelerated MD (AMD) simulation.[23,24] Nevertheless the nature of slow dynamics occurring in folded proteins, a question of fundamental importance for the understanding of a vast range of biochemical processes, remains the subject of much debate. Until now only two structurally homologous α–β proteins have been studied in sufficient detail to allow a quantitative description of slow dynamics from RDCs, severely limiting our understanding of the general nature of these observations. Here, we analyze the conformational dynamics occurring on timescales from picoseconds to milliseconds in a small β-barrel protein. We simultaneously determine the three-dimensional structure and backbone dynamics of the third SH3 domain of CD2-associated protein (CD2AP)[27–30] (SH3C) directly from RDCs using the 3DGAF approach,

[*] Dr. L. Salmon, A. Grimm, Dr. L. Mollica, Dr. M. R. Jensen, Dr. M. Blackledge
Protein Dynamics and Flexibility, Institut de Biologie Structurale Jean-Pierre Ebel, CNRS-CEA-UJF, UMR 5075
41 rue Jules Horowitz, 38027 Grenoble Cedex (France)
E-mail: martin.blackledge@ibs.fr

Dr. L. Pierce, Dr. P. R. L. Markwick, Prof. J. A. McCammon
Dept of Chemistry and Biochemistry, UCSD, San Diego
Howard Hughes Medical Institute
9500 Gilman Dr., La Jolla, CA (USA)

Prof. N. van Nuland
Structural Biology Brussels, Vrije Universiteit Brussel
Pleinlaan, Brussels (Belgium)

Dr. J.-L. Ortega Roldan
Dept of Biochemistry, University of Oxford
South Parks Road, Oxford (UK)

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

and compare the fitted motions to those present in a series of restraint-free AMD simulations. The analyses provide a comprehensive and convergent description of multi-timescale dynamics of this protein, which is a physiological partner of Ub, identifying significant slow conformational fluctuations in the interaction site of the protein.

A total of 1912 RDCs ($^1D_{NH}$, $^2D_{CHN}$ and $^1D_{CaC}$) were extracted from ^{15}N- and ^{13}C-labeled SH3C aligned in 15 alignment media (see Supporting Information). The SECONDA algorithm[31]—a principal component analysis of the RDC covariance matrix—was used to identify data sets that are self-consistent, and therefore show no evidence of perturbing interaction with any medium. Following this, 1358 RDCs from 10 media were retained in the final data set: bicelles,[32] bicelles doped with CTAB (cetyltrimethyl-ammonium bromide), bicelles doped with SDS (sodium dodecylsulfate), polyacrylamide gel,[33] PEG (polyethylene glycol)/hexanol,[34] bacteriophage,[35] bacteriophage in the presence of high salt, purple membrane[36] all measured at pH 6.0 and 35°C. Correlations between $^1D_{NH}$ RDCs measured in these media are shown in Figure 1. Three pairs of media resulted in highly correlated alignment, while the seven others present lower, or negligible correlation. The high level of reproduction measured between the three correlated pairs of RDCs testifies to the average precision of the measurement but these data do not provide additional information.
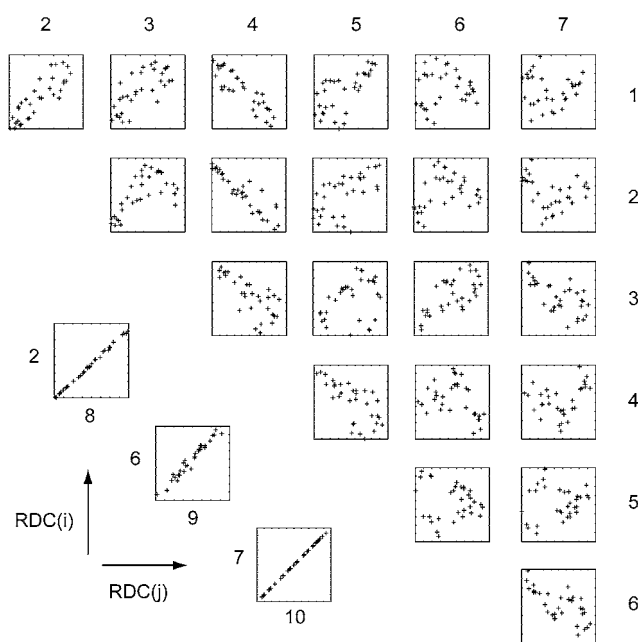
Figure 1. Correlation of the experimental N—^{N}H RDCs from SH3C aligned in 10 alignment media (full details in Supporting Information). RDCs are normalized to maximum and minimum experimental values.

The Dynamic-Meccano approach[37,38] exploits the 1DGAF model combined with the Meccano algorithm,[39] to determine the average solution structure that explicitly accounts for simultaneously determined backbone dynamics. This ab initio structure determination, using only RDCs to construct the peptide chain from GAF-averaged orientations

of peptide planes, compares very closely to the nOe and RDC-based NMR structure of SH3C and the X-ray structure of the SH3C from Cin85 (Figure 2).
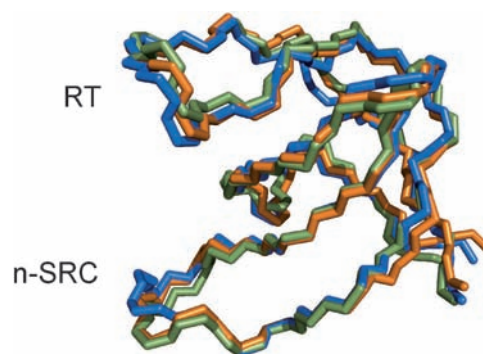

Figure 2. Dynamic Meccano structure of SH3C determined using only RDCs (green) compared to the NMR (orange) structure of SH3C from CD2AP (2kro, backbone rmsd 1.0 Å) and Cin85 (blue) (2ydl, rmsd 0.85 Å), respectively.

The 3DGAF approach that models RDCs in terms of diffusive motions around three orthogonal axes attached to each plane[40] has been shown to quantitatively describe protein conformational dynamics.[22] The approach is independent of any structural model, because the average conformation of each peptide plane is determined simultaneously to its dynamic modes. The use of an anisotropic motional model, in combination with RDCs sampling different directions in the peptide plane, allows for an accurate ab initio determination of the alignment tensors.

The 3DGAF analysis was applied as previously described (Supporting Information).[22] Calculation of randomly selected RDCs (10%) that were not included in the analysis testifies to the predictive value of the approach (Figure 3 a–c) and provides significantly better reproduction of these data than a static structure determination. The 3DGAF motions are dominated by the γ-mode, representing fluctuations about the C^α–C^α axis of each peptide plane of $(14.8 \pm 6.5)°$ (see Supporting Information). The N—^{N}H bond order parameters determined from the 3DGAF analysis (S^2_{3DGAF}) are shown in Figure 3 d, in comparison to order parameters derived from ^{15}N relaxation (S^2_{rel}). The distribution of motions occurring on timescales between nano- and milliseconds is manifest as differences between S^2_{3DGAF} and S^2_{rel}, and reveals a similar pattern to that determined in Ub and GB3.[22,26] Such motions are not ubiquitous throughout the protein. They are indeed found to be negligible in secondary structural elements, whereas in loop regions, in particular the n-SRC (35–43) and RT loops (18–20), significant additional slower motions are apparent. We note that these regions mediate physiological interaction with Ub.[28]

In parallel, we have applied the AMD approach[41] to the interpretation of the experimental data from SH3C. AMD does not use an experimental pseudo-potential, and is therefore restraint-free. Acceleration is achieved by scaling the potential energy landscape by a constant factor (α), for all terms below a given threshold, thereby enhancing the escape
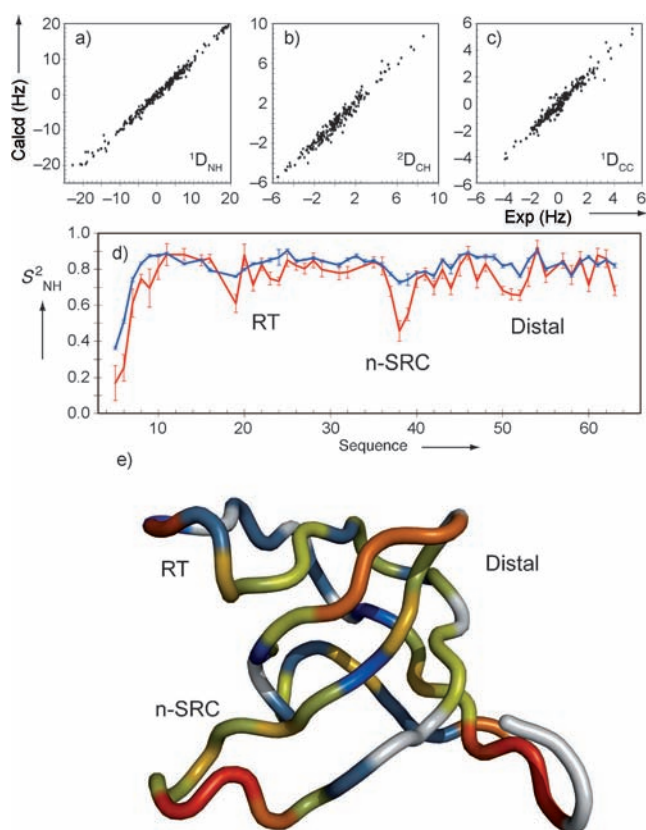
Figure 3. 3DGAF analysis of RDC data. a–c) Reproduction of 10% randomly selected experimental RDCs that were not used in the 3DGAF analysis (10 calculations are shown). d) 3DGAF N—^{N}H order parameters (S^2_{3DGAF}, red) compared to fast motional order parameters (S^2_{rel}, blue). Error bars from noise-based Monte Carlo simulation. e) S^2_{3DGAF} shown on a ribbon representation of the Dynamic Meccano structure (from red (<0.7) through yellow (<0.8) to blue (<0.9); white, no value).

rate between low-energy conformational sub-states. On increasing the level of acceleration, the simulation probes more conformational space. Trajectories are re-weighted to obtain a canonical Boltzmann distribution, and a series of short standard MD simulations are seeded from this distribution. The appropriate level of acceleration, and therefore conformational space, is directly estimated by comparing the experimental RDCs to predicted values from ensembles calculated at different levels of acceleration. R factors—measuring the quality of the reproduction of each RDC type—are shown in Figure 4 together with the N—^{N}H order parameters calculated for each acceleration level (S^2_{AMD}). The optimal level of acceleration reproducibly lies around 160 kcal mol^{-1}, a value similar to that found for Ub under similar conditions.[24]

A representative ensemble of structures is shown in Figure 5, together with a comparison of the optimal S^2_{AMD} and S^2_{3DGAF} values. The similarity between the results derived from the two very different approaches, in one case fitting motional modes and amplitudes to the experimental data using mathematical models, and in the other case comparing to restraint free MD simulation, is striking, and substantiates assumptions implicit to both approaches. We have compared
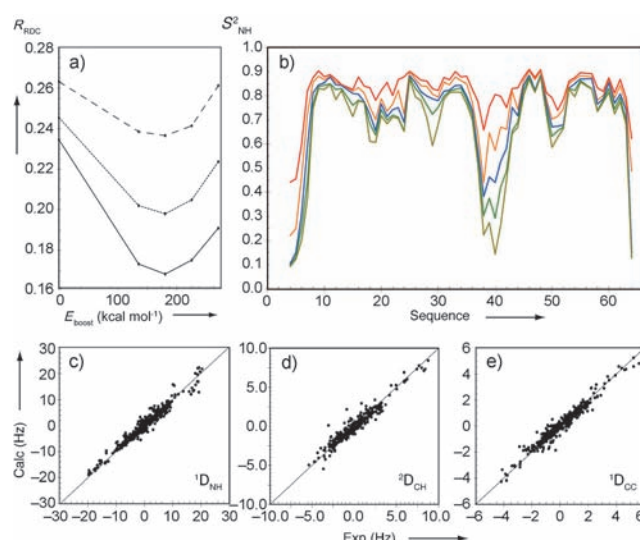
Figure 4. AMD analysis of experimental RDCs. a) Total Q values shown for $^1D_{NH}$ (solid line), $^2D_{CHN}$ (dashed line), and $^1D_{CaC}$ (dotted line). b) NH bond order parameters calculated at different acceleration levels (red, standard MD; orange to green, increasing acceleration). c–e) Reproduction of experimental RDCs at the optimal acceleration level.
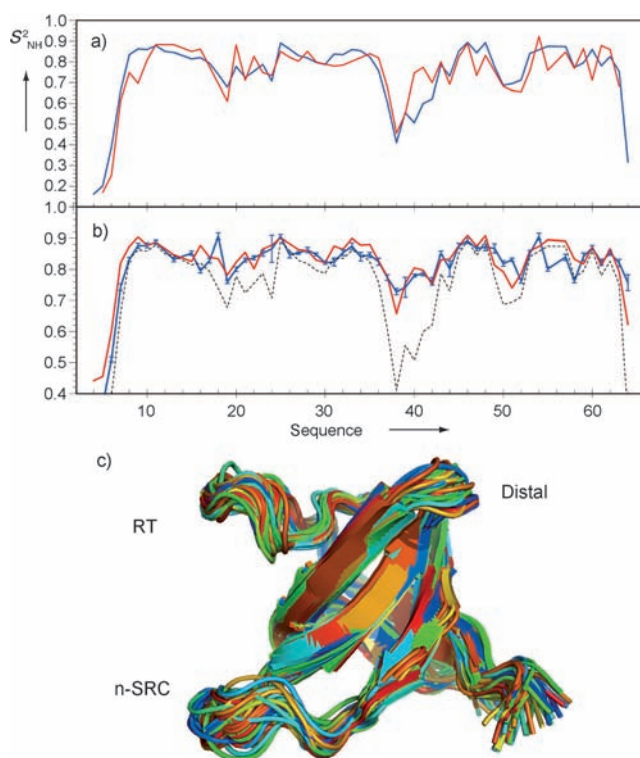
Figure 5. a) AMD-derived (S^2_{AMD}, blue) and 3DGAF N—^{N}H bond vector order parameters (S^2_{3DGAF}, red). b) Fast motional experimental order parameters (S^2_{rel}, blue) compared to S^2 values calculated from fast motions occurring in different sub-states sampled by the AMD. S^2_{AMD} values are shown for comparison (dotted lines). c) Representative AMD ensemble.

S^2_{rel} values with fast motional order parameters derived from the short MD simulations performed from the ensemble of

AMD-derived sub-states (Figure 5 b). The distribution of fast motions calculated over the entire potential energy surface is quantitatively closer to experimental S^2_{rel} values than any of a series of standard MD simulations starting from the RDC structure (Supporting Information). This provides evidence for the dependence of fast motions on the rugosity of the potential energy landscape sampled on longer timescales.

In conclusion we have measured an extensive set of RDCs in the third SH3 domain of CD2AP, a small β-barrel protein. Independent analyses of RDCs, using analytical fitting of the mean orientation and associated modes and amplitudes of each peptide plane, or comparison with restraint-free AMD simulation, resulted in a comprehensive and remarkably convergent description of multi-timescale dynamics in this protein. Extensive cross-validation procedures were used in both cases to guarantee the self-consistency of each analysis. The similarity of order parameters determined using spin relaxation and RDCs throughout the β-sheet demonstrates the absence of significant conformational fluctuations over timescales spanning up to six orders of magnitude (ns–ms). Importantly these methods both identify large-amplitude slow motions that are localized in the functionally important n-SRC loop that mediates the interaction of SH3C with Ub. These motions are invisible to other biophysical techniques, including NMR spin relaxation. Further studies of the complex between SH3C and Ub will be required to provide additional insight into the role of these slow dynamics, that are localized in the interaction sites of both proteins, in the mechanism of intermolecular recognition.

[1] H. Frauenfelder, S. G. Sligar, P. G. Wolynes, *Science* **1991**, *254*, 1598–1603.

[2] K. Henzler-Wildman, M. Lei, V. Thai, S. Kerns, M. Karplus, D. Kern, *Nature* **2007**, *450*, 913-U27.

[3] A. Palmer, *Curr. Opin. Struct. Biol.* **1997**, *7*, 732–737.

[4] E. Eisenmesser, O. Millet, W. Labeikovsky, D. Korzhnev, M. Wolf-Watz, D. Bosco, J. Skalicky, L. Kay, D. Kern, *Nature* **2005**, *438*, 117–121.

[5] A. J. Baldwin, L. E. Kay, *Nat. Chem. Biol.* **2009**, *5*, 808–814.

[6] J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror, D. E. Shaw, *Curr. Opin. Struct. Biol.* **2009**, *19*, 120–127.

[7] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Nat. Struct. Biol* **1997**, *4*, 292–297.

[8] J. H. Prestegard, H. M. al-Hashimi, J. R. Tolman, *Q. Rev. Biophys.* **2000**, *33*, 371–424.

[9] A. Bax, A. Grishaev, *Curr. Opin. Struct. Biol.* **2005**, *15*, 563–570.

[10] M. Blackledge, *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 23–61.

[11] J. Meiler, J. J. Prompers, W. Peti, C. Griesinger, R. Brüschweiler, *J. Am. Chem. Soc.* **2001**, *123*, 6098–6107.

[12] N. A. Lakomek, T. Carlomagno, S. Becker, C. Griesinger, J. Meiler, *J. Biomol. NMR* **2006**, *34*, 101–115.

[13] K. Briggman, J. Tolman, *J. Am. Chem. Soc.* **2003**, *125*, 10164–10165.

[14] P. Bernadó, M. Blackledge, *J. Am. Chem. Soc.* **2004**, *126*, 4907–4920.

[15] G. M. Clore, C. D. Schwieters, *J. Am. Chem. Soc.* **2004**, *126*, 2923–2938.

[16] K. Lindorff-Larsen, R. Best, M. DePristo, C. Dobson, M. Vendruscolo, *Nature* **2005**, *433*, 128–132.

[17] B. Hess, R. M. Scheek, *J. Magn. Reson.* **2003**, *164*, 19–27.

[18] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471–1475.

[19] T. S. Ulmer, B. E. Ramirez, F. Delaglio, A. Bax, *J. Am. Chem. Soc.* **2003**, *125*, 9179–9191.

[20] S. A. Showalter, R. Brüschweiler, *J. Am. Chem. Soc.* **2007**, *129*, 4158–4159.

[21] R. Bryn Fenwick, S. Esteban-Martin, X. Salvatella, *J. Phys. Chem. Lett.* **2010**, *1*, 3438–3441.

[22] L. Salmon, G. Bouvignies, P. Markwick, N. Lakomek, S. Showalter, D.-W. Li, K. Walter, C. Griesinger, R. Brüschweiler, M. Blackledge, *Angew. Chem.* **2009**, *121*, 4218–4221; *Angew. Chem. Int. Ed.* **2009**, *48*, 4154–4157.

[23] L. Salmon, G. Bouvignies, P. Markwick, M. Blackledge, *Biochemistry* **2011**, *50*, 2735–2747.

[24] P. R. L. Markwick, G. Bouvignies, L. Salmon, J. A. McCammon, M. Nilges, M. Blackledge, *J. Am. Chem. Soc.* **2009**, *131*, 16968–16975.

[25] P. Bernadó, M. Blackledge, *J. Am. Chem. Soc.* **2004**, *126*, 7760–7761.

[26] G. Bouvignies, P. Bernadó, S. Meier, K. Cho, S. Grzesiek, R. Brüschweiler, M. Blackledge, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13885–13890.

[27] J. L. Ortega Roldan, M. Romero, A. Ora, E. Ab, O. Mayorga, A. Azuaga, N. van Nuland, *J. Biomol. NMR* **2007**, *39*, 331–336.

[28] J. L. Ortega-Roldan, M. R. Jensen, B. Brutscher, A. I. Azuaga, M. Blackledge, N. A. J. van Nuland, *Nucleic Acids Res.* **2009**, *37*, e70.

[29] J. L. Ortega Roldan, M. Blackledge, N. A. J. van Nuland, A. I. Azuaga, *J. Biomol. NMR* **2011**, *50*, 103–117.

[30] L. Salmon, J.-L. Ortega Roldan, E. Lescop, A. Licinio, N. van Nuland, M. R. Jensen, M. Blackledge, *Angew. Chem.* **2011**, *123*, 3839–3843; *Angew. Chem. Int. Ed.* **2011**, *50*, 3755–3759.

[31] J. Hus, R. Bruschweiler, *J. Biomol. NMR* **2002**, *24*, 123–132.

[32] N. Tjandra, A. Bax, *Science* **1997**, *278*, 1111–1114.

[33] H. Sass, G. Musco, S. Stahl, P. Wingfield, S. Grzesiek, *J. Biomol. NMR* **2000**, *18*, 303–309.

[34] M. Rückert, G. Otting, *J. Am. Chem. Soc.* **2000**, *122*, 7793–7797.

[35] M. Hansen, L. Mueller, A. Pardi, *Nat. Struct. Biol.* **1998**, *5*, 1065–1074.

[36] J. Sass, F. Cordier, A. Hoffmann, A. Cousin, J. Omichinski, H. Lowen, S. Grzesiek, *J. Am. Chem. Soc.* **1999**, *121*, 2047–2055.

[37] G. Bouvignies, P. R. L. Markwick, M. Blackledge, *Proteins Struct. Funct. Genet.* **2008**, *71*, 353–363.

[38] G. Bouvignies, S. Meier, S. Grzesiek, M. Blackledge, *Angew. Chem.* **2006**, *118*, 8346–8349; *Angew. Chem. Int. Ed.* **2006**, *45*, 8166–8169.

[39] J. C. Hus, D. Marion, M. Blackledge, *J. Am. Chem. Soc.* **2001**, *123*, 1541–1542.

[40] T. Bremi, R. Bruschweiler, *J. Am. Chem. Soc.* **1997**, *119*, 6672–6673.

[41] D. Hamelberg, J. Mongan, J. A. McCammon, *J. Chem. Phys.* **2004**, *120*, 11919–11929.

[42] P. R. L. Markwick, G. Bouvignies, M. Blackledge, *J. Am. Chem. Soc.* **2007**, *129*, 4724–4730.