Protein Folding

Deutsche Ausgabe: DOI: 10.1002/ange.201606613
Internationale Ausgabe: DOI: 10.1002/anie.201606613

Conformational Dynamics of apo-GlnBP Revealed by Experimental and Computational Analysis

Yitao Feng[+], Lu Zhang[+], Shaowen Wu, Zhijun Liu, Xin Gao, Xu Zhang, Maili Liu, Jianwei Liu, Xuhui Huang,* and Wenning Wang**

Abstract: *The glutamine binding protein (GlnBP) binds* L*-glutamine and cooperates with its cognate transporters during glutamine uptake. Crystal structure analysis has revealed an open and a closed conformation for apo- and holo-GlnBP, respectively. However, the detailed conformational dynamics have remained unclear. Herein, we combined NMR spectroscopy, MD simulations, and single-molecule FRET techniques to decipher the conformational dynamics of apo-GlnBP. The NMR residual dipolar couplings of apo-GlnBP were in good agreement with a MD-derived structure ensemble consisting of four metastable states. The open and closed conformations are the two major states. This four-state model was further validated by smFRET experiments and suggests the conformational selection mechanism in ligand recognition of GlnBP.*

Molecular recognition plays an important role in many biological processes.[1] Protein conformational dynamics are crucial in molecular recognition, during which ligand binding usually causes local structure rearrangements or large-scale interdomain movements. A typical system for studying the relationship between protein conformational dynamics and ligand binding consists of the periplasmic binding proteins (PBPs). PBPs belong to the ATP-binding cassette (ABC) transporter system and are responsible for substrate uptake.[2] The substrate-binding sites of PBPs are located at the interface between two domains, and proteins can undergo domain reorientation from the open conformation to the closed one. For many PBPs, crystallographic studies have identified two conformations: a ligand-free (or apo) open state and a ligand-bound (or holo) closed state. For a few PBPs, apo closed conformations have also been identified during crystal structure analysis.[3] NMR paramagnetic relaxation enhancement (PRE) studies[4] and MD simulations[5] have revealed a semi-closed state in the ligand recognition process of maltose binding protein (MBP). Furthermore, Markov state model (MSM) based MD simulations have elucidated the binding mechanism for two other PBPs.[6] These results suggest that the conformational dynamics of PBPs are more complicated than a two-state model and that a single technique might not be sufficient for deciphering the underlying conformational space of PBPs.

E. coli GlnBP is a PBP that specifically binds L-glutamine. Previous crystallography studies identified two conformational states of GlnBP, an apo open state[7] (Figure 1 A) and a holo closed state[8] (see the Supporting Information, Figure S1 A). Discrepancies exist in previous experimental dynamics studies of GlnBP. Phosphorescence spectroscopy studies detected interdomain dynamics in both apo- and holo-GlnBP in solution,[9] but NMR PRE studies did not specifically support the existence of the closed conformation for apo-GlnBP.[10] Therefore, one of the problems under debate is whether apo-GlnBP can also adopt the closed conformation or multiple conformations in solution. Detailed conformational dynamics studies are not only crucial for understanding the ligand-binding mechanism, but are also important for the protein design of GlnBP as a biosensor.[11]

Herein, we combined NMR residual dipolar coupling (RDC) analysis, MD simulations, and single-molecule Förster resonance energy transfer (smFRET) methods to explore the conformational dynamics of apo-GlnBP. RDC measurements are a powerful method for studying protein dynamics.[12]

[*] Y. Feng,[+] S. Wu, J. Liu, W. Wang
Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Department of Chemistry, and Institutes of Biomedical Sciences, Fudan University, Shanghai (P.R. China)
E-mail: jwliu@fudan.edu.cn
wnwang@fudan.edu.cn
L. Zhang,[+] X. Huang
Department of Chemistry
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon (Hong Kong)
E-mail: xuhuihuang@ust.hk
Z. Liu
National Center for Protein Science
Shanghai Institute of Biochemistry and Cell Biology
Chinese Academy of Sciences, Shanghai (China)
X. Gao
King Abdullah University of Science and Technology (KAUST)
Computational Bioscience Research Center (CBRC)
Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, Thuwal, 23955 (Saudi Arabia)
X. Zhang, M. Liu
Key Laboratory of Magnetic and Resonance in Biological Systems
State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Centre for Magnetic Resonance
Wuhan Institute of Physics and Mathematics
Chinese Academy of Sciences, Wuhan (China)
X. Huang
Division of Biomedical Engineering
Center of Systems Biology and Human Health
Institute for Advance Study and School of Science
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon (Hong Kong)
[+] These authors contributed equally to this work.
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: http://dx.doi.org/10.1002/anie.201606613.

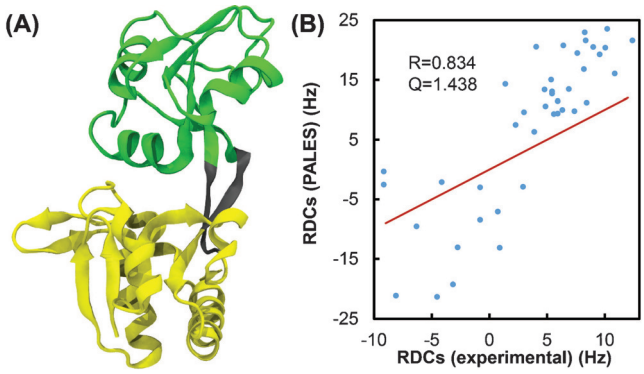

Figure 1. A) Ribbon diagram of the crystal structure of ligand-free GlnBP. The large domain, small domain, and the linker region are shown in yellow, green, and dark gray, respectively. B) Agreement between the apo-GlnBP experimental RDCs and the RDCs predicted based on the open crystal structure.

However, the interpretation of RDC data in the context of substantial conformational motions at atomic resolution is challenging. Here, we resorted to MD simulations and MSM analysis to construct an equilibrium structural ensemble to interpret the RDC data. Furthermore, smFRET studies were conducted for cross-validation. With this combination of techniques, we identified four metastable states co-existing for apo-GlnBP in solution. The two major states are similar to the open and closed states previously revealed by crystal structure analysis whereas the two minor states adopt distinct conformations with various interdomain arrangements.

We first performed spin relaxation measurements for apo-GlnBP. The backbone order parameters derived from a model-free analysis showed that some of the loop regions exhibit mobility at the picosecond time scale, but that the linker between the two domains does not show significant flexibility (Figure S2). Considering the large time-scale coverage of RDC experiments for protein dynamics, we performed backbone ^{1}H–^{15}N RDC measurements of apo-GlnBP (Table S1) and calculated the alignment tensor and RDCs for the apo open crystal structure from first principles using the steric prediction method.[13] The agreement between the predicted and experimental RDCs is very poor, with a Q value of 1.438 and an R value of 0.834 (Figure 1B), suggesting that the open conformation could not interpret the apo-GlnBP RDC data. We also calculated the RDCs for the closed crystal structure, but the agreement was even worse with Q values of 1.82 (Figure S1B). Therefore, neither the open nor the closed conformation alone could describe the conformational state of apo-GlnBP in solution.

We then combined the open and closed conformations to calculate RDCs by varying their populations. The optimal fitting to the apo-GlnBP RDC data needed 30–40% population of the closed conformation (Figure S3). However, two conformations were still not sufficient to interpret the apo-GlnBP RDC data as the lowest Q value was 1.215 (Figure S3).

As the two-state model could not describe the RDC data, it raised the possibility that there are additional conformations distinct from the open and closed crystal structures. To further elucidate the conformational ensemble of apo-GlnBP in solution, we performed large-scale atomic MD simulations (Figure S4). Principal component analysis revealed two dominant motions of GlnBP in MD simulations, namely opening and twisting, which correspond to the first and second eigenvectors, respectively (Figures S5 and S6). MSM analysis based on the MD conformations (Figure S7; see also the General Methods Section in the Supporting Information for model construction and validation) revealed four metastable conformations that include two previously unidentified intermediate conformations (states S1 and S4) in addition to the open and closed structures (states S2 and S3; Figure 2A,B). In particular, the S1 state has a slightly more open conformation than the open crystal structure (S2 state), and the relative orientation of the two domains differs from that in the crystal structure with a twisting angle ("open-twist" conformation). The S4 state has the cleft partially closed with respect to the open conformation without twisting ("pseudo-closed" conformation). It is noteworthy that the closed state (S3) has a remarkable population of about 40%, which is comparable to that of the open state (ca. 47%), indicating that GlnBP can close the cleft in the absence of substrate (Figure 2C).

We further showed that the RDCs calculated from first principles based on the MD conformations (see the Methods Section in the Supporting Information for details)[13,14] agree well with the experimental RDCs of apo-GlnBP (Q value: 0.499; R value: 0.912; Figure 2D). The results suggest that the MD conformational ensemble can well describe the conformational dynamics of apo-GlnBP in solution. Moreover, we also optimized the populations of the four macrostates against the experimental RDCs using a Bayesian statistical framework (see the Supporting Information for details).[15] The resulting state populations only deviate slightly from those directly obtained from the MD simulations (Figure 2C). Comparison with the experimental RDCs gives an improved Q value of 0.412 (Figure 2E). The improvement of the Q value is mainly due to the consideration of interdomain variability in GlnBP (Figures S8–S10).

The confidence level of the agreement between the MD simulations and the experimental RDC data was then explored. Figure 3 indicates that the Q factor remains below 0.43 for a range of populations. Except for the pseudo-closed state, the other three conformations have to have a population of at least 10% for the data to agree with the apo-GlnBP RDCs (Figure 3).

To further validate the existence of multiple conformational states, we performed smFRET experiments for apo-GlnBP. The large and small domains were labeled with a pair of fluorescent dyes (Figure 4A). The FRET efficiency distributions are broad (Figure 4B), suggesting that there are multiple conformational states in apo-GlnBP. The statistics of the FRET efficiencies against the donor and acceptor intensities indicate four regions that might be assigned to four conformational states (Figure S11). We then performed a hidden Markov model (HMM) analysis of the smFRET trajectories.[16] This analysis suggested that the optimal number of states for GlnBP systems is four (Figure S12). The four states were denoted as states 1, 2, 3, and 4 with increasing FRET efficiencies (Figures 4C and S11). State 2

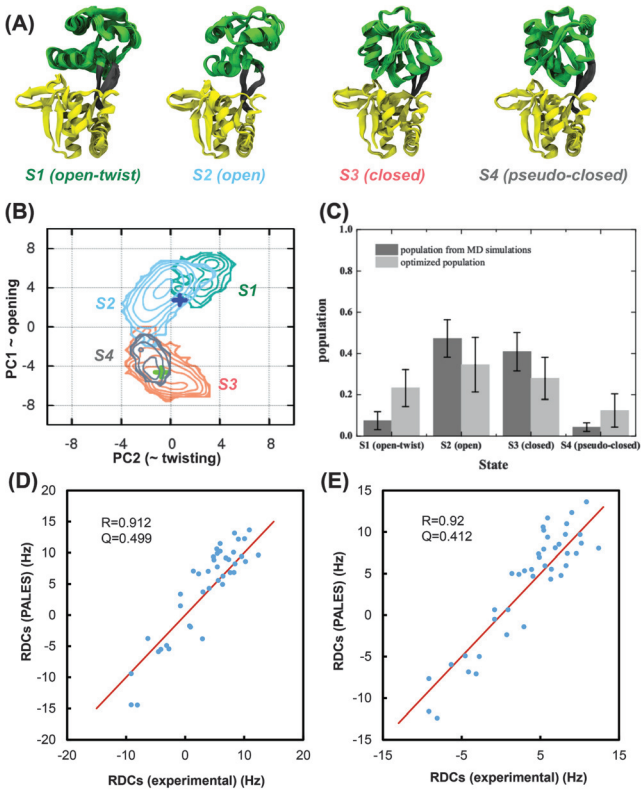

Figure 2. A) Representative conformations of the four macrostates derived from the MD simulations of apo-GlnBP. B) Projection of the four macrostates onto the conformational space described by the first (opening motion) and second (twisting motion) eigenvector of PCA analysis. C) Populations of the four macrostates derived from MD simulations and optimized populations (from the last 50000 optimization steps with a step size of 10) according to the experimental RDC data using a Bayesian analysis method.[15] D) Comparison between the experimental RDCs and calculated RDCs from MD simulations for apo-GlnBP. E) Comparison of the experimental RDCs and the calculated RDCs based on the averaged optimized populations through Bayesian analysis.

and state 3 are the major states with a total population of more than 75% whereas state 1 and state 4 are minor states with populations of around 10% each (Figure 4D). Thereby, it is reasonable to assign state 2 and state 3 to the open (S2) and closed (S3) conformations, respectively. As state 1 has the lowest FRET level and the longest distance between the two fluorophores, it was assigned to the open-twist state (S1).

As for the assignment of state 4, its relatively small population is consistent with the small probability of conformation S4. However, the smFRET analysis showed that state 4 has an even shorter fluorophore–fluorophore distance than state 3 (closed) while macrostate S4 (pseudo-closed) derived from the MD simulation has a slightly longer average Thr59–Thr130 (fluorophore label sites) distance than S3 (closed; 31.9 Å vs. 31.2 Å). It is worth noting that the smFRET efficiency is not only determined by the distance between donor and acceptor, but also by their relative orientation and the donor quantum yield, which in turn is sensitive to the local conformation.[17] All of these photophysical properties may result in the higher FRET efficiency of macrostate S4 with a pseudo-closed conformation relative to S3.

Overall, the above results strongly suggest the existence of multiple conformational states for ligand-free GlnBP, which was the main concern of previous studies.[9,10] Analysis of RDC measurements showed that the apo-GlnBP data could not be well interpreted by existing crystal structures (Figure S3). The conformational ensemble derived from MD simulations provides a much better model for RDC data interpretation. Previous computational studies have also revealed the conformational flexibility of GlnBP[18] and the two major modes of conformational movement, namely hinge bending and twist motions. The >40% population of the closed conformation in apo-GlnBP suggests that the ligand-binding process includes a conformation-selection mechanism, obviously facilitating ligand binding, and multiple conformational states provide even more possibilities for the manipulation of the binding affinity.[19] For example, changing the relative populations of the open and closed conformations by mutating Ile329 alters the ligand affinity of MBP.[20] This reciprocal logic for ligand affinity manipulation could also be important for biosensor engineering of PBPs.[2b] On the other hand, the conformational flexibility of GlnBP may facilitate its interaction with the cognate transporter, as in the case of the MBP-MalFGK$_2$ transporter. We notice that convergent developments in various spectroscopic and computational techniques may open up the possibility for integrative studies of protein dynamics at the atomic level. For example, a recent study combined NMR, SAXS, and smFRET to unravel

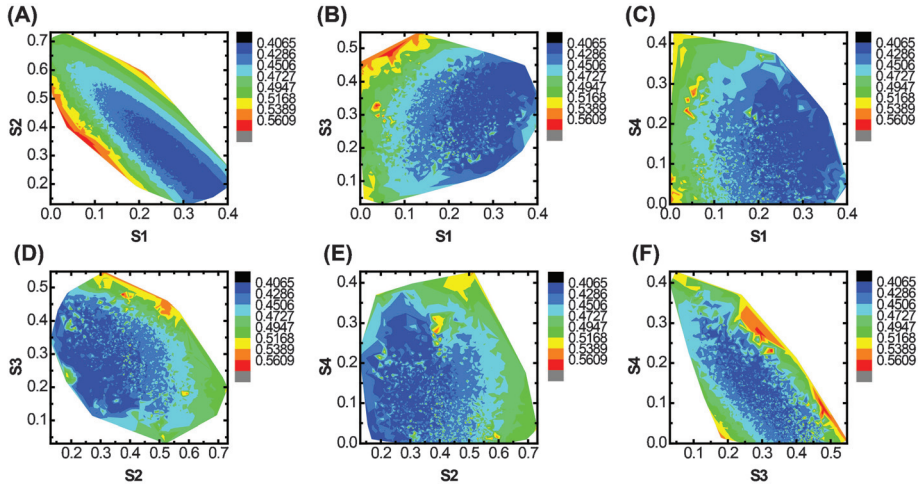

Figure 3. Q factor values for the agreement of the RDC data of apo-GlnBP with the structure ensemble as a function of the relative populations of macrostates S1 and S2 (A), S1 and S3 (B), S1 and S4 (C), S2 and S3 (D), S2 and S4 (E), and S3 and S4 (F).

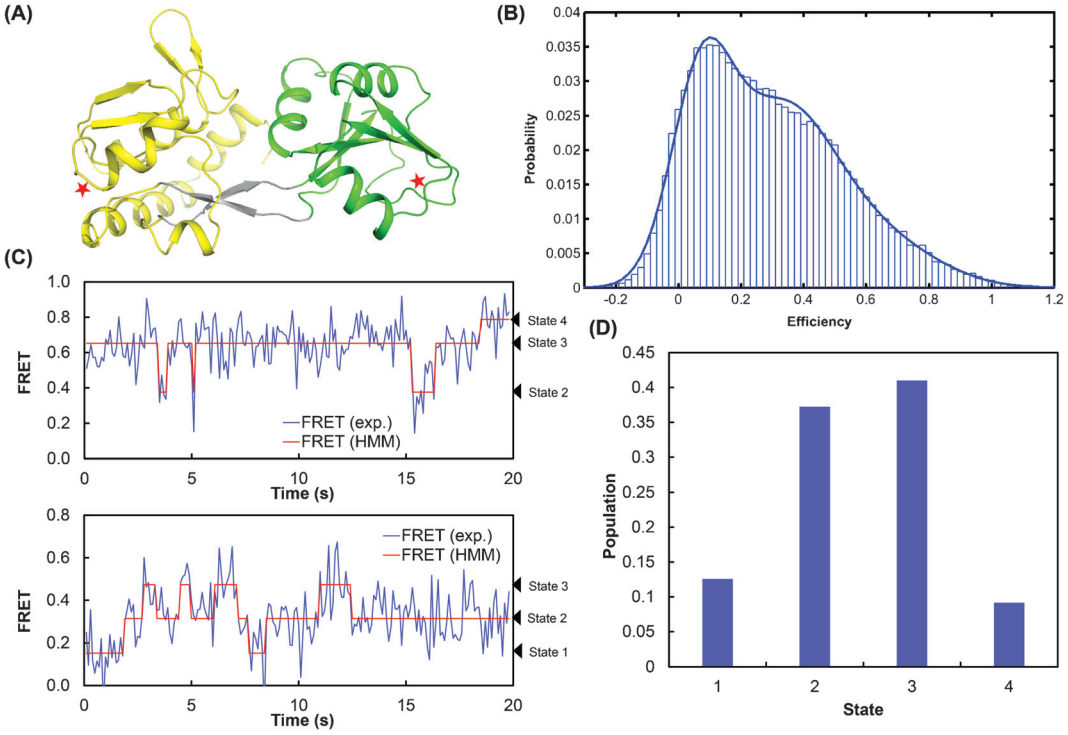

Figure 4. Single-molecule FRET experiments with apo-GlnBP. A) Diagram showing the positions of the fluorescent donor–acceptor labels (Thr59 and Thr130) on GlnBP. B) smFRET efficiency histogram of apo-GlnBP, plotting the photon bursts of each time bin. C) Sample smFRET time traces analyzed with an empirical Bayesian HMM method.[16] Top: FRET signals showing transitions between states 2, 3, and 4. Bottom: FRET signals showing transitions between states 1, 2, and 3. D) Populations of the four states in apo-GlnBP.

the large-scale conformational dynamics crucial for protein–protein interactions.[21] Herein, we have shown that MSM MD simulations can provide a reliable structure ensemble for RDC data interpretation. The combined approach employed here could become widely applicable for investigating the conformational dynamics of biological macromolecules. It is worth noting that by relying on the quality of the equilibrium ensemble derived from large-scale MD simulations and MSM, our strategy avoids the underestimated problems associated with both maximum-entropy and large-weight approaches of ensemble construction.[22] Nevertheless, the accuracy of the force field is always a limitation in MD simulations and thereby of our structure ensemble.

Acknowledgements

This work was supported by the National Major Basic Research Program of China (2016YFA0501702), the National Science Foundation of China (21473034, 21273188), and the Specialized Research Fund for the Doctoral Program of Higher Education (20130071140004). X.H. acknowledges the Hong Kong Research Grants Council (M-HKUST601/13, 609813, 16302214, 16304215, and HKUST C6009-15G). X.G. acknowledges funding from King Abdullah University of Science and Technology (KAUST). We thank Dr. Charles D. Schwieters for insightful discussions and the facility team members for their help with NMR experiments at the National Center for Protein Science Shanghai (NCPSS). This research made use of the computer cluster resources at KAUST.

Keywords: conformational dynamics · FRET · molecular dynamics · NMR spectroscopy · protein folding

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 13990–13994
Angew. Chem. **2016**, *128*, 14196–14200

[1] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789.
[2] a) A. Moussatova, C. Kandt, M. L. O'Mara, D. P. Tieleman, *Biochim. Biophys. Acta Biomembr.* **2008**, *1778*, 1757; b) M. A. Dwyer, H. W. Hellinga, *Curr. Opin. Struct. Biol.* **2004**, *14*, 495.
[3] a) M. M. Flocco, S. L. Mowbray, *J. Biol. Chem.* **1994**, *269*, 8931; b) C. Oswald, S. H. Smits, M. Hoing, L. Sohn-Bosser, L. Dupont, D. Le Rudulier, L. Schmitt, E. Bremer, *J. Biol. Chem.* **2008**, *283*, 32848.
[4] C. Tang, C. D. Schwieters, G. M. Clore, *Nature* **2007**, *449*, 1078.
[5] a) D. Bucher, B. J. Grant, P. R. Markwick, J. A. McCammon, *PLoS Comput. Biol.* **2011**, *7*, e1002034; b) D. Bucher, B. J. Grant, J. A. McCammon, *Biochemistry* **2011**, *50*, 10530.
[6] a) D. A. Silva, G. R. Bowman, A. Sosa-Peinado, X. H. Huang, *PLoS Comput. Biol.* **2011**, *7*, e1002054; b) S. Gu, D. A. Silva, L. Meng, A. Yue, X. Huang, *PLoS Comput. Biol.* **2014**, *10*, e1003767.
[7] C. D. Hsiao, Y. J. Sun, J. Rose, B. C. Wang, *J. Mol. Biol.* **1996**, *262*, 225.
[8] Y. J. Sun, J. Rose, B. C. Wang, C. D. Hsiao, *J. Mol. Biol.* **1998**, *278*, 219.

[9] S. D'Auria, M. Staiano, A. Varriale, M. Gonnelli, A. Marabotti, M. Rossi, G. B. Strambini, *Proteins Struct. Funct. Bioinf.* **2008**, *71*, 743.
[10] G. A. Bermejo, M. P. Strub, C. Ho, N. Tjandra, *Biochemistry* **2010**, *49*, 1893.
[11] a) R. M. de Lorimier et al., *Protein Sci.* **2002**, *11*, 2655; b) L. De Stefano, L. Rotiroti, I. Rendina, L. Moretti, V. Scognamiglio, M. Rossi, S. D'Auria, *Biosens. Bioelectron.* **2006**, *21*, 1664; c) Y. Tian, M. J. Cuneo, A. Changela, B. Höcker, L. S. Beese, H. W. Hellinga, *Protein Sci.* **2007**, *16*, 2240.
[12] A. K. Mittermaier, L. E. Kay, *Trends Biochem. Sci.* **2009**, *34*, 601.
[13] M. Zweckstetter, *Nat. Protoc.* **2008**, *3*, 679.
[14] M. Zweckstetter, A. Bax, *J. Am. Chem. Soc.* **2000**, *122*, 3791.
[15] X. Xiao, N. Kallenbach, Y. K. Zhang, *J. Chem. Theory Comput.* **2014**, *10*, 4152.
[16] J. W. van de Meent, J. E. Bronson, C. H. Wiggins, R. L. Gonzalez, Jr., *Biophys. J.* **2014**, *106*, 1327.
[17] a) T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, S. Weiss, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6264; b) R. Roy, S. Hohng, T. Ha, *Nat. Methods* **2008**, *5*, 507.
[18] a) A. Pang, Y. Arinaminpathy, M. S. Sansom, P. C. Biggin, *FEBS Lett.* **2003**, *550*, 168; b) J. G. Su, X. Jiao, T. G. Sun, C. H. Li, W. Z. Chen, C. X. Wang, *Biophys. J.* **2007**, *92*, 1326; c) T. G. Sun, J. P. Hu, C. H. Li, W. Z. Chen, C. X. Wang, *J. Mol. Struct.: THEOCHEM* **2005**, *725*, 9; d) H. H. Loeffler, A. Kitao, *Biophys. J.* **2009**, *97*, 2541.
[19] J. S. Marvin, H. W. Hellinga, *Nat. Struct. Biol.* **2001**, *8*, 795.
[20] O. Millet, R. P. Hudson, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12700.
[21] E. Delaforge et al., *J. Am. Chem. Soc.* **2015**, *137*, 15122.
[22] E. Ravera, L. Sgheri, G. Parigi, C. Luchinat, *Phys. Chem. Chem. Phys.* **2016**, *18*, 5686.

Received: July 7, 2016
Revised: August 28, 2016
Published online: October 12, 2016

