VIP Structural Biology

DOI: 10.1002/anie.201105648

Solution NMR Structure of Proteorhodopsin**

*Sina Reckel, Daniel Gottstein, Jochen Stehle, Frank Löhr, Mirka-Kristin Verhoefen, Mitsuhiro Takeda, Robert Silvers, Masatsune Kainosho, Clemens Glaubitz, Josef Wachtveitl, Frank Bernhard, Harald Schwalbe, Peter Güntert, and Volker Dötsch**

About ten years ago the characterization of an uncultivated marine bacterium revealed a new type of retinal-binding, integral membrane protein called proteorhodopsin (PR).[1] Many variants of PR exist that are spectrally tuned to the light condition in their environment and can be classified into two major groups, the blue- and green-absorbing forms.[2] Functional assays confirmed their ability to pump protons in a light-dependent manner similar to other microbial rhodopsins.[3] The high abundance of bacteria living in oceanic surface waters makes PR highly interesting because of its potential role in non-chlorophyll-based phototrophy in oceanic carbon cycling and energy flux.[4] The green-absorbing variant of PR and in particular its retinal-binding pocket has been intensively investigated by mutational and spectroscopic analysis,[3,5] solid-state NMR spectroscopy,[6,7] and homology modeling.[8] In addition to the retinal-binding site K231, other functionally important residues include the primary proton acceptor D97, the Schiff base counterions R94 and D227 and the primary proton donor E108. Remarkably, D97 possesses an unusually high pK_a value of about 7.5 which is stabilized by H75 near the photoactive center.[6,9] A similar Asp–His cluster has also been observed in xanthorhodopsin.[10] Influenced by the protonation state of D97 the absorption maximum of the retinal cofactor in PR is highly sensitive to changes in pH, ranging from 520 to 540 nm between pH 10 and 4.[3] Furthermore, the direction of proton pumping switches in response to the pH value between an outward directed transport at alkaline pH and an inward directed transport at acidic pH.[11] In contrast to the functional analysis, structural data on PR are still sparse[8,12,13] mostly because of the lack of well-diffracting three-dimensional crystals. Recently, the potential of solution NMR spectroscopy to solve structures of helical membrane proteins has been reported[14,15] and we show herein the de novo structure of the green variant of proteorhodopsin solved by solution NMR spectroscopy.

The structure of PR (Figure 1) was solved in the short-chain lipid diC7PC (diheptanoyl-phosphocholine) combining long-range NOEs with restraints derived from paramagnetic relaxation enhancement (PRE) and residual dipolar couplings (RDCs). The seven transmembrane helices are connected by short loops. Instead of the anti-parallel β-sheet that is observed between helices B and C in other microbial rhodopsins, torsion angles derived from the protein backbone dihedral angle prediction program TALOS + suggest that PR residues G87–P90 form a short β-turn. The loop between helices D and E is longer than predicted by the secondary structure prediction program TMHMM.[16] In contrast, the loop region connecting helices E and F is shorter than predicted as residues E170–N176 form a helical extension (E′) of helix E. Helix E′ is connected to helix E through a slight helical distortion at G169. Without the extension, helix E has approximately the same length as its neighboring helix D and is thus significantly shorter than the other five helices. Transmembrane helix F is slightly kinked around P201 ending in the longest and most dynamic loop of PR connecting helices F and G. Helix G contains a kink at residue N230 similar to the π-bulge observed in other microbial retinal-binding proteins.[13,17]

[*] S. Reckel,[+] D. Gottstein,[+] Dr. F. Löhr, Prof. Dr. C. Glaubitz, Dr. F. Bernhard, Prof. Dr. P. Güntert, Prof. Dr. V. Dötsch
Institute of Biophysical Chemistry
and Center for Biomolecular Magnetic Resonance
Goethe University Frankfurt
Max-von-Laue Str. 9, 60438 Frankfurt (Germany)
E-mail: vdoetsch@em.uni-frankfurt.de

J. Stehle, R. Silvers, Prof. Dr. H. Schwalbe
Institute for Organic Chemistry and Chemical Biology
and Center for Biomolecular Magnetic Resonance
Goethe University Frankfurt
Max-von-Laue Str. 7–9, 60438 Frankfurt (Germany)

Dr. M.-K. Verhoefen, Prof. Dr. J. Wachtveitl
Institute of Physical and Theoretical Chemistry
Goethe University Frankfurt
Max-von-Laue Str. 7, 60438 Frankfurt (Germany)

Dr. M. Takeda, Prof. Dr. M. Kainosho
Structural Biology Research Center, Nagoya University
Furo-cho, Chikusa-ku, 464-8601 (Japan)

Prof. Dr. P. Güntert
Frankfurt Institute for Advanced Studies
Goethe University Frankfurt
Ruth-Moufang-Str.1, 60438 Frankfurt am Main (Germany)

Prof. Dr. M. Kainosho, Prof. Dr. P. Güntert
Center for Priority Areas, Tokyo Metropolitan University
Hachioji, Tokyo 192-0397 (Japan)

[+] These authors contributed equally to this work.

[**] This work was supported by the Deutsche Forschungsgemeinschaft (grant number DO545/7-1 and SFB 807), the European Drug Initiative on Channels and Transporters (EDICT) contract number HEALTH-F4-2007-201924, the NIH (grant number U54 GM094608), the Center for Biomolecular Magnetic Resonance (BMRZ) and the Cluster of Excellence Frankfurt (Macromolecular Complexes). The [11,20-$^{13}C_2$]retinal was a kind gift of Johan Lugtenburg, Peter Verdegem (Leiden University), Neville McLean, Malcolm H. Levitt and Richard C. D. Brown (Southampton University). P.G. gratefully acknowledges financial support by the Lichtenberg program of the Volkswagen Foundation and by a Grant-in-Aid for Scientific Research of the Japan Society for the Promotion of Science.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201105648.

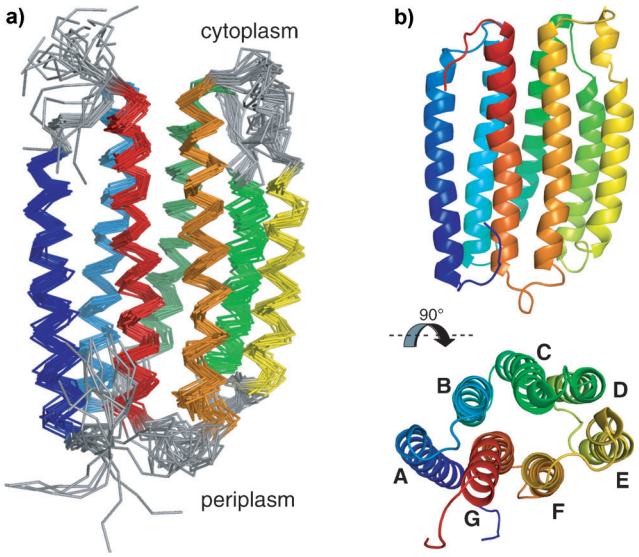

Figure 1. Structure of PR. a) A bundle of 20 conformers with lowest CYANA target function obtained from structure calculations. The helices are color-coded from helix A in dark blue to helix G in red. b) Conformer with the lowest CYANA target function seen from the side and from the top. In the lower panel helices are additionally labeled A–G.

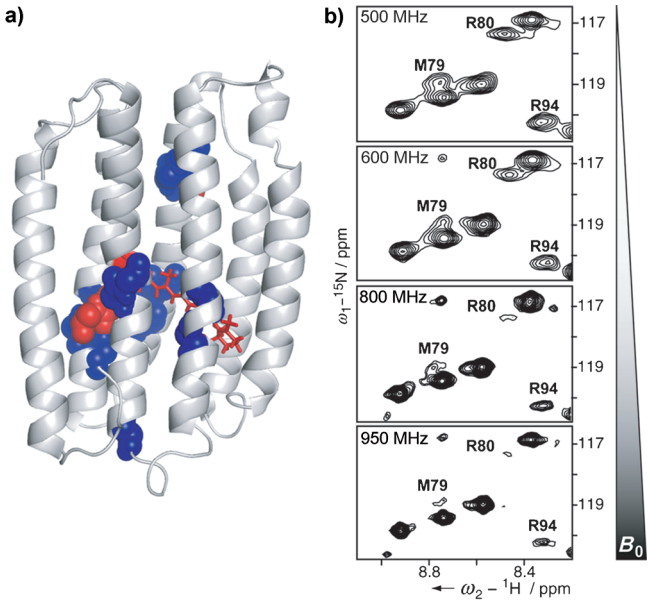

Figure 2. Assignment and field-dependent line broadening of PR. a) 96% of the backbone resonances of PR were assigned (grey). Residues that could not be assigned are depicted as red spheres (Y76, M77, C107) and those with partial backbone assignments as blue spheres. b) The field-dependent line broadening is shown for three representative resonances. [^{15}N,^{1}H]-TROSY spectra of selectively ^{15}N-MRY-labeled PR were measured at different B_o field strengths as indicated in each spectrum. Resonances of R80, M79, and R94 disappear almost completely at 950 MHz, whereas most resonances remain unaffected by the field strength.

To facilitate selective labeling strategies and sample preparation, the structure determination of PR relied on a cell-free expression system.[18] Sample analysis showed that PR was stable and monodisperse in the diC7PC micelle and additional functional evidence was obtained from flash photolysis measurements monitoring the PR photocycle (see Figure S1 in the Supporting Information). Optimization of the NMR conditions allowed the detection of almost all NH backbone resonances (see Figure S2 in the Supporting Information). Nonetheless, the large size of the proteo-micelle, the limited sensitivity of three-dimensional, triple resonance experiments and the signal overlap required combinatorial labeling approaches[19] which, together with uniform labeling, allowed us to assign 96% of the backbone resonances. Conformational exchange broadening, in particular for residues forming the retinal-binding pocket and those involved in the proton pumping mechanism, however, left gaps in the assignment (see Figure 2 and Figure S3 in the Supporting Information). Based on the backbone chemical shifts, the seven-transmembrane-helix topology of PR was confirmed using the programs TALOS + [20] and CSI[21] (see Figure S4 in the Supporting Information). Additionally, {^{1}H}^{15}N-heteronuclear NOE measurements provided information about structured regions of PR and showed highly flexible N- and C-termini and moderately increased mobility only in the loops between helices C and D as well as F and G (see Figure S4 in the Supporting Information).

Because of the α-helical structure of PR and the large size of the proteo-micelle, NOE-derived distance restraints to determine the tertiary fold of PR were difficult to obtain. Long-range NOE information was exclusively derived from methyl groups and aromatic side chains, whereas measurement of backbone H^N–H^N NOEs provided only short-range information within the same helix. The side-chain assignment based on uniform ^{15}N-, ^{13}C-labeling was complicated by fast transverse ^{1}H and ^{13}C relaxation and labeling of methyl groups of Ile, Leu, and Val based on the use of metabolic precursors,[22] was not applicable, because enzymes for this specific precursor metabolism are missing in our cell-free expression system. However, using the cell-free system we could employ SAIL variants of Leu and Val[23] with the advantage that only one of the prochiral methyl groups is detectable and signal overlap is considerably reduced (see Figure S5 in the Supporting Information). SAIL permitted the side-chain assignment of all Val and 60% of the Leu residues. In addition, selective labeling enabled the assignment of 60 out of 65 Ala, Ile, Met, and Thr residues as well as side-chain resonances for eight out of ten Trp residues. Altogether the side-chain assignments covered 44% of the transmembrane region (see Figure S4 in the Supporting Information) and we were able to extract 137 medium- to long-range NOEs from a 4D [^{13}C,^{13}C]-separated NOESY spectrum taking advantage of non-uniform sampling (Figure 3). The side-chain NOEs were essential for the structure determination and played a key role in positioning the helices relative to each other. The majority of long-range distances were, however, derived by paramagnetic relaxation enhancement. In total, 13 single-cysteine mutants were used for PRE measurements with samples that were selectively labeled to minimize signal overlap. Distances were derived by a combination of the two-time-point measurement as described by Iwahara et al.[24] and the method used by Battiste

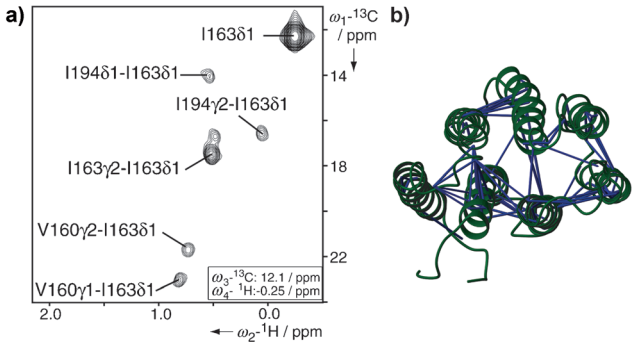

Figure 3. Distance information for the structure calculation of PR. a) Representative plane of the 4D [^{13}C,^{13}C]-separated NOESY spectrum recorded with non-uniform sampling. b) Altogether 137 medium- to long-range NOEs were obtained for methyl groups and Trp side chains of an AILMTVW-labeled sample connecting the individual transmembrane helices.

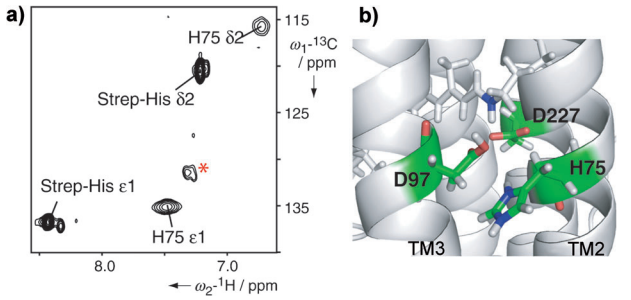

Figure 4. Schiff base environment in PR. a) [^{13}C,^{1}H]-SOFAST-HMQC of ^{15}N,^{13}C-His-labeled PR. Only one additional His residue in Strep-tagged PR enables the assignment of H75 at pH 5. The spectrum shows the low signal-to-noise ratio of the imidazole-ring resonances of H75. Natural abundance background signals are labeled by the red asterisk. b) Close-up of the retinal-binding pocket with retinal shown as gray sticks, the Schiff base nitrogen highlighted in blue and residues that were included into the modeling colored by the element (TM = transmembrane helix).

and Wagner,[25] resulting in 290 upper and 716 lower distance limits (see Figure S6 in the Supporting Information). To increase the structural accuracy, measurements of residual dipolar couplings provided additional restraints for the structure calculation. Selectively labeled PR samples were therefore aligned in 4% strained polyacrylamide gels, which allowed us to derive 81 backbone NH RDC restraints within the transmembrane region of PR. The structure was calculated with CYANA (see Table S1 in the Supporting Information). It combined backbone torsion angle restraints obtained from TALOS+, restraints for α-helical hydrogen bonds as well as sequential backbone NOEs and medium- to long-range side-chain NOEs together with PRE- and RDC-derived restraints. Structural validation relied on titration with paramagnetic agents that are either water- or detergent-soluble (see Figure S7 in the Supporting Information).

To investigate the retinal-binding pocket of PR, distance information positioning the cofactor within the protein is essential. Selective observation of cofactor resonances could be achieved with [11,20-$^{13}C_2$]-labeled retinal. The C20 methyl group had a reasonable signal-to-noise ratio and no overlap, whereas the olefinic group in position C11 showed only a very weak and overlaid signal. We thus identified six retinal-to-protein NOEs which were used in the structure calculation (see Figure S8 in the Supporting Information). Notably, the Cδ1 methyl group of L105 is one out of two methyl groups showing a NOE to the C20 methyl group of retinal. This residue is a key determinant of the spectral properties of the two main variants of PR, green- and blue-absorbing versions[26] and our structural data now confirm its position close to the Schiff base. As the retinal-to-protein distance restraints could only be derived from position C20 located in proximity to the Schiff base, accurate alignment of the cofactor within the protein was not achieved. Also, because of significant line broadening of residues in the retinal-binding pocket, the structural information in this region is limited. While the side-chain resonances of the characteristic H75 residue could be assigned upon removal of the His-tag, the inherently low signal-to-noise ratio prevented attempts to derive long-range distance information (Figure 4). To better reflect the surrounding residues of the retinal, distance constraints based on published biochemical evidence[6,27] (see Figure S8 in the Supporting Information) were used to calculate a structure with a modeled binding site. In particular the side-chain atoms of D97 and D227 were given an upper distance limit of 5 Å to the Schiff base and the H75 side chain was positioned within 3.5 Å of D97 to build the hydrogen bond network (Figure 4). In addition, the proximity of Y200 to the ring structure of the retinal was introduced with an upper limit of 7 Å. The resulting structure overlays well with the one based solely on experimental NMR data differing only in the side-chain orientations of the restrained residues.

In conclusion, the solution NMR structure of PR reveals differences from its homologues such as the absence of an anti-parallel β-sheet between helices B and C (Figure 5).

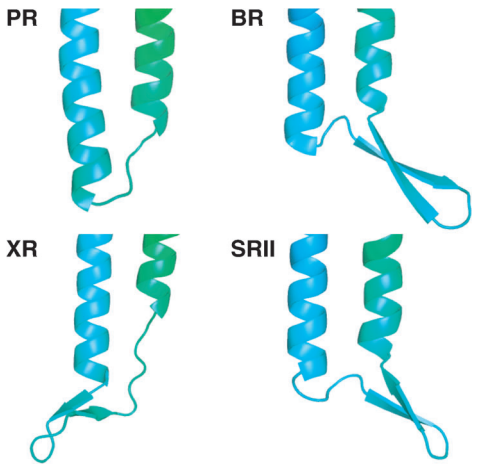

Figure 5. B–C loop of PR and its three homologues bacteriorhodopsin (BR), sensory rhodopsin II (SRII), and xanthorhodopsin (XR; PDB ID: 1C3W, 2KSY, and 3DDL, respectively). In PR an anti-parallel β-sheet between helices B and C as observed in the other three structures is not present. Sequence alignment further supports these findings as PR possesses a significantly shortened sequence in this region which favors the formation of a β-turn rather than two extended β-strands (see Figure S8 in the Supporting Information).

Structural data on retinal-binding proteins indicated that the B–C loop interacts with other extracellular loop regions and is important to maintain protein function and stability. In PR, the F–G loop might partially replace the shielding properties of the missing antiparallel β-sheet. Another interesting element in the structure of PR constitutes the unexpectedly short E–F loop that comprises only four residues including a proline whereas residues E170–N176 form an extension of helix E. The conformational restrictions in this region provide an explanation why site-directed mutagenesis at position 178 (A178R) influenced the spectroscopic properties of the retinal even though it is distant from the retinal-binding site itself.[28] Although the side-chain assignment of PR covers about 44% of the transmembrane region, the number of experimental long-range NOE information remains limited and the structural accuracy cannot compete with the recently determined NMR structure of sensory rhodopsin II.[15] Including biochemical data into the PR structure calculation has shown, however, that this biochemical information is consistent with our structure. Certainly, within the retinal-binding protein family, PR remains a special case because its crystal structure remains elusive despite considerable efforts. The relatively short loops and the absence of the extended β-sheet may contribute to these difficulties to produce three-dimensional crystals because a protein that is well-buried in the micelle has less possibility to form stabilizing crystal contacts. This underscores the importance of the present solution NMR structure of PR.

Experimental Section

The green-absorbing PR was cloned into a pIVEX2.3d vector and expressed in a S30-based continuous-exchange cell-free system. Stable-isotope-labeled amino acids (Cambridge Isotope Laboratories) or SAIL amino acids (SAIL Technologies, Inc.) were added directly to the reaction mixture. PR was expressed in the detergent mode in the presence of 0.6 mM all-*trans* retinal (Sigma), containing 0.4% digitonin (Sigma) mixed with diC7PC (Avanti Polar Lipids) in a 4:1 molar ratio. Ni-affinity purification was necessary to remove impurities and exchange the detergent to diC7PC. The final NMR buffer conditions were 25 mM NaOAc, pH 5 with 2 mM DTT. The final detergent concentration was approximately 2%. Protein concentrations typically ranged from 0.3 to 0.5 mM. The side-chain assignment and NOESY experiments required the use of deuterated detergents. PRE experiments to derive long-distance restraints were performed using *S*-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) labeled single-cysteine mutants. RDCs were obtained upon alignment in 4% polyacrylamide gels using a molar ratio of acrylamide to *N*,*N*′-methylenebisacrylamide of 150:1. Structure calculations were performed with CYANA.[29] The atomic coordinates, chemical shifts, and restraints have been submitted to the Protein Data Bank, www.pdb.org (PDB ID code 2l6x) and the BioMagnResBank, www.bmrb.wisc.edu (accession code 17327). Further details are provided in the Supporting Information.

Received: August 9, 2011
Published online: October 27, 2011

Keywords: membrane proteins · NMR spectroscopy · proteorhodopsin · structural biology

[1] O. Beja, L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich, E. F. DeLong, *Science* **2000**, *289*, 1902.
[2] D. Man, W. Wang, G. Sabehi, L. Aravind, A. F. Post, R. Massana, E. N. Spudich, J. L. Spudich, O. Beja, *Embo J.* **2003**, *22*, 1725.
[3] T. Friedrich, S. Geibel, R. Kalmbach, I. Chizhov, K. Ataka, J. Heberle, M. Engelhard, E. Bamberg, *J. Mol. Biol.* **2002**, *321*, 821.
[4] E. F. DeLong, O. Beja, *PLoS Biol.* **2010**, *8*, e1000359.
[5] a) A. K. Dioumaev, L. S. Brown, J. Shih, E. N. Spudich, J. L. Spudich, J. K. Lanyi, *Biochemistry* **2002**, *41*, 5348; b) A. K. Dioumaev, J. M. Wang, Z. Balint, G. Varo, J. K. Lanyi, *Biochemistry* **2003**, *42*, 6582; c) M.-K. Verhoefen, K. Neumann, I. Weber, C. Glaubitz, J. Wachtveitl, *Photochem. Photobiol.* **2009**, *85*, 540.
[6] a) F. Hempelmann, S. Hölper, M.-K. Verhoefen, A. C. Woerner, T. Köhler, S. A. Fiedler, N. Pfleger, J. Wachtveitl, C. Glaubitz, *J. Am. Chem. Soc.* **2011**, *133*, 4645; b) N. Pfleger, A. C. Woerner, J. Yang, S. Shastri, U. A. Hellmich, L. Aslimovska, M. S. Maier, C. Glaubitz, *Biochim. Biophys. Acta.* **2009**, *1787*, 697.
[7] a) N. Pfleger, M. Lorch, A. C. Woerner, S. Shastri, C. Glaubitz, *J. Biomol. NMR* **2008**, *40*, 15; b) S. Shastri, J. Vonck, N. Pfleger, W. Haase, W. Kuehlbrandt, C. Glaubitz, *Biochim. Biophys. Acta Biomembr.* **2007**, *1768*, 3012.
[8] R. Rangarajan, J. F. Galan, G. Whited, R. R. Birge, *Biochemistry* **2007**, *46*, 12679.
[9] V. B. Bergo, O. A. Sineshchekov, J. M. Kralj, R. Partha, E. N. Spudich, K. J. Rothschild, J. L. Spudich, *J. Biol. Chem.* **2009**, *284*, 2836.
[10] H. Luecke, B. Schobert, J. Stagno, E. S. Imasheva, J. M. Wang, S. P. Balashov, J. K. Lanyi, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16561.
[11] E. Lörinczi, M.-K. Verhoefen, J. Wachtveitl, A. C. Woerner, C. Glaubitz, M. Engelhard, E. Bamberg, T. Friedrich, *J. Mol. Biol.* **2009**, *393*, 320.
[12] a) L. Shi, M. A. Ahmed, W. Zhang, G. Whited, L. S. Brown, V. Ladizhansky, *J. Mol. Biol.* **2009**, *386*, 1078; b) J. Yang, L. Aslimovska, C. Glaubitz, *J. Am. Chem. Soc.* **2011**, *133*, 4874.
[13] L. Shi, E. M. Lake, M. A. Ahmed, L. S. Brown, V. Ladizhansky, *Biochim. Biophys. Acta Biomembr.* **2009**, *1788*, 2563.
[14] a) W. D. Van Horn, H. J. Kim, C. D. Ellis, A. Hadziselimovic, E. S. Sulistijo, M. D. Karra, C. Tian, F. D. Sönnichsen, C. R. Sanders, *Science* **2009**, *324*, 1726; b) S. Sobhanifar, B. Schneider, F. Löhr, D. Gottstein, T. Ikeya, K. Mlynarczyk, W. Pulawski, U. Ghoshdastider, M. Kolinski, S. Filipek, P. Güntert, F. Bernhard, V. Dötsch, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9644; c) J. A. Butterwick, R. MacKinnon, *J. Mol. Biol.* **2010**, *403*, 591; d) M. J. Berardi, W. M. Shih, S. C. Harrison, J. J. Chou, *Nature* **2011**, *476*, 109.
[15] A. Gautier, H. R. Mott, M. J. Bostock, J. P. Kirkpatrick, D. Nietlispach, *Nat. Struct. Mol. Biol.* **2010**, *17*, 768.
[16] A. Krogh, B. Larsson, G. von Heijne, E. L. Sonnhammer, *J. Mol. Biol.* **2001**, *305*, 567.
[17] J. P. Cartailler, H. Luecke, *Structure* **2004**, *12*, 133.
[18] D. Schwarz, F. Junge, F. Durst, N. Frölich, B. Schneider, S. Reckel, S. Sobhanifar, V. Dötsch, F. Bernhard, *Nat. Protoc.* **2007**, *2*, 2945.
[19] a) F. Hefke, A. Bagaria, S. Reckel, S. J. Ullrich, V. Dötsch, C. Glaubitz, P. Güntert, *J. Biomol. NMR* **2010**, *49*, 75; b) S. Sobhanifar, S. Reckel, F. Junge, D. Schwarz, L. Kai, M. Karbyshev, F. Löhr, F. Bernhard, V. Dötsch, *J. Biomol. NMR* **2010**, *46*, 33; c) N. Trbovic, C. Klammt, A. Koglin, F. Löhr, F. Bernhard, V. Dötsch, *J. Am. Chem. Soc.* **2005**, *127*, 13504.
[20] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213.
[21] D. S. Wishart, B. D. Sykes, *J. Biomol. NMR* **1994**, *4*, 171.
[22] V. Tugarinov, V. Kanelis, L. E. Kay, *Nat. Protoc.* **2006**, *1*, 749.

[23] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono, P. Güntert, *Nature* **2006**, *440*, 52.
[24] J. Iwahara, C. Tang, G. M. Clore, *J. Magn. Reson.* **2007**, *184*, 185.
[25] J. L. Battiste, G. Wagner, *Biochemistry* **2000**, *39*, 5355.
[26] J. M. Kralj, E. N. Spudich, J. L. Spudich, K. J. Rothschild, *J. Phys. Chem. B* **2008**, *112*, 11770.
[27] a) D. Ikeda, Y. Furutani, H. Kandori, *Biochemistry* **2007**, *46*, 5365; b) J. Y. Jung, A. R. Choi, Y. K. Lee, H. K. Lee, K.-H. Jung, *FEBS Lett.* **2008**, *582*, 1679.
[28] a) M. Yoshitsugu, M. Shibata, D. Ikeda, Y. Furutani, H. Kandori, *Angew. Chem.* **2008**, *120*, 3987; *Angew. Chem. Int. Ed.* **2008**, *47*, 3923; b) M. Yoshitsugu, J. Yamada, H. Kandori, *Biochemistry* **2009**, *48*, 4324.
[29] P. Güntert, *Prog. Nucl. Magn. Reson. Spectrosc.* **2003**, *43*, 105.

