Solid-State NMR Spectroscopy

DOI: 10.1002/anie.201201160

In Situ Solid-State NMR Spectroscopy of Protein in Heterogeneous Membranes: The Baseplate Antenna Complex of *Chlorobaculum tepidum***

Natalia V. Kulminskaya, Marie Ø. Pedersen, Morten Bjerring, Jarl Underhaug, Mette Miller, Niels-Ulrik Frigaard, Jakob T. Nielsen, and Niels Chr. Nielsen**

Structural biology is currently booming, as seen clearly by the approximately 170 new structures deposited each week in the protein data bank (PDB) with the vast majority of these deposited structures being derived from either X-ray crystallography (about 87%) or solution NMR spectroscopy (about 12%).[1] In light of this rapid development, it is becoming increasingly evident that a major up-coming challenge is to study proteins in their native environment, rather than relying on either crystals or buffered solutions as used in those two classical methods, respectively. However, this is far from a trivial task, as the native environment often is extremely heterogeneous containing many different types of binding partners, such as sugars, lipids, and co-factors, constraining the protein structure. This is especially true when studying membrane proteins, and for this group of proteins the typical laboratory protocol involves a reconstitution step in which a membrane-mimicking environment is used to solubilize the protein. This often unnaturally simple environment might severely perturb the folding state of the protein, and important details on the functionally relevant structure may be lost.

Currently, only few techniques, including cryo-electron microscopy (cryo-EM)[2] and solid-state NMR spectroscopy,[3] offer the possibility to obtain atomic resolution structure information for proteins in native heterogeneous environments. Cryo-EM enables single-particle characterization with resolution down to about 4 Å, while solid-state NMR spectroscopy relies on ensemble measurements with substantially higher resolution. The potential of solid-state NMR spectroscopy for in situ characterization has been demonstrated by analysis of bacteriorhopsin (bR) in purple membranes,[4] GvpA proteins in collapsed gas vesicles,[3a,5] and the LR11 (SorLA) transmembrane domain,[3c] Mistic,[6] and PagL[7] in native *Escherichia coli* membranes. These studies relied on large expression levels of the protein either through very high abundance in 2D arrayed arrangements (bR and GvpA), or overexpression (LR11, Mistic, PagL), the latter with the risk of introducing bias by the condensed environment. In the present study, we address a more heterogeneous system through structural characterization of the photosynthetic antenna protein, CsmA, from a phototrophic prokaryote. CsmA is present in the lipid environment of chlorosomes, and we have employed solid-state NMR spectroscopy using whole organelle preparations to extract structural knowledge of this protein.

Common to all photosynthetic systems is the use of a light-harvesting antenna, which increases the number of photons accessible to the photochemical reaction centers.[8] The largest known antenna system is the chlorosome found in green sulphur bacteria (*Chlorobi*), filamentous anoxygenic phototrophs (*Chloroflexi*), and the aerobic phototroph, *Candidatus Chloracidobacterium termophilum*.[9] This antenna has an exceptional ability of extracting energy from low intensity light because of a very high number of pigments per reaction center. The chlorosomes of *Chlorobaculum* (*Cba.*) *tepidum*, a common model organism for green sulphur bacteria, appears as oblong bodies, attached to the cytoplasmic side of the plasma membrane through a water-soluble protein, the Fenna–Matthews–Olson (FMO) protein (Figure 1).[10] The interior of the chlorosomes is formed by long rods of self-aggregated bacteriochlorophyll (BChl) *c* held together by pigment–pigment interactions[10] in a structure recently determined by solid-state NMR spectroscopy.[11] The chlorosomes are surrounded by a lipid-protein monolayer holding ten different polypeptides.[9b] At the attachment surface between the chlorosome and the FMO protein, the envelope has a crystalline-like baseplate containing multiple CsmA proteins (6.2 kDa) in complex with BChl *a* and carotenoids.[12] This CsmA–BChl *a* baseplate is vital for the organism as it transfers energy from the chlorosome interior to the FMO protein, which is the last antenna before the energy reaches the reaction center.[9b]

[*] M. Sc. N. V. Kulminskaya, Dr. M. Ø. Pedersen, Dr. M. Bjerring, Dr. J. Underhaug, Dr. J. T. Nielsen, Prof. Dr. N. C. Nielsen
Center for Insoluble Protein Structures (inSPIN)
Interdisciplinary Nanoscience Center (iNANO) and
Department of Chemistry, University of Aarhus
Langelandsgade 140, 8000 Aarhus C (Denmark)
E-mail: jtn@chem.au.dk
ncn@inano.au.dk
Assoc. Prof. Dr. M. Miller
Department of Biochemistry and Molecular Biology
University of Southern Denmark (Denmark)
Assoc. Prof. Dr. N.-U. Frigaard
Department of Biology, University of Copenhagen (Denmark)
Dr. M. Ø. Pedersen
Novo Nordisk A/S, Måløv (Denmark)
Dr. J. Underhaug
Department of Biomedicine, University of Bergen (Norway)

[**] We acknowledge support from the Danish National Research Foundation, the Danish Centre for Scientific Computing, and the BIONMR FP7 program.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201201160.

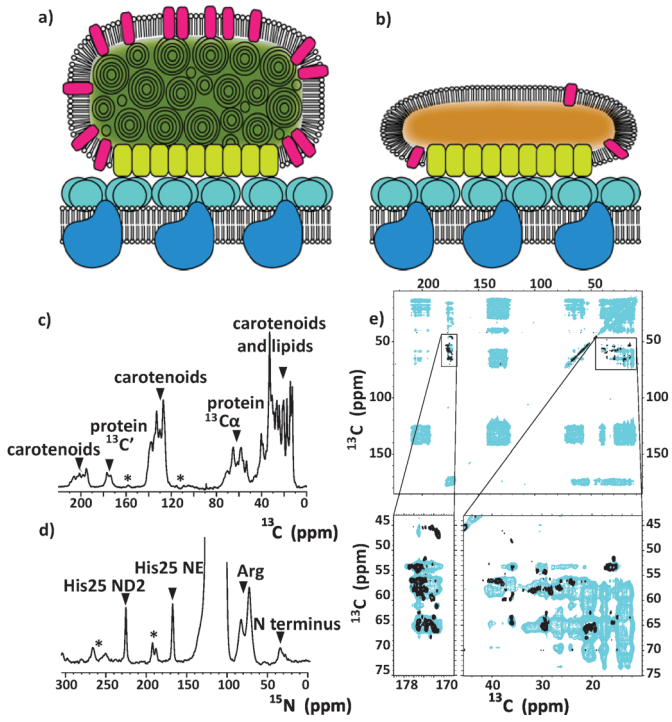

Figure 1. a) The chlorosome (modified from Ref. [13]) consists of BChl *c* self-aggregates (dark green) along with carotenoids, surrounded by a lipid monolayer containing nine types of polypeptides (magenta). The baseplate consists of CsmA proteins (light green) binding BChl *a*. The FMO protein (light blue) is bound to the baseplate and to the reaction centers (dark blue) in the cytoplasmic membrane.[13] b) The carotenosome having a lipid envelope containing fewer polypeptides (magenta) and only carotenoids (orange) but not BChl *c*. c) ^{13}C CP-MAS spectrum of uniformly ^{13}C,^{15}N-labeled carotenosomes. d) ^{15}N CP-MAS spectrum of the same sample. Arrows indicate the most significant resonances from the carotenosomes and stars represent BChl *a* signals. e) Overlaid ^{13}C,^{13}C (cyan) and ^{15}N-filtered ^{13}C,^{13}C (black) DARR spectra. Boxes below show regions with cross-peaks between $^{13}C\alpha$–$^{13}C'$ and between $^{13}C\alpha$ and side-chain carbon nuclei. Additional information is found in the Supporting Information.

The structure of CsmA has previously been determined in organic solvents using liquid-state NMR spectroscopy.[13] Since the organic environment represents individually dissolved proteins and prevents attachment of BChl *a*, and a reconstitution protocol failed to give soluble species[14] we continued our investigation using a whole organelle approach, where the CsmA–BChl *a* baseplate could be studied in its native heterogeneous environment using solid-state NMR spectroscopy. To obtain samples containing sufficiently large amounts of protein for NMR characterization, we based our study on simplified chlorosomes, the so-called carotenosomes (Figure 1b), derived from a *bchK* mutant of *Cba. tepidum* devoid of BChl *c*.[15] Carotenosomes are obviously less efficient in harvesting light than wild-type chlorosomes, but they remain functional.[16] The carotenosomes contain lipids, carotenoids, BChl *a*, and the proteins CsmA and CsmD (CsmA:CsmD ≈ 9:1). All other chlorosome proteins normally found in wild-type chlorosomes were found only in trace amounts or were not detected.[16] These proteins have, however, not been ascribed any pivotal functions for the system, as seen from knock-out studies.[17] It was observed that the CsmA protein still forms remarkably stable oligomeric CsmA–BChl *a* baseplate complexes in the carotenosomes, and that these were indistinguishable spectroscopically from those of wild-type chlorosome.[16] As drawn in Figure 1b, the carotenosomes are flattened in structure as compared to chlorosomes because of the absence of BChl *c*, and the isolated organelles appear brightly orange because of the lack of BChl *c*.[16] Cells of a *Cba. tepidum bchK* mutant for isolation of ^{13}C,^{15}N-isotope labelled carotenosomes were grown as described by Frigaard et al.[15] Details are given in the Supporting Information.

Structural characterization of uniformly ^{13}C,^{15}N-labelled carotenosomes was carried out using multidimensional magic-angle spinning (MAS) solid-state NMR spectroscopy. One-dimensional ^{13}C cross-polarization (CP) MAS spectra (Figure 1c) revealed signals from large amounts of lipids and carotenoids in addition to CsmA and BChl *a*. The relatively weak ^{13}C protein signals and the strong signals from the other constituents resonated mainly in disjoint regions (Figure 1c,e), although there was some overlap between intense carotenoid signals and aromatic protein signals (120–140 ppm region) as well as between aliphatic lipid chains and carotenoid methyl signals and those of aliphatic protein side chains. Weak signals around 95 and 160 ppm were identified as BChl *a* signals. The ^{15}N CP-MAS spectrum (Figure 1d) almost exclusively contained signals from CsmA and BChl *a*, which were clearly distinguishable. The exclusivity of the ^{15}N signals suggests the use of ^{15}N–^{13}C dipolar-coupling filtering as a means to obtain high-resolution spectra for the baseplate alone, as demonstrated in Figure 1e overlaying a 2D ^{13}C,^{13}C DARR spectrum (blue) with a similar ^{15}N-filtered 2D spectrum obtained by initially transferring the polarization through the amide ^{15}N nuclei (black).

^{15}N-edited NCACX, NCOCX, and CONCA 3D experiments enabled almost complete resonance assignment of the non-terminal region of the protein while most resonances from BChl *a* were assigned based on ^{13}C,^{13}C DARR spectra. Figure 2a demonstrates a sequential walk through backbone resonances from the V22-M21-V20-E19-F18 stretch. All backbone resonances for CsmA were assigned unambiguously (see Table S2 in the Supporting Information) except the terminal residues M1-G5 and G52-S59, for which we judge signals to be absent because of increased mobility. Likewise, with a few exceptions all side-chain resonances were assigned unambiguously for all residues. Only one signal was observed for each nucleus indicating that only a single conformation of CsmA exists in the baseplate, and thus the proteins of the entire baseplate must be part of a fully symmetric oligomeric complex. At room temperature, the ^{13}C signals from BChl *a* were relatively weak. This may be ascribed to low $^{1}H \rightarrow ^{13}C$ CP efficiency because of a low density of directly attached protons and high flexibility of BChl *a* in the carotenosomes. It appeared that cooling of the sample to −9 °C improved the pigment signals leading to the assignments in Table S2 in the Supporting Information, and our assignments are in good agreement with those of Egorova-Zachernyuk et al.[18] for pure BChl *a* in the solid state. Besides CsmA and BChl *a*, the

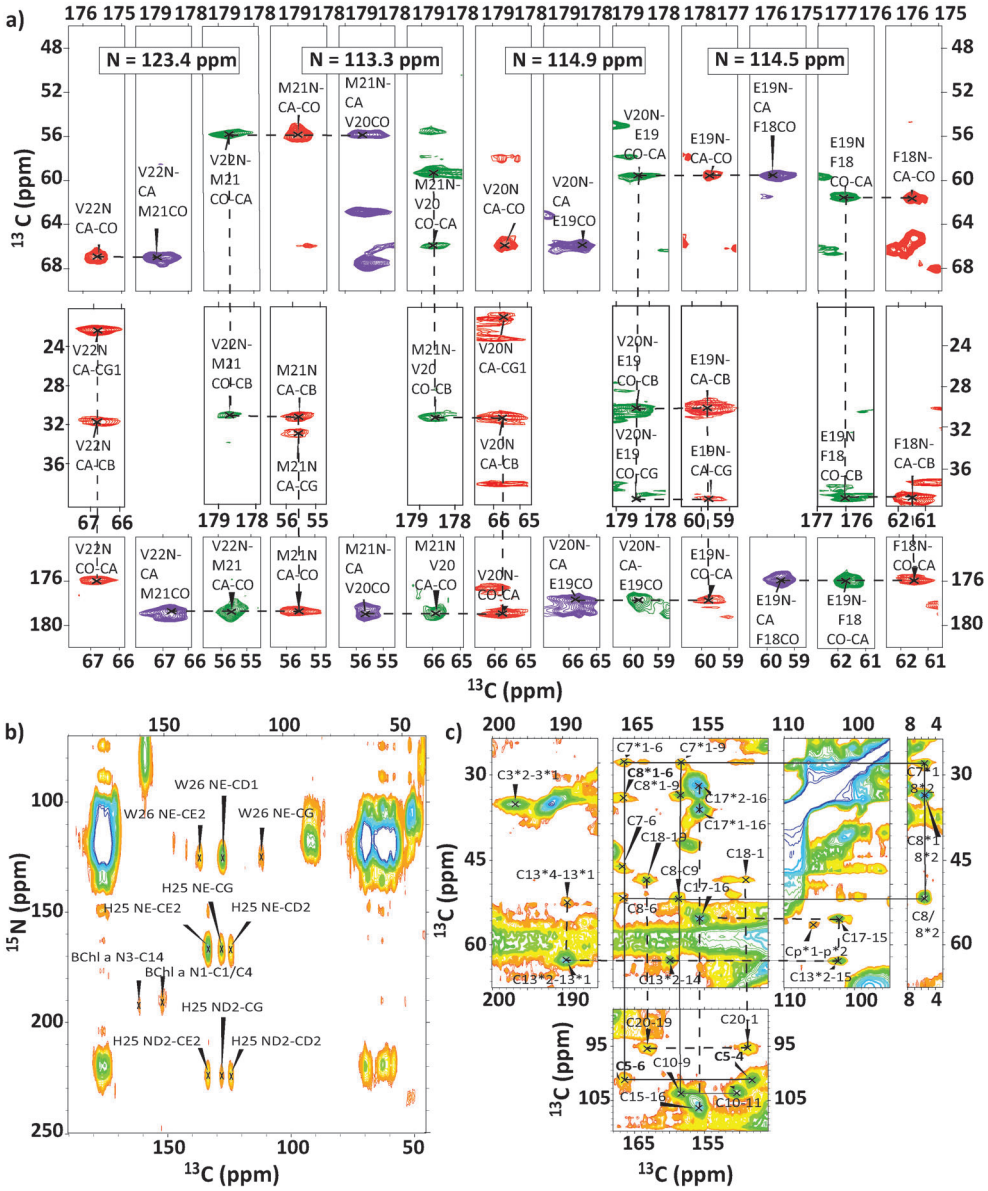

Figure 2. Representative assignments of solid-state NMR spectra of *U*-^{13}C,^{15}N-carotenosomes. a) Strip plots from ^{15}N,^{13}C,^{13}C 3D correlation spectra showing assignments for residues V22-F18 obtained from NCACX (red), NCOCX (green), and CONCA (magenta) experiments. The strips were taken at the ^{15}N chemical shifts indicated at the top of the strips and viewed as ^{13}Cα–^{13}C′ (top bar), $^{13}C_{sidechain}$–^{13}C′/^{13}Cα (middle bar) and ^{13}C′-^{13}Cα (bottom bar) slices. Horizontal lines represent intraresidual and sequential connectivities while vertical lines represent intraresidual connectivities within the same experiment and, thus, outline the sequential assignment pathway. b) ^{15}N,^{13}C correlation spectra showing representative assignments of Bchl *a*, His, and Trp side-chain resonances. c) ^{13}C,^{13}C DARR spectrum (−9 °C) showing sequential assignments of BChl *a*.

baseplate contains significant amounts of carotenoids.[19] Despite that the carotenoids give large signals in the ^{13}C,^{13}C DARR spectra, we were not able to identify cross-peaks to CsmA or BChl *a*. This may be ascribed to heterogeneous ordering of the carotenoid molecules and/or high degree of flexibility.

Taking advantage of the fact that ^{13}C and ^{15}N chemical shifts are sensitive to the protein backbone conformation,[20] Figure 3a compares secondary chemical shifts for the signals observed by solid-state NMR spectroscopy with the corresponding secondary shifts observed for CsmA dissolved in a 1:1 chloroform:methanol solution. In support of the previous solution structure of CsmA,[13] largely all liquid-state NMR secondary shifts, except for S2, and all solid-state NMR secondary shifts were strongly supportive of an α-helical structure.

Statistical analysis of the assigned chemical shifts using TALOS+[21] may provide detailed inference on the backbone conformation as demonstrated by predictions of the dihedral ϕ and ψ angles in Figure S4 in the Supporting Information from the solid-state NMR chemical shifts. The majority of predicted ϕ and ψ angles of the protein fall within a narrow range of $<20°$ indicating a uniform backbone conformation of CsmA. This narrow distribution of predicted torsion angles corresponds to the most highly energetically favorable α-helical region of the Ramachandran plot (Figure 3b). Hence, TALOS+ predicts a long uniform canonical α-helix over residues V6-G49 for CsmA in the baseplate. The solution structure of the CsmA monomer is also α-helical[13] but more scatter in the backbone torsion angles (see the Supporting Information) and, most notably, a kink in the helix between residues 34–38 was observed.[13] This apparent deviation from a continuous α-helical structure could be because of the absence of any long-range distance constraints in the structure determination of the solution structure (and fewer medium-range for that particular stretch). Here we used TALOS+ to predict the backbone conformation from the solution chemical shifts. Indeed, contrasting the derived structure, all TALOS+ predictions fall inside the same narrow distribution as for the solid-state predictions (see the Supporting Information). Thus, all chemical shift data indicated that the backbone conformation

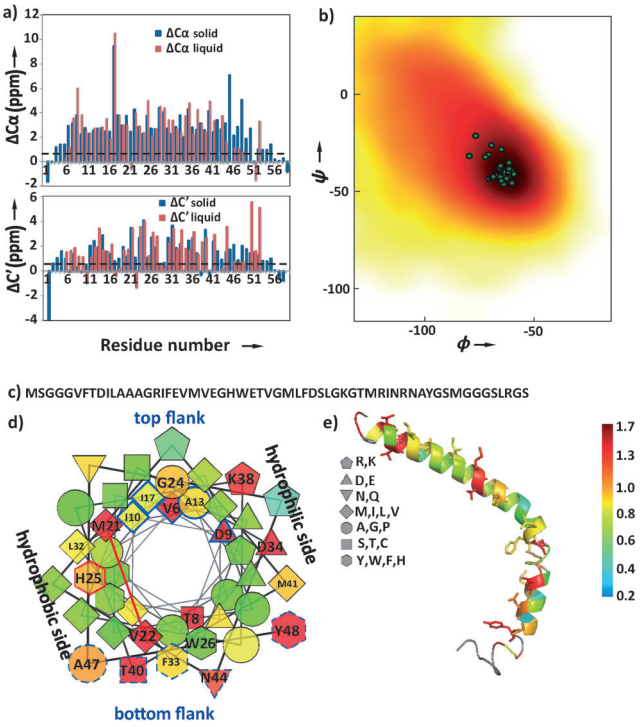

Figure 3. Structural information based on chemical shifts. a) Secondary chemical shifts (observed minus random coil value) from liquid- and solid-state NMR spectra. Cut-off values from the chemical shift index (CSI)[20] corresponding to a significant indication of α-helical structure are shown as dashed lines. b) Predicted dihedral angles by TALOS + for CsmA shown as green dots overlayed on the α-helical region of the Ramachandran density plot based on statistical backbone dihedral angle data for 679 proteins. The brown, red, yellow, and white regions represent progressively less populated regions. c) Amino acid sequence of CsmA. d) Helical projection of the α-helical structure. e) The solution structure of CsmA highlighting residues with substantial differences in chemical shift with sticks. Colors in (d) and (e) correspond to scaled rms values of chemical shift deviations for each residue between solid- and liquid-state chemical shifts as defined in the Supporting Information. In all panels SHIFTX2[22] was used to recalibrate a chemical shift offset for ^{13}C and ^{15}N for the two datasets to compensate for differences in the sample and buffer conditions.

is very similar for the CsmA monomer and the CsmA baseplate oligomeric structure.

Based on the chemical shift analysis, we believe that the differences between liquid- and solid-state NMR data are mainly due to changes in side-chain conformation or tertiary structure, that is, contacts between CsmA monomers in the baseplate or between CsmA and BChl *a* or carotenoids. To quantify the observed differences in chemical shifts, a similarity measure was calculated as a weighted root-mean-square deviation (rmsd) between liquid- and solid-state chemical shifts for N, C′, Cα, and Cβ for each residue (the definition is given in the legend to Figure S3 in the Supporting Information). Based on this similarity measure we can pinpoint the residues having the largest combined differences in chemical shifts as visualized in Figures 3 d,e, and S3. Residues with large differences appear in four separate patterns (Figure 3 d,e) having differences significantly larger than expected within a secondary element and larger than expected due the differences in the solvent (on the order of about 0.2 ppm).[23] The largest differences are found at the N and C termini, which could be due to quenched mobility by confinement in the baseplate. Indeed, prediction of the order parameter S^2 by the RCI method[20] implemented in TALOS + shows a rigid central region of CsmA in both the monomeric and oligomeric form, but for the solution structure a significantly higher mobility is predicted at the termini (see Figure S5 in the Supporting Information). Furthermore, the largest differences within these termini and extending along the helix were located at a particular side of the helix as highlighted on the helix-projection (top and bottom flanks in Figure 3 d) suggesting interfacial binding. The largest difference outside the termini was found for V22 and further differences for M21, G24, and H25. We propose that this could be due to coordination of BChl *a* to the histidine in the putative binding motif G24-H25-W26,[13] which places the ligand in the proximity of M21 and V22, possibly facilitated by the flexible G24 backbone. Lastly, a significant difference was found for D34 and K38, which are part of the putative flexible linker in the solution structure that might indicate a small difference in the backbone conformation at this part of the helix.

We have presented a solid-state NMR analysis of the CsmA–BChl *a* protein–pigment complex in the baseplate antenna complex of *Chlorobaculum tepidum*. This study demonstrates the applicability of solid-state NMR spectroscopy to study proteins in their native and highly heterogeneous environment. By using different combinations of 2D and 3D spectra, 90 % of the CsmA resonances were assigned. Analysis of the assigned chemical shift data indicated a canonical α-helical structure for CsmA and a fully symmetric conformation of the CsmA molecules in the baseplate. Comparison with the liquid-state chemical shifts provided clues for the relative orientation of the CsmA monomers and interaction with the BChl *a* ligand. In more general terms, our study demonstrates that NMR spectroscopy provides unprecedented opportunities to compare the structure of a protein in its isolated form and in its native environment containing all natural lipids and cofactors or to manipulate environmental conditions and subsequently study conformational changes of proteins.

Received: February 12, 2012
Revised: May 16, 2012
Published online: June 8, 2012

Keywords: NMR spectroscopy · photosynthesis · proteins · solid-state structures

[1] http://www.rcsb.org/pdb/general_information/news_publications/annual_reports/annual_report_year_2011.pdb.
[2] a) G. Zanetti, J. A. Briggs, K. Grunewald, Q. J. Sattentau, S. D. Fuller, *PLoS Pathog.* **2006**, *2*, e83; b) F. Brandt, L. A. Carlson, F. U. Hartl, W. Baumeister, K. Grunewald, *Mol. Cell* **2010**, *39*, 560 – 569.
[3] a) A. C. Sivertsen, M. J. Bayro, M. Belenky, R. G. Griffin, J. Herzfeld, *J. Mol. Biol.* **2009**, *387*, 1032 – 1039; b) A. Goldbourt, B. J. Gross, L. A. Day, A. E. McDermott, *J. Am. Chem. Soc.*

2007, *129*, 2338 – 2344; c) R. Fu, X. Wang, C. Li, A. N. Santiago-Miranda, G. J. Pielak, F. Tian, *J. Am. Chem. Soc.* **2011**, *133*, 12370 – 12373.
[4] a) T. Vosegaard, M. Kamihira-Ishijima, A. Watts, N. C. Nielsen, *Biophys. J.* **2008**, *94*, 241 – 250; b) M. Kamihira, T. Vosegaard, A. J. Mason, S. K. Straus, N. C. Nielsen, A. Watts, *J. Struct. Biol.* **2005**, *149*, 7 – 16.
[5] A. C. Sivertsen, M. J. Bayro, M. Belenky, R. G. Griffin, J. Herzfeld, *Biophys. J.* **2010**, *99*, 1932 – 1939.
[6] T. Jacso, W. T. Franks, H. Rose, U. Fink, J. Broecker, S. Keller, H. Oschkinat, B. Reif, *Angew. Chem.* **2012**, *124*, 447 – 450; *Angew. Chem. Int. Ed.* **2012**, *51*, 432 – 435.
[7] M. Renault, R. Tommassen-van Boxtel, M. P. Bos, J. A. Post, J. Tommassen, M. Baldus, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4863 – 4868.
[8] a) R. E. Blankenship, *Molecular mechanisms of photosynthesis*, Blackwell Science Oxford, Malden, MA, **2002**; b) D. A. Bryant, N. U. Frigaard, *Trends Microbiol.* **2006**, *14*, 488 – 496.
[9] a) D. A. Bryant, A. M. Costas, J. A. Maresca, A. G. Chew, C. G. Klatt, M. M. Bateson, L. J. Tallon, J. Hostetler, W. C. Nelson, J. F. Heidelberg, D. M. Ward, *Science* **2007**, *317*, 523 – 526; b) N.-U. Frigaard, D. Bryant, *Microbiol. Monogr.* **2006**, *2*, 79 – 114.
[10] R. E. Blankenship, K. Matsuura, *Light-Harvesting Antennas in Photosynthesis*, Kluver Academic Publishers, Dordrecht, **2003**, pp. 195 – 217.
[11] S. Ganapathy, G. T. Oostergetel, P. K. Wawrzyniak, M. Reus, A. Gomez Maqueo Chew, F. Buda, E. J. Boekema, D. A. Bryant, A. R. Holzwarth, H. J. de Groot, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8525 – 8530.
[12] G. A. Montaño, H. M. Wu, S. Lin, D. C. Brune, R. E. Blankenship, *Biochemistry* **2003**, *42*, 10246 – 10251.
[13] M. Ø. Pedersen, J. Underhaug, J. Dittmer, M. Miller, N. C. Nielsen, *FEBS Lett.* **2008**, *582*, 2869 – 2874.
[14] M. Ø. Pedersen, L. Pham, D. B. Steensgaard, M. Miller, *Biochemistry* **2008**, *47*, 1435 – 1441.
[15] N.-U. Frigaard, G. D. Voigt, D. A. Bryant, *J. Bacteriol.* **2002**, *184*, 3368 – 3376.
[16] N.-U. Frigaard, H. Li, P. Martinsson, S. K. Das, H. A. Frank, T. J. Aartsma, D. A. Bryant, *Photosynth. Res.* **2005**, *86*, 101 – 111.
[17] N.-U. Frigaard, H. Li, K. J. Milks, D. A. Bryant, *J. Bacteriol.* **2004**, *186*, 646 – 653.
[18] T. Egorova-Zachernyuk, B. van Rossum, C. Erkelens, H. de Groot, *Magn. Reson. Chem.* **2008**, *46*, 1074 – 1083.
[19] T. B. Melø, N.-U. Frigaard, K. Matsuura, K. Razi Naqvi, *Spectrochim. Acta Part A* **2000**, *56*, 2001 – 2010.
[20] D. S. Wishart, B. D. Sykes, *J. Biomol. NMR* **1994**, *4*, 171 – 180.
[21] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213 – 223.
[22] B. Han, Y. Liu, S. W. Ginzinger, D. S. Wishart, *J. Biomol. NMR* **2011**, *50*, 43 – 57.
[23] M. L. Tremblay, A. W. Banks, J. K. Rainey, *J. Biomol. NMR* **2010**, *46*, 257 – 270.
[24] T. W. Dillon, D. Lending, T. R. Crews, R. Blankenship, *Comput. Inform. Nurs.* **2003**, *21*, 198 – 205.

