



Published in final edited form as:

Angew Chem Int Ed Engl. 2011 December 9; 50(50): 11942–11946. doi:10.1002/anie.201105648.

Solution NMR Structure of Proteorhodopsin

Sina Reckel⁺,

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632

Daniel Gottstein⁺,

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632

Jochen Stehle,

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. Frank Löhr,

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632

Dr. Mirka-Kristin Verhoefen,

Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue Str. 7, 60438 Frankfurt, Germany

Dr. Mitsuhiro Takeda,

Structural Biology Research Center, Nagoya University, Furo-cho, Chikusa-ku, 464-8601, Japan

Robert Silvers,

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

Prof., Dr. Masatsune Kainosho,

Structural Biology Research Center, Nagoya University, Furo-cho, Chikusa-ku, 464-8601, Japan; Center for Priority Areas, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

Prof., Dr. Clemens Glaubitz,

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632

Prof., Dr. Josef Wachtveitl,

Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue Str. 7, 60438 Frankfurt, Germany

Dr. Frank Bernhard,

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632

⁺These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

Prof., Dr. Harald 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

Prof., Dr. Peter Güntert, and

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632; Frankfurt Institute for Advanced Studies, Goethe University Frankfurt, Ruth-Moufang-Str.1, 60438 Frankfurt am Main, Germany; Center for Priority Areas, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

Prof., Dr. Volker Dötsch

Institute of Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Max-von-Laue Str. 9, 60438 Frankfurt, Germany, Fax: (+49)-69-798-29632

Volker Dötsch: vdoetsch@em.uni-frankfurt.de

Keywords

proteorhodopsin; membrane proteins; NMR spectroscopy; cell-free expression; structural biology

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[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 ~ 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 pH 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 due to the lack of well-diffracting three-dimensional crystals. Recently, the potential of solution NMR to solve structures of helical membrane proteins has been demonstrated[14, 15] and we present in this publication the de novo structure of the green variant of proteorhodopsin solved by solution NMR spectroscopy.

The structure of PR 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) (Figure 1). The seven transmembrane helices are connected by short loops. Instead of the anti-parallel β -sheet that is observed between helix B and C in other microbial rhodopsins, TALOS+-derived torsion angles suggest that PR residues G87-P90 form a short β -turn. The loop between helix D and E is longer than predicted by the secondary structure prediction program TMHMM.[16] In contrast, the loop region connecting helix 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 via 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]

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 (Supporting Information Figure S1). Optimization of NMR conditions allowed the detection of almost all NH backbone resonances (Supporting Information Figure S2). Nonetheless, the large size of the proteo-micelle, limited sensitivity of three-dimensional, triple resonance experiments and 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 (Figure 2 and Supporting Information Figure S3). Based on the backbone chemical shifts, the seven-transmembrane-helix topology of PR was confirmed using the programs TALOS+[20] and CSI[21] (Supporting Information Figures S4). Additionally, $\{^1\text{H}\}^{15}\text{N}$ -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 (Supporting Information Figure S4).

Due to 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. While measurement of backbone $\text{H}^{\text{N}}\text{-H}^{\text{N}}$ NOEs provided only short-range information within the same helix, long-range NOE information was exclusively derived from methyl groups and aromatic side chains. Side chain assignment based on uniform ^{15}N , ^{13}C labeling was complicated by fast transverse ^1H and ^{13}C relaxation and labeling of Ile, Leu and Val methyl groups, based on the use of metabolic precursors,[22] was not applicable, since 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, thus considerably reducing signal overlap (Supporting Information Figure S5). SAIL permitted the side chain assignment of all Val and 60% of the Leu. 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 8 out of 10 Trp residues. Altogether the side chain assignments covered 44% of the transmembrane region (Supporting Information Figure S4) and we were able to extract 137 medium- to long-range NOEs from a 4D [^{13}C , ^{13}C]-separated NOESY 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 and Wagner,[25] resulting in 290 upper and 716 lower distance limits (Supporting Information Figure S6). 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 (Supporting Information Table S1). 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 (Supporting Information Figure S7).

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 -labeled retinal. While the olefinic group in position C_{11} showed only a very weak and overlaid signal, the C_{20} methyl group had a reasonable signal-to-noise ratio and no overlap. We thus identified 6 retinal-to-protein NOEs which were used in the structure calculation (Supporting Information Figure S8). Notably, the $\text{C}_{\delta 1}$ methyl group of L105 is one out of two methyl groups showing an NOE to the C_{20} 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. Since the retinal-to-protein distance restraints could only be derived from position C_{20} located in proximity to the Schiff base, accurate alignment of the cofactor within the protein was not achieved. Also, due to 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] (Supporting Information Figure S8) 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). 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 anti-parallel β -sheet. Another interesting element in the structure of PR constitutes the unexpectedly short E-F loop that comprises only 4 residues including a proline while 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] While 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, however, demonstrated that this biochemical information is consistent with our structure. Certainly, within the retinal-binding protein family, PR remains a special case as despite considerable efforts, its crystal structure remains elusive. 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

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. 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. Side chain assignment and NOESY experiments required the use of deuterated detergent. PRE experiments to derive long-distance restraints were performed using MTSL labeled monocysteine 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.wis.edu (accession code 17327). Further details are provided in the Supporting Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DO545/7-1 and SFB 807), the European Drug Initiative on Channels and Transporters (EDICT) Contract No. HEALTH-F4-2007-201924, NIH (U54 GM094608), the Center for Biomolecular Magnetic Resonance (BMRZ) and the Cluster of Excellence Frankfurt (Macromolecular Complexes). The 11,20-13C-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.

References

1. Beja O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF. *Science*. 2000; 289:1902. [PubMed: 10988064]
2. Man D, Wang W, Sabehi G, Aravind L, Post AF, Massana R, Spudich EN, Spudich JL, Beja O, Embo J. 2003; 22:1725. [PubMed: 12682005]
3. Friedrich T, Geibel S, Kalmbach R, Chizhov I, Ataka K, Heberle J, Engelhard M, Bamberg E. *J Mol Biol*. 2002; 321:821. [PubMed: 12206764]
4. DeLong EF, Beja O. *PLoS Biol*. 2010; 8:e1000359. [PubMed: 20436957]
5. a) Dioumaev AK, Brown LS, Shih J, Spudich EN, Spudich JL, Lanyi JK. *Biochemistry*. 2002; 41:5348. [PubMed: 11969395] b) Dioumaev AK, Wang JM, Balint Z, Varo G, Lanyi JK. *Biochemistry*. 2003; 42:6582. [PubMed: 12767242] c) Verhoefen MK, Neumann K, Weber I, Glaubitz C, Wachtveitl J. *Photochem Photobiol*. 2009; 85:540. [PubMed: 19192201]
6. a) Hempelmann F, Hölper S, Verhoefen MK, Woerner AC, Köhler T, Fiedler SA, Pflieger N, Wachtveitl J, Glaubitz C. *J Am Chem Soc*. 2011; 133:4645. [PubMed: 21366243] b) Pflieger N, Woerner AC, Yang J, Shastri S, Hellmich UA, Aslimovska L, Maier MS, Glaubitz C. *Biochim Biophys Acta*. 2009; 1787:697. [PubMed: 19268651]
7. a) Pflieger N, Lorch M, Woerner AC, Shastri S, Glaubitz C. *J Biomol NMR*. 2008; 40:15. [PubMed: 17968661] b) Shastri S, Vonck J, Pflieger N, Haase W, Kuehlbrandt W, Glaubitz C. *Biochim Biophys Acta*. 2007; 1768:3012. [PubMed: 17964280]
8. Rangarajan R, Galan JF, Whited G, Birge RR. *Biochemistry*. 2007; 46:12679. [PubMed: 17927209]
9. Bergo VB, Sineshchekov OA, Kralj JM, Partha R, Spudich EN, Rothschild KJ, Spudich JL. *J Biol Chem*. 2009; 284:2836. [PubMed: 19015272]
10. Luecke H, Schobert B, Stagno J, Imasheva ES, Wang JM, Balashov SP, Lanyi JK. *Proc Natl Acad Sci U S A*. 2008; 105:16561. [PubMed: 18922772]
11. Lörinczi E, Verhoefen MK, Wachtveitl J, Woerner AC, Glaubitz C, Engelhard M, Bamberg E, Friedrich T. *J Mol Biol*. 2009; 393:320. [PubMed: 19631661]
12. a) Shi L, Ahmed MA, Zhang W, Whited G, Brown LS, Ladizhansky V. *J Mol Biol*. 2009; 386:1078. [PubMed: 19244620] b) Yang J, Aslimovska L, Glaubitz C. *J Am Chem Soc*. 2011; 133:4874. [PubMed: 21401038]
13. Shi L, Lake EM, Ahmed MA, Brown LS, Ladizhansky V. *Biochim Biophys Acta*. 2009; 1788:2563. [PubMed: 19799854]
14. a) Van Horn WD, Kim HJ, Ellis CD, Hadziselimovic A, Sulistijo ES, Karra MD, Tian C, Sönnichsen FD, Sanders CR. *Science*. 2009; 324:1726. [PubMed: 19556511] b) Sobhanifar S, Schneider B, Löhr F, Gottstein D, Ikeya T, Mlynarczyk K, Pulawski W, Ghoshdastider U, Kolinski M, Filipek S, Güntert P, Bernhard F, Dötsch V. *Proc Natl Acad Sci U S A*. 2010; 107:9644. [PubMed: 20445084] c) Butterwick JA, MacKinnon R. *J Mol Biol*. 2010; 403:591. [PubMed: 20851706] d) Berardi MJ, Shih WM, Harrison SC, Chou JJ. *Nature*. 2011; 476:109. [PubMed: 21785437]
15. Gautier A, Mott HR, Bostock MJ, Kirkpatrick JP, Nietlispach D. *Nat Struct Mol Biol*. 2010; 17:768. [PubMed: 20512150]
16. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. *J Mol Biol*. 2001; 305:567. [PubMed: 11152613]
17. Cartailler JP, Luecke H. *Structure*. 2004; 12:133. [PubMed: 14725773]

18. Schwarz D, Junge F, Durst F, Frölich N, Schneider B, Reckel S, Sobhanifar S, Dötsch V, Bernhard F. *Nat Protoc.* 2007; 2:2945. [PubMed: 18007631]
19. a) Hefke F, Bagaria A, Reckel S, Ullrich SJ, Dötsch V, Glaubitz C, Güntert P. *J Biomol NMR.* 2010; 49:75. [PubMed: 21170670] b) Sobhanifar S, Reckel S, Junge F, Schwarz D, Kai L, Karbyshev M, Löhr F, Bernhard F, Dötsch V. *J Biomol NMR.* 2010; 46:33. [PubMed: 19680602] c) Trbovic N, Klammt C, Koglin A, Löhr F, Bernhard F, Dötsch V. *J Am Chem Soc.* 2005; 127:13504. [PubMed: 16190707]
20. Shen Y, Delaglio F, Cornilescu G, Bax A. *J Biomol NMR.* 2009; 44:213. [PubMed: 19548092]
21. Wishart DS, Sykes BD. *J Biomol NMR.* 1994; 4:171. [PubMed: 8019132]
22. Tugarinov V, Kanelis V, Kay LE. *Nat Protoc.* 2006; 1:749. [PubMed: 17406304]
23. Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Mei Ono A, Güntert P. *Nature.* 2006; 440:52. [PubMed: 16511487]
24. Iwahara J, Tang C, Clore GM. *J Magn Reson.* 2007; 184:185. [PubMed: 17084097]
25. Battiste JL, Wagner G. *Biochemistry.* 2000; 39:5355. [PubMed: 10820006]
26. Kralj JM, Spudich EN, Spudich JL, Rothschild KJ. *J Phys Chem B.* 2008; 112:11770. [PubMed: 18717545]
27. a) Ikeda D, Furutani Y, Kandori H. *Biochemistry.* 2007; 46:5365. [PubMed: 17428036] b) Jung JY, Choi AR, Lee YK, Lee HK, Jung KH. *FEBS Letters.* 2008; 582:1679. [PubMed: 18435930]
28. a) Yoshitsugu M, Shibata M, Ikeda D, Furutani Y, Kandori H. *Angew Chem Int Ed Engl.* 2008; 47:3923. [PubMed: 18404767] b) Yoshitsugu M, Yamada J, Kandori H. *Biochemistry.* 2009; 48:4324. [PubMed: 19334675]
29. Güntert P. *Prog Nucl Magn Reson Spectrosc.* 2003; 43:105.

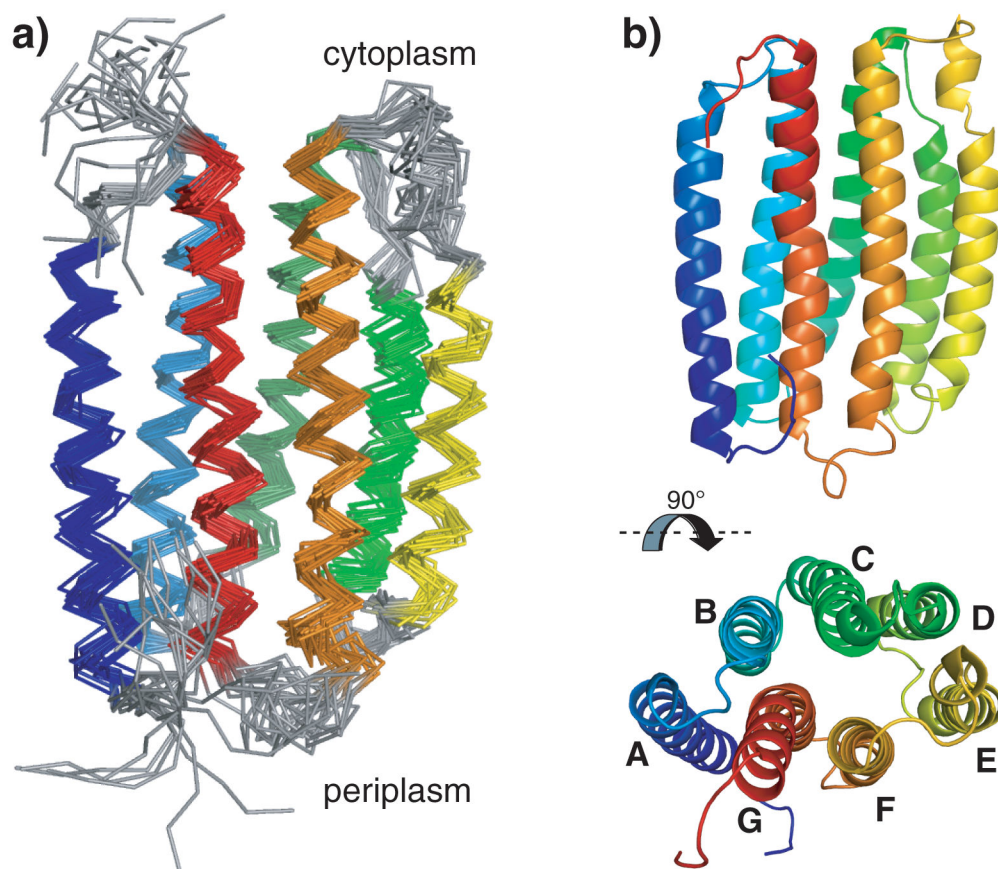


Figure 1. Structure of PR. a) Bundle of the 20 conformers with lowest CYANA target function obtained from structure calculation. Helices are color-coded from helix A in dark blue to helix G in red. b) Cartoon representation of the conformer with the lowest CYANA target function seen from the side and from the top. In the lower panel helices are additionally labeled A-G.

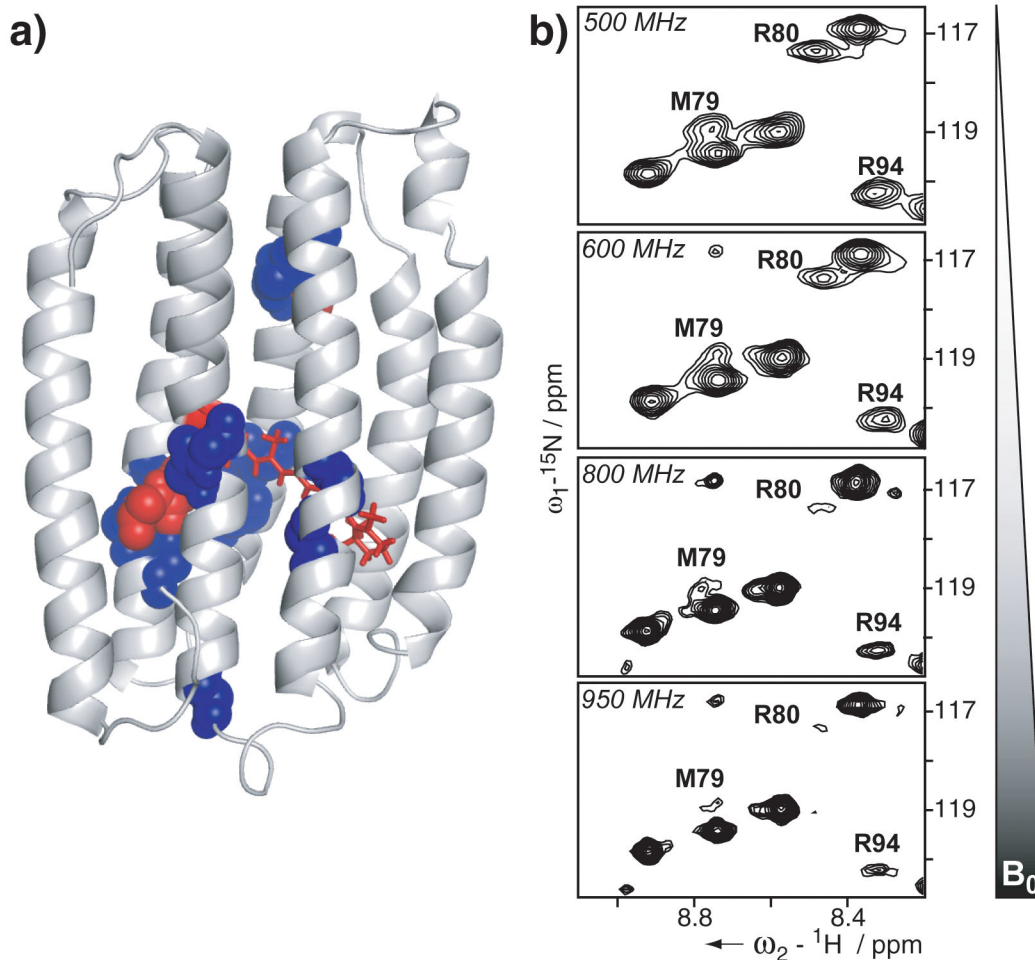


Figure 2.

Assignment and field-dependent line broadening of PR. a) 96% of the backbone resonances of PR were assigned (grey color). Residues that could not be assigned are depicted in red spheres (Y76, M77, C107) and those with partial backbone assignments in blue spheres. b) The field-dependent line broadening is illustrated for three representative resonances. $[\text{^{15}N, ^1H}]$ -TROSY spectra of selectively ^{15}N -MRY-labeled PR were measured at different B_0 field strengths as indicated in each spectrum. While most resonances remain unaffected by the field strength, resonances of R80, M79 and R94 disappear almost completely at 950 MHz.

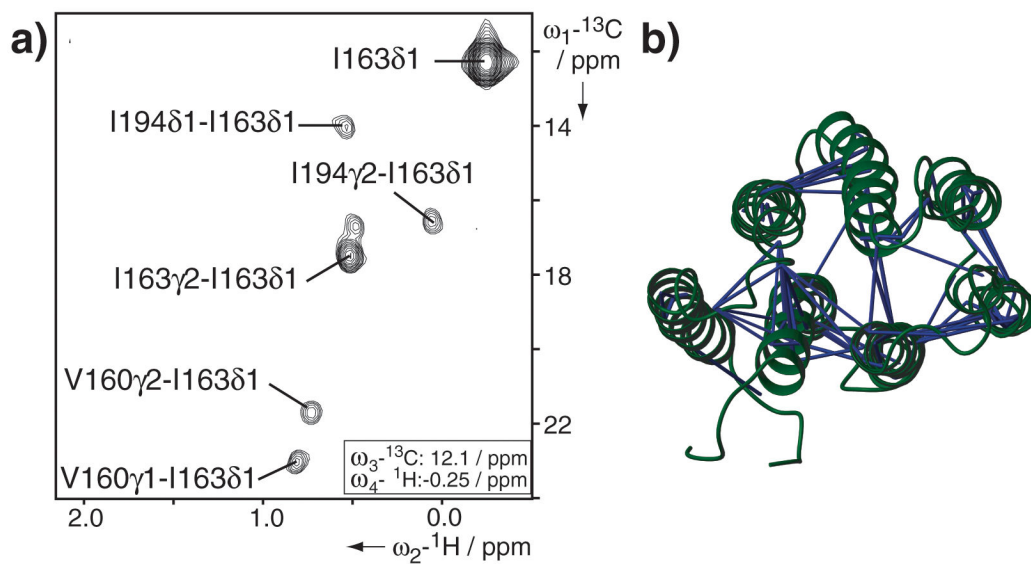


Figure 3.

Distance information for the structure calculation of PR. a) Representative plane of the 4D [^{13}C , ^{13}C]-separated NOESY 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.

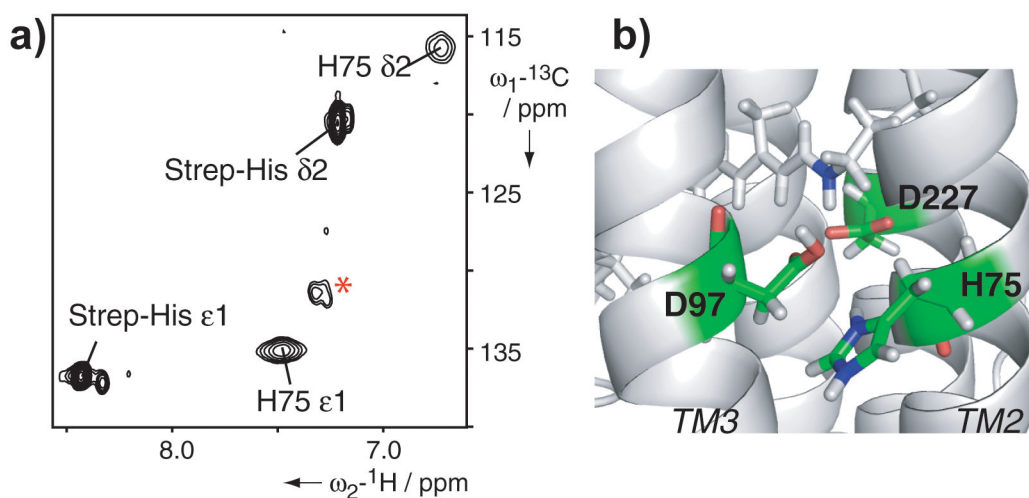


Figure 4.

Schiff base environment in PR. a) [$^{13}\text{C}, ^1\text{H}$]-SOFAST-HMQC of $^{15}\text{N}, ^{13}\text{C}$ -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 in gray sticks, the Schiff base nitrogen highlighted in blue and residues that were included into the modeling colored by element.

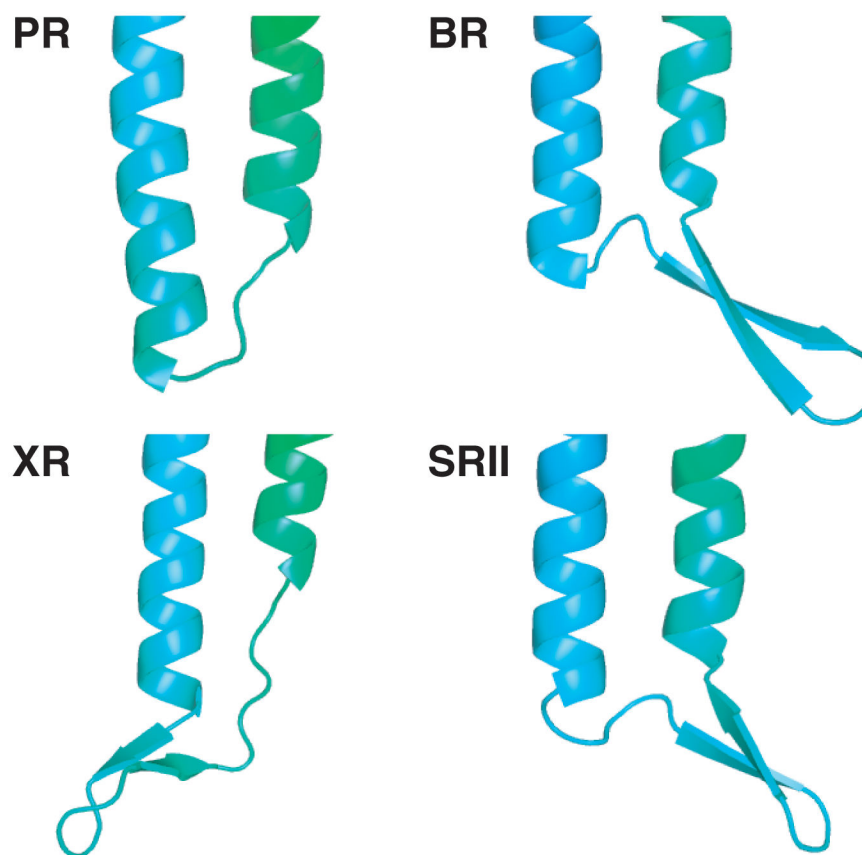


Figure 5. B-C loop of PR and its three homologues bacteriorhodopsin (BR), sensory rhodopsin II (SRII) and xanthorhodopsin (XR) (PDB ID: 1C3W, 2KSY and 3DDL respectively). In PR an anti-parallel β -sheet between helices B and C as observed in the other three structures is not present. Sequence alignment 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 (Supporting Information Figure S8).