Published in final edited form as:

J Struct Funct Genomics. 2013 September ; 14(3): 119–126. doi:10.1007/s10969-013-9159-5.

Solution NMR structures provide first structural coverage of the large protein domain family PF08369 and complementary structural coverage of dark operative protochlorophyllide oxidoreductase complexes

Surya VSRK Pulavarti,

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Yunfen He,

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Erik A. Feldmann,

Department of Chemistry and Biochemistry, Miami University, and Northeast Structural Genomics Consortium, Oxford, OH 45056, USA

Alexander Eletsky,

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Thomas B. Acton,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Rong Xiao,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

John K. Everett,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Gaetano T. Montelione,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA; Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, UMDNJ, Piscataway NJ 08854, USA

Michael A. Kennedy, and

Department of Chemistry and Biochemistry, Miami University, and Northeast Structural Genomics Consortium, Oxford, OH 45056, USA

Thomas Szyperski

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Michael A. Kennedy: kennedm4@muohio.edu; Thomas Szyperski: szypersk@buffalo.edu

Abstract

High-quality NMR structures of the C-terminal domain comprising residues 484-537 of the 537-residue protein Bacterial chlorophyll subunit B (BchB) from *Chlorobium tepidum* and residues 9-61 of 61-residue Asr4154 from *Nostoc sp.* (strain PCC 7120) exhibit a mixed α/β fold comprised of three α -helices and a small β -sheet packed against second α -helix. These two proteins share 29 % sequence similarity and their structures are globally quite similar. The structures of BchB(484-537) and Asr4154(9-61) are the first representative structures for the large protein family (Pfam) PF08369, a family of unknown function currently containing 610 members in bacteria and eukaryotes. Furthermore, BchB(484-537) complements the structural coverage of the dark-operating protochlorophyllide oxidoreductase (DPOR).

Keywords

BchB; DPOR; Asr4154; PF08369; PCP-red; structural genomics

Introduction

The C-terminal domain of 537-residue protein BchB from *Chlorobium tepidum* (UniProtKB accession number Q9F715) and 61-residue protein Asr4154 (Q8YPN9) from *Nostoc sp.* (strain PCC 7120) belong to Pfam [1] protein family PF08369 ('PCP-red domain'), which currently contains 610 members from bacteria and eukaryotic organisms participating in 21 different unique domain organizations (architectures). However, the majority of the members (554) are found in only two of the 21 domain organizations: 484 sequences belong to PCP-red domains located C-terminally in a nitrogenase component 1 type oxidoreductase subunit (Pfam family PF00148), specifically a protochlorophyllide reductase subunit, and 70 sequences are expressed as single PCP-red domain proteins. Notably, the genome of *C. tepidum* encodes three paralogs of PCP-red domains [UniProtKB Q9F715(484-537), Q8KAN1(416-468) and Q8KCD9(170-214)], while the genome of *Nostoc sp.* encodes two paralogs [Q8YRK5(459-503) and Q8YPN9)].

In anoxygenic photosynthetic bacteria, light-independent protochlorophyllide reductases, also known as dark operative protochlorophyllide oxidoreductase (DPOR), catalyze the stereo specific reduction of the C17-C18 double bond of ring D of porphyrin in protochlorophyllide to form chlorophyllide [2]. In these bacteria, such as *C. tepidum*, DPOR is a hetero-octameric complex containing the three proteins BchL, BchN, and BchB [2] (in chlorophyll-synthesizing organisms the corresponding proteins are named ChlL, ChlN and ChlB [3]) which are arranged such that two BchL₂ homodimers (functioning as an ATP-dependent electron carriers comprising [4Fe-4S] clusters [3, 4, 5]) interact with one heterotetrameric (BchN/BchB)₂ complex containing two [4Fe-4S]-clusters and two protochlorophyllide binding sites, respectively.

The X-ray crystal structures of the (ChlN-ChlB)₂, (BchN-BchB)₂- and [(BchN/BchB)₂BchL₂]₂ complexes from, respectively, *Thermosynechococcus elongatus* [5], *Rhodobacter capsulatus* [6] and *Prochlorococcus marinus* [7] shed light on the mechanism for the enzymatic reduction of protochlorophyllide to chlorophyllide, as well as their evolutionary relationship to nitrogenases [5]; electrons are transferred from BchL₂ such that

they reach the active site for protochlorophyllide reduction [3]. For the NMR structure of PCP-red domain BchB(484-537) presented in this paper, it is important to note that (i) the C-terminal residues of ChlB and BchB in the X-ray structures of (ChlN-ChlB)₂ and (BchN-BchB)₂, which belong to a 52-residue linker and the PCP-red domain, could not be located in the electron density maps [3, 6] and that (ii) the structure of the PCP-red domain in [(BchN/BchB)₂BchL₂]₂ differs slightly from the solution structure of BchB(484-537) (see below).

Proteins BchB(484-537) and Asr4154(9-61), which share 29 % sequence identity, were selected as targets of the Protein Structure Initiative (PSI) and then assigned to the Northeast Structural Genomics Consortium (NESG; <http://www.nesg.org>) for structure determination (NESG target IDs CtR69A and NsR143, respectively) as part of a cooperative inter-center effort aimed at providing structural coverage of large, uncharacterized protein domain families [8]. The two proteins belong to the two major domain organizations alluded to above (i.e., PCP-red domains located C-terminally to ChlB/BchB, or PCP-red domains expressed as single proteins). Since no atomic resolution structures were available when these targets were chosen, high modeling leverage [9] could be anticipated. Here, we report the high-quality solution NMR structures of BchB(484-537) and Asr4154(9-61), which represent the first structures available for Pfam family PF08369.

Materials and Methods

BchB(484-537) and Asr4154(9-61) were cloned, expressed, and purified following protocols [10, 11, 12] established by NESG (see Supplementary Material for details). The constructs included short C-terminal hexa-His tag facilitating protein purification (LEHHHHHH). The corresponding pET expression vectors (NESG CtR69A-484-537-21.4 and NsR143-9-61-21.16) have been deposited in the PSI Materials Repository (<http://psimr.asu.edu/>). Protein samples were prepared at ~1 mM and ~0.8 mM concentrations for BchB(484-537) and Asr4154(9-61), respectively in 90% H₂O/10% D₂O, in buffer solutions containing 20 mM MES, along with 100 mM NaCl, 5 mM CaCl₂, 10 mM DTT, 50 mM DSS and 0.02% NaN₃ at pH 6.5. A [5% ¹³C; U-¹⁵N]-labeled samples enabled stereospecific assignments of the methyl groups of Val and Leu residues [13]. Isotropic overall rotational correlation times of ~4.5 ns and ~5.6 ns were inferred, respectively, from ¹⁵N spin relaxation times for BchB(484-537) and Asr4154(9-61) (Supplementary Material; <http://www.nmr2.buffalo.edu/nescg.wiki>), indicating that both proteins are monomeric in solution. This finding was confirmed by analytical gel-filtration (Supplementary Figs. S1 and S2).

NMR data were acquired at 25 °C on Varian INOVA 600 and 750 spectrometers for BchB(484-537) and at 20 °C on Varian INOVA and Bruker AVANCE 850 spectrometers for Asr4154(9-61). The VARIAN spectrometers were equipped with cryogenic 1H{¹⁵N,¹³C} probes. The total measurement times for BchB(484-537) and Asr4154(9-61) were ~4.5 days and ~29 days, respectively. Nearly complete sequence-specific ¹H, ¹³C and ¹⁵N resonance assignments (Table 1; Supplementary Figs. S3 and S4) were obtained from conventional triple-resonance NMR experiments (Supplementary Material) using the programs AutoAssign [14, 15] and PINE [16], followed by manual assignment for side-chain resonances. Assignments were validated by AVS software suite [17]. Chemical shifts, NOESY peak lists and time domain NMR data were deposited in the BioMagResBank [accession numbers 16649 and 17035 for BchB(484-537) and Asr4154(9-61), respectively].

Structure calculations were performed using standardized methods of the NESG consortium [18, 19] (<http://www.nmr2.buffalo.edu/nescg.wiki>) and consensus analysis of automated NOESY cross peak assignments provided by the programs CYANA [20, 21] and AutoStructure [22] based on ¹H-¹H NOE-derived upper limit distance constraints, and back-

bone and dihedral angle constraints derived from chemical shifts using the program TALOS + [23] for residues located in well-defined regular structural elements. Stereospecific assignments for methylene protons for BchB(484-537) were performed with the GLOMSA module of CYANA. The final structure calculation were performed on CYANA and XPLOR-NIH 2.20 for BchB(484-537) and Asr4154(9-61), respectively, followed by refinement in an 'explicit water bath' using the program CNS [24]. Structure validation was accomplished using the Protein Structure Validation Software (PSVS) server 1.3 [25], and agreement of structures and NOESY peak lists was verified using the AutoStructure RPF 2.2.2 package [18]. The novel modeling leverage of the NMR structures of BchB(484-537) and Asr4154(9-61) was estimated using the program PSI-BLAST [26] as described in the Supplementary Material.

Results and Discussion

High-quality solution NMR structures of BchB(484-537) and Asr4154(9-61) were obtained (Table I) and the coordinates were deposited in the Protein Data Bank [27] on 12/22/2009 (accession code 2KRU) and 06/30/2010 (accession code 2L09), respectively. Both structures exhibit a mixed α/β fold (Fig. 1), consisting of three α -helices and a short two-stranded parallel β -sheet. The location of secondary structure elements are: β -strand 1(487-488 in BchB/11-12 in Asr4154), α -helix I (490-497/14-22), α -helix II (501-518/27-42), β -strand 2(522-523/46-47), α -helix III (525-536/49-61). Notably, a 3_{10} -helix was observed for residues 24-28 in Asr4154(9-61) where Phe 25 and Phe 26, located right before α -helix II, adopt the backbone dihedral angles of a 3_{10} -helix ($=-60$, $=-30$ for the $i+1$ and $i+2$ residues, respectively), and a hydrogen bond is observed between the carbonyl oxygen of Pro24 and the NH of Arg 28, as expected for a 3_{10} -helix, in all 20 members of the ensemble. The relative orientation of the α -helices resembles a bundle with the β -sheet being packed against α -helix II. As expected for 29 % sequence identity, the two structures are quite similar: the root mean square deviation (RMSD) calculated between the mean coordinates of the backbone heavy atoms N, C $^{\alpha}$, and C' of the regular secondary structure elements is 2.0 Å.

A search of the PDB [27] for structurally similar proteins using the DALI server [28] yielded a large number of hits spanning a quasi-continuum of Z-scores > 4.0 (i.e., > 122 and > 322 for BchB(484-537) and Asr4154(9-61), respectively). This finding alone indicates that the rather simple three-helix bundle motif (Fig. 1) emerged several times independently during evolution. For BchB(484-537), the highest Z-scores were obtained for RuvB like proteins, i.e domain II of holliday junction DNA resolvase RuvB from *Thermatoga maritima* (PDB ID: 1IN7, chain-A; Z-score of 7.3) and variants, domain II of *Campylobacter jejuni* holliday junction DNA helicase (PDB ID: 3PFI) and variants, and *Pyrococcus horikoshii* DNA polymerase II subunit (PDB ID: 2KXE). Likewise, for Asr4154(9-61) the highest scores were obtained for RuvB like proteins, i.e domain III of human TIP49b (RuvB-like 2, PDB code 3UK6, chain-E; Z-score of 8.4), domain II of human RUVB-like1 (PDB ID: 2C9O) and its variants, and *Pyrococcus horikoshii* DNA polymerase II subunit (PDB ID: 2KXE). Such RuvB like proteins are classified as AAA+ ATPases, which are chaperonine-like ATPases associated with various cellular functions [29]. AAA+ ATPases typically exhibit an ATP hydrolyzing N-terminal domain and a C-terminal domain with the interface between the two domains forming a binding pocket for ATP [29]. Specifically, α -helices of domain II of *Thermatoga maritima* RuvB [30] and domain III of human TIP49b [31] align structurally with α -helices I-III of BchB(484-537) and Asr4154(9-61) respectively. However, visual inspection shows that both proteins lack a polypeptide segment that could match the location of α -helix I in domain II of RuvB and domain III of TIP49b (Supplementary Fig. S5a and S5b). Hence, it appears likely that BchB(484-537) and Asr4154(9-61) do not bind ATP and are not functionally related to

AAA+ ATPases. This conjecture was confirmed by the observation that no backbone amide chemical shift perturbations are registered upon titration of BchB(484-537) with ATP (Supplementary Fig. S6). Taken together, the DALI-based search for structurally similar protein did not result in additional insights into function.

Electrostatic surface potential calculations show that BchB(484-537) exhibits a mixed surface charge distribution (Fig. 1e), and analysis using Mark-Us/SCREEN2 [32, 33] reveals the presence of a surface cleft 'C' (surface area 42 Å²) formed between the three α-helices formed by residues Val 499, Val 503, Lys 506, Val 507, Lys 532, Leu 535, and Gly 536 (Fig. 1d). Likewise, Asr4154(9-61) exhibits a mixed electrostatic surface distribution (Fig. 1k) which, however, is quite different from BchB(484-537). Moreover, the surface of Asr4154(9-61) does not exhibit any larger surface clefts. These findings indicate that BchB(484-537) and Asr4154(9-61) may function quite differently.

A structure based prediction of protein-protein interaction sites [34, 35] using the server PredUs revealed that BchB(484-537) may interact with other proteins (Fig 1f). Since it represents the C-terminal domain of full-length BchB forming (BchN-BchB)₂ in *C. tepidum*, it appears likely that BchB(484-537) interacts with the remainder of the (BchN-BchB)₂ complex. This view is supported by the X-ray crystal structures of the (BchN-BchB)₂ and [(BchN/BchB)₂BchL₂]₂ complexes from, respectively, *R. capsulatus* [6] (Fig. S7) and *P. marinus* [7], when also considering that the PCP-red domains in these complexes exhibit 42% and 44% sequence identity with BchB(484-537). First, prediction of protein-protein interaction sites based on the X-ray structure of (BchN-BchB)₂ from *R. capsulatus* (which does not contain the PCP-red domain) not only highlights surfaces involved in the hetero-tetramer formation but also additional smaller patches that might interact with the C-terminal PCP-red domain BchB(484-537) (patch 'P1' in Fig. S7) and/or BchL₂ (patch 'P2' in Fig. S7). Second, inspection the [(BchN/BchB)₂BchL₂]₂ complex structure from *P. marinus* (Fig. 2a) shows that (i) the corresponding patch 'P2' (Fig. S7) does indeed interact with the BchL₂ dimer, and that (ii) corresponding patch 'P1' is in close spatial proximity to the PCP-red domain albeit not identical with the actual interaction site. Notably, Moser et al. suggest that the PCP-red domain closes the substrate binding site at the BchN-BchB interface to possibly prevent Pchl_{ide}-induced photodynamic damage [7]. When considering the flexible nature of the linker (not located in the electron density map) attaching the PCP-red domain to the remainder of BchB, it may thus be that the PCP-red domain adopts different orientations at different stages of the catalytic cycle [3, 4, 5]. Moreover, superposition of the PCP-red domain structure of [(BchN/BchB)₂BchL₂]₂ with the NMR structure of BchB(484-537) (Fig. 2b) reveals that they are quite similar: the RMSD calculated for the backbone heavy atoms N, C^α, and C' of the -helices is only 1.15 Å. The only minor structural difference is related to the small -sheet present in BchB(484-537) but not in the PCP-red domain from *P. marinus* (Fig. 2c). This might be a consequence of the two proteins having different amino acid sequences (Fig. S8).

Finally, the novel structural leverage, that is, the number of new protein structures that can be reliably modeled [9] using the experimental structures BchB(484-537) and Asr4154(9-61) is 391 and 260, respectively, when considering the 515 unique sequences of PF08369. Since the modeling families for the two structures overlap, they result in an estimated structural coverage of ~80 % for PF08369.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Donald Petrey, Columbia University, for helpful discussions. This work was supported by the National Institutes of Health, grant number: U54 GM094597 (T.S., M.K. and G.T.M.).

References

1. Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A. The Pfam protein families database. *Nucleic Acids Res.* 2010; 38:D211–D222. [PubMed: 19920124]
2. Reinbothe C, Bakkouri ME, Buhr F, Muraki N, Nomata J, Kurisu G, Fujita Y, Reinbothe S. Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. *Trends Plant Sci.* 2010; 15:614–624. [PubMed: 20801074]
3. Bröcker MJ, Virus S, Ganskow S, Heathcote P, Heinz DW, Schubert WD, Jahn D, Moser J. ATP-driven reduction by dark-operative protochlorophyllide oxidoreductase from *Chlorobium tepidum* mechanistically resembles nitrogenase catalysis. *J Biol Chem.* 2008; 283:10559–10567. [PubMed: 18252716]
4. Sarma R, Barney BM, Hamilton TL, Jones A, Seefeldt LC, Peters JW. Crystal structure of the L protein of *Rhodobacter sphaeroides* light-independent protochlorophyllide reductase with MgADP bound: a homologue of the nitrogenase Fe Protein. *Biochemistry.* 2008; 47:13004–13015. [PubMed: 19006326]
5. Bröcker MJ, Schomburg S, Heinz DW, Jahn D, Schubert WD, Moser J. Crystal structure of the nitrogenase-like dark operative protochlorophyllide oxidoreductase catalytic complex (ChlN/ChlB)₂. *J Biol Chem.* 2010; 285:27336–27345. [PubMed: 20558746]
6. Muraki N, Nomata J, Ebata K, Mizoguchi T, Shiba T, Tamiaki H, Kurisu G, Fujita Y. X-ray crystal structure of the light-independent protochlorophyllide reductase. *Nature.* 2010; 465:110–114. [PubMed: 20400946]
7. Moser J, Lange C, Krausze J, Rebelein J, Schubert WD, Ribbe MW, Heinz DW, Jahn D. Structure of ADP-aluminium fluoride-stabilized protochlorophyllide oxidoreductase complex. *Proc Natl Acad Sci U S A.* 2013; 110:2094–2098. [PubMed: 23341615]
8. Dessailly BH, Nair R, Jaroszewski L, Fajardo JE, Kouranov A, Lee D, Fiser A, Godzik A, Rost B, Orengo C. PSI-2: structural genomics to cover protein domain family space. *Structure.* 2009; 17:869–881. [PubMed: 19523904]
9. Liu JF, Montelione GT, Rost B. Novel leverage of structural genomics. *Nat Biotechnol.* 2007; 25:850–853.
10. Acton TB, Gunsalus KC, Xiao R, Ma LC, Aramini J, Baran MC, Chiang YW, Climent T, Cooper B, Denissova NG, Douglas SM, Everett JK, Ho CK, Macapagal D, Rajan PK, Shastry R, Shih LY, Swapna GVT, Wilson M, Wu M, Gerstein M, Inouye M, Hunt JF, Montelione GT. Robotic cloning and protein production platform of the Northeast Structural Genomics Consortium. *Methods Enzymol.* 2005; 394:210–243. [PubMed: 15808222]
11. Acton TB, Xiao R, Anderson S, Aramini J, Buchwald WA, Ciccocanti C, Conover K, Everett J, Hamilton K, Huang YJ, Janjua H, Kornhaber G, Lau J, Lee DY, Liu GH, Maglaqui M, Ma LC, Mao L, Patel D, Rossi P, Sahdev S, Shastry R, Swapna GVT, Tang YF, Tong SC, Wang DY, Wang H, Zhao L, Montelione GT. Preparation of protein samples for NMR structure, function, and small-molecule screening studies. *Fragment-Based Drug Design: Tools, Practical Approaches, and Examples.* 2011; 493:21–60.
12. Xiao R, Anderson S, Aramini J, Belote R, Buchwald WA, Ciccocanti C, Conover K, Everett JK, Hamilton K, Huang YJ, Janjua H, Jiang M, Kornhaber GJ, Lee DY, Locke JY, Ma LC, Maglaqui M, Mao L, Mitra S, Patel D, Rossi P, Sahdev S, Sharma S, Shastry R, Swapna GVT, Tong SN, Wang D, Wang H, Zhao L, Montelione GT, Acton TB. The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. *J Struct Biol.* 2010; 172:21–33. [PubMed: 20688167]
13. Neri D, Szyperki T, Otting G, Senn H, Wuethrich K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434

- repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry*. 1989; 28:7510–7516. [PubMed: 2692701]
14. Moseley HNB, Monleon D, Montelione GT. Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. *Methods Enzymol*. 2001; 339:91–108. [PubMed: 11462827]
 15. Zimmerman DE, Kulikowski CA, Huang YP, Feng WQ, Tashiro M, Shimotakahara S, Chien CY, Powers R, Montelione GT. Automated analysis of protein NMR assignments using methods from artificial intelligence. *J Mol Biol*. 1997; 269:592–610. [PubMed: 9217263]
 16. Bahrami A, Assadi AH, Markley JL, Eghbalian HR. Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy. *PLoS Comput Biol*. 2009; 5:e1000307. [PubMed: 19282963]
 17. Moseley HNB, Sahota G, Montelione GT. Assignment validation software suite for the evaluation and presentation of protein resonance assignment data. *J Biomol NMR*. 2004; 28:341–355. [PubMed: 14872126]
 18. Huang YJ, Powers R, Montelione GT. Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. *J Am Chem Soc*. 2005; 127:1665–1674. [PubMed: 15701001]
 19. Liu G, Shen Y, Atreya HS, Parish D, Shao Y, Sukumaran DK, Xiao R, Yee A, Lemak A, Bhattacharya A, Acton TA, Arrowsmith CH, Montelione GT, Szyperski T. NMR data collection and analysis protocol for high-throughput protein structure determination. *Proc Natl Acad Sci U S A*. 2005; 102:10487–10492. [PubMed: 16027363]
 20. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol*. 1997; 273:283–298. [PubMed: 9367762]
 21. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol*. 2002; 319:209–227. [PubMed: 12051947]
 22. Huang YJ, Tejero R, Powers R, Montelione GT. A topology-constrained distance network algorithm for protein structure determination from NOESY data. *Proteins*. 2006; 62:587–603. [PubMed: 16374783]
 23. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR*. 1999; 13:289–302. [PubMed: 10212987]
 24. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr Sect D Biol Crystallogr*. 1998; 54:905–921. [PubMed: 9757107]
 25. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. *Proteins*. 2007; 66:778–795. [PubMed: 17186527]
 26. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997; 25:3389–3402. [PubMed: 9254694]
 27. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res*. 2000; 28:235–242. [PubMed: 10592235]
 28. Holm L, Sander C. Dali: a network tool for protein structure comparison. *Trends Biochem Sci*. 1995; 20:478–480. [PubMed: 8578593]
 29. Neuwald AF, Aravind L, Spouge JL, Koonin EV. AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*. 1999; 9:27–43. [PubMed: 9927482]
 30. Putnam CD, Clancy SB, Tsuruta H, Gonzalez S, Wetmur JG, Tainer JA. Structure and mechanism of the RuvB holliday junction branch migration motor. *J Mol Biol*. 2001; 311:297–310. [PubMed: 11478862]
 31. Petukhov M, Dagkessamanskaja A, Bommer M, Barrett T, Tsaneva I, Yakimov A, Quéval R, Shvetsov A, Khodorkovskiy M, Käs E, Grigoriev M. Largescale conformational flexibility determines the properties of AAA+ TIP49 ATPases. *Structure*. 2012; 20:1321–1331. [PubMed: 22748767]

32. Nayal M, Honig B. On the nature of cavities on protein surfaces: application to the identification of drug-binding sites. *Proteins*. 2006; 63:892–906. [PubMed: 16477622]
33. Petrey D, Fischer M, Honig B. Structural relationships among proteins with different global topologies and their implications for function annotation strategies. *Proc Natl Acad Sci U S A*. 2009; 106:17377–17382. [PubMed: 19805138]
34. Zhang QC, Deng L, Fisher M, Guan J, Honig B, Petrey D. PredUs: a web server for predicting protein interfaces using structural neighbors. *Nucleic Acids Res*. 2011; 39:W283–W287. [PubMed: 21609948]
35. Zhang QC, Petrey D, Norel R, Honig BH. Protein interface conservation across structure space. *Proc Natl Acad Sci U S A*. 2010; 107:10896–10901. [PubMed: 20534496]

Abbreviations

BchB	Bacterial chlorophyll subunit B
BchN	Bacterial chlorophyll subunit N
ChlN	Chlorophyll subunit N
ChlB	Chlorophyll subunit B
DPOR	Dark-operative protochlorophyllide oxidoreductase
DSS	4,4-dimethyl-4-silapentane-1-sulfonate sodium salt
DTT	Dithiothreitol
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
NESG	Northeast Structural Genomics Consortium
NOE	Nuclear Overhauser effect
PCP-red	Protochlorophyllide reductase
PDB	Protein Data Bank
RMSD	Root mean square deviation

Fig. 1.

a Stereo view of the 20 conformers representing the solution structure of BchB(484-537) obtained after superposition of the C α atoms of the regular secondary structure elements for minimal RMSD. Residues 484-486 and 538-544 of the disordered N- and C-terminal polypeptide segments were omitted for clarity and the termini are labeled as 'N' and 'C'. **b** Ribbon diagram of residues 487-537 of the lowest-energy conformer of BchB(484-537): α -helices are shown in red and yellow, β -strands are depicted in cyan, other polypeptide segments are in gray. **c** Sausage representation of backbone and superposition of the conformation of the best defined side chains (Table 1). A spline curve was drawn through the mean positions of C α atoms of residues 487-537 with the thickness proportional to the mean global displacement of C α atoms in the 20 conformers superimposed in **(a)**. **d** Surface and space-filling representations of the lowest-energy conformer of BchB(484-537) colored according to the degree of residue conservation. The default ConSurf color scheme for residue conservation is employed: burgundy for the strongest conservation and cyan for the highest variability. The surface cavity is labeled with a 'C'. **e** The electrostatic surface potential map of lowest energy NMR conformer. Positively and negatively charged regions are colored in blue and red, respectively, and the neutral regions are colored in white. **f** Surface and space-filling representations of the lowest-energy conformer of BchB(484-537) colored according scores from PredUs prediction of regions potentially involved in protein-protein interactions: the default PredUs color scheme for residues with scores larger than

zero are shown from light red to red with increasing score. **g-l** Show the same as in **a-f** but for Asr4154(9-61) (terminal residues 9-10 and 61-69 were excluded for clarity).

Fig. 2.

a The (BchN-BchB)₂ from *R. capsulatus* (PDB ID: 3AEK) was superimposed on the (BchN-BchB)₂ unit of *P. marinus* DPOR complex (PDB ID: 2YNM) and is shown as a space filling representation colored according to scores from PredUs prediction for regions potentially involved in protein-protein interaction (residues with scores larger than zero are shown from light red to red with increasing score; see Fig. S7). The patches labeled with 'P1' are close to Pchl_a binding site and might interact with the PCP-red domains at different stages of the catalytic cycle (see text). The PCP-red domains of the *P. marinus* DPOR complex are displayed as ribbon drawings colored in gold, while other domains in the structure of the *P. marinus* DPOR complex are in gray. **b** Ribbon representation of the lowest energy conformer of BchB(484-537) (PDB ID: 2KRU; α -helices and β -sheet are represented in blue and cyan respectively) superimposed on crystal structure of PCP-red domain from *P. marinus* (2YNM, chain D; gold). The structures were superimposed according to DALI alignment by minimizing the RMSD of the C $^{\alpha}$ atoms of the following residues: 490-497, 501-517, 525-534 of BchB(484-537) and 483-490, 494-510, 518-527 of *P. marinus* PCP-red domain. **c** Comparison of the backbone dihedral angles ϕ and ψ observed in the BchB(484-537) solution NMR structure and the *P. marinus* PCP-red domain structure. The values of the ensemble of NMR conformers are shown in blue with boxes drawn around the observed range, and the values observed in the crystal structure are shown in red. While the polypeptide segment corresponding to β -strand I in BchB(484-537) is likewise in an extended conformation in the *P. marinus* PCP-red domain, this is not the case

for the segment corresponding to β -strand I in BchB(484-537). The amino acid residue numberings for BchB(484-537) and the PCP-red domain from *P. marinus* are provided in black (bottom) and red (top), respectively.

Table I
Statistics for NMR Structures of BchB(484-537) and Asr4154(9-61)

	BchB(484-537)	Asr4154(9-61)		
Completeness of resonance assignments ^b				
Backbone (%)	99.0	99.2		
Side chain (%)	98.7	97.0		
Aromatic (%)	91.3	84.2		
Stereospecific methyl (%)	100	100		
Conformationally-restricting constraints ^c				
Distance constraints				
Total	1013	1237		
intra-residue (<i>i</i> = <i>j</i>)	219	276		
sequential ($ i-j = 1$)	259	264		
medium range ($1 < i - j < 5$)	273	333		
long range ($ i - j \geq 5$)	262	364		
Dihedral angle constraints	70	86		
Hydrogen bond constraints	0	0		
No. of long range constraints per residue	4.8	6.6		
Residual constraint violations ^c				
Average no. of distance violations per structure:				
0.1 - 0.2 Å	3.45	1.2		
0.2 - 0.5 Å	0.25	0		
> 0.5 Å	0	0		
Average no. of dihedral angle violations per structure:				
1 - 10°	1.15	0.45		
> 10°	0	0		
Model Quality ^c				
RMSD backbone atoms (Å) ^d	0.4	0.3		
RMSD heavy atoms (Å) ^d	1.0	0.9		
RMSD bond lengths (Å)	0.003	0.004		
RMSD bond angles (°)	0.5	0.6		
MolProbity Ramachandran statistics ^{c,d}				
Most favored regions (%)	99.80	97.20		
Allowed regions (%)	0.20	2.80		
Disallowed regions (%)	0.00	0.00		
Global quality scores (Raw/Z-score) ^c				
Verify3D	0.30	-2.57	0.29	-2.73
ProsaII	0.54	-0.45	0.81	0.66
Procheck (phi-psi) ^d	0.33	1.61	0.33	1.61
Procheck (all) ^d	0.25	1.48	0.23	1.36
MolProbity clash score	16.60	-1.32	20.59	-2.01

	BchB(484-537)		Asr4154(9-61)	
RPF Scores ^e				
Recall/Precision	0.993	0.876	0.887	0.955
F-measure/DP-score	0.931	0.860	0.919	0.793
<i>BMRB accession number</i>	16649		17035	
<i>PDB ID</i>	2KRU		2L09	

^a Structural statistics computed for the ensemble of 20 deposited structures.

^b Computed using AVS software [17] from the expected number of resonances, excluding: highly exchangeable protons (N-terminal, Lys, and Arg amino groups, hydroxyls of Ser, Thr, Tyr), carboxyls of Asp and Glu, non-protonated aromatic carbons, and the C-terminal His6 tag.

^c Calculated using PSVS 1.4 [25]. Average distance violations were calculated using the sum over r^{-6} .

^d Based on ordered residue ranges [S(phi) + S(psi) > 1.8], Residues 487-534 for BchB(484-537) and 11-61 for Asr4154(9-61).

^e RPF scores [18] reflecting the goodness-of-fit of the final ensemble of structures (including disordered residues) to the NOESY data and resonance assignments.