Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Dimerization properties of the *Rp*BphP2 chromophore-binding domain crystallized by homologue-directed mutagenesis

Dom Bellini and Miroslav Z. Papiz*

Institute of Integrative Biology, Liverpool University, Biosciences Building, Crown Street, Liverpool L69 7ZB, England

Correspondence e-mail: mpapiz@liverpool.ac.uk

Received 2 March 2012
Accepted 7 May 2012

PDB Reference:
*Rp*BphP2-CBD*, 4e04.

Bacteriophytochromes (BphPs) are biliverdin IXα-containing photoreceptors that photoconvert between red (Pr) and far-red (Pfr) absorbing states. BphPs are one half of a two-component system that transmits a light signal to a histidine kinase domain and then to a gene-response regulator. In *Rhodopseudomonas palustris*, synthesis of a light-harvesting complex (LH4) is controlled by two BphPs (*Rp*BphP2 and *Rp*BphP3). Despite their high sequence identity (52%), their absorption spectra are very different. The spectra of *Rp*BphP2 exhibit classic Pr-to-Pfr photoconversion, whereas *Rp*BphP3 quenches and a high-energy Pnr state emerges [Giraud *et al.* (2005), *J. Biol. Chem.* **280**, 32389–32397]. Crystallization of the chromophore-binding domain (CBD) of *Rp*BphP2 (*Rp*BphP2-CBD) proved to be difficult and the structure of *Rp*BphP3-CBD was used to crystallize *Rp*BphP2-CBD* using homologue-directed mutagenesis. The structure shows that dimerization is an important factor in successful crystallization of *Rp*BphP2-CBD* and arises from an N136R mutation. Mutations at this site correlate with an ability to dimerize in other truncated BphPs and may also be important for full-length dimer formation. Comparison of the *Rp*BphP3-CBD and *Rp*BphP2-CBD* biliverdin IXα pockets revealed that the former has additional hydrogen bonding around the B and D pyrrole rings that may constrain photoconversion to Pfr, resulting in a strained photoexcited Pnr state.

1. Introduction

Bacteriophytochromes (BphPs) are bacterial homologues of the phytochrome family of photoreceptors which control a variety of photomorphogenic responses. They are part of a two-component signalling system, and in canonical BphPs the light-dependent signal controls autophosphorylation and phosphotransfer to a response regulator which mediates the differential expression of target genes (Bhoo *et al.*, 2001; Giraud *et al.*, 2002; Hughes *et al.*, 1997; Lamparter *et al.*, 2002; Davis *et al.*, 1999). BphPs can be divided into two families: those with a Pr dark-stable state and those with a Pfr dark-stable state (the latter are sometimes called bathy-BphPs; Rottwinkel *et al.*, 2010; Giraud & Verméglio, 2008). The N-terminal photosensory core domain (PCD) is a conserved structure within BphPs and is composed of a Per–Arnt–Sim (PAS) domain followed by a cyclic di-GMP phosphodiesterase/adenyl cyclase/Fhla (GAF) domain and a phytochrome-associated (PHY) domain. The biliverdin IXα (BV) chromophore is contained in the GAF domain and is covalently attached by an auto-lyase reaction to a Cys residue close to the N-terminus of the PAS domain. For this reason,

the PAS and GAF domains are together referred to as the chromophore-binding domain (CBD). The canonical C-terminal output domain is a homologue of the cytoplasmic histidine kinase (HK) domain which forms homodimers through the Dhp dimerization domain (Marina *et al.*, 2005; Yamada *et al.*, 2009). Current research has concentrated on the initial photosensory event and structures have been determined of the photosensory core domain (PCD; Essen *et al.*, 2008; Yang *et al.*, 2008) and the smaller CBD (Ulijasz *et al.*, 2010; Wagner *et al.*, 2005; Yang *et al.*, 2007) in the Pr or Pfr dark-stable states. Although a great deal of progress has been made, a definitive explanation of the spectral properties and the mechanism for signal transmission remains to be given. The photosynthetic bacterium *Rhodopseudomonas palustris* is rich in BphPs of different types and strain CGA009 contains six: one of these is *Rp*BphP1, which controls the expression of a large cluster of photosynthetic genes (Giraud *et al.*, 2002), while *Rp*BphP2 and *Rp*BphP3 (Evans *et al.*, 2005; Giraud *et al.*, 2005) specifically control the expression of $pucAB_d$ genes that encode a light-harvesting complex LH4 which is optimized for low light conditions (Hartigan *et al.*, 2002; Tharia *et al.*, 1999). These two BphPs are contiguous in the genome and probably arose by gene duplication, as indicated by their high sequence identity (52%). Despite their similarity, they exhibit very different red/far-red spectra (Figs. 1*a* and 1*b*) and both are required for the biosynthesis of LH4 (Giraud *et al.*, 2005). It has been proposed that *Rp*BphP2 measures the distribution of light between 700 and 750 nm in a water environment, where solar light intensity quickly falls away at wavelengths longer than 700 nm, while *Rp*BphP3 senses light intensity at around 700 nm. They therefore monitor different properties of the solar radiation arriving at the bacterium (Evans *et al.*, 2005): the ratio of solar intensities [*I*(708 nm):*I*(750 nm)] and the absolute intensity at ∼704 nm. Structure determination of *Rp*BphP3-CBD (Yang *et al.*, 2007) has guided the investigation of the different spectral properties of *Rp*BphP2-PCD and *Rp*BphP3-PCD; an attempt was made to interconvert between the two spectra by site-directed mutation (the larger PCD construct rather than the CBD was used in these studies to ensure native spectra during photoconversion between dark and illuminated states). Because these BphPs are closely related, comparing their structures could provide a deeper understanding of their spectral properties; however, exhaustive crystallization trials failed to produce crystals of *Rp*BphP2-CBD. In such cases crystals can sometimes be obtained by modification of surface residues to improve protein solubility and/or crystal contacts (Dale *et al.*, 2003). It was reasoned that as these proteins share a high amino-acid sequence identity, crystal contacts could be engineered in *Rp*BphP2-CBD to mimic those in the *Rp*BphP3-CBD crystals. The crystal asymmetric unit of *Rp*BphP3-CBD is a dimer and so, for completeness, NCS-dimer contacts should also be modified. By modifying these two types of contacts, it is also possible to assess the relative contributions of crystallographic contacts and NCS contacts to crystal formation. This approach proved successful in the crystallization and structure determination of the modified *Rp*BphP2-CBD*. Dimer formation was found to be an important factor for crystallization and has general

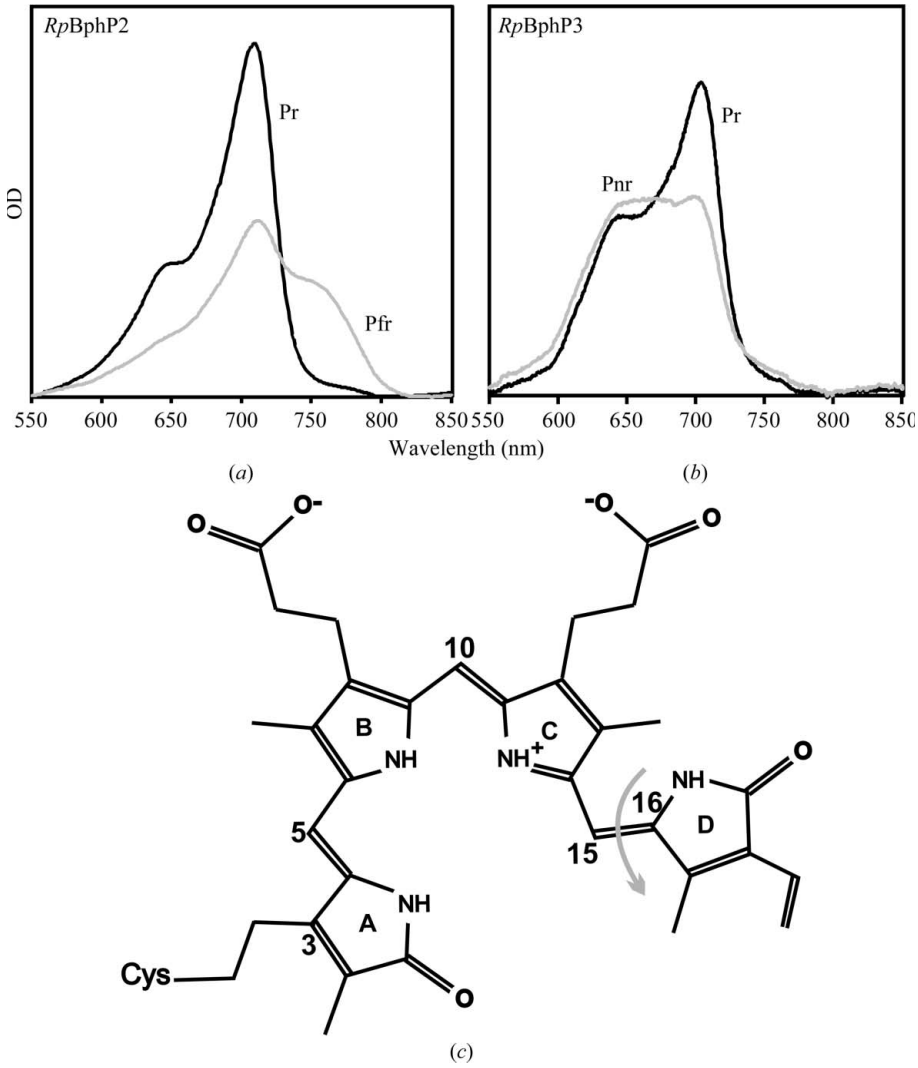

Figure 1
Absorption spectra of full-length (*a*) *Rp*BphP2 and (*b*) *Rp*BphP3 molecules from *R. palustris*. Solid spectra are for the dark-stable Pr states and grey spectra are for the Pfr and Pnr states obtained by photoconversion with red (690 nm) light. (*c*) shows the structure of biliverdin IXα; the methine carbon-bridge atoms C_5, C_{10} and C_{15} link pyrrole rings A, B, C and D. Propionate groups connect to pyrroles B and C. Isomerization, denoted by an arrow, takes place by rotation about the $C_{15}{=}C_{16}$ double bond, which breaks during photoconversion.

Table 1
Modified amino-acid residues in *Rp*BphP2-CBD* based on the structure of *Rp*BphP3-CBD.

Contact type indicates the crystallographic hydrogen-bond contacts (X*n*) and noncrystallographic hydrogen-bond contacts (NCS*n*).

Residues	*Rp*BphP3-CBD (position)	*Rp*BphP2-CBD* (position)	Contact type
PDGERAFN	112–119	98–105†	X1
RY	140–141	126–127	X2
SVR	148–150	134–136	NCS1
R	250	236	X3
EV	309–310	295–296	NCS2

† This region contains an amino-acid insertion at Asn105 which displaces RpBphP2-CBD* by +1 residue after 105. We retain the native protein sequence numbering here to aid the comparison of *Rp*BphP2-CBD* with *Rp*BphP2-CBD and other BphPs.

implications for the shapes of other BphPs. Finally, a comparison of the BV binding sites was carried out to determine future strategies for investigating their very different absorption spectra.

2. Materials and methods

2.1. Design of the *Rp*BphP2-CBD* construct for crystallography

The sequence of *Rp*BphP2-CBD was modified in order to resemble the crystal structure of *Rp*BphP3-CBD (PDB entry 2ool; Yang *et al.*, 2007) with respect to crystal and NCS contacts, giving *Rp*BphP2-CBD*. The program *PyMOL* (http://www.pymol.org) was used to visualize the molecular packing. The space group of the *Rp*BphP3-CBD crystal was *P*321, with an asymmetric unit composed of a dimer. Major contacts are found between crystallographic dimers on a crystal twofold axis and include residues 112–119, which form a β-sheet with a neighbouring asymmetric unit (X1). Hydrogen bonds are also made between crystallographic dimers: from Arg140 NH1 to Arg250 O (X2) and from Tyr141 OH to Arg250 NH2 (X3). Residues involved in NCS dimer interfaces that differed in the two proteins were also identified for modification. A total of 16 amino acids were identified, of which five were involved in NCS contacts (Table 1).

2.2. Cloning and purification

The DNA sequence of RPA3015 encoding bacteriophytochrome *Rp*BphP2 from *R. palustris* strain CGA009 (GenBank ID BX571963) was used over the region encoding amino-acid residues 1–319. Two genes were designed *in silico* to have an *Nde*I restriction site at the 5′ end and an *Xho*I site at the 3′ end: one was unmodified and the other was modified to replace amino-acid residues as in Table 1. Constructs were made by GenScript and inserted into the vector pUC57. Genes were subcloned, using the STRU cloning protocol (Bellini *et al.*, 2011), into pET24a using *Nde*I and *Xho*I sites to create pDB36 and pDB37 plasmids which produce C-terminally His_6-tagged unmodified (*Rp*BphP2-CBD) and modified (*Rp*BphP2-CBD*) proteins, respectively. *Escherichia coli* BL21 (DE3) cells transformed with either plasmid were induced with 1 m*M* IPTG and left to grow overnight at 291 K. The harvested cell pellets were resuspended in standard Ni^{2+} affinity chromatography binding buffer containing 10 m*M* bilivirdin IXα (Frontier Scientific Inc.) and protease-inhibitor tablets (Roche) and the cells were disrupted using a French press. The protein was purified on His-Trap HP according to the manufacturer's instructions and was gel-filtered using a HiLoad 26/60 Superdex 200 column (GE Healthcare) in 5 m*M* Tris–HCl pH 8, 10 m*M* NaCl.

2.3. Gel-filtration molecular-weight analysis

Purified protein (200 µl at approximately 3 mg ml^{-1}) was loaded onto a Superose 12 10/300 GL column (GE Healthcare) using an ÄKTAexplorer (GE Healthcare). The column was calibrated using a high-molecular-weight calibration kit (GE Healthcare).

2.4. UV–visible spectroscopy

UV–visible spectra were recorded at room temperature on a Perkin Elmer Lambda 35 UV–Vis spectrometer. Spectra were recorded either in the dark or after illuminating the sample with light passed through an interference filter centred at 680 $\pm$ 10 nm (Knight Optical Ltd). Photoconversion measurements were carried out in 50 µl volume UVette cells (Eppendorf).

2.5. Crystallization and data collection

Protein at a concentration of 10 mg ml^{-1} was used in crystallization trials with Qiagen crystallization screens. Crystals were viewed using an interference filter (Knight Optical Ltd) centred at 500 $\pm$ 20 nm, which is a BphP non-absorbing region, enabling crystals to be harvested and cryoprotected with 25% ethylene glycol. Diffraction data were collected using a Pilatus 6M-F detector on beamline I03 at Diamond Light Source from crystal samples cryocooled at a temperature of 100 K. Images were indexed, integrated and scaled using *xia*2, *XDS* (Kabsch, 2010) and *SCALA* (Winn *et al.*, 2011).

2.6. Structure determination and refinement

Crystal phases were obtained by molecular replacement with the program *Phaser* (McCoy *et al.*, 2007). The structure was modelled in *Coot* v.6.2 (Emsley & Cowtan, 2004) and was refined with *REFMAC*5 v.5.5 (Murshudov *et al.*, 2011) using isotropic atomic temperature factors augmented by a TLS thermal model consisting of each protomer vibrating anisotropically described by a ten-term tensor. NCS restraints were not applied during refinement. *PyMOL* and *CCP*4*mg* (Winn *et al.*, 2011) were used to illustrate the structure. Superimpositions and comparisons of structures were made with the program *RAPIDO* (Mosca & Schneider, 2008).

Table 2
Summary of data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Molecules per asymmetric unit	2
Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 52.33, b = 79.82, c = 149.85, $\alpha = \beta = \gamma$ = 90
Data resolution range (Å)	54.6–1.79 (1.84–1.79)
Total unique reflections	59311
Data multiplicity	4.3 (4.3)
Completeness (%)	99.8 (99.5)
R_{merge} (%)	2.3 (53.7)
$\langle I/\sigma(I)\rangle$	24.9 (2.3)
Refinement resolution range	15.0–1.79
Reflections used in refinement	56814
R_{cryst} (%)/R_{free}† (%)	18.1/23.2
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.019
Bond angles (°)	2.35
Chiral volumes (Å^3)	0.14
Average B factors	
Protein (Å^2)	49.7
Water (Å^2)	60.0

† $R_{cryst}/R_{free} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. R_{free} was calculated for 5% of reflections randomly excluded from refinement.

3. Results

3.1. Crystal structure of *Rp*BphP2-CBD* and crystal packing

More than 2000 unique crystallization conditions failed to produce crystals from the native construct *Rp*BphP2-CBD, but crystals grew readily in the first round of screening for the *Rp*BphP2-CBD* construct (Table 1). The crystallization conditions were optimized to 18% PEG 3350, 0.2 m*M* $MgCl_2$ or 0.2 m*M* $CaCl_2$, 0.1 *M* Tris–HCl pH 8.5. These conditions failed to crystallize the wild-type protein. Data were collected and structure phases were estimated by molecular replacement using *Rp*BphP3-CBD (PDB entry 2ool) as the template search model. The crystal structure, which belong to space group $P2_12_12_1$ and contain a dimer in the asymmetric unit, was refined to a resolution of 1.8 Å (Table 2). The dimer is related by an NCS twofold axis which lies approximately in the *bc* plane, making an angle of ∼40° with the *b* axis. The structure was refined with residues 13–320 and 11–320 in chains *A* and *B*, respectively, two BV molecules and 534 water molecules. The model closely resembles the structures of *Rp*BphP3-CBD and *Dr*BphP-CBD (both in the Pr state), with C^{α}-atom r.m.s. deviations of 1.8 and 2.2 Å, respectively, between these structures and *Rp*BphP2-CBD*. The structure exhibits a standard CBD fold with BV covalently bound to residue Cys15 in the PAS domain and enclosed in a GAF-domain

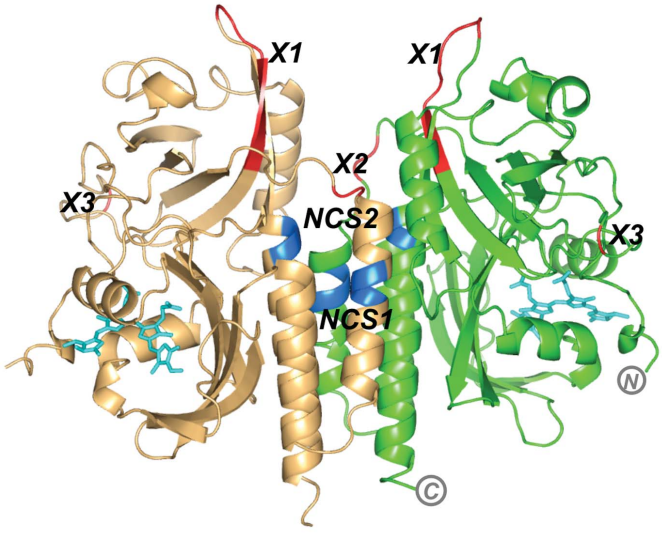

Figure 2
The structure of the *Rp*BphP2-CBD* dimer. The positions of crystal contacts designed by homologue-directed mutagenesis based on the structure of *Rp*BphP3-CBD are labelled X1, X2 and X3 and coloured red. The noncrystallographic hydrogen-bonding contact mutations NCS1 and NCS2 are coloured blue. Biliverdin IXa is coloured cyan.

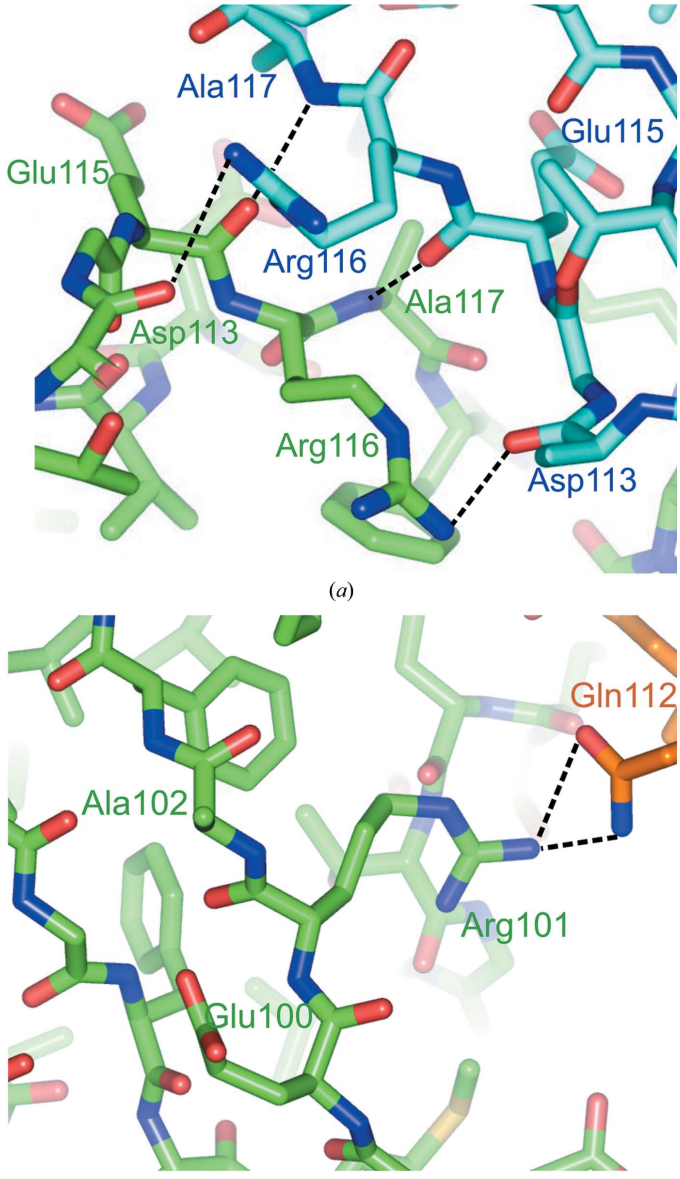

Figure 3
Crystal contacts (*a*) across a crystallographic twofold axis in *Rp*BphP3-CBD at amino-acid region X1 (112–119) and (*b*) the equivalent residues in *Rp*BphP2-CBD* showing different lattice contacts. Different protein chains are coloured green, cyan and orange.

pocket. The N-terminal segment 1–27 is threaded through loop 215–228 in the GAF domain, forming an unusual knot (Fig. 2). The PAS domain is a five-stranded β-sheet (31–36, 39–45, 88–94, 103–111 and 114–120) with strands 2 and 3 joined by three α-helices (47–52, 56–59 and 70–81). The PAS domain is connected to the GAF domain by α-helices hA (132–145) and hB (149–164). The topology of the GAF domain is similar to that of the PAS domain and is composed of a β-sheet (167–172, 180–187, 196–199, 266–275 and 278–288) with strands 2 and 3 joined by an α-helix (202–214), a short β-strand (217–220) and an α-helix (253–263). The short β-strand connects to strand 2, making the core of the GAF domain a six-stranded β-sheet. α-Helix hE (294–320) is at the C-terminus and together with hA forms the dimer interface. The crystal symmetry of *Rp*BphP2-CBD* and *Rp*BphP3-CBD results in very different packing. For example, the main crystal contacts between *Rp*BphP3-CBD dimers (space group *P*321) are on the crystallographic twofold axis involving residues 112–119 in a cross-dimer β-sheet (X1; Fig. 3*a*), whereas the space group of *Rp*BphP2-CBD* is $P2_12_12_1$ and therefore does not possess an equivalent dyad region. In *Rp*BphP2-CBD* the X1 regions do not make contact with each other; residue Arg102 in region X1 forms a single hydrogen bond to Gln112. The latter amino acid is slightly removed from X1 in the adjacent subunit located in a turn between two β-strands (Fig. 3*b*). *Rp*BphP2-CBD* also contains a disordered region with high thermal parameters close to X2 in chains *A* and *B*. In *Rp*BphP3-CBD the equivalent region is disordered in chain *B* but is ordered in chain *A* through a Tyr141–Arg250 hydrogen bond to an asymmetric unit related by a crystallographic threefold. The equivalent Tyr127 in *Rp*BphP2-CBD* is instead involved in hydrogen bonding to Glu71 and Gln74, strengthening the NCS dimer (Fig. 4*d*). This is of interest as the mutation was introduced as a possible crystallographic contact but is instead involved in NCS contacts.

3.2. NCS contacts in *Rp*BphP2-CBD*

Gel-filtration chromatography indicates that *Rp*BphP2-CBD is a monomer in solution while *Rp*BphP2-CBD* is a dimer (Fig. 4*a*) and suggests that dimerization is an important factor in the crystallogenesis of *Rp*BphP2-CBD*. The NCS interface in *Rp*BphP2-CBD* is made at hA and hE, but the details of the contacts in chains *A* and *B* differ, revealing non-exact NCS dyad contacts (Figs. 4*b* and 4*c*). Arg136 in chain *A* (Arg136*A*) makes a salt bridge with Glu301*B* and forms hydrogen bonds to Gln298*B* O and Gln305*B* NE2 through Arg136*A* O, while Arg136*B* makes a salt bridge to Glu301*A* that utilizes different atoms and forms hydrogen bonds to Gln298*A* OE1 but not to Gln305*A* NE2. This asymmetry is also found in the hydrogen bonds between Tyr127*A* and Gln74*B* and Glu71*B* (Fig. 4*d*), while Tyr127*B* follows a partially disordered region and makes a hydrogen bond to Asn105*A* ND2. As has previously been stated, the mutation was made as a potential crystal contact but is involved in NCS contacts in both cases. Interestingly, in *Rp*BphP3-CBD the salt bridge is made by Arg150*A* but not by Arg150*B*, which appears to be more affected by the disorder in the neighbouring region 138–143. Dimerization asymmetry can also be observed in the structure of the larger contruct *Pa*BphP-PCD from *Pseudomonas aeruginosa*, which contains a PHY domain as well as a CBD. It may be that there is an important contribution to dimer symmetry from the missing HK domain; however, lattice contacts could also be involved in breaking of symmetry if there is significant protein flexibility. Of the five

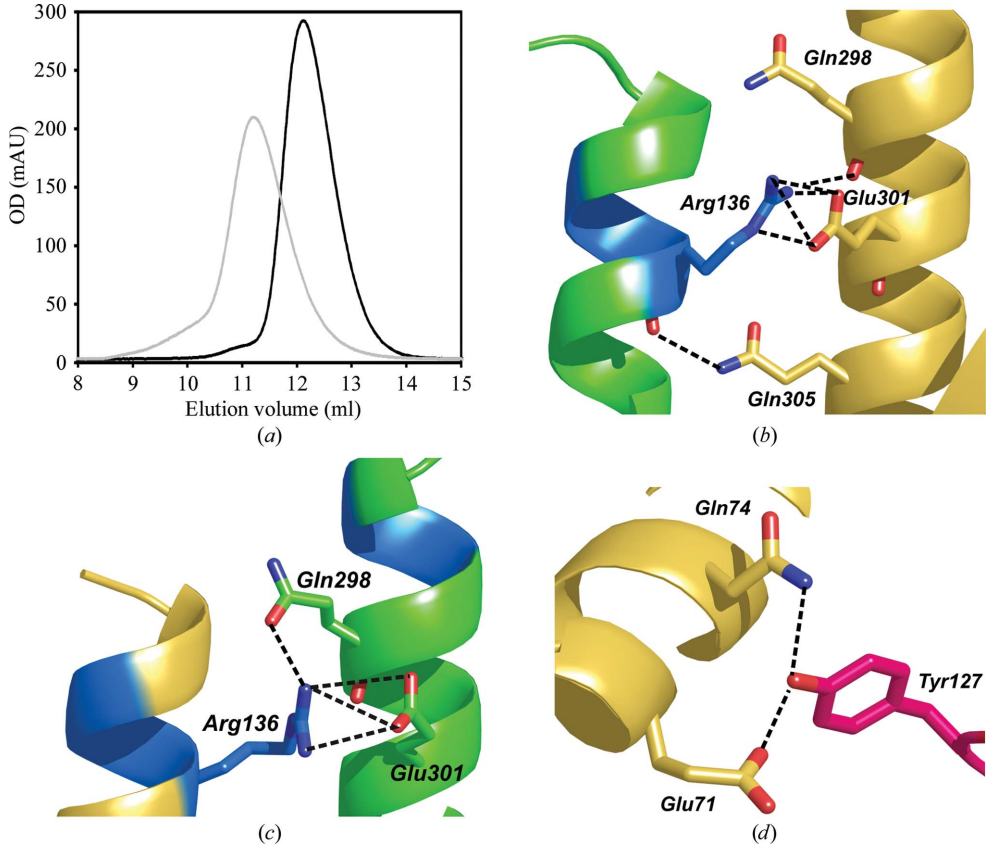

Figure 4
(*a*) Gel-filtration elution plots from a Superose 12 10/300 GL column: the black curve corresponds to *Rp*BphP2-CBD eluting as a monomer of apparent molecular weight 37 000 kDa and the grey curve corresponds to *Rp*BphP2-CBD* eluting as a dimer with an apparent molecular weight of 71 000 kDa. (*b*) and (*c*) present the two NCS1 helix interfaces between chain *A* (green) and chain *B* (brown) (*i.e.* *A*–*B* and *B*–*A*) showing non-exact NCS interactions at the mutated residue Arg136. (*d*) A new NCS interaction, not found in *Rp*BphP3-CBD, between Tyr127 and Glu71 and Gln74.

NCS mutations in *Rp*BphP2-CBD*, only Arg136 makes significant hydrogen-bond contacts and it also makes the sole hydrogen-bonding contacts between the dimer-interfacing helices. The hydrogen-bonding partners Gln299, Glu301 and Gln305 are conserved in both structures. Therefore, it appears that the N136R mutation in *Rp*BphP2-CBD* is sufficient to cause dimerization. An additional contribution of the Tyr127 mutation to NCS contacts can be discounted as it is disordered in one of the *Rp*BphP3-CBD subunits and this residue is Glu127 in wild-type *Rp*BphP2-CBD, which can provide similarly efficient hydrogen bonding to Glu71 and Gln74.

3.3. Biliverdin IX*α* pocket

The biliverdin IX*α* (BV) pocket of *Rp*BphP2-CBD* is enclosed by 32 amino acids that interact with BV *via* van der Waals or hydrogen-bond contacts. The r.m.s. deviation between C^{α} atoms of only those residues in RpBphP2-CBD* and RpBphP3-CBD that make contacts with BV is 0.31 Å, indicating that mutations introduced elsewhere in *Rp*BphP2-CBD* have little impact on the BV site. This is also confirmed in the similarity between the absorption spectra of *Rp*BphP2-CBD and *Rp*BphP2-CBD*, which are very sensitive to structural differences at the BV site (Supplementary Fig. 1[1]). It has been shown that *Rp*BphP3–521 (truncated after PHY) with mutations K183M, S297A, L207Y and D209G, which introduce the corresponding amino acids found in the homologous *Rp*BphP2 construct (*Rp*BphP2–505), goes some way towards reproducing *Rp*BphP2–505-like absorption spectra, with the largest contribution provided by the L207Y mutation. However, the Pnr spectra were not reproduced by the reverse mutation of the residues surrounding pyrrole D from those in *Rp*BphP2-CBD to those in *R*pBphP3-CBD (Yang *et al.*, 2007). In the bathy-BphP *Pa*BphP-PCD from *P. aeruginosa*, the residues that directly stabilize the Pfr state by hydrogen bonding to the BV D-ring carbonyl can be identified. Because these residues are conserved in all BphPs and therefore are not the mutations made to *Rp*BphP3–521, the mutations that were introduced made subtle changes to the BV environment that resulted in a Pfr-like spectrum. X-ray diffraction experiments on the mixed Pfr/Pr states of *Pa*BphP-PCD have established that BV undergoes extensive changes in conformation during photoconversion, namely *E*-to-*Z* isomerization of the D pyrrole about the $C_{15}{=}{=}C_{16}$ bond accompanied by a global rotation of BV about an axis approximately normal to ring A (Yang *et al.*, 2009). More recently, temperature-scan cryocrystallography between temperatures of 100 and 180 K captured three photoconversion states; the upper temperature of 180 K indicates that part of the photocycle, up to the Lumi-F state, was trapped but not the Meta-Fa, Fb and Pr states. Nevertheless, progressive torsion twisting of BV methine carbon bond bridges, which join the pyrrole rings, and a rearrangement of the neighbouring protein residues was observed (Yang *et al.*, 2011). It is clear that the photoconversion event is not localized to D-ring isomerization and that an extensive reorganization of the BV backbone as well as movement of amino-acid residues must be involved across the full span of the BV pocket.

To identify other contacts that may contribute to global rearrangement of BV, an analysis of the BV pockets of the *Rp*BphP2-CBD* and *Rp*BphP3-CBD structures was made. Additional hydrogen bonds are present in *Rp*BphP3-CBD which are not found in RpBphP2-CBD*, making the BV environment more rigid, especially around the B and D rings (Figs. 5*a* and 5*b*). It has been established using IR time-resolved spectroscopy that only 6% of the photoconversion is accounted for by the rupture of hydrogen bonds responsible for $C_{15}{=}{=}C_{16}$-bond isomerization and that the dominant de-excitation pathway (89%) is from excited-state proton transfer

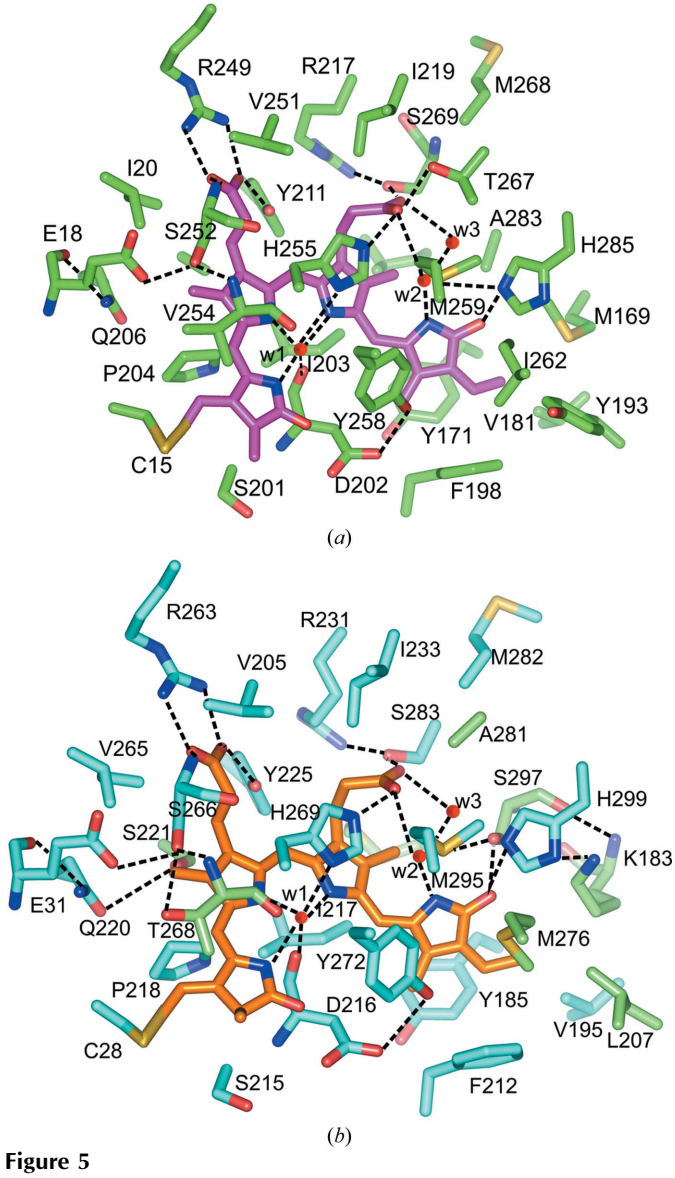

Figure 5
Biliverdin IX*α* pocket and hydrogen-bonding network in (*a*) *Rp*BphP2-CBD* and (*b*) *Rp*BphP3-CBD. The residues that differ between the two structures are coloured light green in (*b*). Waters involved in the network are labelled w1, w2 and w3. More hydrogen bonds are seen around pyrroles B and D in *Rp*BphP3-CBD than in *Rp*BphP2-CBD*.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MH5063). Services for accessing this material are described at the back of the journal.

between pyrrole N atoms. This suggests that other structural groups are involved across the whole of the BV and a fuller description of hydrogen-bonding networks in both proteins is required. Both structures possess four hydrogen-bond clusters which contribute to stabilizing (i) the B-ring propionate, (ii) the C-ring propionate, (iii) the N atoms of pyrrole rings A, B and C and (iv) the D-ring pyrrole carbonyl (Supplementary Fig. 2).

Cluster 1. Within *Rp*BphP2-CBD*, the B-ring propionate hydrogen bonds to Arg249, Tyr211 and Ser252. The main-chain carbonyl of Ser252 is stabilized by hydrogen bonds to the Glu18 carboxyl and Val254 N. In *Rp*BphP3-CBD these residues are Arg263, Tyr225 and Ser266, but Val254 is now Thr268 and contributes an extra side-chain hydrogen bond to Ser266. Ser221, which is Ala207 in *Rp*BphP2-CBD*, also adds an extra side-chain hydrogen bond to the cluster in *Rp*BphP3-CBD.

Cluster 2. The propionate of ring C makes one additional hydrogen bond to Thr267 of *Rp*BphP2-CBD*, which is Ala281 in *Rp*BphP3-CBD. This is the only additional hydrogen bond in *Rp*BphP2-CBD* which is not found in *Rp*BphP3-CBD.

Cluster 3. Both homologues have identical hydrogen bonding at the N atoms of pyrrole rings A, B and C mediated by water W1.

Cluster 4. The D-ring carbonyl of *Rp*BphP2-CBD* is hydrogen bonded to His285 NE2, and the D pyrrole N atom is hydrogen bonded to the C-ring propionate mediated by waters W2 and W3. In addition to these hydrogen bonds, *Rp*BphP3-CBD contains a hydrogen bond from the D-ring carbonyl to Ser297 OG; secondary hydrogen bonds are found between Lys183 NZ and His285 ND1. Finally, main-chain hydrogen bonds are formed between Lys183 N and Ser297 O. This tight cluster of hydrogen bonds in *Rp*BphP3-CBD stabilizes His285 and Ser297, which make the primary hydrogen bonds to the carbonyl on ring D.

It has already been established that the hydrogen bonds surrounding pyrrole D in *Rp*BphP3-CBD prevent efficient isomerization about the $C_{15}{=}{=}C_{16}$ bond (Toh *et al.*, 2011), but our work shows that a cluster of hydrogen bonds around pyrrole B may also be important and could constrain the torsion twisting of the BV methine carbon bond bridges during photoconversion, which may also prevent photoconversion to Pfr.

4. Discussion

4.1. Crystallogenesis of *Rp*BphP2-CBD*

The formation of *Rp*BphP2-CBD* crystals is not caused by mutations aimed at mimicking *Rp*BphP3-CBD crystallographic contacts, hence the difference in their space groups. Instead, crystallization is correlated with dimer formation in solution and the observation that the engineered mutant N136R makes several key hydrogen bonds at the NCS dimer interface of *Rp*BphP2-CBD*. This residue is Arg150 in *Rp*BphP3-CBD and is Arg156 in *Dr*BphP-CBD, while in *Pa*BphP-PCD it is Gln128, which has a shorter side chain but not as short as that of Asn in *Rp*BphP2-CBD. In the case of *Pa*BphP-PCD the full-length helix hE joins the CBD to the PHY domain and contributes to the dimer interface along its full length. It appears that Asn136 in *Rp*BphP2-CBD is sufficient to ensure a predominantly monomeric form of CBD in solution which inhibits crystal formation. When a single mutation N136R is inserted into *Rp*BphP2-CBD it also forms dimers rather than monomers. It is interesting to note that Cph1-PCD crystallizes as a monomer and also has Asn in the equivalent position (Essen *et al.*, 2008). We have recently determined the structure of *Rp*BphP1-PCD, which contains Ala at the same position, to a resolution of 2.5 Å (unpublished work) and this also crystallizes as a monomer. Dimer formation could be important for function in canonical BphPs because trans-autophosphorylation can proceed in HK domains between opposing subunits (Marina *et al.*, 2005; Ninfa *et al.*, 1993; Yang & Inouye, 1991). This is not true for all HK domains; it has recently been shown that (*cis*) kinase action is also possible (Casino *et al.*, 2009, 2010). The low-resolution SAXS structure of the complete *Rp*BphP2 indicates a Y-shaped dimer with the CBD domains forming the upper limbs of the 'Y' (Evans *et al.*, 2006), which is consistent with our observations that *Rp*BphP2-CBD cannot dimerize. The recognized dimerization domain is in the C-terminal HK domain, where the Dhp domain forms a four-helix bundle. In the case of *Dr*BphP, electron-microscopic images of the complete molecule indicate dimerization along the whole length of the BphP (Li *et al.*, 2010), which is consistent with the observation that *Dr*BphP-CBD alone can dimerize. Taking this information together it suggests that in the BphP family there are separate dimer interfaces at the CBD, PHY and HK domains and that in some BphPs the CBD plays no part in dimer formation, reflecting structural diversity amongst the family. A case can be made that the Arg/Asn residue is important for dimer formation/inhibition not only in CBD constructs but also in PCDs (which contain the PHY domain), and that in the case of Cph1 and *Rp*BphP2 dimerization may depend mostly on the Dhp four-helix bundle of the HK domain. Although Cph1-PCD crystallizes as a monomer, closer inspection shows a strongly interacting antiparallel dimer on the crystallographic twofold axis. It may be of interest to determine whether it exists as a nonphysiological dimer in solution which aids crystallogenesis. The monomeric crystals of *Rp*BphP1-PCD have been difficult to reproduce, probably because of their high solvent content (75%), and this exception to the rule perhaps points to the importance of dimer formation in BphP crystallization. More generally, if oligomeric assemblies are important for comparative interspecies crystallographic studies, strengthening these contacts may be a generally applicable technique when problems are encountered (in particular when protein domains rather than the whole molecule are under study and interactions between the domains are sufficiently weakened to form monomers or a mixture of monomers and dimers). If a homologous structure is available, the NCS contacts can be used to model stronger contacts. Engineering crystallographic contacts should also be attempted: although they did not help in the current example,

neither did they disrupt crystallogenesis despite comprising 3.5% of the molecule. However, in other structures they may contribute to crystal growth.

4.2. Differences between RpBphP2-CBD* and RpBphP3-CBD photoconversion

The absorption spectrum of *Rp*BphP3 is unusual because it does not photoconvert to a Pfr spectrum on red-light illumination; rather, the 704 nm Pr peak quenches and a high-energy Pnr peak is observed at ∼650 nm. With the determination of the *Rp*BphP3-CBD structure, some residues around pyrrole D were identified which modify the environment of the BV pocket and produce a Pfr-like spectrum (Yang *et al.*, 2007). Subsequent time-resolved spectroscopy also focused on hydrogen bonding around pyrrole D as well as the deprotonation of pyrrole N atoms (Toh *et al.*, 2010). In neither case were the origins of the *Rp*BphP3-CBD Pnr state explained or reproduced by the mutation of *Rp*BphP2-PCD. The X-ray work on the mixed Pr/Pfr states of *Pa*BphP-PCD confirms that photoconversion is not simply an isomerization about the $C_{15}{=}{=}C_{16}$ bond but progresses through global changes of the BV backbone (Yang *et al.*, 2009, 2011). Other important intermediates have been identified by time-resolved spectroscopy (the Lumi-R, Meta-Ra and Meta-Rc states) and involve the deprotonation and reprotonation of pyrrole N atoms on a timescale of picoseconds to milliseconds (Bischoff *et al.*, 2001; Borucki *et al.*, 2005; Foerstendorf *et al.*, 2001; Kneip *et al.*, 1999; von Stetten *et al.*, 2007; Toh *et al.*, 2011). If the additional hydrogen bonding surrounding pyrroles B and D in *Rp*BphP3 restricts full BV methine carbon bridge bond rotation at C atoms C_5, C_{10} and C_{15}, this could cause a strained excited BV conformation. The structure of *Rp*BphP3-CBD may provide a good starting model for studying intermediate photoconversion states by time-resolved spectroscopic experiments. Conjugate-bonded molecules such as bacteriochlorophyll *a*, carotenoids and biliverdin have an initial excited state that is delocalized across all double bonds. In photosynthetic complexes the absorbed energy is dissipated by energy transfer to other nearby chromophores/complexes or is dissipated as heat under excessive light conditions. In BphPs, in which the chromophores are isolated and fluorescence is relatively low, absorbed energy can only relax by a process of thermal reorganization of the kind observed by cryocrystallography (Yang *et al.*, 2011). The hydrogen-bonding network in *Rp*BphP3 may prevent complete thermal relaxation to the Pfr state, resulting in a strained BV backbone and the observed Pnr high-energy (∼650 nm) absorption state.

We would like to thank the Biotechnology and Biological Sciences Research Council (BB/E002609/2) and the Science and Technology Facilities Council for supporting this work.

References

Bellini, D., Fordham-Skelton, A. P. & Papiz, M. Z. (2011). *Mol. Biotechnol.* **48**, 30–37.
Bhoo, S.-H., Davis, S. J., Walker, J., Karniol, B. & Vierstra, R. D. (2001). *Nature (London)*, **414**, 776–779.
Bischoff, M., Hermann, G., Rentsch, S. & Strehlow, D. (2001). *Biochemistry*, **40**, 181–186.
Borucki, B., von Stetten, D., Seibeck, S., Lamparter, T., Michael, N., Mroginski, M. A., Otto, H., Murgida, D. H., Heyn, M. P. & Hildebrandt, P. (2005). *J. Biol. Chem.* **280**, 34358–34364.
Casino, P., Rubio, V. & Marina, A. (2009). *Cell*, **139**, 325–336.
Casino, P., Rubio, V. & Marina, A. (2010). *Curr. Opin. Struct. Biol.* **20**, 763–771.
Dale, G. E., Oefner, C. & D'Arcy, A. (2003). *J. Struct. Biol.* **142**, 88–97.
Davis, S. J., Vener, A. V. & Vierstra, R. D. (1999). *Science*, **286**, 2517–2520.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Essen, L. O., Mailliet, J. & Hughes, J. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 14709–14714.
Evans, K., Fordham-Skelton, A. P., Mistry, H., Reynolds, C. D., Lawless, A. M. & Papiz, M. Z. (2005). *Photosynth. Res.* **85**, 169–180.
Evans, K., Grossmann, J. G., Fordham-Skelton, A. P. & Papiz, M. Z. (2006). *J. Mol. Biol.* **364**, 655–666.
Foerstendorf, H., Benda, C., Gärtner, W., Storf, M., Scheer, H. & Siebert, F. (2001). *Biochemistry*, **40**, 14952–14959.
Giraud, E., Fardoux, J., Fourrier, N., Hannibal, L., Genty, B., Bouyer, P., Dreyfus, B. & Verméglio, A. (2002). *Nature (London)*, **417**, 202–205.
Giraud, E. & Verméglio, A. (2008). *Photosynth. Res.* **97**, 141–153.
Giraud, E., Zappa, S., Vuillet, L., Adriano, J. M., Hannibal, L., Fardoux, J., Berthomieu, C., Bouyer, P., Pignol, D. & Verméglio, A. (2005). *J. Biol. Chem.* **280**, 32389–32397.
Hartigan, N., Tharia, H. A., Sweeney, F., Lawless, A. M. & Papiz, M. Z. (2002). *Biophys. J.* **82**, 963–977.
Hughes, J., Lamparter, T., Mittmann, F., Hartmann, E., Gärtner, W., Wilde, A. & Börner, T. (1997). *Nature (London)*, **386**, 663.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F. & Schaffner, K. (1999). *Biochemistry*, **38**, 15185–15192.
Lamparter, T., Michael, N., Mittmann, F. & Esteban, B. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 11628–11633.
Li, H., Zhang, J., Vierstra, R. D. & Li, H. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 10872–10877.
Marina, A., Waldburger, C. D. & Hendrickson, W. A. (2005). *EMBO J.* **24**, 4247–4259.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mosca, R. & Schneider, T. R. (2008). *Nucleic Acids Res.* **36**, W42–W46.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Ninfa, E. G., Atkinson, M. R., Kamberov, E. S. & Ninfa, A. J. (1993). *J. Bacteriol.* **175**, 7024–7032.
Rottwinkel, G., Oberpichler, I. & Lamparter, T. (2010). *J. Bacteriol.* **192**, 5124–5133.
Stetten, D. von, Seibeck, S., Michael, N., Scheerer, P., Mroginski, M. A., Murgida, D. H., Krauss, N., Heyn, M. P., Hildebrandt, P., Borucki, B. & Lamparter, T. (2007). *J. Biol. Chem.* **282**, 2116–2123.
Tharia, H. A., Nightingale, T. D., Papiz, M. Z. & Lawless, A. M. (1999). *Photosynth. Res.* **61**, 157–167.
Toh, K. C., Stojković, E. A., Rupenyan, A. B., van Stokkum, I. H. M., Salumbides, M., Groot, M. L., Moffat, K. & Kennis, J. T. M. (2011). *J. Phys. Chem. A*, **115**, 3778–3786.
Toh, K. C., Stojković, E. A., van Stokkum, I. H. M., Moffat, K. & Kennis, J. T. M. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 9170–9175.

Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L. & Vierstra, R. D. (2010). *Nature (London)*, **463**, 250–254.
Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. (2005). *Nature (London)*, **438**, 325–331.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yamada, S., Sugimoto, H., Kobayashi, M., Ohno, A., Nakamura, H. & Shiro, Y. (2009). *Structure*, **17**, 1333–1344.
Yang, Y. & Inouye, M. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 11057–11061.
Yang, X., Kuk, J. & Moffat, K. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 14715–14720.
Yang, X., Kuk, J. & Moffat, K. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 15639–15644.
Yang, X., Ren, Z., Kuk, J. & Moffat, K. (2011). *Nature (London)*, **479**, 428–432.
Yang, X., Stojkovic, E. A., Kuk, J. & Moffat, K. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 12571–12576.