

Light-induced complex formation of bacteriophytochrome *RpBphP1* and gene repressor *RpPpsR2* probed by SAXS

Miroslav Z. Papiz^{1,2}, Dom Bellini¹, Kate Evans³, J Günter Grossmann^{1,2} and Tony Fordham-Skelton²

1 Institute of Integrative Biology, University of Liverpool, UK

2 STFC Daresbury Laboratory, Warrington, UK

3 Pharmacy and Biomolecular Sciences, Liverpool John Moores University, UK

Keywords

bacteriophytochrome; complex formation; photo-induced changes; photosynthesis; SAXS

Correspondence

M. Z. Papiz, Institute of Integrative Biology, University of Liverpool, Biosciences Building, Liverpool L69 7ZB, UK
Tel: +44 1925266860
E-mail: mpapiz@liverpool.ac.uk

(Received 8 August 2018, revised 4 March 2019, accepted 24 June 2019)

doi:10.1111/febs.14973

Bacteriophytochrome proteins (BphPs) are molecular light switches that enable organisms to adapt to changing light conditions through the control of gene expression. Canonical type 1 BphPs have histidine kinase output domains, but type 3 *RpBphP1*, in the bacterium *Rhodospseudomonas palustris* (*Rps. palustris*), has a C terminal PAS9 domain and a two-helix output sensor (HOS) domain. Type 1 BphPs form head-to-head parallel dimers; however, the crystal structure of *RpBphP1* Δ HOS, which does not contain the HOS domain, revealed pseudo anti-parallel dimers. HOS domains are homologs of Dhp dimerization domains in type 1 BphPs. We show, by applying the small angle X-ray scattering (SAXS) technique on full-length *RpBphP1*, that HOS domains fulfill a similar role in the formation of parallel dimers. On illumination with far-red light, *RpBphP1* forms a complex with gene repressor *RpPpsR2* through light-induced structural changes in its HOS domains. An *RpBphP1*:*RpPpsR2* complex is formed in the molecular ratio of 2 : 1 such that one *RpBphP1* dimer binds one *RpPpsR2* monomer. Molecular dimers have been modeled with Pfr and Pr SAXS data, suggesting that, in the Pfr state, stable dimeric four α -helix bundles are formed between HOS domains, rendering *RpBphP1* functionally inert. On illumination with light of 760 nm wavelength, four α -helix bundles formed by HOS dimers are disrupted, rendering helices available for binding with *RpPpsR2*.

Introduction

Bacteriophytochrome proteins (BphPs) are photoreceptor proteins that transmit light-dependent signals to downstream response regulators which affect a variety of light-dependent physiological functions through the control of gene expression. The photo-sensing molecule is a tetrapyrrole biliverdin IX α which is covalently attached to a cysteine residue near to the N-terminus. Biliverdin IX α absorbs light and switches reversibly

between the far-red absorbing Pfr and red absorbing Pr state by isomerization of the D ring pyrrole about the C₁₄=C₁₅ bond. All BphPs have photosensory core domains (PCDs) at their N termini, comprising PAS-GAF-PHY domains, but can have variable C-terminus output domains. PCD is the minimal structural unit necessary for stable and reversible photo-conversion. Canonical type 1 BphPs have a histidine kinase output

Abbreviations

AppA, Blue light photo-sensor protein; BphP, Bacteriophytochrome; CBD, chromophore-binding domain; c-di-GMP, bis-(3'-5')-cyclic dimeric guanosine monophosphate; DGC, diguanyl cyclase; Dhp, histidine kinase dimerization domain; GAF, GMP-specific phosphodiesterases-adenylyl cyclases-FhIA domain; HK, histidine kinase domain; HOS, two-helix output sensor domain; HTH, helix-turn-helix motif; PAS, Per-Arnt-Sim domain; PCD, photosensory core domain; Pfr, far-red absorbing state; PHY, photochrome specific domain; Pr, red absorbing state; SAXS, small angle X-ray scattering; SEC, size exclusion chromatography.

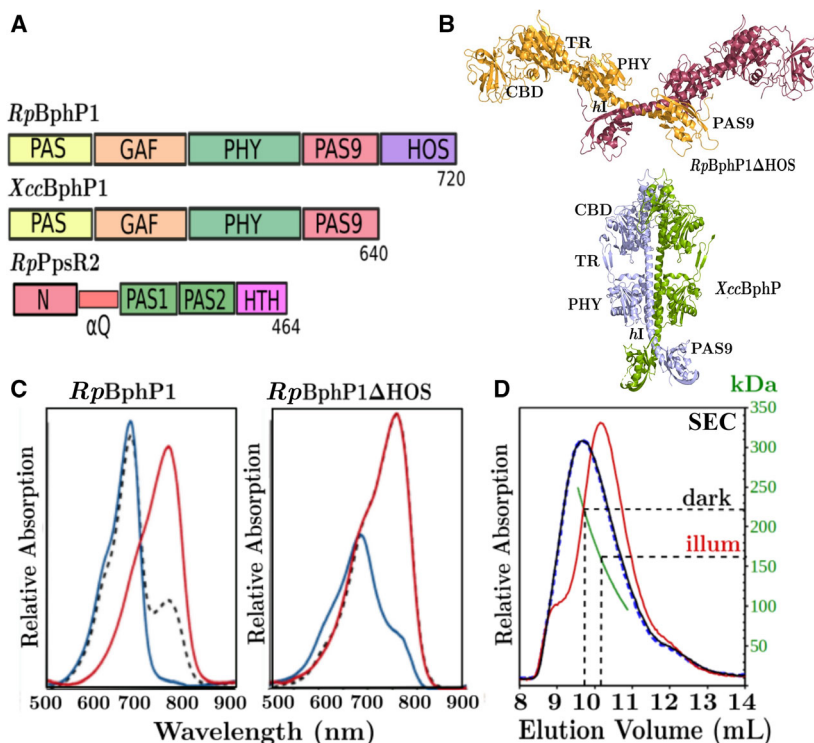


Fig. 1. Domain organization, tertiary/quaternary structure, UV-Vis spectra, and SEC chromatography. Domain organizations of (A) bacteriophytochrome *RpBphP1* from *Rhodospseudomonas palustris*, bacteriophytochrome, *XccBphP1* from *Xanthomonas campestris pv. campestris* and the gene repressor *RpPpsR2* from *Rhodospseudomonas palustris*. *RpBphP1* and *XccBphP1* are composed of the chromophore binding domain comprising PAS-GAF domains (CBD) or the larger photo-sensory core domain PAS-GAF-PHY (PCD), which is the minimal unit for stable photo-conversion, and C-terminal output domains PAS9-HOS and PAS9 respectively. Gene repressor *RpPpsR2* is composed of the PAS domain variants N, PAS1, and PAS2. Domains N and PAS1 are connected by a glutamate rich Q-linker α -helix (α Q) and the C terminus terminated by the helix-turn-helix DNA binding motif HTH. (B) High resolution X-ray structures of *RpBphP1* Δ HOS (excluding HOS domain, 4GW9) and *XccBphP1* (5AKP). Quaternary structures are pseudo anti-parallel dimers for *RpBphP1* Δ HOS and parallel dimers for *XccBphP1*, TR is a tongue region capping the biliverdin chromophore pocket, PHY is a phytochrome specific domain and PAS9 is a type 9 PAS domain. (C) *RpBphP1* UV-Vis spectra of Pfr dark state (red) and the photo-converted Pr state (blue), a spectrum was also measured 30 min after the start of dark reversion (black dashed line). *RpBphP1* dark reversion half-life is 50 min and for *RpBphP1* Δ HOS is estimated to be less than 1 min. (D) size exclusion chromatography (SEC) of *RpBphP1*; apparent dimer molecular weights are ~ 220 kDa for the Pfr dark state (blue) and ~ 160 kDa for the Pr state (red) which was prepared by illumination with 760 nm light. A fully dark reverted Pfr protein (dashed-blue) is shown to elute as the original Pfr state demonstrating that the process is reversible.

module that consists of a dimerization domain (Dhp) and a kinase domain [1]. *RpBphP1* from *Rps. palustris* (Fig. 1A) is a type 3 BphP, with a different output domain, that controls the expression of a large cluster of genes involved in photosynthesis [2]. The output domain is composed of type 9 PAS domain (PAS9) followed by a two-helix output sensor (HOS) [3]. Although type 1 and type 3 output domains are functionally and structurally different, HOS domains are distant homologs of Dhp, suggesting some structural similarity. *RpBphP1*, in the red-absorbing Pr state, modifies the behavior of a cognate response regulator *RpPpsR2* [3–5] and we show here that these two molecules form complexes under the influence of far-red light. Most BphPs form parallel head-to-head dimers,

as has been shown by cryo-electron microscopy on full-length type 1 *DrBphP* from *Deinococcus radiodurans* [6]. High resolution X-ray structures of the chromophore binding domain (CBD) [7–9], PCD [10–14], and PCD-PAS9 (*RpBphP1* Δ HOS) [3] have been determined, providing important insights into light-induced structural changes. The full-length type 3 *XccBphP* has also been determined (Fig. 1A) from *Xanthomonas campestris pv. campestris* (*Xcc*); however, this BphP possesses C terminus PAS9 domains only [15]. Its biological function is also different, which is to suppress plant infection by attenuation of bacterial virulence systems [16]. The X-ray structures of *XccBphP* and *RpBphP1* Δ HOS have been determined in their Pr and Pfr states, respectively, offering an opportunity to

investigate light-induced changes in type 3 BphPs. A major structural difference in *XccBphP* is that the long helix hI, which joins the PHY domain to PAS9, is bowed and the PAS9 domain is rotated outwards from the dimer twofold axis by $\sim 30^\circ$ [3,15]. Another surprising feature of *XccBphP* and *RpBphP1ΔHOS* is their different quaternary organizations (Fig. 1B), head-to-head parallel dimer for *XccBphP* and a pseudo anti-parallel dimer in *RpBphP1ΔHOS* with a cross-over angle between protomers of $\sim 170^\circ$. Yet another full-length BphP structure is type 5 *IsPadC* from *Idiomarina* sp. This BphP allows red-light regulation of the second messenger *bis*-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) by diguanylyl cyclase activity in the C terminus GGDEF domain [17,18]. Despite the differences at the C terminus, all BphP types share common light-induced structural elements that transmit conformation changes to their C-terminal domains through a long helix (hI).

The target protein of *RpBphP1* is the transcription regulator *RpPpsR2*. This protein is found only in purple bacteria and is responsible for controlling the expression of photosynthetic genes. In purple bacteria, PpsR are loosely conserved, reflecting a high evolutionary divergence of properties [19]. Experiments on PpsR, of *Rhodobacter sphaeroides* (*Rba. sphaeroides*) and *Rhodobacter capsulatus* (*Rba. capsulatus*), have revealed a common mechanism of binding to a pair of palindromic DNA sequences (TG₁₂ACA) under oxidizing conditions [20,21]. Formation of a disulfide bond in the C-terminal HTH domain of PpsR was shown to stimulate binding to DNA [22,23]. However, whereas in *Rba. capsulatus*, PpsR is only redox sensing in *Rba. sphaeroides*, it is also antagonized by a blue light sensing protein AppA [24,25]. In *Rps. palustris* and the closely related *Bradyrhizobium* sp., there are two such proteins, PpsR1 and PpsR2, but in most strains only PpsR1 has a Cys residue, whereas it is missing in PpsR2, indicating different redox behavior [19]. An exception to this finding is the *Rps. palustris* strain CGA009, a laboratory strain that has been extensively used to study this bacterial species. CGA009 is not a typical strain because PpsR2 has the residue C428 and thus a similar redox behavior to PpsRs of *Rba. capsulatus* and *Rba. sphaeroides*. In CGA009, *RpPpsR2* is under dual control, redox and far-red light, through interaction with *RpBphP1*. However, all wild-type strains studied so far have an R428 within *RpPpsR2* and are not redox sensors. They are under the sole control of far-red light, as will be discussed later [26].

Small angle X-ray scattering (SAXS) experiments on full-length *RpBphP1* and complex formation with

RpPpsR2 will be presented here. Experimental findings indicate that full-length *RpBphP1* forms parallel dimers consistent with canonical BphPs and at odds with the structure of *RpBphP1ΔHOS*, which we conclude is a pathological dimer caused by HOS domain deletion. Based on the SAXS data, measured in the dark and under illumination with 760 nm light, models of *RpBphP1*-Pfr and *RpBphP1*-Pr are presented, and are subsequently employed to explain the role played by the HOS domains in the formation of native dimers and *RpBphP1*:*RpPpsR2* complexes.

Results

Spectral properties of *RpBphP1* and *RpBphP1ΔHOS* dimers

Samples of *RpBphP1* and 70 kDa *RpBphP1ΔHOS* fragment were prepared. The latter was purified by Ni affinity chromatography of protein, and is a natural degradation product when HOS domains are cleaved after approximately 3-week period storage. Illuminating full-length *RpBphP1* with light of wavelength 760 nm causes complete photo-conversion to the red Pr state, which dark-reverts to the Pfr state with a half-life of ~ 50 min. The *RpBphP1ΔHOS* fragment exhibits incomplete photo-conversion with a half-life much shorter than 60 s (Fig. 1C) and completely dark-reverts to a Pfr state within 30 min. In comparison, in the same time, the dark-reverted population of *RpBphP1* states is $\sim 65:35\%$ Pr/Pfr. Molecular weights determined by size exclusion chromatography (SEC) are ~ 220 and 160 kDa for Pfr and Pr states respectively (Fig. 1D). These states have also been shown to be reversible on dark reversion by SEC. In contrast, no difference is observed between Pfr or Pr states in the SEC profiles of *RpBphP1ΔHOS*. SEC molecular weight of *RpBphP1*-Pfr is anomalously high, whereas the molecular weight of Pr is close to the value calculated from the amino acid composition, implying that the Pr shape is more isotropic compared to Pfr. Greater than expected molecular weights can occur for elongated molecules because they tend to reside in the void volume of the SEC column [27] and this kind of behavior has also been observed for monoclonal antibodies [28]. Taken together, SEC data and UV-Vis spectra suggest that HOS domains play an important role in the creation of stable and reversible Pr states within the chromophore pocket. It is clear that the Pr state, as observed by UV-Vis spectroscopy, is associated with a reversible change in protein conformation responsible for a relatively large variation in apparent molecular weight, as observed in dark/light SEC data.

Dark and illuminated SAXS data collection

The structure of full-length *RpBphP1*, in both Pfr and Pr states, was investigated via the Small Angle X-ray Scattering (SAXS) method. Given that *RpBphP1* holo-protein degrades to a 70 kDa fragment with the loss of HOS domains (Fig. 2A, lane 2), it was important to obtain a mono-dispersed sample and thus avoid SAXS data artifacts. A purification protocol was devised to minimize protein degradation by performing every purification step on ice and in darkness. Moreover, both purification and SAXS data collection were completed within 24 h of rupturing bacterial cells. SDS/PAGE of a sample taken directly after Pr SAXS data collection confirmed that the protein is full-length and mono-dispersed *RpBphP1* (Fig. 2A, lane 1). In addition, gels show only one ~80 kDa band and a complete absence of the 70 kDa *RpBphP1ΔHOS* fragment. To ensure a 100% Pfr, all SAXS measurements were performed in darkness on dark-purified protein. A 100% Pr state was prepared and maintained in a darkened SAXS experimental beam line by continuous sample illumination with monochromatic light of 760 nm wavelength. Data collection protocols were assessed offline by UV-Vis spectroscopy to ascertain whether a pure Pr state is attainable using similar conditions to those on the SAXS beam line (i.e., identical optics, light source, sample/light geometry, sample cell thickness, and protein concentration). A sample of 4 mg·ml⁻¹ protein concentration placed in a 2 mm UV-Vis cell path length was shown to undergo photo-conversion to a 100% Pr state when exposed to $\lambda = 760$ nm light within 60 s (Fig. 2B) and this state was maintained by continuous illumination, as evident from the second spectrum (dotted line) measured 30 min later. An *RpBphP1ΔHOS* UV-Vis absorption spectrum obtained from an illuminated sample indicates only partial photo-conversion, with both Pr and Pfr bands observed in the spectrum (Fig. 1C). This property was used to estimate *RpBphP1ΔHOS* impurity by comparing a SAXS sample to the spectrum comprising of pure *RpBphP1* and pure *RpBphP1ΔHOS* spectra combined in known proportions. Sample impurity could be modeled with the equation $A_{760}/A_{680} = 0.84\Phi + 0.0075$ where A_{760}/A_{680} is the absorption ratio measured at wavelengths 760 and 680 nm, and Φ is the fraction of *RpBphP1ΔHOS* impurity. The estimated SAXS sample impurity of $\Phi = 0.01 \pm 0.008$ is close to the error limit of measurement; hence, for the purposes of SAXS experiments, the sample can be considered > 99% pure. SAXS data revealed significant differences between states with the radii of gyration (Rg) 47.8/42.8 and Dmax 165 Å/140 Å for Pfr and Pr,

respectively (Fig. 3) suggesting that the Pfr molecule is more elongated than the Pr. This is also consistent with the SEC data, which indicates an elongated molecule, as observed by an anomalously high Pfr molecular weight. Molecular masses were determined from Pfr and Pr SAXS data by applying the SAXS MOW2 program [29], which yielded 163 ± 3.1 kDa for Pfr and 158 ± 2.5 kDa for Pr, indicating that the sample was pure and mono-dispersed, thus meeting the method requirements. Linear Guinier plots indicate absence of sample aggregation. The Guinier region for *RpBphP1* SAXS data is $S = 0.014\text{--}0.030$ Å⁻¹, corresponding to a lower real spatial limit of 450 Å which is ~3 times greater than the longest molecular dimension of 160 Å (Fig. 3). *Ab-initio* Pfr and Pr envelopes were reconstructed (Fig. 2C) using the data range $S = 0.0\text{--}0.5$ Å⁻¹ assuming molecular dimers but not molecular anisotropy, as this enabled the automatic *ab-initio* algorithm to create a two-fold axis along or across the longest molecular dimension, thus allowing for the possibility of parallel or pseudo anti-parallel dimers. Calculated envelopes show dimer axes along the longest molecular dimension, which is consistent with parallel dimers, while ruling out pseudo anti-parallel dimers. Ten Pfr and Pr envelopes were averaged with normalized spatial discrepancies (NSD) 0.65 and 0.82 respectively. The dimensions of Pfr and Pr dimers were 170 and 145 Å, respectively, tapering from 90 to 40 Å, and from 95 to 60 Å (Fig. 2C). In the third dimension, Pfr is 40 Å and Pr is 55 Å, indicating that the Pr dimer is more isotropic in shape than Pfr. Pr and Pfr intensity difference ranged from ~30 000 to 1000 photon counts in the scattering range $S = 0.0\text{--}0.15$ Å⁻¹ (Fig. 2J). The value $S = 0.15$ Å⁻¹ corresponds to a real spatial dimension of ~40 Å and is consistent with the observed Pfr and Pr envelope dimensions, which differ by ~30 Å, implying relatively large structural changes.

SAXS modeling of *RpBphP1*-Pfr structure

RpBphP1 models were developed and optimized to the SAXS data. The core of the model was based on the X-ray structure of *RpBphP1ΔHOS* in the Pfr state. Moreover, PCD-PAS9 domains were built into the parallel dimer structure of *PaBphP1*-PCD (3C2W) which is also a bathy-BphP with 32% sequence identity. To complete a full-length structure, a HOS domain was built with the program MODELLER using four sequence-aligned Dhp homologue X-ray structures (Fig. 4). The HOS model begins with a random coil (631–643) followed by a short α -helix (hJ', 643–651), which is disrupted at residue 651 by the sequence

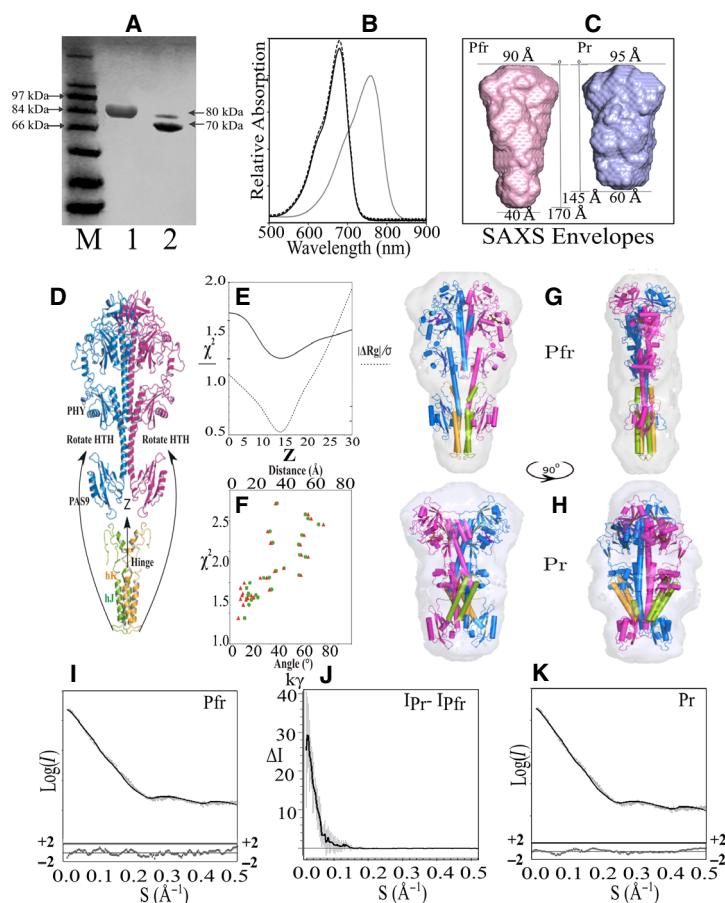


Fig. 2. Pfr and Pr SAXS data collection statistics and structure modeling. SDS/PAGE gel (A) of *RpBphP1* protein taken from the SAXS sample cell immediately after Pr SAXS data collection, the sample was created by continuous illumination with 760 nm light; M protein markers, lane 1–80 kDa full length protein, lane 2 – by comparison a 3 week old sample showing degradation to a 70 kDa *RpBphP1*ΔHOS fragment; (B) photo-conversion UV-Vis spectrum of *RpBphP1* measured offline using the same configuration as on the SAXS beamline; a white light source placed 4 cm from the sample with an intervening interference filter producing 760 nm wavelength light, UV-Vis optical path-length was 2 mm and protein concentration 4 mg·mL⁻¹. Pfr dark state (gray) and Pr (black) spectrum measured 1 min after illuminating with 760 nm light and 30 min after continuous illumination with 760 nm light (dashed). (C) SAXS *ab-initio* calculated envelopes, Pfr (pink) and Pr (blue); Pfr has a dimer axis of 170 Å, tapering from 90 to 40 Å; Pr has dimer axis 145 Å, tapering from 95 to 60 Å. (D) A full length Pfr parallel dimer was assembled from *RpBphP1*ΔHOS (4GW9) monomers and HOS domain dimers were created in Modeller from four sequence aligned Dhp homolog structures, Pr was assembled from a modified *XccBphP* and HOS dimers in which HTH motifs were bent at the hinge toward the PHY domain, relative orientations were optimized in CRY SOL. (E) CRY SOL Pfr optimization statistics; χ^2 (solid) as a function of HOS translation along the common Z twofold axis. The function $|\Delta Rg|/\sigma$ (dashed) was used to test the quality of calculated model Rg with respect to the observed Rg. The function is the modulus difference between observed and calculated Rg divided by its standard deviation σ ; (F) Pr structures were tested by rotating HTH motifs into Pr envelope ‘bulge’ region between PAS9 and PHY domains. CRY SOL χ^2 values are plotted with respect to ‘Angle’ (red triangles) which is an angle defined by the line from the hinge region, through the HTH helices, relative to the fixed line between hinge and helix hF in PHY, ‘Distance’ (green circles) is the corresponding separation in Å between T in HTH and hF. Pfr (G) and Pr (H) models positioned in their respective SAXS envelopes, front and side views. PCD-PAS9 domains are colored blue/magenta and HOS domains are colored green/orange, the same subunit is colored blue/orange and magenta/green. (I) and (K) are Log(I) versus scattering vector $S(\text{Å}^{-1})$ plots for Pfr and Pr states respectively. Solid lines are model calculated SAXS curves superimposed on SAXS experiment scattering points (gray), reduced residuals $(I(S)_{\text{exp}} - k \cdot I(S)_{\text{cal}})/\sigma(S)$ are shown which are differences between experimental and calculate intensities divided by experimental standard deviation $\sigma(S)$, k is a scaling factor; (J) shows intensity differences (ΔI) between illuminated Pr and dark Pfr SAXS data versus scattering vector S. ΔI values are expressed in kilo photons (k γ) units.

Pro-Gly, after which it continues into a helix-turn-helix (HTH) motif, formed by α -helices hJ and hK, and ends in a random coil at residues 720–730.

Several Modeller runs with different sets of HOS homologue structures consistently produced a break between hJ' and hJ helices, resulting in a range of

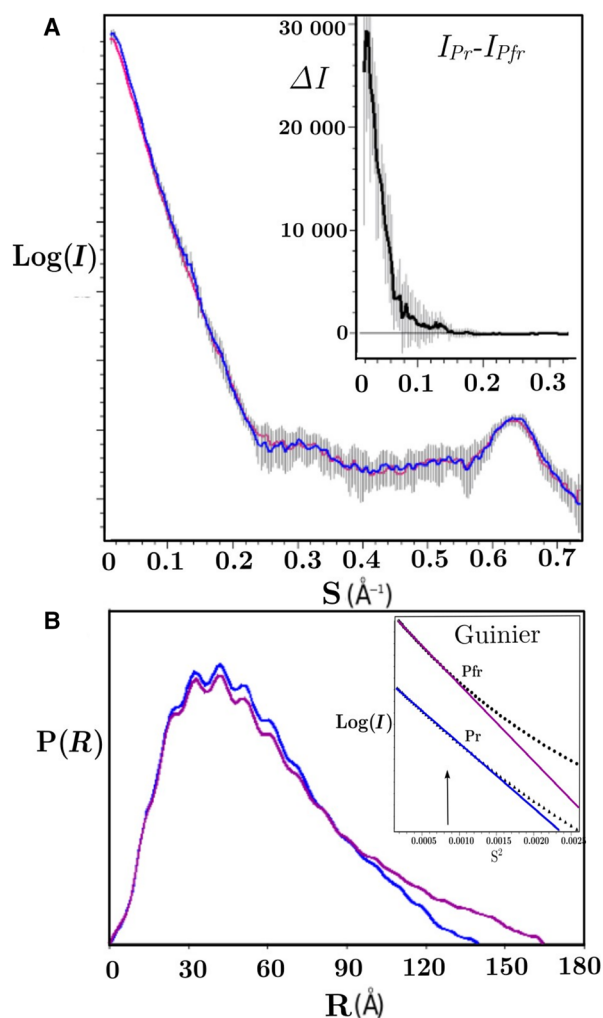


Fig. 3. Raw SAXS data $\log(I(S))$, $P(R)$, intensity differences and Guinier plots. Experimental SAXS scattering curves $\log(I(S))$ versus $S(\text{\AA}^{-1})$ with error bars (A), Pfr (red) and Pr (blue), inset are intensity differences ($I_{Pr}-I_{Pfr}$) between illuminated and dark SAXS data with superimposed error bars. (B) Pair-distance distribution function $P(R)$ versus distance $R(\text{\AA})$, Pfr D_{\max} is 160 and 140 \AA for Pr. Ripples in $P(R)$ originated from inter helix packing of $\sim 10 \text{\AA}$ which can also be seen as a scattering curve peak at $S = 0.65 \text{\AA}^{-1}$. Inset are Guinier plots of Pfr and Pr up to the maximum Guinier resolution $S_{\max} = 1.3/Rg$ where Rg is the radius of gyration 47.8 and 42.8 \AA for Pfr and Pr respectively. The Guinier plots are linear down to $S = 0.014 \text{\AA}^{-1}$ which corresponds to a real space dimension of 450 \AA (~ 3 times D_{\max}) suggesting the absence of molecular aggregates.

angles between helices of 20–95°, indicating that the 651–653 region behaves like a hinge. Full-length Pfr dimers were created by varying the relative positions of PCD-PAS9 and HOS domain dimers assuming a common twofold axis. A one-dimensional translation search of HOS dimers was conducted along the

Table 1. SAXS data collection and scattering parameters. Parameters are reported for the dark Pfr and illuminated Pr states.

Data collection parameters		Pfr state	Pr state
Instrument	STFC SRS Beam Line 2.1		
Beam geometry	Horizontal 5 mm \times Vertical 1 mm		
Wavelength (\AA)	1.54		
Camera lengths (m)	1.00 and 4.25		
S range (\AA^{-1})	0.014–0.730		
Exposure time ^a	60 \times 1 min		
Concentration ($\text{mg}\cdot\text{mL}^{-1}$)	1 and 4		
Temperature (K)	281		
<hr/>			
Structural parameters			
$I(0)$ from $P(r)$ (cm^{-1})		0.1113 ± 0.0003	0.1330 ± 0.0004
Rg from $P(r)$ (\AA)		47.84 ± 0.14	42.82 ± 0.11
$I(0)$ from Guinier (cm^{-1})		0.1186 ± 0.0005	0.1356 ± 0.0009
Rg from Guinier (\AA)		47.48 ± 0.28	42.47 ± 0.26
D_{\max} (\AA)		165 ± 5	140 ± 5
Porod volume (10^3\AA^3)		144 ± 1.5	136 ± 1.0
Molecular mass calculation			
MoW2 [29] from SAXS data (kDa)		163 ± 3.1	158 ± 2.5
Dimer mass from sequence (kDa)		160	
Software employed			
Data reduction	BSL/OTOKO [50]		
Data processing	GNOM [51]		
<i>ab-initio</i> SAXS envelope analysis	DAMMIN [43]		
Validation and averaging	SUPCOMB/DAMFIL/DAMAVR [44,52]		
Fragment modeling	HHPRED/MODELLER [48,49]		
Computation of model $I(S)$	CRY SOL [45]		
Rigid body refinement	SASREF [46]		
3D graphics presentation	PYMOL [53]		

^a Sixty 2D images were measured on a multi-wire chamber. Each image was exposed to X-rays for 60 seconds. The 60 images were checked for radiation damage and averaged into one image. No significant radiation damage was observed so the effective exposure time of the averaged image was 1 h.

twofold Z axis and was fitted against the SAXS scattering data in the CRY SOL program (Fig. 2D). A χ^2 fit to the SAXS data is shown in Fig. 2E. A second function, $|\Delta Rg|/\sigma$, was devised to evaluate the consistency between the calculated and experimental radii of gyration. It is defined as the modulus difference between the observed and calculated radius of gyration Rg divided by the experimental standard deviation σ (Fig. 2F). Both functions exhibit well-defined minima, $\chi^2 = 1.4$ and $|\Delta Rg|/\sigma = 0.6$, at a Z shift of 13.7 \AA relative to an arbitrary starting position. A 3D search was also employed using the rigid body refinement

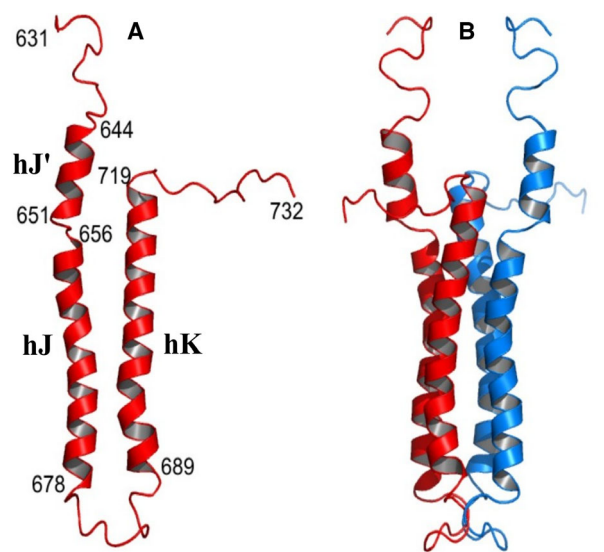


Fig. 4. Model of HOS domain. The HOS domain sequence was used to search for homolog PDB structures with *HHpred* toolkit. Four structures (PDBID, 4MT8, 4Q20, 3A0R, 5IDJ) were found, with probabilities $\sim 97\%$ and E-values in the range 7×10^{-5} – 1.5×10^{-6} , which are histidine kinase dimerization domains (Dhp) and were used as templates to build HOS domain dimers within the program *MODELLER* [49]. (A) is a single HOS domain comprising three α -helices hJ', hJ, and hK, (B) is a four α -helix bundle dimer made from two HOS monomers. Random coils were predicted in the sequence regions 632–643 and 720–732. A small α -helix segment hJ' was predicted at 644–651 with a break between residues 651–654 followed by a helix-turn-helix (HTH) motif which is hJ (656–678), T (679–688), and hK (689–719). In several runs of *Modeller*, using different Dhp template structures, the amino acid sequence 651–654 was observed to behave as a hinge; the angle between helix axes of hJ' and hJ varied between 20° and 95° .

program *SASREF*, which accessed 200 000 models resulting in the best solution with $\chi^2 = 1.6$, which was within 3.5 \AA of the above optimized model. A hydration shell-optimized *CRY SOL* calculation provides a fit to the SAXS data of $\chi^2 = 0.92$ (Fig. 2I) with reduced residuals across the scattering range *S* (0.0 – 0.5 \AA^{-1}) of -1 to $+1$. An important feature of the optimized Pfr model is that HOS hJ' helices must slide between hI helices which can form a short four α -helix bundle, while helices hJ and hK form a longer four α -helix bundle as observed in homolog Dhp dimerization domains of Histidine Kinases (Fig. 2G). It is not possible to see structural details within the SAXS envelopes, but X-ray structure models augment this data, indicating how four α -helix bundles can be best positioned within the envelope. In the case of Pfr helix hJ' must overlap with hI, to form a short four α -helix bundle, so that the longer four α -helix bundle can fully reside within the Pfr envelope.

Modeling of the *RpBphP1*-Pr structure

The *RpBphP1*-Pr model was based on the parallel dimer X-ray structure *XccBphP* in the Pr state, modified to the *RpBphP1* amino acid sequence and rebuilt in *Modeller* [15]. *XccBphP* structure has a shorter hI helix and was extended by three turns. The PAS9 domain in this model is rotated away from the twofold axis by $\sim 30^\circ$, as it is in *XccBphP*, which positions it partially outside the Pr SAXS envelope and reducing PAS9 rotation to $\sim 15^\circ$ positions it completely inside the envelope. Rotation variability is also observed in the structure of *XccBphP*, where PAS9 domains deviate from each other by $\sim 13^\circ$ indicating significant flexibility. As the Pr envelope is shorter than the Pfr envelope along the direction of the dimer axis by $\sim 25 \text{ \AA}$, an extended HOS domain four α -helix bundle cannot be accommodated within the envelope or reproduce the observed experimental *Rg*, which results in the $|\Delta Rg|/\sigma$ of around 6. A major rearrangement of HTH motifs (hJ and hK bundles) toward the N terminus is required to obtain $|\Delta Rg|/\sigma < 1$. An unaccounted-for feature in the Pr envelope is a 'bulge' situated between PHY and PAS9 domains which is capable of accommodating HTH motifs. A systematic search was performed of HTH orientations over the 'bulge' region using rearrangement about the hinge region while leaving hJ' helices anchored to hI. The behavior of a number of models is illustrated in Fig. 2F, which depicts a plot of *CRY SOL* χ^2 values for each model as a function of 'Angle' and 'Distance'. The angular displacement 'Angle' is defined by a line through the moving length of HTH relative to a fixed line joining the hinge region to helix hF in the PHY domain. 'Distance' is the corresponding spatial displacement of the turn T in HTH to hF. The top five models have $\chi^2 = 1.4$ – 1.6 with HTH motifs approaching hF with 'Angle' in the 6° – 12° range and 'Distance' in the 11 \AA – 16 \AA range. The best Pr model analyzed in *CRY SOL* with the optimized hydration shell parameters yields $\chi^2 = 0.94$. A more extensive rigid body refinement of the HTH position, accessing ~ 200 000 models, provides a solution of $\chi^2 = 0.7$ which differs from manually refined models by a slightly larger HTH 'Angle' displacement of around 20° . Figure 2K indicates that the best model fit to the SAXS data has reduced residuals between -1 and $+1$ across the entire *S* range. It is clear from exhaustive modeling that the best models have HTH motifs bent toward the PHY domain so that T of HTH approaches helix hF. Such models account for the 'bulge' region and the $|\Delta Rg|/\sigma$ and χ^2 values are reproduced to values that are less than 1.0. However,

accurate positioning of HTH motifs in the ‘bulge’ region is not possible, as reasonable solutions can be obtained over an ‘Angle’ range of 6°–25°. The uncertainty in positioning HTH within the ‘bulge’ is to be expected because these motifs contribute only 14% to the whole structure and thus to the SAXS data.

***RpBphP1-RpPpsR2* complex**

Bacterium *Rps. palustris* photosynthesizes anaerobically and, in strain CGA009, low oxygen conditions have been shown to weaken the binding of photosynthesis repressor proteins PpsR to specific DNA promoter regions. It has also been shown that far-red light antagonizes binding of *RpPpsR2* to DNA promoter regions on photo-conversion of *RpBphP1* to the active Pr state [2,4]. A combination of redox and light control ensures that photosynthetic apparatus is switched on under optimal conditions [4,30]. Under reducing conditions, *RpPpsR2* is observed to be predominantly a dimer in solution (Fig. 5F) and not a tetramer, as observed in *Rba. sphaeroides* [22]. SDS/PAGE experiments (Fig. 5D) show that, in a 5 mM DTT buffer, ~80% of intermolecular disulfide bonds are broken and dimers are free to monomerize. Ni-affinity chromatography was employed to monitor complex formation between untagged full-length *RpBphP1* or untagged *RpBphP1ΔHOS* and N-terminal His₆-tagged *RpPpsR2* (Fig. 5A–C). Ni-affinity experiments were performed in the dark and after stimulation with monochromatic 760 nm light. Passing dark adapted mixtures of untagged *RpBphP1* (B) and His₆-tagged *RpPpsR2* (P) through a Ni²⁺ column shows that B does not bind, eluting at buffer imidazole concentration of 10 mM, while P elutes as two peaks, at imidazole concentrations P1 = 210 mM and P2 = 378 mM, corresponding to monomeric and dimeric protein respectively. A 4 : 1 ratio B : P mixture, exposed to 760 nm light, elutes at 10 mM imidazole as a slightly reduced B peak and a BP peak at 280 mM imidazole concentration (Fig. 5A). BP is clearly a complex formed between B and P, as can be seen from eluted fractions monitored with 400 nm light, which is solely a biliverdin absorption wavelength (Fig. 5B). By comparison, *RpBphP1ΔHOS* (B₇₀) cannot form complexes with *RpPpsR2*, as confirmed by the absence of a 400 nm absorbing band at high imidazole concentrations (Fig. 5C). SDS/PAGE of peak BP shows that it separates into two peaks of molecular weight ~80 kDa (*RpBphP1*) and ~50 kDa (*RpPpsR2*), respectively, as shown in Fig. 5E. A linear SDS/PAGE absorption scan indicates that the *RpBphP1:RpPpsR2* complex has a protein ratio of

~2 : 1 (Fig. 5E) suggesting that a complex is formed between one *RpBphP1* dimer and one *RpPpsR2* monomer. A complex of two *RpBphP1* dimers and one *RpPpsR2* dimer cannot be ruled out. However, BP elutes at imidazole concentration of 280 mM which is closer to monomeric P1 than to dimeric P2, indicating that BP complexes probably bind to a Ni-affinity column through only one His₆ tag ligand. Moreover, because 20% of *RpPpsR2* disulfide bonds remain oxidized, a 110 kDa, SDS/PAGE peak would be expected if dimers were involved and this is not observed.

Discussion

Light-dependent properties of full-length bacteriophytochromes

Canonical type 1 BphPs contain two dimerization interfaces, a hydrophobic N-terminus interface, formed between helices hA, hB, and hE in the chromophore binding domain (CBD) and the C-terminus Dhp dimerization domain [31]. In *RpBphP1*, as the N-terminus interface is hydrophilic, only the C-terminus dimer interface exists. SAXS data suggest that, on photo-conversion to the Pr state, a structural rearrangement occurs at the C terminus of *RpBphP1*, which indicates weakening of the dimer interface, but dimers are still observed, and this is likely to be because the short four α -helix bundle remains. Light-induced structural changes, common to all BphP, are the structural rearrangements within the biliverdin pocket which are transmitted through the PCD domain and the long hI helix. Further structural changes are dependent on the type of BphP and the specific nature of C-terminus domains. It is therefore of interest to discuss how this differs amongst various full-length BphPs.

Bacteriophytochrome *DrBphP* from *Deinococcus radiodurans* is a type 1 BphP exhibiting a Pr dark state. The photo-conversion of *DrBphP* from Pr to Pfr states has been investigated by time resolved SAXS over the range 10 μ s to 500 ms [32] and significant differences were observed in Pfr and Pr SAXS envelopes even though radii of gyration values changed by <0.5 Å. The SAXS envelopes showed a rotation of the catalytic kinase domains about the dimer axis of ~50°. An alternative picture has emerged via single particle negative stain electron microscopy [33]. Two species of photo-converted particle were observed, Pfr and Pfr', which are shortened (143 Å to 124 Å) and broadened by a rotation of PHY domains by 35° away from the two-fold axis. Other differences were observed at the HK domains, whereas in the Pfr

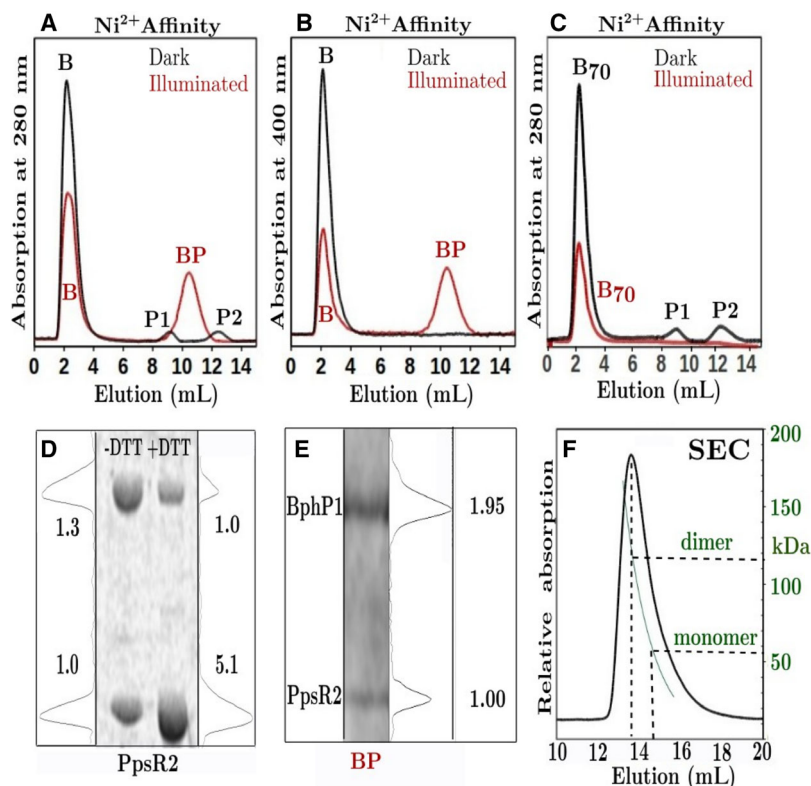


Fig. 5. Complex formation by Ni(II)-affinity chromatography and SDS/PAGE. Ni(II)-affinity plots (A–C) of samples incubated in the dark (black) or after illumination with light of wavelength 760 nm for 5 min (red). Samples were of untagged *RpBphP1* or untagged *RpBphP1* Δ HOS mixed with N-terminal His₆-tagged *RpPpsR2* in the protein ratio 4 : 1. Profiles (A) and (C) were monitored at the protein absorption wavelength 280 nm. (B) was measured at the biliverdin IX α absorption wavelength 400 nm. Protein peak B is full length *RpBphP1* and B₇₀ the 70 kDa fragment *RpBphP1* Δ HOS and in the dark both elute at imidazole concentrations of 10 mM, while peak P1 and peak P2 are the monomeric and dimeric His₆-tagged *RpPpsR2* eluting at 212 and 378 mM imidazole respectively. On illumination with 760 nm light the *RpBphP1*-*RpPpsR2* complex peak BP is observed eluting at 280 mM, the peak monitored at 400 nm (B) shows that this peak contains *RpBphP1*. (D) SDS/PAGE gels of *RpPpsR2* without (-DTT) and with (+DTT) 5 mM dithiothreitol. DTT breaks the disulfide bond between HTH domains. Linear scans of -DTT gels quantify *RpPpsR2* with disulfide bonds and free cysteine in the ratio of ~ 1.3 : 1 and ~ 1 : 5.1 for +DTT gels. (E) An SDS/PAGE gel of peak BP shows the presence two proteins within the complex in the protein ratio ~ 2 : 1, *RpBphP1* of molecular weight ~ 80 kDa and *RpPpsR2* of molecular weight ~ 50 kDa. (F) A SEC plot of *RpPpsR2* in 5 mM DTT indicating a dimer of mass ~ 117 kDa and a small tail which may indicate monomers at ~ 50 kDa.

particles, a partial disruption and separation of HK dimers was observed in the Pfr' particle HK dimers appear to be completely flexible and are practically invisible in EM micrographs. The observed disruption of HK domains suggests a change in the cross-over angle of hI helices forcing HK domains apart. This is in contrast with what is observed in *RpBphP1*, where it is the Pr state that is destabilized, but this may only be a reflection of the fact that it is the dark states that are structurally stable in bathy BphPs.

IsPadC has a diguanylyl cyclase (DGC) C-terminus domain that modifies the second messenger molecule c-di-GMP. Structural asymmetry is observed in the X-ray structure of *IsPadC*_{reg2} which has one monomer in the Pr state and the other in the Pfr state. Mutations

that stabilize the mixed state were engineered within the 25-residue coiled-coil region (Reg2). In the Pfr protomer, Reg2 mutations have the effect of straightening the coiled-coil, pushing its DGC domain closer to the two-fold axis. This mutant is catalytically very active, suggesting that asymmetry may be important to its function. Comparing the structures of *IsPadC* and *RpBphP1* shows that the Reg2 region maps directly onto the last 25 residues of helix hI in *RpBphP1*. Bowing of hI helices was observed in *XccBphP1*-Pr suggesting that changes in conformation here is also important in type 3 BphPs. It has been proposed that asymmetry brings DGC monomers closer together, which renders the DGC dimer more active. The movement of Pfr DGC domain toward the two-fold axis is

similar to that observed for *RpBphP1* where PAS9 domains are also closer to the two-fold axis in Pfr than in Pr. It is interesting that *IsPadC* SAXS data of native protein show large changes in the Rg and Dmax, which are 49–56 and 160–170 Å for the dark Pr and the light Pfr states respectively. The SAXS data indicate that, in both *IsPadC* and *RpBphP1*, Pfr states form more elongated dimers than for Pr. Based on the UV-Vis absorption spectra, which show incomplete photo-conversion, a mixed Pr/Pfr state would be expected in the SAXS light-illuminated data. However, a SAXS envelope has not been shown for this data. Such a calculation, if possible, would provide important information on the degree of asymmetry in the native protein.

Although comparison between different BphP types is useful, there are structural differences within the PCD that complicate matters. For example, the canonical dark-stable Pr state, in type 1 BphPs, has bowed helices joining PAS-GAF to PHY domains resulting in a large hole, whereas in other BphP types these helices are relatively straight, causing closer contacts within the dimer spine. It is interesting that the canonical *DrBphP* has shorter Pfr dimers but Pfr dimers in both *IsPadC* and *RpBphP1* are more elongated. This is despite *IsPadC* and *DrBphP* both being Pr dark stable, whereas *IsPadC* and *RpBphP1* have Pr and Pfr dark stable states, respectively, yet exhibit similar SAXS Pfr properties. Structural differences in Pr and Pfr states probably also depend on the very different functional requirements of C terminus domains which are; for type 1 *DrBphP* a movement of Kinase domains toward the Dhp, in type 5 *IsPadC* it is the bringing together of DGC domains to make catalytic active dimers and in type 3 *RpBphP1* a separation of HOS domains to facilitate the formation of complexes with *RpPpsR2*.

SAXS experiments were designed to obtain pure Pfr and Pr samples and to enable easier data analysis and *ab-initio* reconstruction of envelopes. However, in a real-world situation, light will be distributed across the full red to far-red energy spectrum. There is, therefore, no reason to suppose that mixed states could not exist for *RpBphP1*. One possible model of the complex will be presented below, which could apply to both pure Pr dimers and mixed Pfr:Pr dimers.

HOS dimerization interface

HOS domains are essential for the assembly of head-to-head parallel dimers through the formation of a C-terminus four α -helix bundle that has the same dimer formation function as do Dhp domains in type 1 BphPs. This similarity may point to a common

molecular ancestor, for example, type 3 BphP may have evolved from a type 1 BphP by insertion of a PAS9 domain and the loss of a kinase domain with additional point mutations that change Dhp domains to HOS domains. An important difference is that α -helix hI is continuous with Dhp helices in type 1 BphPs [14] whereas an intervening PAS9 domain disrupts this helix in *RpBphP1*. Although precise details on HOS contacts cannot be determined using the SAXS data, modeling nevertheless suggests that hJ' helices must overlap with hI helices so that the long four α -helix bundle can fall within the Pfr envelope. An interesting feature of our Pfr model is that helix hJ is brought in line with helix hI, creating a spatial helical topology that is similar to type 1 BphP, albeit with a break between these helices caused by an intervening PAS9 domain. The Pr model best fits the SAXS data when helices hJ and hK occupy the bulge region between PHY and PAS9 domains and this can be achieved through the hinge region which allows helices hJ and hK to articulate independently while leaving helices hJ' in situ to form a short four helix bundle with hI helices. The Pfr model is consistent with the observation that the Pfr state is functionally inert, and we propose that this is because the long four α -helix bundle confers structural stability to the HOS dimer while making HTH motifs unavailable for complex formation with *RpPpsR2*. Alternatively, our Pr model has free hJ and hK helices that can interact with *RpPpsR2*. The corollary of this is that an HOS deleted construct is incapable of forming complexes with *RpPpsR2*.

A mechanism for HOS domain activation

There are a number of studies characterizing photo-induced structural changes in the PCD domain which can be used to propose a tentative mechanism for HOS domain activation in the Pr state. Light excitation of biliverdin IX α causes isomerization of the D pyrrole ring and rearrangement of amino acid residues in the biliverdin pocket [34,8]. These changes are transmitted to a tongue motif (466–476) capping the biliverdin pocket where an α -helix unfolds into a β -strand on photo-conversion from the Pfr to Pr state. A two-stranded β -sheet is formed with β -strand 439–448, which elongates the tongue motif, pushing the PHY domain toward the C terminus [11,35]. These first light-induced structural changes characterize the transition from Pfr to Pr states and appear to be common to all BphPs. In type 3 *XccBphP-Pr*, the movement of PHY domain toward the C terminus causes curvature of α -helix hI and a rotation of the PAS9 domain away from the twofold axis [15]. A consequence of this

motion could be the weakening of interactions between HTH motifs. Hence, we propose that they separate and move independently toward their respective PHY domains through a rearrangement of conformation in the hinge region. Although SAXS data do not permit precise modeling of four α -helix bundle contacts, large movements of HTH motifs can be observed in the data which suggests that they must separate. Figure 6 provides a description of a possible mechanism for the transition from Pfr to Pr states. Whereas in our model a short and long four α -helix bundle stabilizes the Pfr dimer, making it functionally inert (Fig. 6A,B), the Pr dimer is only stabilized by a shorter four α -helix bundle (hJ' - hI)₂, as shown in Fig. 6E,F. To illustrate how the rotation of PAS9 may cause HTH separation, an intermediate conformation, Pr*, is shown assuming that HTH domains follow the movement of PAS9 domains. The intermediate structure can be calculated from transformation matrices defining PAS9 movement from Pfr to Pr positions and applying these to their respective Pfr HTH motifs (Fig. 6C,D). The motions of HTH motifs can be described as a radial angular separation of $\theta \sim 20^\circ$ and a shearing motion of $\psi \sim 15^\circ$, resulting in a separation between HTH motif turns (T) of ~ 20 Å. The formation of four α -helix bundles is energetically favorable due to hydrophobic and helix dipole interactions. However, if HTH motifs are forced apart, other factors can dominate, such as the minimization of amino acid conformation energies within the hinge region, which in the absence of any other constraints can result in large movements of HTH motifs. A feature of the Pr model is that HTH motifs are released from association within a four α -helix bundle so that their sticky inner helix surfaces are exposed and able to form a complex with *RpPpsR2*.

An *RpBphP1*-*RpPpsR2* complex model

SAXS and biochemical data suggest one possible *RpBphP1*:*RpPpsR2* complex model that is consistent with experimental results, although the model presented here should be treated with caution until hard structural data become available. Nevertheless, it is worthwhile discussing the constraints that observed data can impose on complex models. These constraints are: (a) the formation of a complex with a protein ratio *RpBphP1*:*RpPpsR2* of 2:1, (b) a central role played by HOS domains in the formation of complexes which also explains why *RpBphP1* Δ HOS cannot make complexes, and (c) complexes must be able to form with the Pr and not with the Pfr form of *RpBphP1*. A model was built based on these

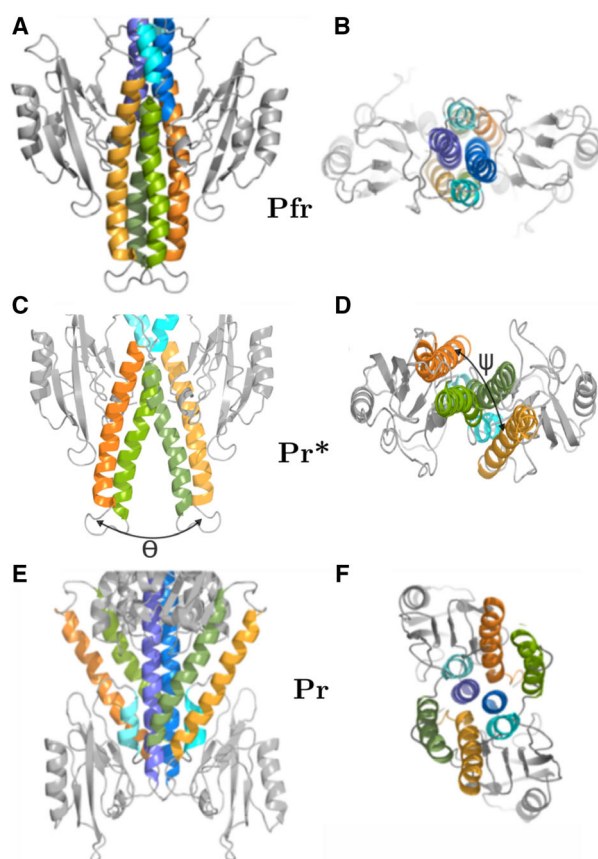


Fig. 6. HOS domain transition from Pfr to Pr state. The proposed HOS domain transition from Pfr (A and B) to Pr (E and F) via an intermediate Pr* (C and D) model (front and top views). Pfr-HOS dimers form a four α -helix bundle (hJ - hK)_A-(hJ - hK)_B where A and B are the dimer protein chains. The intermediate model Pr* shows the beginning of α -helix bundle separation and was generated from transformation matrices calculate by superimposition of PAS9 of Pfr onto Pr and applied to each Pfr PAS9-HTH structure in turn. Transition from Pfr to Pr* can be described as a two angle motion of HTH motifs (hJ - hK)_A and (hJ - hK)_B; the first is a separation angle $\theta \sim 20^\circ$ and the second a shearing angle $\psi \sim 15^\circ$. We propose that this is facilitated by bending at the hinge. PAS9 domains (gray), helix hI (blue), short α -helix hJ' (cyan), hJ (green) and hK (orange). The two subunits A and B are colored in darker and lighter shades respectively.

constraints and from the full-length structures of *RpBphP1* and *RpPpsR2*. A full-length model of *RpPpsR2* was based on a homolog partial structure *RsPpsR* Δ HTH (PDBID 4HH2) from *Rba. sphaeroides* which represents $\sim 83\%$ of the full-length molecule and has $\sim 28\%$ sequence identity with *RpPpsR2*. The missing part of the structure is the HTH domain at the C terminus belonging to the FIS family of DNA-binding proteins which form four α -helix bundles. A number of homologous HTH domains were found

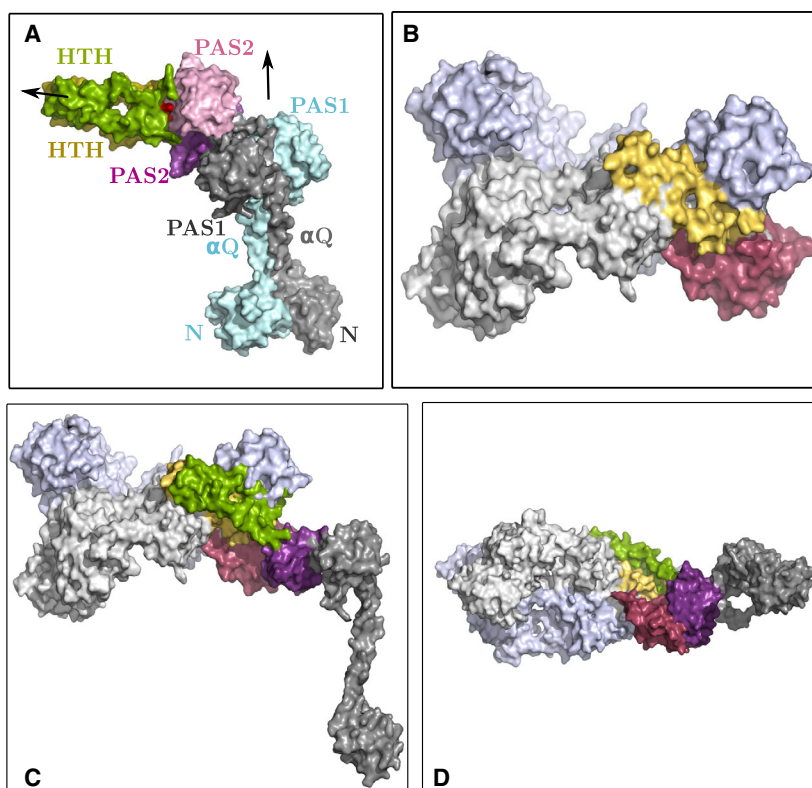


Fig. 7. Space filling model showing a possible $(RpBphP1)_2RpPpsR2$ complex. (A) space filling model of the full length *RpPpsR2* dimer synthesized in Modeller using a sequence modified PpsR homolog structure, 4HH2, from *Rba. sphaeroides* missing the C terminus HTH domain. The missing HTH domain was built in Modeller from PAS-HTH and HTH homolog structures, 4GCZ, 3A0R, 4I5S, 1NTC and 3E7L. The PAS domain of PAS-HTH overlaps with the PAS2 domain of 4HH2 facilitating alignment of the full-length structure. Arrows in (A) indicate two-fold axes relating N- α Q-PAS1 and PAS2-HTH respectively the latter is approximately at right angles to the first. One monomer is colored (gray, purple, green) the other (cyan, pink, yellow). HTH domain disulfide bond is colored red. (B) Space filling model of *RpBphP1*-Pr dimer, PAS9 (red) and HOS (yellow) are part of one monomer the rest of *RpBphP1* is colored pale gray or pale blue for each subunit. (C) front view of hetero-complex $(RpBphP1)_2-RpPpsR2$ in which contacts are made between HTH motif of HOS (yellow), HTH of an *RpPpsR2* monomer (green), contacts are also made between PAS9 (red) and PAS2 (purple) domains. (D) is as in (C) but viewed from the bottom.

using the online toolkit *HHpred* and these were used to build *RpPpsR2* in Modeller (Fig. 7A). Structures were of the HTH or PAS-HTH type, and the latter proved useful for alignment because the PAS domains overlap with the PAS2 domain of *RsPpsR* Δ HTH. The *RpPpsR2* model has two distinct twofold axes, the first relating N- α Q-PAS1 domains and the second approximately perpendicular to the first relating PAS2-HTH domains.

A model of the complex that satisfies the aforementioned constraints has an *RpBphP1*-Pr dimer forming a complex with a single *RpPpsR2* monomer. Combined constraints of a 2 : 1 complex protein ratio and HOS domains that play a direct role in complex formation require that HTH motifs form hetero-dimer four α -helix bundles such that one HTH is donated from an *RpBphP1* dimer (HTH_B) and the other from an *RpPpsR2* monomer (HTH_P). Two forces are

involved in the formation of four α -helix bundles, electrostatic anti-parallel helix dipoles and hydrophobic forces between inner α -helix surfaces. These inner surfaces are usually lined with small hydrophobic residues, such as Leu, Ala, Ile, and Val. However, homo dimers must separate for these hydrophobic surfaces to be available for interaction and so that hetero dimers are formed. This is the case in the *RpBphP1*-Pr model as HTH motifs are bent back whereby hydrophobic surfaces point outwards (Fig. 7B), as well as for *RpPpsR2*, which can be monomerized when the inter-molecular disulfide bond is broken. A model of the complex is shown in Fig. 7C,D. An unexpected feature of the model is that the PAS9 domain of *RpBphP1* and the PAS2 domain of *RpPpsR2* naturally come into contact with one another. Proteins with dimeric PAS-HTH domain architecture are interesting because they are known to form both homo and

hetero dimers. In these structures, HTH domains serve as dimerization domains, while PAS domains provide stability and specificity to the dimer [36]. Heterodimers are well-documented, such as the ARNT protein heterodimer with AHR [37] and mammalian hypoxia-inducible factor [38] or dioxin receptor, which makes homo dimers as well as heterodimers with ARNT [39]. A feature of the proposed complex model is that it can work with pure Pr dimers and mixed Pfr/Pr *RpBphP1* dimers.

Rps. palustris CGA009 is a model strain that has been extensively used for genomic sequencing and other biochemical experiments. Unfortunately, in this strain, gene *bphP1* is frame shifted and not expressed. Thus, for this work, *bphP1* was PCR amplified from strain 2.1.6 which has an identical gene sequence as that of *bphP1* in strain CGA009 if the frame shift is repaired [40]. In fact it has been shown that a repaired *bphP1* gene is functionally a far-red light sensor and that it targets the repressor protein *RpPpsR2* in a redox sensing way [4]. However, strain CGA009 *RpPpsR2* differs from all other DNA sequenced wild type strains (> 10) which do not have a redox Cys in the HTH domain and are thus non-redox sensing forms of PpsR. Consequently, in most strains, *RpPpsR2* is normally only under the control of light by interaction with *RpBphP1*. It has been suggested that CGA009 is a laboratory strain which has acquired a frame-shifted *bphP1* and an R428C *RpPpsR2* mutation over time due to laboratory selection of pigmented bacteria under certain laboratory conditions, which highlights the pitfalls of using laboratory strains [41]. The work presented here on complex formation between a spectrally active *RpBphP1* and an *RpPpsR2* in reducing conditions is therefore a good model of complex formation in wild-type strains. Although wild strains of *RpPpsR2* are not redox sensors, in *Rps. palustris*, photosynthesis remains under redox control through other means. For example, the cytochrome *c2* encoding gene *cycA* transfers electrons to the Reaction Centre from outside the cell and the redox controlled transcription regulator encoding gene *aerR* has promoter regions which are recognized by PpsR2. Also, PpsR2 controls the expression of PpsR1 which is a redox sensor and binds to a number of promoter regions within the major cluster of photosynthetic genes. There is therefore a complex interplay between light-dependent antagonistic action of *RpBphP1* on *RpPpsR2* and redox potential on *RpPpsR1* but not on *RpPpsR2* [19]. A fine modulation and control of photosynthesis has therefore been achieved in this bacterium by partitioning light- and redox-dependent gene expression between two different PpsR molecules.

Materials and methods

Cloning and purification

Gene *rpbphP1* was designed *in silico* to be used in spectroscopy, size exclusion chromatography and Ni²⁺ affinity column binding experiments. The gene was synthesized by GenScript (Piscataway, NJ, USA) in plasmid pUC57 with *NdeI* restriction site at the 5' end and sites *HindIII* and *XhoI* at the 3'. The gene was sub-cloned, using STRU-cloning protocol [42], into pET28a (*NdeI* and *HindIII*), pET24a (*NdeI* and *HindIII*) and pET24a (*NdeI* and *XhoI*) to obtain expression constructs for N-terminal His-tag, C-terminal His-tag and untagged *RpBphP1* protein respectively. For SAXS experiments gene *rpbphP1* was PCR amplified from genomic DNA of bacterium *Rhodospseudomonas palustris* strain 2.1.6 and sub-cloned into pET28a. Transformed *E. coli* BL21 (DE3) cells were grown to a density of OD₆₀₀ = 0.6, induced with 0.02 mM IPTG and incubated overnight at 18 °C. Cells were harvested and re-suspended in chromatography buffers containing 10 mM Biliverdin IX α (Sigma-Aldrich Company Ltd, Dorset, UK), protease inhibitor tablets (Roche Diagnostics GmbH, Mannheim, Germany), DNaseI (10 $\mu\text{g}\cdot\text{mL}^{-1}$), Triton X-100 (0.5%) and disrupted by French Press. Supernatants were loaded onto an affinity His-TrapTM HP or ion exchange HiTrap QFF columns (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA). Finally, proteins were gel filtered using a HiLoad 26/60 Superdex 200 column (GE Healthcare) in a 5 mM TrisHCl pH 8 and 10 mM NaCl buffer. Plasmid pET28a with incorporated gene *ppsR2* from *Rhodospseudomonas palustris* was kindly provided by Shinji Masuda (Tokyo Institute of Technology) and Tom Beatty (University of British Columbia). Protein was expressed in *E. coli* BL21 (DE3), using the same protocol as *RpBphP1*, and purified on a His-Trap HP column followed by gel filtration and stored in a final buffer of 20 mM Tris-HCl pH 8, 300 mM NaCl, and 5 mM dithiothreitol.

UV-visible spectroscopy

UV-Visible spectra were recorded at room temperature in a Perkin Elmer Lambda 35 UV/VIS spectrometer. Spectra were recorded in the dark or after illumination with monochromatic light produced by passing white light through an interference filter with a bandwidth of 10 nm centered on wavelength 760 nm (Knight Optical Ltd, Harrietsham, Kent, UK). The fiber optic/filter system was placed ~ 4 cm from the sample which provided a photon flux density of ~ 250 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ at the sample. There was sufficient light to photo-convert a Pfr dark state into a 100% Pr state within 1 min. The monochromatic light was passed through the top of the UV-Vis cuvette, with a sample path-length of 2 mm, and so full photo-conversion to

the Pr state was achieved over a wide range of protein concentrations, 0.04–4 mg·mL⁻¹.

Gel filtration analysis

Purified protein was loaded onto a Superose 12 10/300 GL in 200 µL volumes at concentrations between 1 and 4 mg·mL⁻¹ in 5 mM TrisHCl pH 8 and 10 mM NaCl. For *RpBphP1* this was either in the dark or after illumination with 760 nm light for 5 min. The column was calibrated with high molecular weight calibration kit (GE Healthcare). Data were plotted on a logarithmic scale using K_{av} , to obtain estimates of protein molecular weight in solution.

RpBphP1-*RpPpsR2* complex formation analysed by Ni(II) affinity chromatography

N-terminal His₆-tagged *RpPpsR2* and untagged *RpBphP1* or untagged *RpBphP1*ΔHOS was mixed in a ratio 1 : 4 respectively. Samples were pre-incubated in the dark or illuminated for 5 min with light of wavelength 760 nm to achieve a full Pr state. Chromatography was carried out over 10–20 min using a 1 mL His-Trap HP column and an ÄKTA Explorer. The elution was monitored at the protein absorption wavelength 280 nm and the biliverdin IX α absorption a wavelength of 400 nm. *RpBphP1* can only be eluted at high imidazole concentrations when in complex with His₆-*RpPpsR2* and monitoring absorption at these two wavelengths allows peaks to be differentiated between those that contain His₆-*RpPpsR2* only and those also in complex with *RpBphP1*.

SAXS data measurement

Small-angle X-ray scattering (SAXS) data were collected at STFC SRS facility beam line 2.1 at the wavelength $\lambda = 1.54 \text{ \AA}$. A SAXS X-ray camera was assembled at two different lengths, 1.00 and 4.25 m, to cover scattering ranges $0.052 < S < 0.730 \text{ \AA}^{-1}$ and $0.014 < S < 0.183 \text{ \AA}^{-1}$ respectively (Table 1). Samples were washed several times with 5 mM Tris buffer (pH 8) in a Vivaspin 500 centrifugal filter unit (Satorius AG, Goettingen, Germany) to remove chemicals such as salt and imidazole. The sample was concentrated to $\sim 4 \text{ mg}\cdot\text{mL}^{-1}$ and spin column flow through buffer was retained for SAXS data background subtraction. For the longer camera length, the sample was further diluted to $1 \text{ mg}\cdot\text{mL}^{-1}$. Samples were injected into a 50 µL volume brass cell with X-ray path length of 1 mm covered on both sides by thin mica windows and chilled to 8 °C by water-circulation around the brass cell. Data were collected as 60, 1-min exposure, images on a time resolved multi-wire area detector, assessed for time dependant radiation damage and averaged. A Pfr dark state was prepared by protein purification in the dark and pre-incubation in the dark

overnight. Pr state sample was induced by continuous illumination with white light passing through a 10 nm bandwidth interference filter centered on 760 nm (Knight Optical (UK) Ltd). The radius of gyration (R_g) was evaluated using the Guinier approximation $I(S) = I(0)\exp(-S^2R_g^2/3)$ for $S\cdot R_g < 1.3$ and was also calculated with Dmax by the indirect Fourier-transform program GNOM. Molecular weights were determined with the program SAXS mow2 [29]. For accurate molecular weight the method assumes a mono-dispersed sample and is also a check on the purity of the sample. The monochromatic light source system, protein concentration, sample path length, and photon flux density used on the SAXS beam line was the same as used in UV-Vis spectroscopy experiments so that a full photo-conversion to the Pr state was guaranteed.

SAXS data processing and modeling

The EMBL program suite ATSAS (<http://www.embl-hamburg.de/biosaxs/software.html>) was used for data analysis. *Ab-initio* envelopes of dark Pfr and illuminated Pr states were calculated from SAXS data in 10 runs, for each state, with the program DAMMIN [43]. The programs SUPCOMB, DAMFIL and DAMAVER were used to calculate average envelopes [44] which were isotropically reduced so that maximum envelope dimensions were the same as Dmax derived from pair distribution functions P(R). Models were initially fitted into SAXS envelopes using the graphics program PYMOL equipped with the plugin CRY SOL to access the quality of model fit to SAXS data [45]. Domain orientations were fine-tuned using χ^2 as a measure of fit between calculated and measured SAXS scattering curves. A rigid body refinement program SASREF was also used to perform an automated 3D search of domain alignments [46]. Discrepancies between experimental and model scattering data were assessed with reduced residuals $(I(S)_{\text{exp}} - k \cdot I(S)_{\text{cal}}) / \sigma(S)$, as a function of S , where $I(S)_{\text{exp}}$ are measured intensities and $I(S)_{\text{cal}}$ are intensities calculated from atomic models, k a scale factor and $\sigma(S)$ are SAXS data standard deviations [47]. A second target function, $|\Delta R_g| / \sigma$, was also used which calculates the consistency between calculated and experimental radii of gyration (R_g), and is defined as the modulus difference between R_g values divided by the experimental standard deviation σ and should be less than 1.

Domain modeling

Atomic models of *RpBphP1* HOS and *RpPpsR2* HTH domains were built so that full length models of *RpBphP1* and *RpPpsR2* could be created. Homolog proteins were found with the toolkit *HHPred* [48] (<https://toolkit.tuebingen.mpg.de>) and sequence aligned to their X-ray structure sequences. Best homolog structures based on hit probabilities, usually 4 to 6 structures, were selected as PDB

templates to automatically build 3D structures with the program MODELLER [49].

Acknowledgements

We thank the staff of STFC Daresbury Synchrotron Radiation Source for beam line access and assistance. Dom Bellini was supported on the BBSRC grant BB/E002609/2 and Kate Evans was supported by a Daresbury Laboratory/John Moores University (Liverpool, UK) PhD scholarship.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

MP- conception and design of experiments, writing and critical revision of the manuscript, SAXS data collection, *ab-initio* envelope determination, modeling of SAXS data and measurement of UV-Vis spectra. DB- cloning and sample purification for SEC measurements, Ni(II)-affinity chromatography measurement of complex formation, critical reading of manuscript and measurement of UV-Vis spectra. KE- PCR cloning of SAXS RpBphP1 sample, protein purification and characterization of purity by biochemical methods and UV-Vis spectroscopy. GG- collection of SAXS dark and light induced state data, processing of raw images, reduction and analysis of SAXS data and revision of the manuscript. AF-S- assessing quality of Pfr and Pr samples for SAXS experiments, checking SAXS sample purity by SDS/PAGE electrophoresis, UV-Vis spectroscopy and assisting in SAXS data collection.

References

- Rottwinkel G, Oberpichler I & Lamparter T (2010) Bathy phytochromes in rhizobial soil bacteria. *J Bacteriol* **192**, 5124–5133.
- Giraud E, Fardoux J, Fourier N, Hannibal L, Genty B, Bouyer P, Dreyfus B & Vermeglio A (2002) Bacteriophytochrome controls photosystem synthesis in anoxygenic bacteria. *Nature* **417**, 202–205.
- Bellini D & Papiz MZ (2012) Structure of a bacteriophytochrome and light-stimulated protomer swapping with a gene repressor. *Structure* **20**, 1436–1446.
- Braatsch S, Johnson JA, Noll K & Beatty JT (2007) The O₂-responsive repressor PpsR2 but not PpsR1 transduces a light signal sensed by the BphP1 phytochrome in *Rhodospseudomonas palustris* CGA009. *FEMS Microbiol Lett* **272**, 60–64.
- Bruscella P, Eraso JM, Roh JH & Kaplan S (2008) The use of chromatin immunoprecipitation to define PpsR binding activity in *Rhodobacter sphaeroides* 2.4.1. *J Bacteriol* **190**, 6817–6828.
- Li H, Zhang J & Vierstra RD (2010) Quaternary organization of a phytochrome dimer as revealed by cryoelectron microscopy. *Proc Natl Acad Sci USA* **107**, 10872–10877.
- Bellini D & Papiz MZ (2012) Dimerization properties of the RpBphP2 chromophore-binding domain crystallized by homologue-directed mutagenesis. *Acta Crystallogr D* **68**, 1058–1066.
- Wagner JR, Brunzelle JS, Forest KT & Vierstra RD (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325–331.
- Yang X, Stojkovic EA, Kuk J & Moffat K (2007) Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. *Proc Natl Acad Sci USA* **104**, 12571–12576.
- Burgie ES, Bussell AN, Walker JM, Dubiel K & Vierstra RD (2014) Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome. *Proc Natl Acad Sci USA* **111**, 10179–10184.
- Essen LO, Mailliet J & Hughes J (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc Natl Acad Sci USA* **105**, 14709–14714.
- Nagano S, Scheerer P, Zubow K, Michael N, Inomata K, Lamparter T & Krauss N (2016) The crystal structures of the N-terminal photosensory core module of *Agrobacterium* phytochrome Agp1 as parallel and anti-parallel dimers. *J Biol Chem* **291**, 20674–20691.
- Yang X, Kuk J & Moffat K (2008) Crystal structure of *Pseudomonas aeruginosa* bacteriophytochrome: photoconversion and signal transduction. *Proc Natl Acad Sci USA* **105**, 14715–14720.
- Yang X, Stojkovic EA, Ozarowski WB, Kuk J, Davydova E & Moffat K (2015) Light signaling mechanism of two tandem bacteriophytochromes. *Structure* **23**, 1179–1189.
- Otero LH, Klinke S, Rinaldi J, Velazquez-Escobar F, Mroginski MA, Fernandez Lopez M, Malamud F, Vojnov AA, Hildebrandt P, Goldbaum FA *et al.* (2016) Structure of the full-length bacteriophytochrome from the plant pathogen *Xanthomonas campestris* provides clues to its long-range signaling mechanism. *J Mol Biol* **428**, 3702–3720.
- Bonomi HR, Toum L, Sycz G, Sieira R, Toscani AM, Gudesblat GE, Leskow FC, Goldbaum FA, Vojnov AA & Malamud F (2016) *Xanthomonas campestris* attenuates virulence by sensing light through a

- bacteriophytochrome photoreceptor. *EMBO Rep* **17**, 1565–1577.
- 17 Gourinchas G, Etzl S, Gobl C, Vide U, Madl T & Winkler A (2017) Long-range allosteric signaling in red light-regulated diguanylyl cyclases. *Sci Adv* **3**, e1602498.
 - 18 Gourinchas G, Heintz U & Winkler A (2018) Asymmetric activation mechanism of a homodimeric red light-regulated photoreceptor. *eLife* **7**. <https://doi.org/10.7554/eLife.34815.001>
 - 19 Jaubert M, Zappa S, Fardoux J, Adriano JM, Hannibal L, Elsen S, Lavergne J, Vermeglio A, Giraud E & Pignol D (2004) Light and redox control of photosynthesis gene expression in *Bradyrhizobium*: dual roles of two PpsR. *J Biol Chem* **279**, 44407–44416.
 - 20 Gomelsky M & Kaplan S (1995) Genetic evidence that PpsR from *Rhodobacter sphaeroides* 2.4.1 functions as a repressor of puc and bchF expression. *J Bacteriol* **177**, 1634–1637.
 - 21 Ponnampalam SN & Bauer CE (1997) DNA binding characteristics of CrtJ. A redox-responding repressor of bacteriochlorophyll, carotenoid, and light harvesting-II gene expression in *Rhodobacter capsulatus*. *J Biol Chem* **272**, 18391–18396.
 - 22 Gomelsky M, Horne IM, Lee HJ, Pemberton JM, McEwan AG & Kaplan S (2000) Domain structure, oligomeric state, and mutational analysis of PpsR, the *Rhodobacter sphaeroides* repressor of photosystem gene expression. *J Bacteriol* **182**, 2253–2261.
 - 23 Masuda S, Dong C, Swem D, Setterdahl AT, Knaff DB & Bauer CE (2002) Repression of photosynthesis gene expression by formation of a disulfide bond in CrtJ. *Proc Natl Acad Sci USA* **99**, 7078–7083.
 - 24 Gomelsky M & Kaplan S (1997) Molecular genetic analysis suggesting interactions between AppA and PpsR in regulation of photosynthesis gene expression in *Rhodobacter sphaeroides* 2.4.1. *J Bacteriol* **179**, 128–134.
 - 25 Masuda S & Bauer CE (2002) AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in *Rhodobacter sphaeroides*. *Cell* **110**, 613–623.
 - 26 Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, Land ML, Pelletier DA, Beatty JT, Lang AS *et al.* (2004) Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodospseudomonas palustris*. *Nat Biotechnol* **22**, 55–61.
 - 27 Iakoucheva LM, Kimzey AL, Masselon CD, Smith RD, Dunker AK & Ackerman EJ (2001) Aberrant mobility phenomena of the DNA repair protein XPA. *Protein Sci* **10**, 1353–1362.
 - 28 Plath F, Ringler P, Graff-Meyer A, Stahlberg H, Lauer ME, Rufer AC, Graewert MA, Svergun D, Gellermann G, Finkler C *et al.* (2016) Characterization of mAb dimers reveals predominant dimer forms common in therapeutic mAbs. *mAbs* **8**, 928–940.
 - 29 Fischer H, de Oliveira Neto M, Napolitano HB, Polikarpov I & Craievich AF (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J Appl Crystallogr* **43**, 101–109.
 - 30 Braatsch S, Bernstein JR, Lessner F, Morgan J, Liao JC, Harwood CS & Beatty JT (2006) *Rhodospseudomonas palustris* CGA009 has two functional ppsR genes, each of which encodes a repressor of photosynthesis gene expression. *Biochemistry* **45**, 14441–14451.
 - 31 Takala H, Bjorling A, Linna M, Westenhoff S & Ihalaainen JA (2015) Light-induced changes in the dimerization interface of bacteriophytochromes. *J Biol Chem* **290**, 16383–16392.
 - 32 Bjorling A, Berntsson O, Lehtivuori H, Takala H, Hughes AJ, Panman M, Hoernke M, Niebling S, Henry L, Henning R *et al.* (2016) Structural photoactivation of a full-length bacterial phytochrome. *Sci Adv* **2**, e1600920.
 - 33 Burgie ES, Wang T, Bussell AN, Walker JM, Li H & Vierstra RD (2014) Crystallographic and electron microscopic analyses of a bacterial phytochrome reveal local and global rearrangements during photoconversion. *J Biol Chem* **289**, 24573–24587.
 - 34 Lagarias JC & Lagarias DM (1989) Self-assembly of synthetic phytochrome holoprotein in vitro. *Proc Natl Acad Sci USA* **86**, 5778–5780.
 - 35 Ulijasz AT, Cornilescu G, Cornilescu CC, Zhang J, Rivera M, Markley JL & Vierstra RD (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. *Nature* **463**, 250–254.
 - 36 Taylor BL & Zhulin IB (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* **63**, 479–506.
 - 37 Sakurai S, Shimizu T & Ohto U (2017) The crystal structure of the AhRR-ARNT heterodimer reveals the structural basis of the repression of AhR-mediated transcription. *J Biol Chem* **292**, 17609–17616.
 - 38 Hogenesch JB, Gu YZ, Jain S & Bradfield CA (1998) The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proc Natl Acad Sci USA* **95**, 5474–5479.
 - 39 Pongratz I, Antonsson C, Whitelaw ML & Poellinger L (1998) Role of the PAS domain in regulation of dimerization and DNA binding specificity of the dioxin receptor. *Mol Cell Biol* **18**, 4079–4088.
 - 40 Evans K, Fordham-Skelton AP, Mistry H, Reynolds CD, Lawless AM & Papiz MZ (2005) A bacteriophytochrome regulates the synthesis of LH4 complex in *Rhodospseudomonas palustris*. *Photosynth Res* **85**, 169–180.
 - 41 Giraud E, Zappa S, Jaubert M, Hannibal L, Fardoux J, Adriano JM, Bouyer P, Genty B, Pignol D &

- Vermeglio A (2004) Bacteriophytochrome and regulation of the synthesis of the photosynthetic apparatus in *Rhodospseudomonas palustris*: pitfalls of using laboratory strains. *Photochem Photobiol Sci* **3**, 587–591.
- 42 Bellini D, Fordham-Skelton AP & Papiz MZ (2011) STRU-cloning: a fast, inexpensive and efficient cloning procedure applicable to both small scale and structural genomics size cloning. *Mol Biotechnol* **48**, 30–37.
- 43 Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* **76**, 2879–2886.
- 44 Kozin MB & Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **34**, 33–41.
- 45 Svergun D, Barberato C & Koch MHJ (1995) CRY SOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768–773.
- 46 Petoukhov MV & Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* **89**, 1237–1250.
- 47 Trehwella J, Duff AP, Durand D, Gabel F, Guss JM, Hendrickson WA, Hura GL, Jacques DA, Kirby NM, Kwan AH *et al.* (2017) 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. *Acta Crystallogr D* **73**, 710–728.
- 48 Soding J (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**, 951–960.
- 49 Sali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779–815.
- 50 Koch MHJ & Bendall P (1981) INSCOM: an interactive data evaluation program for multichannel analysis-type data. *Proc Dig Equip Users Soc*, 13–16.
- 51 Svergun D (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25**, 495–503.
- 52 Volkov VV & Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *Acta Crystallogr D* **36**, 860–864.
- 53 The Pymol Molecular Graphics System, Version 2.1.0. Schrodinger, LLC.