Published in final edited form as:

J Mol Biol. 2008 November 7; 383(2): 403–413. doi:10.1016/j.jmb.2008.08.034.

SOLUTION STRUCTURE OF A CYANOBACTERIAL PHYTOCHROME GAF DOMAIN IN THE RED LIGHT-ABSORBING GROUND STATE

Gabriel Cornilescu^{1,†}, Andrew T. Ulijasz^{2,†}, Claudia C. Cornilescu³, John L. Markley^{1,3}, and Richard D. Vierstra^{2,*}

¹ National Magnetic Resonance Facility at Madison, University of Wisconsin Madison, Wisconsin 53706 USA

² Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706 USA

³ Center for Eukaryotic Structural Genomics, University of Wisconsin Madison, Wisconsin 53706 USA

Summary

The unique photochromic absorption behavior of phytochromes (Phys) depends on numerous reversible interactions between the bilin chromophore and the associated polypeptide. To help define these dynamic interactions, we determined by NMR spectroscopy the first solution structure of the chromophore-binding cGMP phosphodiesterase/adenylcyclase/FhlA (GAF) domain from a cyanobacterial Phy assembled with phycocyanobilin (PCB). The three dimensional NMR structure of *Synechococcus* OS-B' Cph1 in its red light-absorbing Pr state revealed that PCB is bound to Cys-138 of the GAF domain via the A-ring ethylidene side chain and is buried within the GAF domain in a *ZZZsyn,syn,anti* configuration. The D ring of the chromophore sits within a hydrophobic pocket and is tilted approximately 80° relative to the B/C rings by contacts with Lys-52 and His-169. The solution structure revealed remarkable flexibility for PCB and several adjacent amino acids, indicating that the Pr chromophore has more freedom in the binding pocket than anticipated. The propionic acid side chains of rings B and C and Arg-101 and Arg-133 nearby are especially mobile and can assume several distinct and energetically favorable conformations. Mutagenic studies on these arginines, which are conserved within the Phy superfamily, revealed that they have opposing roles with Arg-101 and Arg-133 helping stabilize and destabilize Pfr, respectively. Given the fact that the *Synechococcus* OS-B' GAF domain can by itself complete the Pr→Pfr photocycle, it should now be possible to determine the solution structure of the far-red light-absorbing Pfr chromophore and surrounding pocket using this Pr structure as a framework.

Keywords

phytochrome; GAF domain; bilin chromophore; red light-absorbing form; NMR; protein structure

*Corresponding author: Dr. Richard D. Vierstra, Department of Genetics, 425-G Henry Mall, University of Wisconsin-Madison, Madison, WI 53706, Tel: 608-262-8215, Fax: 608-262-2976, vierstra@wisc.edu.

[†]Both authors contributed equally to this work.

[§]Atomic coordinates and structural constraints have been deposited in the Protein Data Bank (PDB ID code 2K2N) and the NMR data in the BioMagResBank (BMRB access code 15717).

Introduction

Microorganisms and plants have evolved a collection of photoreceptors that help them adapt both physiologically and developmentally to the ambient light environment. One of the most influential is the phytochrome (Phy) superfamily, a large collection of diverse photoreceptors that use bilins (or linear tetrapyrroles) as the chromophore.^{1,2,3} Through unique interactions between the bilin and its binding pocket, Phys can exist in two stable photointerconvertible states, a red light (R)-absorbing Pr form that is typically the ground state, and a far-red light (FR)-absorbing Pfr form that is often biologically active. Through this photochromicity, Phys act as reversible switches in various photosensory cascades. Processes regulated by these switches include seed germination, chloroplast development, shade avoidance and flowering time in plants³, and phototaxis and photosynthetic potential in bacteria.¹

Phylogenetic analyses revealed that individual Phys are composed of a series of defined structural domains that are combinatorially arranged to impart distinct attributes to the photoreceptors.^{1,2,4} The photosensing portion typically includes an N-terminal Per/Arndt/Sim (PAS) domain immediately followed by a cGMP phosphodiesterase/adenylcyclase/FhlA (GAF) domain. The GAF domain provides the bilin lyase activity, and together with the bound chromophore, generates most of the unique R/FR photochromic absorption properties of Phys. For bacterial (BphPs) and fungal Phys, biliverdin IX α (BV) is the chromophore,¹ which is attached via a conserved cysteine upstream of the PAS domain. For cyanobacterial (Cph) and plant Phys, phycocyanobilin (PCB) and phytochromobilin are the chromophores, respectively, which are bound via a conserved cysteine in the GAF domain.² This switch in bilin and attachment site appears to have been advantageous to photosynthetic organisms by better enabling their Phys to detect shading by competitors.⁵ Downstream of the GAF domain is a structurally related phytochrome (PHY) domain that helps complete the Pr $\rightarrow$ Pfr photocycle and stabilizes the Pfr form. The C-terminal end of Phys contains one or more motifs that promote signal transmission and dimerization.^{1,2} The most common output motifs are histidine kinase or related domains that enable Phys to act as photoreversible phosphotransferases.

Despite intense effort, the molecular details related to how Phys photointerconvert between Pr and Pfr and how this photochromicity directs signal transmission remain largely unresolved. Important insights were made recently by us^{5,6} and others⁷ through the x-ray crystallographic analysis of PAS-GAF fragments from two BphPs in the Pr conformation. These static three-dimensional (3D) structures revealed that the BV chromophore is buried deeply within the GAF domain and arranged in a *ZZZsyn,syn,anti* configuration as Pr. A large hydrophobic cavity surrounds the D pyrrole ring,⁵ which may facilitate the proposed *Z* to *E* isomerization of the C15=C16 methylene bridge and concomitant rotation of the D ring during Pr $\rightarrow$ Pfr photoconversion.² The PAS domain does not contact the chromophore directly but instead is connected via a figure-of-eight knot to the bilin and the GAF domain.⁶ The interface between the bilin and the GAF domain includes a number of amino acids highly conserved within the Phy superfamily, many of which have been shown to be photochemically important. Examples include an aspartic acid within an invariant DIP motif that participates in a proton exchange cycle of the bilin during photoconversion, a histidine that torsionally strains the D ring of the Pr chromophore, and a number of amino acids that help the bilin dock with its binding pocket.^{7,8,9,10,11}

Full mechanistic appreciation of Phys still requires one or more 3D structures of the Pfr state to elucidate how the chromophore and protein move during phototransformation. In addition, structures of Cphs and/or plant Phys are needed to confirm expectations that they adopt similar chromophore conformations despite their use of different bilins conjugated at a different site. X-ray crystallographic approaches to the structural determination of Pfr have been particularly challenging. Two main problems are that most Pfr preparations are substantially contaminated

with Pr because of their overlapping absorption, and that Pfr is thermally unstable and will revert non-photochemically back to Pr, attributes which together likely impede crystallization.^{1,2} As an alternative, we are exploring the use of NMR spectroscopy to determine the solution structures of Pr and Pfr. One advantage is that it may be possible to determine the NMR structure of Pfr from a heterogeneous Pr/Pfr population by first solving the Pr structure and then “subtracting” the Pr chemical shifts from those obtained after saturating R. The main technical barrier to this approach has been generating photoactive chromoproteins with sizes below the practical limits of NMR spectroscopy (<30 kDa). Fortunately, we have recently overcome this hurdle with the discovery of a Cph subfamily that is missing the PAS domain and whose 21-kDa GAF domain by itself is almost fully R/FR photochromic.¹²

Here, we exploited one of these “PAS-less” Cphs from the thermotolerant *Synechococcus* OS-B' species^{12,13} to determine the solution NMR structure of a Phy GAF domain as Pr. Despite the absence of the PAS domain and the adjacent knot region, the PCB chromophore of this Cph (designated *SyB*-Cph1) assumes a *ZZZsyn,syn,anti* configuration identical to that of its BphP counterparts, implying that the GAF domain by itself dictates bilin orientation in Phys. This highly resolved model also revealed that the bound bilin is remarkably flexible, indicating that the Pr chromophore is less restrained in the binding pocket than previously appreciated. This flexibility, in turn, may be critical to the chromophore movements expected to occur during Pr→Pfr photoconversion.⁸ Finally, the *SyB*-Cph1 model identified two mobile conserved arginines that are close to the PCB propionate side chains and differentially affect the stability of the Pfr state.

Results and Discussion

Overall Topology and Dynamics of *SyB*-Cph1(GAF)

Using previously described methods,¹² we prepared ample quantities of recombinant *SyB*-Cph1 GAF domain assembled with PCB in which either the polypeptide or chromophore was labeled individually with ¹³C and/or ¹⁵N. ¹H-¹⁵N and ¹H-¹³C heteronuclear single quantum coherence (HSQC) spectra of these samples revealed well-defined chemical shifts, indicating that a NMR model could be generated for Pr and possibly later for Pfr. From the analysis of [¹³C-¹⁵N]-*SyB*-Cph1(GAF) preparations assembled with PCB in the Pr state, chemical shifts could be assigned to individual amino acids, the amide nitrogen identities of which can be found in the ¹H-¹⁵N HSQC NMR spectra shown in Suppl. Figure 1.

By using a variety of NMR spectroscopic methods, including residual dipolar couplings (RDC) in several alignment media, we solved the solution structure of *SyB*-Cph1 GAF in the Pr state. The final NMR model (Protein Data Bank (PDB) code **2K2N**) is absent the C-terminal 6His tag and the first 30 amino acids, which include the RIT consensus motif present in many members of the PAS-less Cph subfamily.¹² Both regions are highly dynamic and largely unstructured, which broadened out most of their NMR signals. From 100 refined structures, a subset of 20 low-energy conformers with the fewest restraint violations was selected to represent the solution structure (Fig. 1a). These conformers had a root mean square deviation of 0.43 Å for backbone heavy atoms over the most structured regions (residues 31–110, 136–171, and 183–202). None of the constraints derived from nuclear Overhauser effects (NOEs) in the final set of conformers was violated by more than 0.5 Å in more than 35% of the calculated models. PROCHECK¹⁴ analysis of the lowest energy model showed that 89% of the residues were within the most favored region and 11% were within the allowed regions of the Ramachandran map. A summary of the agreement between experimental constraints and calculated structures is provided in Supplemental Table 1.

The 3D model of the *SyB*-Cph1 GAF domain consists of a fanned, five-stranded anti-parallel β-sheet (β1–β2, β4–β6) and five distinct α helices, two of which form a crevice that extensively

buries the PCB chromophore (Fig. 1c). In addition, a short β strand that contributes to the PCB D-ring scaffold (β 3, residues 81–82) was detected. The overall fold of SyB-Cph1(GAF) was highly similar to those determined previously by x-ray crystallography for the BphPs from *Deinococcus radiodurans* (DrBphP, PDB accession codes **1ZTU** and **2O9B**^{5,6}) and *Rhodopseudomonas palustris* (RpBphP3, PDB accession code **2OOL**⁷). The backbone root mean square deviation between the structured regions in SyB-Cph1(GAF) (*i.e.*, residues 31–110, 136–171, and 183–202) and the corresponding regions in DrBphP was only 2.50 Å (Fig. 1b). Such similarity indicated that crystallization did not appreciably perturb the GAF domain structures of DrBphP and RpBphP3 and implied that neither the PAS domain nor the figure-of-eight knot is required to stabilize the GAF fold as we previously considered possible.⁶ The PAS-GAF fragment of DrBphP is homodimeric in solution with the dimerization interface involving the α 4 and α 8 helices.⁵ By contrast, NMR data indicated that the SyB-Cph1(GAF) fragment remained monomeric even at high protein concentrations. Whereas the region comparable to the α 8 helix in DrBphP is also helical in SyB-Cph1(GAF) (α 5 in Suppl. Figs. 2 and 3), the region (residues 22–29) corresponding to the α 4 helix in DrBphP was mostly unstructured, suggesting that a similar dimerization surface is absent in our GAF truncation of SyB-Cph1. In fact, backbone signals from only a few isolated residues within the first 30 amino acids could be detected and assigned (*i.e.*, residues 12–14, 17–18, and 26–28).

In contrast to expectations based on the crystal structures of two BphPs, the solution NMR structure of SyB-Cph1(GAF) revealed flexibility in both main chain and side chain atoms for numerous amino acids (Figs. 1a and 2c,e). Even among the 20 lowest energy structures, some side chains exhibited ample breathing of sampled conformations. For example, the side chain NMR signals of Arg-101 and Arg-133 were undetectable beyond C β , most likely due to conformational exchange. The wide range of energetically allowed conformations for the Arg-101 side chain allowed it to fill its large surrounding van der Waals cavity (Fig. 3). Especially flexible was the loop between the β 4 strand and the α 4 helix (residues 105–134). In DrBphP and RpBphP3, this loop lassoes a sequence upstream of the PAS domain to create a 3-stranded anti-parallel β -sheet at the core of the figure-of-eight knot.^{5,6,7} In the absence of the PAS domain and upstream sequence, this comparable stretch in SyB-Cph1(GAF) is solvent exposed and highly mobile (Figs. 1a and 4). NMR signals for residues 115–122 and 126–129 of this loop could not be detected owing to conformational or/and chemical exchange, while the remaining assigned residues had multiple orientations in the 100 calculated conformers, of which the 20 conformers with the lowest energy are shown in Fig. 1a. A short stretch of assigned residues (123–125) in the loop exhibited an increased spin-spin relaxation time (T_2) and a lower ¹⁵N heteronuclear NOE compared to the average along the structured regions, indicative of flexibility in the picosecond/nanosecond time scale (Fig. 4). These residues form a 3_{10} helix (designated as α 3') in several of the conformers. Residues 59–60 and 106–109 exhibited flexibility on similarly short time scales, and the latter segment was found to form a 3_{10} helix (designated as α 3) in all calculated structures.

Structure of PCB

The solution structure of SyB-Cph1(GAF) confirmed expectations^{2,12} that Cphs use a positionally conserved cysteine in the GAF domain (Cys-138 in SyB-Cph1) to bind PCB via a thioether linkage at the C3¹ position of the A-ring ethylidene side chain (Fig. 2a). Counter to predictions for Cphs based on resonance Raman (RR) spectroscopy,¹⁵ the Pr form of SyB-Cph1(GAF) cradles PCB in a *ZZZsyn,syn,anti* configuration that matches ¹⁵N-NMR predictions¹⁶ and closely resembles the conformation of BV bound to BphPs.^{5,6,7} Relative to the nearly coplanar B and C rings, the bound PCB exhibited a number of contortions of the A and D rings, including an approximately 20° rotation of ring A and an approximately 80° rotation of ring D relative to rings B/C (Fig. 2b,d). The C15=C16 methine bridge in SyB-Cph1(GAF) is substantially more torsionally-strained than the corresponding bridges in the BV

chromophores of *DrBphP* and *RpBphP*.^{5,6,7} This increased non-planarity further shortens the π conjugation system of the bilin and may account for the blue-shifted absorption maximum for *SyB-Cph1* relative to other PCB-containing Cphs.¹²

The most energetically favorable conformer of PCB had a 2(*R*),3(*E*)-PCB configuration, which placed the methyl group of its C2 carbon pointing toward Cys-138 (Fig. 2b,d), an orientation similar to the C2¹ methyl group of BV bound to *DrBphP*.⁵ This orientation was assigned on the basis of a NOE contact between the C2¹ methyl protons and one β -methylene hydrogen of Tyr-142, which is located adjacent to Cys-138, and was supported by methyl RDC data that constrained the direction of the bond between the C2¹ methyl and C2 carbon relative to the molecular frame. These constraints are inconsistent with the configuration determined previously for PCB assembled into photosynthetic phycobiliproteins such as α -C-phycocyanin.^{17,18} In these structures, PCB adopts the 2(*S*),3(*E*) configuration, which places the C2¹ methyl group pointing away from Cys-138. When the 2(*S*),3(*E*) stereoisomer was used to calculate a separate 3D model of *SyB-Cph1*(GAF), the PCB/protein structures were similar overall but less energetically favorable. Which stereoisomer of PCB is correct awaits further structural refinements.

As with regions of the *SyB-Cph1*(GAF) polypeptide, portions of PCB in the bilin-binding pocket exhibited substantial thermal motion, in particular the propionic acid side chains (Fig. 2c,e). Whereas we could easily detect signals from all six PCB methyl groups in the 2D ¹H-¹³C HSQC spectrum of the Pr state in a [¹³C]-PCB *SyB-Cph1*(GAF) sample, we failed to detect signals from the methylene carbons associated with the propionate moieties, indicating that these side chains are highly mobile (¹² and this work). Similar mobility was reported for the propionate moieties of *Syn-Cph1* based on the analysis of 2D ¹H-¹³C correlations observed in spectra of unlabeled protein assembled with uniformly-labeled [¹³C-¹⁵N]-PCB.¹⁹ This mobility observed in solution was not apparent from crystallographic models of BphPs, which strongly suggested that the propionate carboxyl groups of BV are rigidly fixed via multiple polar contacts.^{5,6,7} This newly found flexibility may indicate that the propionate side chains participate in Pr→Pfr photoconversion in addition to their role in docking the bilin in the binding pocket.²⁰

RR spectroscopy first indicated the Phy chromophore as Pr is cationic with all pyrrole ring nitrogens being protonated at neutral pH.^{9,15} Previous 1D ¹⁵N NMR analysis of [¹⁵N]-PCB *SyB-Cph1*(GAF) confirmed this protonation state by detecting four distinct ¹⁵N peaks in the spectral region for protonated nitrogens (<180 ppm).¹² However, a ¹H-¹⁵N correlation spectrum of the same sample displayed a single H^N-N crosspeak, implying that only one of the pyrrole nitrogen-associated protons is tightly bound.¹² The remaining three protons readily exchange with the solvent with one changing its environment during photoconversion. Here, we could assign the tightly bound proton to the D-ring nitrogen (131 ppm) on the basis of NOE contacts to the protein (Suppl. Fig. 4). Because its H^N-N crosspeak does not change position upon R irradiation,¹² we conclude that the D-ring nitrogen in *SyB-Cph1*(GAF) does not change its chemical environment upon Pr→Pfr photoconversion.

This finding contradicts previous proposals that the D-ring rotates around the C15=C16 methine bridge during photoconversion,^{15,21,22} because such a change would be expected to expose this nitrogen to a radically different chemical environment in the bilin binding pocket (⁶ and Fig 2a). One possible reason for this contradiction is that the photocycle of PAS-less Phys is inherently different from those of more canonical Phys and does not undergo the *Z* to *E* isomerization of the C15=C16 double bond. However, both absorption and RR spectroscopy strongly suggest that the GAF domain of *SyB-Cph1* undergoes a Pr→Pfr photocycle similar to those of other Phys.¹² A second possibility is that the *Z* to *E* isomerization occurs in *SyB-Cph1*(GAF) but that event does not significantly change the environment of the D-ring pyrrole

nitrogen, possibly due to commensurate movements of nearby atoms. A third explanation is that another pyrrole ring moves more radically than the D ring during photoconversion. This third scenario is supported by analyses of BphPs assembled with sterically locked bilins^{21,23} and by ¹⁵N NMR spectroscopy of *Syn*-Cph1,²⁴ both of which suggest that the A ring rotates substantially during Pr→Pfr photoconversion.

The Chromophore Binding Pocket

As found in BphPs, the bilin in *SyB*-Cph1 is largely buried within the GAF domain, with the opening guarded by the $\alpha 2$ and $\alpha 4$ helices (Fig. 1c). A possible hydrogen bond between the hydroxyl group of Tyr-142 and the carboxylate of Asp-86 may help clamp the opening shut (Fig. 3a,b), which for *DrBphP* appears to prevent the entry of cyclic porphyrins.⁸ Many of the important chromophore/protein contacts present in *DrBphP* and *RpBphP3*^{5,6,7} are retained in the *SyB*-Cph1(GAF) structure, with most side chains surrounding PCB maintaining rigidity (Fig. 2c,e).

Key to the photochemistry of BphPs and likely other Phy families is an aspartate residue (Asp-86 in *SyB*-Cph1) present within an invariant DIP motif. This aspartate participates in the proton exchange cycle of the Phy chromophore during Pr→Pfr photoconversion, possibly by serving as an immediate proton acceptor.^{5,8,9} The high-resolution model of *DrBphP* showed that the main chain oxygen of this aspartate contributes to an extensive hydrogen bonding network involving a centrally positioned pyrrole water, the pyrrole nitrogens of rings A–C, and the N^{δ1} of an adjacent histidine (His-139 in *SyB*-Cph1).⁵ The partial negative charges of both the main-chain oxygen of Asp-86 and the N^{δ1} of His-139 may stabilize the cationic bilin. Although our model of *SyB*-Cph1(GAF) places His-139 in a similar position near pyrrole rings A–C, Asp-86 is closer to the A ring with its main chain oxygen hydrogen bonding directly to its pyrrole nitrogen (Fig. 2b). At present, it is unknown whether this change reflects differences between BphP- and Cph-type Phys or simply results from the absence of the PAS and PHY domains in our *SyB*-Cph1(GAF) structure. Regardless of the spatial differences, Asp-86 is important to the photochemistry of *SyB*-Cph1. A D86H variant of *SyB*-Cph1 is highly fluorescent and fails to photoconvert from Pr to Pfr in R.¹²

As with the BphP models, the D ring of *SyB*-Cph1(GAF) is surrounded by a spacious pocket and is torsionally-strained relative to the B- and C rings by various contacts. Included are hydrogen bonds between the D-ring carbonyl and N^{ε2} of His-169 and ζ -amino group of Lys-52 (3.2 and 1.9 Å, respectively, in our lowest energy conformer), and possibly by van der Waals interactions with Tyr-54, Leu-77, Phe-82, and Tyr-142 (Fig. 2b,d). Only a few of the described Phys possess a lysine at the position comparable to Lys-52, including several other PAS-less Phys and *RpBphP3* (Suppl. Fig. 2 and 7,12). Instead, most Phys prefer amino acids with aliphatic side chains.^{4,7} The 3D structure of *RpBphP3* revealed that the ζ -amino group of this lysine hydrogen bonds with the D-ring carbonyl, but elimination of this contact had little effect on absorption and photoconversion.⁷ Similarly, a K52M variant of *SyB*-Cph1(GAF-PHY) readily assembled with PCB and had normal Pr and Pfr absorption spectra (Fig. 5c). However, it thermally reverted slightly faster from Pfr back to Pr, indicating that the Lys-52 contact helps stabilize the Pfr chromophore (Fig. 5c).

The propionate side chains of BV in *DrBphP* and *RpBphP3* contact a similar set of amino acids to presumably promote docking of the bilin in the binding pocket.^{5,6,7} Included in *DrBphP* are an arginine/tyrosine pair and a histidine/serine/serine cluster that interact electrostatically with the B- and C-ring propionate oxygens, respectively. Surprisingly, the solution structure of *SyB*-Cph1(GAF) predicted that few of these contacts are retained, presumably owing to the flexibility of the propionate side chains, even though most of these residues are identical (Arg-133, His-139, and Ser-151) or similar (Phe-95) in *SyB*-Cph1 (Suppl. Fig. 2). In fact, the strongest contact between the bilin and protein in BphPs, that involving a double salt bridge

between guanidinium group of Arg-254 and the B-ring propionate oxygens, is ambiguous in the solution structure of *SyB-Cph1*(GAF); the side chain of the comparable arginine (Arg-133) actually points toward the bilin in only a few of the lowest energy structures (Fig. 2d,e). Instead, the propionates are capable of contacting another positionally conserved arginine (Arg-101), which generally points toward the B and C-ring propionates in our structures of *SyB-Cph1*(GAF) but is oriented in the opposite direction in the crystal structures of *DrBphP* and *RpBphP3*.^{5,6,7} Side chain resonances of Arg-101 were broadened out by conformational exchange and could not be accurately modeled. However, Arg-101 lies within a spacious cavity that allows substantial movement of this residue, with some calculated conformers able to contact the propionate side chains (Figs. 2d,e and 3).

To examine the importance of Arg-101 and Arg-133 to *SyB-Cph1* assembly and photochemistry, we analyzed a set of variants at these positions in a GAF-PHY construction. We previously showed that an R254A variant in *DrBphP* (Arg-133 in *SyB-Cph1*) has a substantially altered Pr absorption spectrum and a dramatically increased stability as Pfr, with the variant showing little or no thermal reversion from Pfr to Pr.⁸ Here, we found that the analogous mutation in *SyB-Cph1*(GAF-PHY) did not affect PCB binding and had little effect on Pr and Pfr absorption (Fig. 5a,b). However, as with *DrBphP*, the R133A variant thermally reverted to Pr more slowly than wild-type, indicating that the Pfr state is destabilized by Arg-133 (Fig. 5c). The R101A variant also had little or no effect on PCB binding and Pr and Pfr absorption. But in contrast to R133A, the R101A chromoprotein was appreciably less stable as Pfr and displayed a three-fold increase in Pfr→Pr thermal reversion, thus indicating that Arg-101 helps stabilize Pfr (Fig. 5c). The opposite effects of the Arg-101 and Arg-133 mutations appear to be additive in that the double R101A/R133A variant had a reversion rate intermediate between the two single substitutions (Fig. 5c). Unlike R101A, the R101K substitution did not increase the reversion rate, strongly suggesting that the proposed salt bridge between the guanidinium group of Arg-101 and the B-ring propionate oxygens is critical for Pfr stability. Taken together, these results imply that the stability of the Pfr state of *SyB-Cph1* and likely other Phys is controlled by the opposing actions of Arg-101, which stabilizes Pfr, and Arg-133, which destabilizes Pfr.

Conclusions

The solution structure of the *SyB-Cph1* GAF domain presented here provides the first view of the bilin-binding pocket of a Phy that uses the GAF domain to covalently couple PCB/phytochromobilin-type chromophores. Thus, this structure likely represents an appropriate model for the chromophore-binding pocket of Cphs and plant Phys. Given the absence of the PAS domain and knot, substantial differences in the contacts between the bilin and the surrounding pocket, and the use of a cysteine in the GAF domain versus PAS domain to ligate the bilin, we were surprised to discover that the configuration of PCB (*ZZZsyn,syn,anti*) in *SyB-Cph1*(GAF) is remarkably similar to the topologies reported for BV bound to *DrBphP* and *RpBphP3*.^{5,6,7} Consequently, we conclude that the GAF domain is alone responsible for configuration of the chromophore and provides most of the protein scaffold necessary the unique photochemistry of these photoreceptors. Our Cph model also identified a number of new bilin/protein contacts that may be important for the photochemistry of Cphs/Phys as compared to BphPs. Examples of contacts in *SyB-Cph1*(GAF) not predicted from either or both BphP structures include those between PCB and Lys-52, Asp-86 and Arg-101.

A notable difference between of our solution structure of *SyB-Cph1* and the crystal structures of *DrBphP* and *RpBphP*^{5,6,7} is the apparent flexibility for both the Phy polypeptide and chromophore. These differences could reflect the thermostability of *SyB-Cph1* and/or the inherent characteristics that distinguish x-ray crystallographic models and NMR spectroscopic models. The NMR-generated solution structure likely provides a more accurate representation

of the chromophore and binding pocket. Its described flexibility may be essential for the A- and/or D-ring rotations predicted for the bilin and for subsequent conformational changes predicted for the protein during Pr→Pfr photoconversion. The need for such motions is also supported by the inability of crystallized Phys to photoconvert from Pr to Pfr, presumably because the Pr crystal lattice cannot accommodate such structural rearrangements (^{7,25} unpublished data).

Bilin flexibility may also explain why the Pr state is relatively immune to many amino acid substitutions within the chromophore pocket, with potentially only a few residues actually being essential for the unique absorption properties of Pr (^{7,8,9,11} this report). That Pfr is much more sensitive to many amino acids substitutions implies that this state is more dependent than the Pr state on chromophore/protein contacts, possibly by binding the bilin more tightly within the chromophore pocket. Our structure also identified an unexpected interplay between two conserved and mobile arginines that regulate Pfr stability through their interactions with the B- and C-ring propionate side chains. Using this solution Pr structure of SyB-Cph1(GAF) as a reference, it should now be possible to define the structural changes of the bilin and its surrounding pocket during Pr→Pfr photoconversion by similar NMR analyses of R-irradiated samples. In fact, from the assigned backbone amide nitrogens shown in Suppl. Figure 1 coupled with the chemical shift movements reported by Uliasz *et al.* ¹² for SyB-Cph1 after R irradiation, we have already identified a number of amino acids whose environment changes substantially during photoconversion.

Materials and Methods

Protein Expression and Purification

Unlabeled and uniformly labeled (U) [U -¹⁵N; U -¹³C]-SyB-Cph1(GAF) protein was expressed, assembled with either ¹⁵N or ¹³C-labeled PCB, and purified as described. ¹² To unambiguously assign the histidine side chains, the C-terminal 6His tag in SyB-Cph1 was replaced with GST, and the resulting polypeptide was expressed and uniformly labeled with ¹⁵N or ¹³C, assembled with PCB, purified on a glutathione column, and then liberated from the GST tag by thrombin cleavage (Novagen, Madison, Wisconsin).

SyB-Cph1 samples used for backbone, side chain, NOE, and backbone dynamics experiments contained 1.7 mM [U -¹⁵N; U -¹³C]-SyB-Cph1(GAF) with unlabeled PCB in 10 mM deuterated Tris-HCl (pH 8.5) and 0.03% NaN₃ in 93% H₂O/7% D₂O. Similar samples of 1 mM [U -¹⁵N; U -¹³C]-SyB-Cph1(GAF) supplemented with either 15 mg/mL filamentous pf1 phage (ASLA Biotech, Riga, Latvia) or sodium dodecylsulfate-doped ditetradecyl-phosphatidylcholine/dihexyl-phosphatidylcholine bicelles (molar ratio of 1:30:10 ²⁶) were used to record ¹D_{NH} and ¹D_{Cα Hα} RDCs. A 1 mM sample of unlabeled protein assembled with PCB ¹³C-labeled at all six methyl groups and at the CH₂ carbons adjacent to the carboxylate group of each of the two propionate side chains ¹² in 10 mM deuterated Tris-HCl (pH 8.5), 0.03% sodium azide, and 100% D₂O was used to collect 3D ¹³C NOE data and to measure isotropic methyl ¹J_{CH} couplings. A similar protein sample in anisotropic medium containing 0.5 mM protein and 15 mg/mL filamentous pf1 phage was used to measure methyl RDCs (¹D_{CH}). A 1 mM sample of unlabeled protein assembled with ¹⁵N-labeled PCB was used to record 3D ¹⁵N NOE and 1D ¹⁵N-direct detected experiments. ¹²

NMR Data Collection

All experiments were recorded at 25 °C, except for the measurements of bicelle RDCs, which were recorded at 33 °C. Backbone assignments were obtained from standard 3D CBCA(CO)NH, HNCACB, HNC(O) experiments. The side-chain assignments were derived primarily from 3D C(CO)NH, H(CCO)NH, HBHA(CO)NH, HCCH-COSY, and HCCH-TOCSY data.

Distance constraints were obtained from 3D ^{15}N -edited NOESY ($t_{\text{mix}}=150$ ms) and 3D ^{13}C -edited NOESY ($t_{\text{mix}}=120$ ms) experiments. Backbone dynamics were extracted from ^{15}N T_1 , ^{15}N T_2 , and $^{15}\text{N}\{-^1\text{H}\}$ NOE data. NMR spectra were recorded on 800 and 600 MHz Varian INOVA spectrometers with cryogenic probes. NH and $\text{C}^\alpha\text{H}^\alpha$ couplings were measured from a 3D HNCO antiphase ^1H -coupled in the ^{15}N dimension ²⁷ and a 3D HCA(CO)N antiphase ^1H -coupled in the $^{13}\text{C}^\alpha$ dimension, respectively. PCB methyl RDCs were recorded using a J-modulated [^1H - ^{13}C] HSQC ²⁸ recorded with a 600 MHz Bruker DMX-Avance spectrometer.

Resonance Assignments and Secondary Structure Calculations

^{15}N T_2 measurements with 1.7 mM [$U\text{-}^{15}\text{N}$; $U\text{-}^{13}\text{C}$]-SyB-Cph1(GAF) (Fig. 4), yielded uniform values around 45 ms for the rigid part of the molecule, suggesting that the protein is monomeric in solution under these conditions. Approximately 75% of the backbone and side chain resonances were assigned manually using the PIPP/STAPP software.²⁹ When excluding residues rendered undetectable due to conformational and/or chemical exchange on intermediate NMR timescale (*i.e.* residues P116–S122 and P126–E129), the level of assignment reached approximately 80%. The TALOS program ³⁰ was used to provide 133 pairs of ϕ/ψ backbone torsion angle restraints and to identify the secondary structural elements which were confirmed by local NOEs (Suppl. Fig. 3). The intermolecular NOEs are counted, classified, and summed in Table 1. Hydrogen bond restraints were inferred initially for α helices and later for β strands when the level of structural refinement allowed their unambiguous alignment within the β sheet. Two distance restraints of 1.9 Å and 2.9 Å per involved pair of residues were used to represent hydrogen bonds for $\text{H}^{\text{N}}\text{-O}$ and N-O , respectively.³¹

3D Structure Calculations and Refinements

Structure calculations and refinements made use of the torsion angle molecular dynamics and the internal variable dynamics modules of Xplor-NIH.³² The Dundee PRODRG2 Server ¹⁸ was used to generate the PCB topology and parameters. A separate structure calculation run (100 structures) was used to identify and generously constrain side chain dihedral angles; these were consistent with a unique rotameric state in more than 90% of the structures. Peak intensities in 3D NOESY spectra were assigned using the PIPP/STAPP package and converted into a continuous distribution of 2117 approximate interproton distance restraints, with a uniform 40% distance error applied to take into account spin diffusion. We attempted to measure RDCs in more than one alignment medium. To accommodate consistent measurements at pH 8.5, hydrolysis-resistant dialkyl analogs ²⁶ of the traditional DMPC:DHPC bicelles were exploited.³³ Neutral and positively charged bicelles (doped with CTAB) failed to yield acceptable results because these media interacted with the protein. Negatively charged bicelles doped with SDS offered a second compatible alignment medium, but yielded RDC sets highly correlated with those from filamentous pf1 phage. We used the negatively charged bicelle $^1\text{D}_{\text{NH}}$ and $^1\text{D}_{\text{CaHa}}$ RDC data, which had smaller experimental errors than the phage data, for structural restraints. We also used as structural constraints a few methyl CH RDCs from protein aligned with pf1 phage. The $^1\text{D}_{\text{NH}}$ and $^1\text{D}_{\text{CaHa}}$ RDC sets measured in pf1 were used to obtain the best SVD fitted magnitude and rhombicity of the protein alignment tensor in this medium, but not as direct structural constraints.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Mario Rivera for providing the ^{13}C -isotopically labeled α -aminolevulinic acid, and Drs. Katrina Forest and William M. Westler for helpful advice. This work was supported by a grant from the National Science Foundation (NSF) (MCB 07191530) to R.D.V. and an American Heart Association postdoctoral fellowship to A.T.U. C.C.C. was supported by National Institute of Health (NIH) grant 1U54 GM074901 (NIGMS). This study was a collaboration with the National Magnetic Resonance Facility at Madison, which is supported by National Institute of Health (NIH) grants P41 RR02301 (BRTP/NCRR) and 1P41 GM66326 (NIGMS). Additional equipment was purchased with funds from the University of Wisconsin-Madison, the NIH (RR02781, RR08438), the NSF (DMB-8415048, OIA-9977486, BIR-9214394), and the U.S. Department of Agriculture.

Abbreviations

3D	three dimensional
BphP	bacteriophytochrome
BV	biliverdin IX α
Cph	cyanobacterial phytochrome
FR	far-red light
GAF	cGMP phosphodiesterase/adenylcyclase/FhlA
HSQC	heteronuclear single quantum coherence
NOE	nuclear Overhauser effect
PAS	Per/Arndt/Sim
PCB	phycocyanobilin
Phy	phytochrome
PHY	phytochrome domain
Pr	red light-absorbing state of phytochrome
Pfr	far-red light-absorbing state of phytochrome
PDB	Protein Data Bank
R	red light
RDC	residual dipolar coupling
RR	resonance Raman
SyB	<i>Synechococcus</i> sp. OS-B'
Syn	<i>Synechocystis</i>
T_2	spin-spin relaxation

References

1. Vierstra, RD.; Karniol, B. Phytochromes in microorganisms. In: Briggs, WR.; Spudich, JL., editors. Handbook of Photosensory Receptors. Wiley-VCH Press; Weinheim, Germany: 2005. p. 171-196.
2. Rockwell NC, Su YS, Lagarias JC. Phytochrome structure and signaling mechanisms. Annu Rev Plant Biol 2006;57:837–858. [PubMed: 16669784]

3. Quail PH. Phytochrome photosensory signalling networks. *Nat Rev Mol Cell Biol* 2002;3:85–93. [PubMed: 11836510]
4. Karniol B, Wagner JR, Walker JM, Vierstra RD. Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors. *Biochem J* 2005;392:103–116. [PubMed: 16004604]
5. Wagner JR, Zhang J, Brunzelle JS, Vierstra RD, Forest KT. High resolution structure of *Deinococcus* bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J Biol Chem* 2007;282:12298–12309. [PubMed: 17322301]
6. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD. A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 2005;438:325–331. [PubMed: 16292304]
7. Yang X, Stojkovic EM, Kuk J, Moffat K. Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, *RpBPhP3*, reveals residues that modulate photoconversion. *Proc Natl Acad Sci USA* 2007;104:12571–12576. [PubMed: 17640891]
8. Wagner JR, Zhang Z, von Stetten D, Gunter M, Murgida DH, Mroginski MA, Walker JM, Forest KT, Hildebrandt P, Vierstra RD. Mutational analysis of *Deinococcus radiodurans* bacteriophytochrome reveals key amino acids necessary for the photochromicity and the proton exchange cycle of phytochromes. *J Biol Chem* 2008;283:12212–12226. [PubMed: 18192276]
9. von Stetten D, Seibeck S, Michael N, Scheerer P, Mroginski MA, Murgida DH, Krauss N, Heyn MP, Hildebrandt P, Borucki B, Lamparter T. Highly conserved residues Asp-197 and His-250 in Agp1 phytochrome control the proton affinity of the chromophore and Pfr formation. *J Biol Chem* 2007;282:2116–2123. [PubMed: 17121858]
10. Fischer AJ, Rockwell NC, Jang AY, Ernst LA, Waggoner AS, Duan Y, Lei H, Lagarias JC. Multiple roles of a conserved GAF domain tyrosine residue in cyanobacterial and plant phytochromes. *Biochemistry* 2005;44:15203–15215. [PubMed: 16285723]
11. Hahn J, Strauss HM, Landgraf FT, Gimenez HF, Lochnit G, Schmieder P, Hughes J. Probing protein-chromophore interactions in Cph1 phytochrome by mutagenesis. *FEBS J* 2006;273:1415–1429. [PubMed: 16689929]
12. Uliasz AT, Cornilescu G, von Stetten D, Kaminski S, Mroginski MA, Zhang J, Bhaya D, Hildebrandt P, Vierstra RD. Characterization of two thermostable cyanobacterial phytochromes reveals global movements in the chromophore-binding domain during photoconversion. *J Biol Chem* 2008;283:21251–21266. [PubMed: 18480055]
13. Bhaya D, Grossman AR, Steunou AS, Khuri N, Cohan FM, Hamamura N, Melendrez MC, Bateson MM, Ward DM, Heidelberg JF. Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J* 2007;1:703–713. [PubMed: 18059494]
14. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 1993;26:283–291.
15. Mroginski MA, Murgida DH, von Stetten D, Kneip C, Mark F, Hildebrandt P. Determination of the chromophore structures in the photoinduced reaction cycle of phytochrome. *J Amer Chem Soc* 2004;126:16734–16735. [PubMed: 15612706]
16. Hahn J, Kuhne R, Schmieder P. Solution-state (15)N NMR spectroscopic study of alpha-C-phycocyanin: implications for the structure of the chromophore-binding pocket of the cyanobacterial phytochrome Cph1. *ChemBiochem* 2007;8:2249–2255. [PubMed: 17973280]
17. Nield J, Rizkallah PJ, Barber J, Chayen NE. The 1.45 Å three-dimensional structure of C-phycocyanin from the thermophilic cyanobacterium *Synechococcus elongatus*. *J Struct Biol* 2003;141:149–155. [PubMed: 12615541]
18. Schuttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 2004;60:1355–1363. [PubMed: 15272157]
19. Strauss HM, Hughes J, Schmieder P. Heteronuclear solution-state NMR studies of the chromophore in cyanobacterial phytochrome Cph1. *Biochemistry* 2005;44:8244–8250. [PubMed: 15938613]
20. Hanzawa H, Inomata K, Kinoshita H, Kakiuchi T, Jayasundera KP, Sawamoto D, Ohta A, Uchida K, Wada K, Furuya M. *In vitro* assembly of phytochrome B apoprotein with synthetic analogs of the phytochrome chromophore. *Proc Natl Acad Sci USA* 2001;98:3612–3617. [PubMed: 11248126]

21. Inomata K, Noack S, Hammam MA, Khawn H, Kinoshita H, Murata Y, Michael N, Scheerer P, Krauss N, Lamparter T. Assembly of synthetic locked chromophores with *Agrobacterium* phytochromes Agp1 and Agp2. *J Biol Chem* 2006;281:28162–28173. [PubMed: 16803878]
22. Inomata K, Hammam MA, Kinoshita H, Murata Y, Khawn H, Noack S, Michael N, Lamparter T. Sterically locked synthetic bilin derivatives and phytochrome Agp1 from *Agrobacterium tumefaciens* form photoinsensitive Pr- and Pfr-like adducts. *J Biol Chem* 2005;280:24491–24497. [PubMed: 15878872]
23. Seibeck S, Borucki B, Otto H, Inomata K, Khawn H, Kinoshita H, Michael N, Lamparter T, Heyn MP. Locked 5Zs-biliverdin blocks the Meta-Ra to Meta-Rc transition in the functional cycle of bacteriophytochrome Agp1. *FEBS Lett* 2007;581:5425–5429. [PubMed: 17976380]
24. Rohmer T, Strauss H, Hughes J, de Groot H, Gartner W, Schmieder P, Matysik J. ¹⁵N MAS NMR studies of Cph1 phytochrome: Chromophore dynamics and intramolecular signal transduction. *J Phys Chem B* 2006;110:20580–20585. [PubMed: 17034247]
25. Scheerer P, Michael N, Park JH, Noack S, Forster C, Hammam MA, Inomata K, Choe HW, Lamparter T, Krauss N. Crystallization and preliminary X-ray crystallographic analysis of the N-terminal photosensory module of phytochrome Agp1, a biliverdin-binding photoreceptor from *Agrobacterium tumefaciens*. *J Struct Biol* 2006;153:97–102. [PubMed: 16377207]
26. Ottiger M, Bax A. Bicelle-based liquid crystals for NMR-measurement of dipolar couplings at acidic and basic pH values. *J Biomol NMR* 1999;13:187–191. [PubMed: 10070759]
27. Cornilescu G, Bax A. Measurement of proton, nitrogen, and carbonyl chemical shielding anisotropies in a protein dissolved in a dilute liquid crystalline phase. *J Am Chem Soc* 2000;122:10143–10154.
28. Ottiger M, Delaglio F, Marquardt JL, Tjandra N, Bax A. Measurement of dipolar couplings for methylene and methyl sites in weakly oriented macromolecules and their use in structure determination. *J Magn Reson* 1998;134:365–369. [PubMed: 9761712]
29. Garrett DS, Powers R, Gronenborn AM, Clore GM. A common sense approach to peak picking two-, three- and four-dimensional spectra using automatic computer analysis of contour diagrams. *J Magn Reson* 1991:95.
30. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 1999;13:289–302. [PubMed: 10212987]
31. Wüthrich, K. *NMR of Proteins and Nucleic Acids*. Wiley Interscience; New York: 1986.
32. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 2003;160:65–73. [PubMed: 12565051]
33. Tjandra N, Bax A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* 1997;278:1111–1114. [PubMed: 9353189]
34. Hansen MR, Mueller L, Pardi A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat Struct Biol* 1998;5:1065–1074. [PubMed: 9846877]

Fig. 1.

3D solution structure of the Pr state of the GAF domain of SyB-Cph1 assembled with PCB. The highly dynamic first 30 amino acids and C-terminal 6His tag of the GAF domain of SyB-Cph1 are not modeled. PCB is shown in cyan, β strands are shown in blue, α helices in red, and unstructured regions in yellow. (a) Superposition of the protein backbone from the 20 lowest energy NMR conformers. (b) Superposition of the lowest energy NMR structure for SyB-Cph1(GAF) with the high resolution crystal structure of the GAF domain from DrBphP (PDB ID code **2O9B**) assembled with BV (shown in grey). (c) Stereoview of the protein backbone of SyB-Cph1 in the lowest energy NMR conformer (PDB ID code **2K2N**). Specific α helices and β strands are labeled.

Fig. 2.

Views of the PCB-binding pocket of *SyB-Cph1*(GAF). (a) Overall view of the PCB-binding pocket. (b and d) Side and top views of the binding pocket for the lowest energy NMR conformer. (c and e) Side and top views for 5 of the lowest energy NMR conformers illustrating the mobility of the chromophore and adjacent residues, especially the propionate acid side chains of PCB and the adjacent Arg-101 and Arg-133 side chains. PCB is shown in cyan with the protein backbone shown in green. Key amino acid side chains that contact the chromophore are in red, yellow and orange. Also shown is the thioether linkage between the C3¹ carbon of PCB and the sulfur of Cys-138. The A–D pyrrole rings are indicated. The dashed lines indicate potential polar contacts between PCB and nearby amino acids.

Fig. 3. Mobility of Arg-101 within the pocket adjacent to the B-ring propionic acid side chain (PA-ring B) of PCB. **(a)** Overall view of the SyB-Cph1 GAF domain highlighting Arg-101 (magenta), Arg-133 (yellow), and PA-ring B (cyan). **(b)** and **(c)** Front and side close up views of the 10 lowest energy NMR conformers for Arg-101 in its pocket. The positions of the Arg-101 side chain in the 10 lowest-energy conformers are shown by van der Waals surfaces in panel a and by stick representations in panels b and c.

Fig. 4.

Mobility predictions for amino acid residues 31–202 of *SyB-Cph1*(GAF) assembled with PCB in the Pr state. Plots show the experimental ^{15}N backbone amide relaxation parameters versus the residue number for *SyB-Cph1*(GAF) at 800 Mhz. Data include longitudinal relaxation time (T_1), transverse relaxation time (T_2), T_1/T_2 ratios, and the steady-state ^{15}N - $\{^1\text{H}\}$ NOE. The positions of specific α helices and β strands are indicated above the panels.

Fig. 5. Photochemical analysis of *SyB-Cph1* variants at positions Lys-52, Arg-101 and Arg-133. The recombinant apoproteins encompassing the GAF-PHY region were assembled with PCB. **(a)** Covalent binding of PCB. Samples were subjected to SDS-PAGE and either assayed for the bound bilin by zinc-induced fluorescence (Zn) or stained for protein (Prot). **(b)** UV-vis absorption spectra as Pr (black lines) or following saturating R (dashed lines) and FR-minus-R difference spectra. Difference maxima and minima are indicated. **(c)** Thermal reversion of Pfr back to Pr. Samples were irradiated with saturating R and then kept in the dark at 58 °C. Absorption spectra were recorded at various intervals and used to calculate the % Pfr remaining. WT, wild type *SyB-Cph1*(GAF-PHY).