PNAS
www.pnas.org

Supplementary Information for

A far-red cyanobacteriochrome lineage specific for verdins

Marcus V. Moreno,[a] Nathan C. Rockwell,[a] Manuel Mora,[a] Andrew J. Fisher,[ab] J. Clark Lagarias[a1]
[a]Department of Molecular and Cellular Biology, University of California, Davis, CA 95616, USA
[b]Department of Chemistry, University of California, Davis, CA 95616, USA

[1]J. Clark Lagarias
Email: jclagarias@ucdavis.edu

This PDF file includes:

Supplementary text
Figures S1 to S6
Tables S1 to S2
SI References

Other supplementary materials for this manuscript include the following:

Dataset S1

www.pnas.org/cgi/doi/10.1073/pnas.2016047117

Supplementary Information Text

Supplementary Methods

Bioinformatics. To identify additional far-red/red CBCRs, BLAST (1) searches were performed using JSC1_58120g3 as a query sequence. The resulting CBCRs were aligned with known red/green family CBCRs and were then manually adjusted based on structural analysis. Structures of AnPixJg2 (PDB accession 3W2Z: (2)), slr1393 (5DFX and 5M82: (3)), and NpR6012g4 (6BHN and 6BHO: (4)) were used with DSSP to add secondary structure and solvent accessibility information to the sequence alignment using an in-house Python script that also removed each position having >5% gaps, yielding a final alignment having 146 sequences and 157 characters. The phylogenetic tree was calculated in PhyML-Structure (5) using that input file with an EX-EHO substitution model in partition mode with four substitution rate categories and calculated estimates for the gamma shape distribution parameter and the proportion of invariable sites (command-line settings: -m EX_EHO -M PART -a e -c 4 -v e). Statistical support was evaluated using the default SH-aLRT method implemented in PhyML (6).

Protein Expression and Spectral analysis. Tailed primers with NcoI (5'-ACTTGTCCATGGAGCAGGCGTTAAACCG-3') and SmaI (5'-GTCATGCCCGGGTTGCGCCAAATACTCGGC-3') restriction sites were used to amplify JSC1_58120g3 from Leptolyngbya JSC-1 genomic DNA (generously gifted by Prof. Don Bryant, Penn State University). The genes Mic7113_1903g4, and AFZ15460g4 were purchased as synthetic genes from GenScript (Piscataway, New Jersey, USA) with flanking NcoI and SmaI restriction sites. The synthetic genes and PCR products were digested with NcoI and SmaI and cloned into pBAD-Cph1-CBD (7). P591T variant JSC1_58120g3 was produced via site-directed mutagenesis of pBAD-JSC1_58120g3-CBD using PfuUltra high-fidelity DNA polymerase (Agilent) and appropriate oligonucleotide primers. C-terminal intein-CBD fusion proteins were co-expressed in E. coli strain LMG194 with pPL-PCB (8) or pPL-PΦB (9), which respectively encode enzymes sufficient for biosynthesis of PCB or PΦB. CBCRs were purified on chitin resin (NEB) following previously established protocols, with final dialysis in TKKG buffer (25 mM TES-KOH pH 8.0, 25 mM KCl, 10% (v/v) glycerol) (7). Following purification with chitin resin, JSC1_58120g3 samples for crystallization were dialyzed in Low Salt buffer (25 mM TES-Cl pH 8.0, 25 mM KCl) and purified on a Q-Sepharose Fast Flow ion exchange resin, with elution via gradient with High Salt buffer (25 mM TES-Cl pH 8.0, 1 M KCl). Samples for crystallization were then dialyzed in pure MilliQ water and concentrated to approximately 15 mg/mL, as measured by modified Bradford Assay (Bio-Rad) with bovine serum albumin as standard (Pierce). Absorbance spectra were acquired as previously described (7, 10). For denaturation assays, 100 µl protein was diluted with 1 ml 7 M guanidinium chloride/1% HCl (v/v) (10). Photochemical difference spectra are reported as (*15Z-15E*).

Crystallization and Structure Determination. JSC1_58120g3 expressed with PcyA was crystallized in the dark at 25 °C via sitting drop vapor diffusion in 0.25 M Ammonium Citrate, 25% (w/v) PEG 3350. JSC1_58120g3 expressed with HY2 was crystallized in the dark at 25 °C via sitting drop vapor diffusion in 0.1 M Ammonium Chloride, 35% (w/v) PEG 6000, 0.1 M Sodium Acetate. Drops were mixed with 1 µL 15 mg/mL protein and 1 µL reservoir solution. Crystals were soaked in reservoir solution containing 30% (v/v) ethylene glycol under green safety light and flash frozen in liquid nitrogen. Diffraction data were collected at 100 K on beamline 9-2 at the Stanford Synchrotron Radiation Lightsource and beamline 24-ID-E at the Advanced Photon Source for JSC1_58120g3 co-expressed with PcyA and HY2, respectively. The structure of JSC1_58120g3 co-expressed with PcyA was solved by PhaserMR (11) using a poly-Ala model of AnPixJg2. An initial model was generated using Phenix Autobuild (12, 13), and subsequently refined against the 1.5 Å data set using Refmac5 (14) and phenix.refine (15). The structure of JSC1_58120g3 expressed with HY2 was solved by PhaserMR using the final PcyA-expressed model and subsequently refined against the 2.3 Å data using Refmac5 and phenix.refine. Structural figures were generated using PyMol (16). Conserved residues were identified using Homolmapper (17), with conserved DPYLoar residues defined as being found in at least 75% of DPYLoar CBCRs and fewer than 33% of all CBCRs represented in the present phylogeny. Tyr592 is included as a

DPYLoar residue (found in 100% of DPYLoar CBCRs) despite being common outside of the lineage (56.9% overall) due to its apparent structural role in maintaining the unusual chromophore positioning in JSC1_58120g3. Ala653 (41.2% of DPYLoar CBCRs) also does not meet the stated cutoff but is included because DPYLoar CBCRs also commonly possess a Ser residue at this position (53.9%) that replaces a conserved Pro found in typical XRG CBCRs. Ring tilts were calculated between adjacent rings by calculating the angle between the two planes defined by the three ring heavy atoms in each ring closest to the adjacent ring using an in-house Python script. For example, the ring tilt between rings B and C was defined as the angle between the plane defined by N_B, C9, and C8 (B-ring) and that defined by N_C, C11, and C12 (C-ring).

Supplementary Figures S1-S6

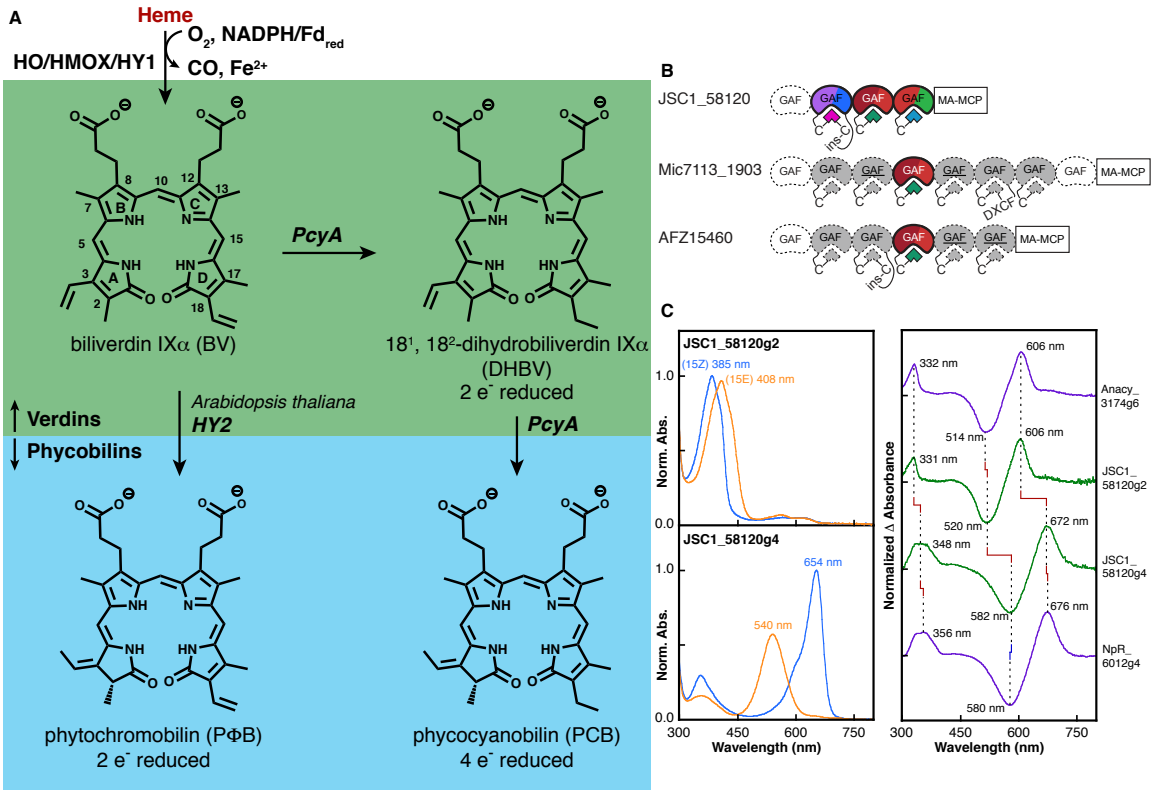

Fig. S1. Bilin chromophore biosynthesis. (A) Heme is converted to biliverdin IXα (BV, top left) by a heme oxygenase. BV can be reduced to phytochromobilin (2-electron reduction, bottom left) or to phycocyanobilin (4-electron reduction, bottom right) by the ferredoxin-dependent bilin reductases HY2 or PcyA. Biliverdin reduction by PcyA is a two-step reaction that proceeds through the 2-electron reduced intermediate $18^1,18^2$ dihydrobiliverdin IXα (DHBV, top right). The chromophores are classified into verdins or phycobilins based on redox state of the bilin A-ring. (B) Domain cartoons are shown for JSC1_58120, Mic7113_1903, and AFZ15460. CBCRs are colored by their photocycles, with far-red absorbance represented by brick red and UV absorbance represented by purple. Chromophores are represented as polygons, with DHBV in green, PCB in light blue, and PVB in fuschia. Dashed, white GAF domains do not contain the first Cys required for chromophore attachment. Dashed, grey GAF domains contain the first Cys, but have not yet been described. Underlined GAF domains are DPYLoar lineage members that have not yet been characterized. (C) (left) Dark-adapted *15Z* (blue) and photoproduct *15E* (orange) absorbance spectra are shown for native JSC1_58120g2 and JSC1_58120g4 expressed with PCB biosynthetic enzymes. (right) *15Z-15E* photochemical difference spectra obtained after denaturation are shown for JSC1_58120g2 and JSC1_58120g4 (green) and for covalent adducts of PVB and PCB standards (violet; Anacy_3174g6 and NpR6012g4, respectively).

Fig. S2. Phylogeny and conserved residue alignment. (A) Alignment of the chromophore-binding region for selected DPYLoar frCBCRs identified in this study to AmBV4 and canonical XRG CBCRs. Conserved DPYLoar residues are highlighted in salmon, conserved AmBV4 residues in violet, and generally conserved residues in grey. (B) Full XRG phylogenetic tree with core DPYLoar and potential transitional groups labeled. CBCR labels are colored based on lineage and/or photocycle designation, when known. SH-aLRT confidence levels are shown as indicated.

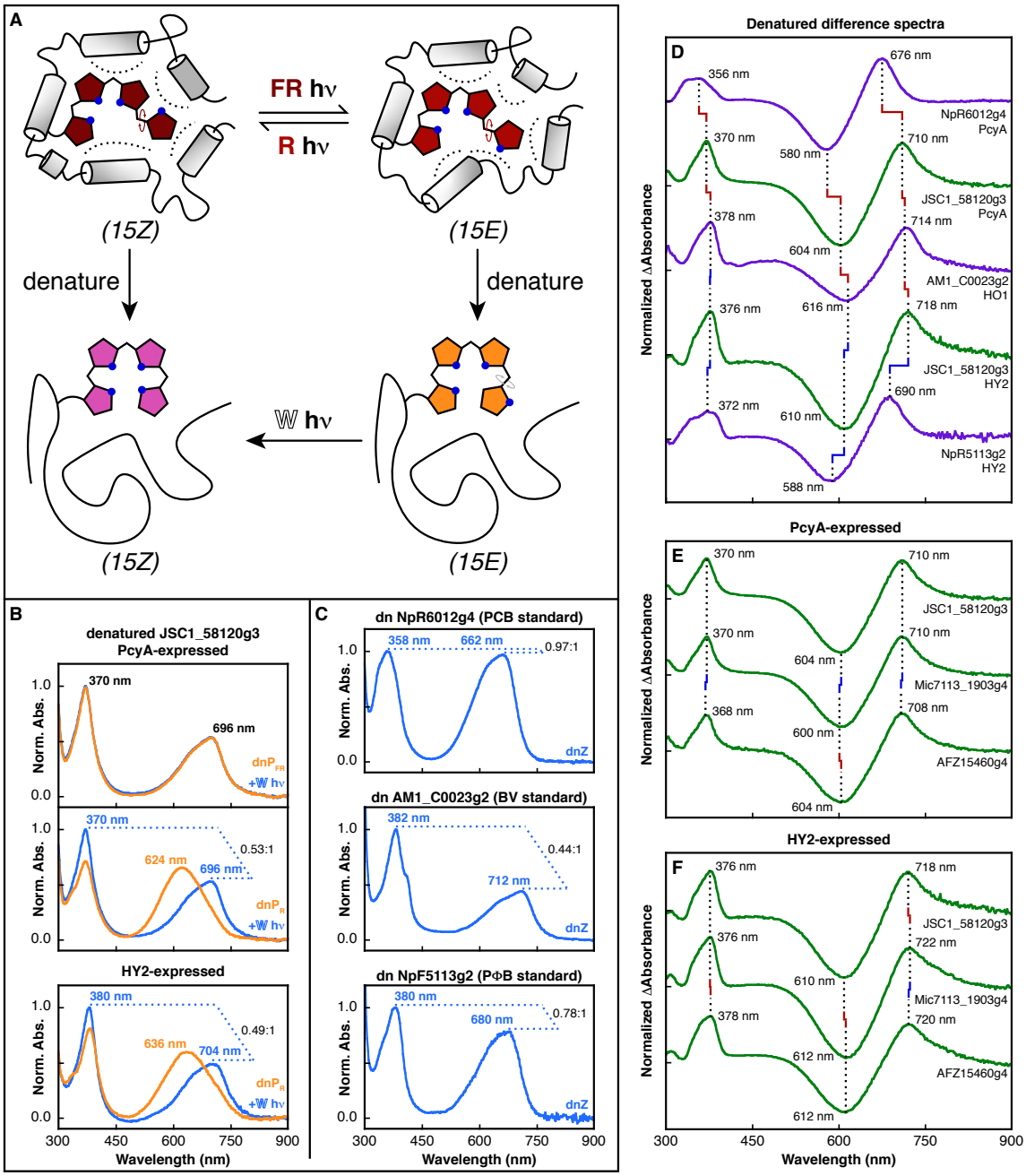

Fig. S3. Characterization of DPYLoar CBCRs under denaturing conditions. (A) Cartoon diagram illustrating CBCR denaturation experiments. Native CBCRs covalently attach to bilin chromophore, tune absorbance through secondary interactions, and reversibly photoconvert between two light-absorbing states via *15Z* to *15E* photoisomerization. Denatured CBCRs lose secondary tuning interactions, resulting in characteristic spectra for a given bilin adduct. *15Z* denatured CBCR are non-photoreactive, while denatured 15E CBCRs photoconvert to *15Z* with exposure to white light. (B) Denatured absorbance spectra for far-red-absorbing *15Z* and red-absorbing *15E* JSC1_58120g3 expressed with PcyA (top, middle) and *15E* JSC1_58120g3 expressed with HY2 (bottom) are shown before (orange) and after white light illumination (blue). Non-reactive denatured dark-adapted state and photoconversion of denatured photoproduct state allows assignment of configuration at C15. S_1:S_2 peak absorbance ratios for *15Z* are indicated. (C) Denatured *15Z* absorbance spectra for covalent adducts of PCB, BV, and PΦB standards (NpR6012g4, AM1_C0023g2, and NpR5113g2, respectively) are shown, with S_1:S_2 peak

absorbance ratios labeled as in B. (D) 15Z-15E photochemical difference spectra obtained after denaturation are shown for JSC1_58120g3 (green) and for PCB, PFB, and BV standards (violet; samples as in panel C). (E-F), DPYLoar CBCRs JSC1_58120g3, Mic7113_1903g4, and AFZ15460g4 expressed with PcyA (E) and with HY2 (F). Peak wavelengths are labeled, and dashed lines compare adjacent spectra.

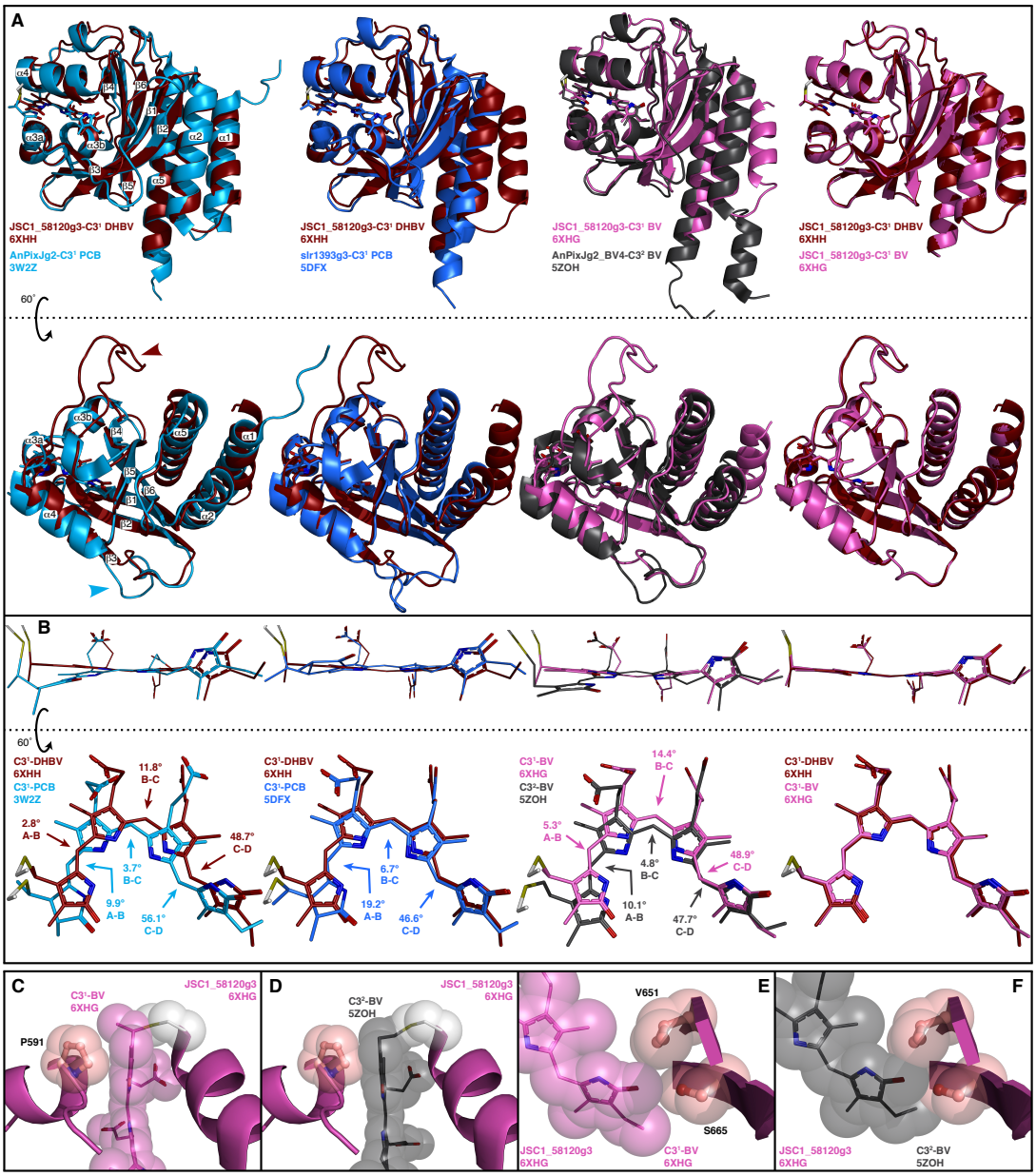

Fig. S4. Alignments of JSC1_58120g3 with XRG and BV4 structures. (A) 60˚ X-rotated backbone alignments of JSC1_58120g3-DHBV to AnPixJg2 (left) or slr1393g3 (center-left), JSC1_58120g3-BV to AnPixJg2_BV4 (center-right), and JSC1_58120g3-DHBV to JSC1_58120g3-BV (right). Protein mainchain is depicted in ribbon view. JSC1_58120g3-DHBV is colored brick red, AnPixJg2 teal, slr1393g3 blue, JSC1_58120g3-BV magenta, and AnPixJg2_BV4 charcoal. Secondary structure is labeled in the JSC1-58120g3-DHBV to AnPixJg2 alignment, and brick red and teal arrowheads indicate extra loops present or absent in JSC1_58120g3, respectively. (B) 60˚ X-rotated bilin-adducts aligned and colored as in panel A. XRG-conserved residues are colored by atom with carbon in light grey. A-B, B-C, and C-D bilin ring tilts are labeled for each chromophore. (C-D) A-ring view of JSC1_58120g3-BV structure with C3^1-attached BV chromophore (C) or superimposed C3^2-attached BV chromophore from AnPixJg2_BV4 structure (D). DPYLoar-conserved Pro591 is colored by atom with carbons in salmon, and Pro591 and chromophore-adducts are depicted in space-filling spheres to show clash between Pro591 and modeled C3^2-BV A-ring. (E-F) D-ring view of JSC1_58120g3-BV structure with C3^1- and modeled C3^2-attached BV as in (C-D). Space-filling representation of chromophore-adduct and DPYLoar-conserved Val651 and Ser665 residues highlight clashes with C3^2-BV D-ring.

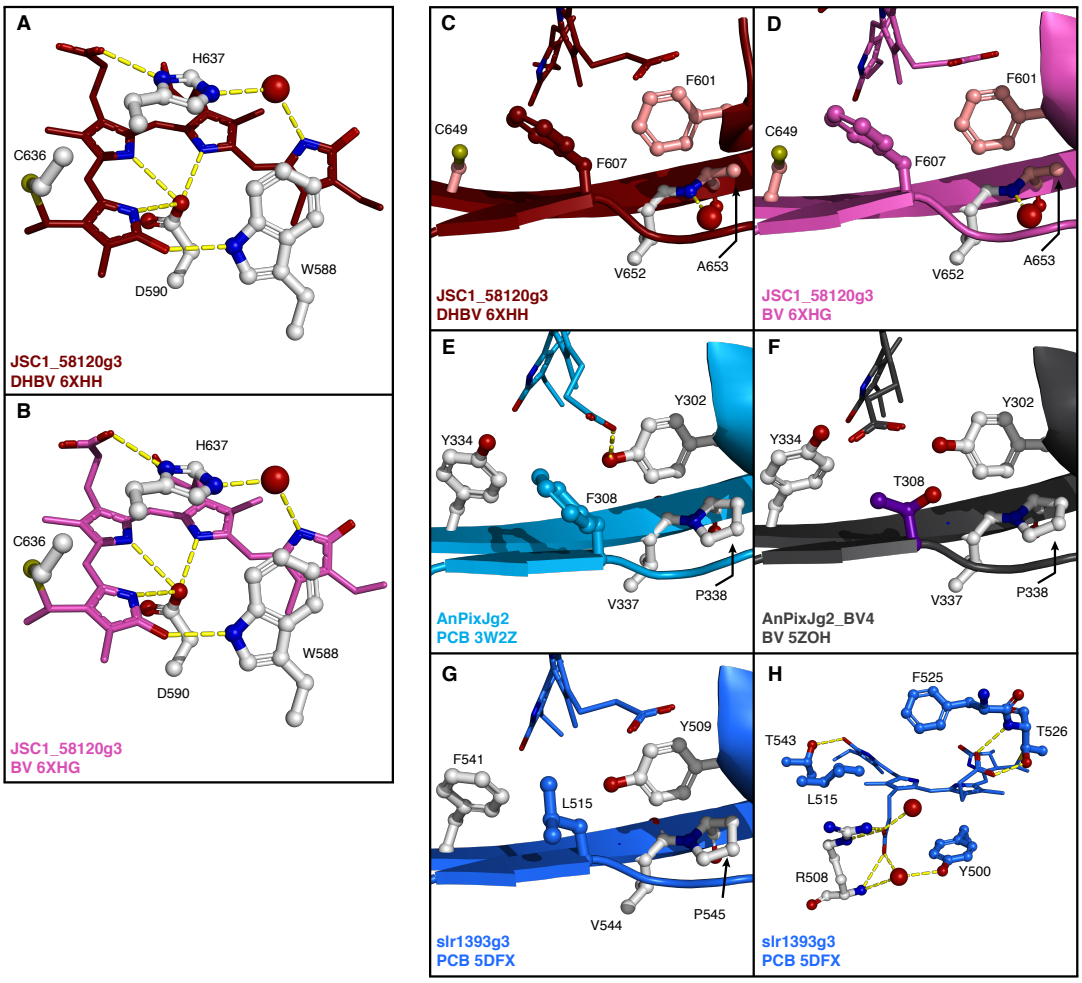

Fig. S5. Structure of conserved residues. (A-B) Structural arrangement of XRG-conserved residues and contacts with chromophore in JSC1_58120g3-DHBV (A) and JSC1_58120g3-BV (B). (C-G) Repositioning of β4 residues with changes to adjacent β5 and α3 residues in JSC1_58120g3-DHBV (C), JSC1_58120g3-BV (D), AnPixJg2 (E), AnPixJg2_BV4 (F), and slr1393g3 (G). (H) Chromophore arrangement and protein contacts in slr1393g3. Coloring as in Fig. S4, with AmBV4-conserved residues colored by atom with carbon in violet. Water molecules are depicted as red spheres.

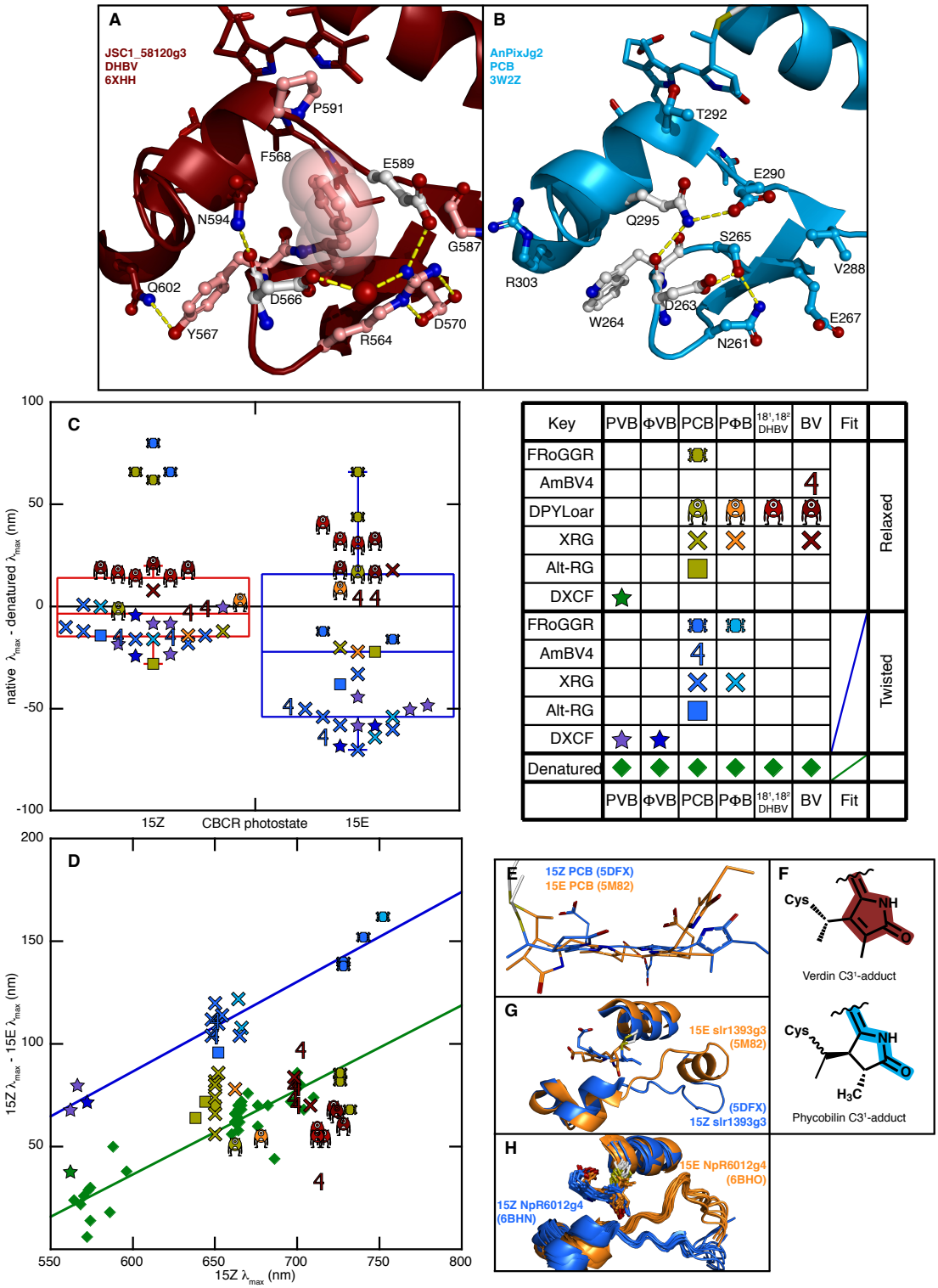

Fig. S6. DPYLoar chromophore tuning. (A) Conserved DPYLoar residues create a network of interactions linking strands β1-β2 with the α3 region and envelop a DPYLoar-conserved phenylalanine in a strained conformation placing the sidechain tightly apposed to the phycobilin-excluding residue Pro591. (B) Equivalent residues in AnPixJg2 form fewer contacts. Coloring as noted in S4 and S5. (C) Degree of chromophore tuning in native CBCRs relative to denatured absorbance is shown in a box plot for *15Z* dark state and *15E* photoproduct. Individual data points

reflect CBCR class and chromophore identity as shown in the key. (D) Photoconversion-induced blue shift is plotted against peak *15Z* absorbance for CBCRs exhibiting twisted, blue-shifted photoproducts, relaxed photoproducts, and denatured CBCRs as shown in the key. Twisted and denatured data were fit by linear regression (r^2 = 0.93 and 0.84, respectively). Relaxed CBCRs follow the denatured trend and are distinct from twisted CBCRs. (E) Photoconversion-induced chromophore twisting as seen in slr1393g3 dark and photoproduct structures. Dark-adapted *15Z* photostate is colored by atom with carbon in blue, photoproduct *15E* state carbon in orange. (F) Chemical representation of verdin and phycobilin $C3^1$-adduct A-rings, with conjugation highlighted in brick red and teal respectively. (G-H) Photoconversion-induced structural rearrangement of loop-helix α3 and chromophore A-ring twist in slr1393g3 (G) and NpR6012g4 NMR ensemble (H). Protein mainchain depicted in ribbon view and colored in blue for dark-adapted *15Z* photostate or orange for photoproduct *15E* photostate.

Supplementary Tables S1 & S2

Table S1. Spectral Properties of Selected Far-Red-Absorbing and XRG CBCRs[a]

	CBCR	Bilin	Reference	Native Z λ_{max}	Native E λ_{max}	Native SAR	Denatured Z λ_{max}	Denatured E λ_{max}	Denatured SAR
DPYLoar	JSC1_58120g3	18¹,18²-DHBV	This Study	376, 712	374, 654	0.91	370, 696	370 624	0.22
		BV		386, 722	386, 654	0.94	380, 704	382, 636	0.29
	P591T_	PCB		358, 662	360, 612	1.36	362, 664	364, 596	0.45
	JSC1_58120g3	PΦB+BV		380, 678	378, 624	1.03	380, 676	382, 616	0.34
	Mic7113_1903g4	18¹,18²-DHBV		376, 716	374, 662	1.70	370, 702	370, 622	0.44
		BV		388, 728	388, 668	1.57	380, 710	382, 636	0.36
	AFZ15460g4	18¹,18²-DHBV		374, 712	372, 658	1.18	370, 698	370, 626	0.45
		BV		388, 724	388, 658	1.14	380, 704	382, 642	0.29
XRG	NpR5113g2	PCB	(18)	354, 650	354, 528		672	576	
	NpR6012g4	PCB	(18)	356, 650	352, 542		358, 662	364, 600	
		PΦB	(18)	368, 666	552		682	592	
		BV		-	-	-	-	-	-
	AnPixJg2	PCB	(19)	648	543		661	598	
	slr1393g3	PCB	(20)	648	536		662	596	
	AM1_1870g3	PCB	(21)	353, 665	561		664	594	
		BV	(21)	375, 706	641		700	620	
BV4	AM1_1557	PCB	(22)	351, 649	545		664	594	
		BV	(22)	378, 697	622		700	620	
	AM1_C0023g2	PCB	(23)	352, 650	539		665	603	
		BV	(23)	377, 699	618		700	660	
	AM1_6305g2	BV	(24)	702	605				
	AnPixJg2_BV4	BV	(24)	699	620				
	AnPixJg4_BV4	BV	(24)	699	629				
	AM1_1870g3_BV4	BV	(24)	713	680[b]				
	NpF2164g5_BV4	BV	(24)	681	-				
FRoGGR	Anacy_4718g3	PCB	(10)	740	590		660	604	
	Anacy_4718g3	PFB	(10)	752	590				
	Anacy_4718g3	BV	(10)	-	-	-	-	-	-
	Sta7437_1656	PCB	(10)	728	642		664	600	

[a]All wavelengths are reported in nanometers. Where two wavelengths are reported, the short wavelength S_0-S_2 (Soret) transition is listed first, followed by the long wavelength S_0-S_1 transition. Single wavelengths are S_0-S_1 transitions. Specific absorbance ratio (SAR) is calculated as the ratio of the peak absorbance of the S_0-S_1transition of the native *15Z* photostate to the aromatic amino acid absorbance at 280 nm and serves as an approximate relative indicator of chromophore binding efficiency.

[b]AM1_1870g3_BV4 photoproduct exhibits substantial overlap with *15Z* dark state.

Table S2. Data collection and refinement statistics of 58120g3$_{PcyA}$ and 58120g3$_{Hy2}$

Data Collection	**58120g3$_{PcyA}$**	**58120g3$_{Hy2}$**
X-ray Source	SSRL 9-2	APS 24-ID-E
Wavelength (Å)	0.9795	0.9792
Temperature (K)	100	100
Space Group	P1	P1
Unit Cell Parameters		
a,b,c (Å)	32.3, 35.8, 74.9	32.2, 36.1, 75.5
α,β,γ (˚)	100.2, 95.4, 89.9	100.5, 95.5, 90.0
Resolution (Å)	73.37 - 1.5 (1.539 - 1.50)	73.86 - 2.3 (2.360 - 2.3)
R_{merge} (%)[a]	5.2 (57.0)	8.3 (42.7)
$CC_{1/2}$	0.999 (0.797)	99.1 (75.2)
I/σ(I)	13.8 (2.2)	6.38 (1.91)
Total reflections	169,907 (4,785)	24,961 (1,881)
Unique reflections	44,374 (1,265)	13,152 (968)
Completeness (%)	84.3 (48.9)	89.0 (91.5)
Redundancy	3.8 (3.8)	1.9 (1.9)
Refinement		
Resolution (Å)	30.52 - 1.5	36.93 - 2.3
No. of reflections (F>0) used in refinement	44,337	13,129
R_{factor} (%)[b]	19.44	23.10
R_{free} (%)[b]	23.13	26.44
RMS bond length (Å)	0.006	0.003
RMS bond angle (˚)	1.004	0.784
Mean B factor (Å^2)	20.71	44.27

[a] Rmerge = $[\sum_h\sum_i|I_h - I_{hi}|/\sum_h\sum_i I_{hi}]$ where I_h is the mean of I_{hi} observations of reflection *h*. Numbers in parenthesis represent highest resolution shell.

[b] R_{factor} and c $R_{free} = \sum||F_{obs}| - |F_{calc}|| / \sum|F_{obs}| \times 100$ for 95% of recorded data (R_{factor}) or 5% data (R_{free})

****

Dataset S1 (separate file). Alignment file used to generate phylogenetic tree. Individual CBCRs are labeled with protein name or GenBank accession number followed by "gX", where X is a positive integer and indicates which sequential (N- to C-terminal) CBCR GAF domain is described.

****

SI References

1. S.F. Altschul*, et al.*, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389-3402 (1997).
2. R. Narikawa*, et al.*, Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 110(3):918-923 (2013).
3. X. Xu*, et al.*, Structural elements regulating the photochromicity in a cyanobacteriochrome. *Proc. Natl. Acad. Sci. U.S.A.* 117(5):2432-2440 (2020).
4. S. Lim*, et al.*, Correlating structural and photochemical heterogeneity in cyanobacteriochrome NpR6012g4. *Proc. Natl. Acad. Sci. U.S.A.* 115(17):4387-4392 (2018).
5. S.Q. Le & O. Gascuel, Accounting for Solvent Accessibility and Secondary Structure in Protein Phylogenetics Is Clearly Beneficial. *Syst. Biol.* 59(3):277-287 (2010).
6. S. Guindon*, et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59(3):307-321 (2010).

7. N.C. Rockwell, L. Shang, S.S. Martin, & J.C. Lagarias, Distinct classes of red/far-red photochemistry within the phytochrome superfamily. *Proc. Natl. Acad. Sci. U.S.A.* 106(15):6123-6127 (2009).
8. G.A. Gambetta GA & J.C. Lagarias, Genetic engineering of phytochrome biosynthesis in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 98(19):10566-10571 (2001).
9. A.J. Fischer*, et al.*, Multiple roles of a conserved GAF domain tyrosine residue in cyanobacterial and plant phytochromes. *Biochemistry* 44(46):15203-15215 (2005).
10. N.C. Rockwell, S.S. Martin, & J.C. Lagarias, Identification of Cyanobacteriochromes Detecting Far-Red Light. *Biochemistry* 55(28):3907-3919 (2016).
11. A.J. McCoy*, et al.*, Phaser crystallographic software. *J. Appl. Crystallogr.* 40(4):658-674 (2007).
12. T.C. Terwilliger*, et al.*, Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* 64(1):61-69 (2008).
13. P.D. Adams*, et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66(2):213-221 (2010).
14. G.N. Murshudov, A.A. Vagin, & E.J. Dodson Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D Biol. Crystallogr.* 53(3):240-255 (1997).
15. P.V. Afonine*, et al.*, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68(4):352-367 (2012).
16. Schrödinger LLC (The PyMOL Molecular Graphics System), Version 2.2.0.
17. N.C. Rockwell & J.C. Lagarias, Flexible mapping of homology onto structure with homolmapper. *BMC Bioinform.* 8:123 (2007).
18. N.C. Rockwell, S.S. Martin, & J.C. Lagarias, Red/Green Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry* 51(48):9667-9677 (2012).
19. R. Narikawa, Y. Fukushima, T. Ishizuka, S. Itoh, & M. Ikeuchi, A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* 380(5):844-855 (2008).
20. J. Zhang*, et al.*, Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chem. Int. Ed.* 49(32):5456-5458 (2010).
21. R. Narikawa, K. Fushimi, W. Ni Ni, & M. Ikeuchi, Red-shifted red/green-type cyanobacteriochrome AM1_1870g3 from the chlorophyll d-bearing cyanobacterium Acaryochloris marina. *Biochem. Biophys. Res. Commun.* 461(2):390-395 (2015).
22. R. Narikawa*, et al.*, A biliverdin-binding cyanobacteriochrome from the chlorophyll d-bearing cyanobacterium Acaryochloris marina. *Sci. Rep.* 5:7950 (2015).
23. K. Fushimi*, et al.*, Photoconversion and Fluorescence Properties of a Red/Green-Type Cyanobacteriochrome AM1_C0023g2 That Binds Not Only Phycocyanobilin But Also Biliverdin. *Front. Microbiol.* 7:588 (2016).
24. K. Fushimi*, et al.*, Rational conversion of chromophore selectivity of cyanobacteriochromes to accept mammalian intrinsic biliverdin. *Proc. Natl. Acad. Sci. U.S.A.* 116(17):8301-8309 (2019).