

Rational Construction of Compact *de Novo*-Designed Biliverdin-Binding Proteins

Molly M. Sheehan,[†] Michael S. Magaraci,[†] Ivan A. Kuznetsov,[†] Joshua A. Mancini,[‡] Goutham Kodali,[‡] Christopher C. Moser,[‡] P. Leslie Dutton,[‡] and Brian Y. Chow^{*,†}

[†]Department of Bioengineering and [‡]Department of Biochemistry and Biophysics, Johnson Research Foundation, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

S Supporting Information

ABSTRACT: We report the rational construction of *de novo*-designed biliverdin-binding proteins by first principles of protein design, informed by energy minimization modeling in Rosetta. The self-assembling tetrahelical bundles bind biliverdin IXa (BV) cofactor autocatalytically *in vitro*, like photosensory proteins that bind BV (and related bilins or linear tetrapyrroles) despite lacking sequence and structural homology to the natural counterparts. Upon identification of a suitable site for ligation of the cofactor to the protein scaffold, stepwise placement of residues stabilized BV within the hydrophobic core. Rosetta modeling was used in the absence of a high-resolution structure to inform the structure-function relationships of the cofactor binding pocket. Holoprotein formation stabilized BV, resulting in increased far-red BV fluorescence. Via removal of segments extraneous to cofactor stabilization or bundle stability, the initial 15 kDa *de novo*-designed fluorescence-activating protein was truncated without any change to its optical properties, down to a miniature 10 kDa “mini”, in which the protein scaffold extends only a half-heptad repeat beyond the hypothetical position of the bilin D-ring. This work demonstrates how highly compact holoprotein fluorochromes can be rationally constructed using *de novo* protein design technology and natural cofactors.

De novo-designed proteins are useful tools for exploring principles of protein folding, assembly, and biochemical functions that build on structure–function and sequence diversity landscapes distinct from those of natural protein scaffolds.^{1–3} Self-assembling tetrahelical bundles,^{4–10} created by binary patterning of hydrophobic and hydrophilic residues with high α -helical propensity,¹¹ comprise the best-established class of *de novo*-designed scaffolds. They provide stable frames for binding cofactors, as protein maquettes^{5–7,12} for rationally engineering artificial holoproteins in which the cofactor-interacting contributions of individual residues are largely isolated from one another (Figure 1a).

Previously reported maquette holoproteins incorporated rigid, planar cofactors such as hemes, chlorins, porphyrins, and flavins. Recently, we reported that they also bind flexible bilins or linear tetrapyrroles and identified determinants for autocatalytic ligation of phycocyanobilin (PCB), namely, a free cysteine and the stabilization of the bilin propionates.⁶

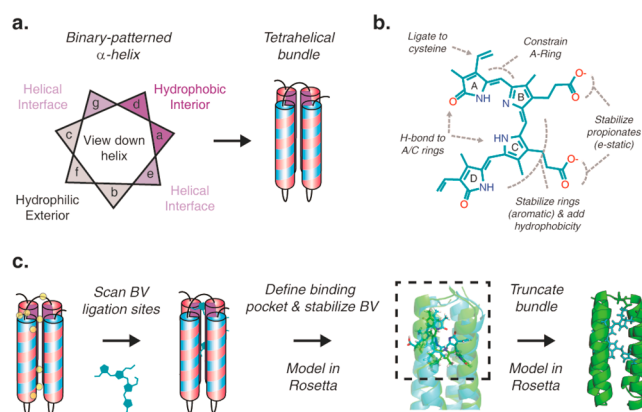


Figure 1. Engineering *de novo*-designed proteins to stabilize biliverdin. (a) Self-assembling single-chain tetrahelical bundles created by binary patterning of hydrophobic and hydrophilic residues with a high α -helix formation propensity, described by the helical wheel. (b) Strategy for stabilizing biliverdin within the core. (c) Holoprotein stepwise construction.

Here, we report the rational construction of compact *de novo*-designed proteins that bind biliverdin (BV), the optically active cofactor in bacteriophytochromes (Bph) and Bph-derived protein tools.^{13–16} Energy minimization modeling in Rosetta^{8,9,17,18} informed the placement of residues for stabilizing BV, which increased its far-red fluorescence. Despite lacking sequence or structural homology to natural biological fluorochromes, fluorescent bili-proteins were successfully forward-engineered with molecular weights as small as 10 kDa or half that of a minimal fluorescent domain engineered from a Bph.¹⁴

RESULTS AND DISCUSSION

Rational Design and Construction Strategy. Fluorescent proteins (FPs) have been engineered by directed evolution of Bph,^{13–16} phytochromes (Phys),^{19,20} allophycocyanins^{21,22} (APs), and fatty acid-binding muscle proteins.²³ These engineered proteins are generally rigidified (i) to stabilize the cofactor in a fluorescent conformation, (ii) to limit access of the solvent and oxygen to the cofactor, and (iii) to

Received: October 9, 2018

Revised: November 15, 2018

Published: November 23, 2018

prevent protein structural rearrangements intrinsic to their signaling roles.

Structural insights from Bph- and Phy-derived FPs^{15,16,19} led to a design strategy for stabilizing the bilin by hydrogen bonding to the BV propionates and A-ring and adding hydrophobic core bulk around the D-ring (Figure 1b). In our rational construction strategy (Figure 1c), we first experimentally identified a suitable cofactor attachment site on a scaffold, which was derived from maquettes with molten globular cores^{5,24} that accommodate a range of cofactor types and sizes. BV-stabilizing residues were subsequently introduced stepwise to define a pocket within the apoprotein core. In the absence of a high-resolution structure for this scaffold, the binding pocket structure-function analysis was informed by energy minimization modeling using Rosetta, given its reported ability to predict helical bundle topologies and binding sites for rigid, planar cofactors.^{8–10}

Cysteine Ligation Scanning. Bilin-containing holoproteins can be reconstituted *in vitro* by autocatalytic ligation of bilin to cysteine.^{6,25,26} To identify suitable ligation positions around which to construct a binding pocket, we scanned cysteine sites for BV covalent attachment efficiency to purified apoproteins *in vitro* (Figure 2). All core residues (heptad repeat positions a and d) were leucines to limit potential contributions to bilin stabilization by structured interactions within the core.

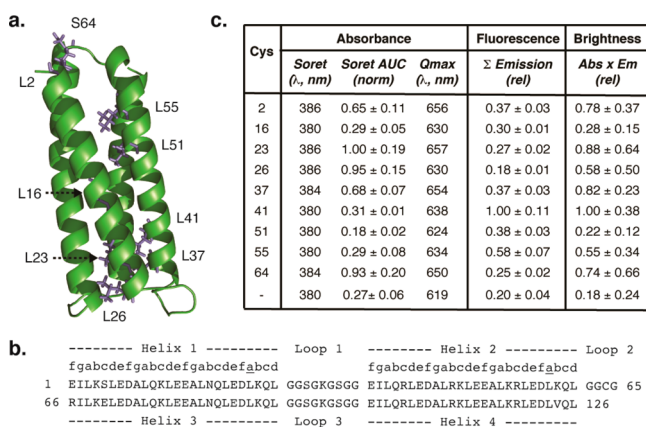


Figure 2. Cysteine scanning for biliverdin (BV) attachment. (a) Rosetta-generated Pymol model of the scaffold, with candidate attachment sites. (b) Scaffold sequence. (c) Relative BV attachment, fluorescence, and brightness summary (mean \pm the standard error). Fluorescence measured at fixed holoprotein concentrations ($\lambda_{\text{ex}} = 600$ nm). Brightness calculated as absorbance \times fluorescence. Abbreviations: AUC, area under the curve; Abs, absorbance; Em, emission; Q, Q-band.

BV attachment levels trended with cysteine solvent exposure, where those in the solvent-exposed B-loop (S64C) or near the termini (L23C) provided good relative balances of appreciable cofactor attachment and baseline fluorescence from partitioning into the hydrophobic core (Figure 2c). In selecting construction starting points, we prioritized BV attachment efficiency given the reported challenges in the uptake of the cofactor by Bph-FPs.^{13,16} Subsequent engineering proceeded faster with the loop-bound S64C maquette, the starting point of proteins hereafter. Rosetta modeling suggested a favored BV placement in the core where an

existing arginine (R119) and lysine (K77) of the scaffold stabilize the cofactor propionates.

Rational Cofactor Stabilization. Stepwise modifications had the intended hierarchical effects of increasing the BV attachment efficiency, enhancing the fluorescence quantum yield (ϕ_F), and sharpening absorbance Q-bands (Figure 3 and Figure S1), with the latter two events indicative of bilin rigidification.²⁶ The helix 4 terminus adjacent to the BV-binding cysteine (C64) was rigidified and made more hydrophobic by placing a valine (K124V) at the interfacial b-position of the last heptad repeat (build step 2). Cofactor stabilization and placement continued from the A-ring, the most constrained pyrrole from covalent attachment, by introduction of a serine (L55) intended to hydrogen bond the A-ring nitrogen (build step 3).

The cofactor B-, C-, and D-rings were further immobilized by positioning histidines (L75H and F120H) to π -stack with the pyrroles and to provide hydrophobic core bulk that restricts protein movement and core water access (build step 4), which has stabilized tetrapyrroles effectively in previous maquettes.^{5,6} Rosetta modeling of the final product suggests S5 may further constrain the A-ring by hydrogen bonding to both the A-ring oxygen and H71 (Figure 3a).

The resultant 15 kDa monomer fluoresced modestly in the far-red spectrum ($\lambda_{\text{ex}} = 648$ nm; $\lambda_{\text{em}} = 662$ nm; $\phi_F = 1.58\%$). The quantum yield is similar to that of sanderycyanin, a natural BV-binding fish pigment ($\phi_F = 1.6\%$),²⁷ and is smaller than those of Bph- and AP-derived directed evolution products ($\phi_F \sim 7\text{--}18\%$).^{13–16,21} For the sake of simplicity, we hereafter call this *de novo*-designed fluorescence-activating protein (“dFP”).

dFP was predominantly monomeric in analytical ultracentrifugation (AUC) assays (Figure S1). Circular dichroism measurements confirmed the bundle helicity and showed that cofactor binding enhanced the overall protein thermal stability ($T_{\text{m-apo}} = 44.7$ °C, and $T_{\text{m-holo}} = 50.8$ °C) (Figure S2). Mass spectrometry and zinc acetate staining of denaturing protein gels confirmed covalent bilin attachment (Figure S3a–c). Noncovalently adsorbed BV was sufficiently removed by filtration on desalting columns (Figure S3d).

Biliverdin formed a thioether bond between its vinyl group and C64, based on acidic denaturation studies in guanidinium chloride (Figure S4a–d). The dFP(C64S) mutation destabilized BV within the core, which is evident by the diminished level of uptake, Q-band absorbance, and fluorescence quantum yield ($\phi_F = 0.8\%$) (Figure S4b). BV was stripped from dFP(C64S) upon denaturation and column filtration and, thus, noncovalently bound to this mutant. Similarly, cysteine-containing dFP bound less mesobiliverdin (meso-BV), which differs from BV by its reduced vinyl side chains (to ethyl), and was stripped from the holoprotein upon denaturation (Figure S4c).

Biliverdin of denatured dFP did not appreciably photoconvert upon stimulation ($\lambda = 610 \pm 5$ or >650 nm) based on the difference spectrum (Figure S5), suggesting that its D-ring adopts a 15Z configuration²⁸ as designed (Figures 1a and 3b) and not the photoconverting 15E configuration.²⁸ A future high-resolution structure would greatly inform dFP structure-function analysis.

Miniature 10 kDa Bili-Protein (“mini”). The 15 kDa dFP contains extraneous heptad repeats for cofactor binding and stabilization needs. We sought to engineer a miniature dFP (hereafter called “dFP-mini” or “mini”) (Figure 4a,b) by truncating the scaffold down to where homology modeling

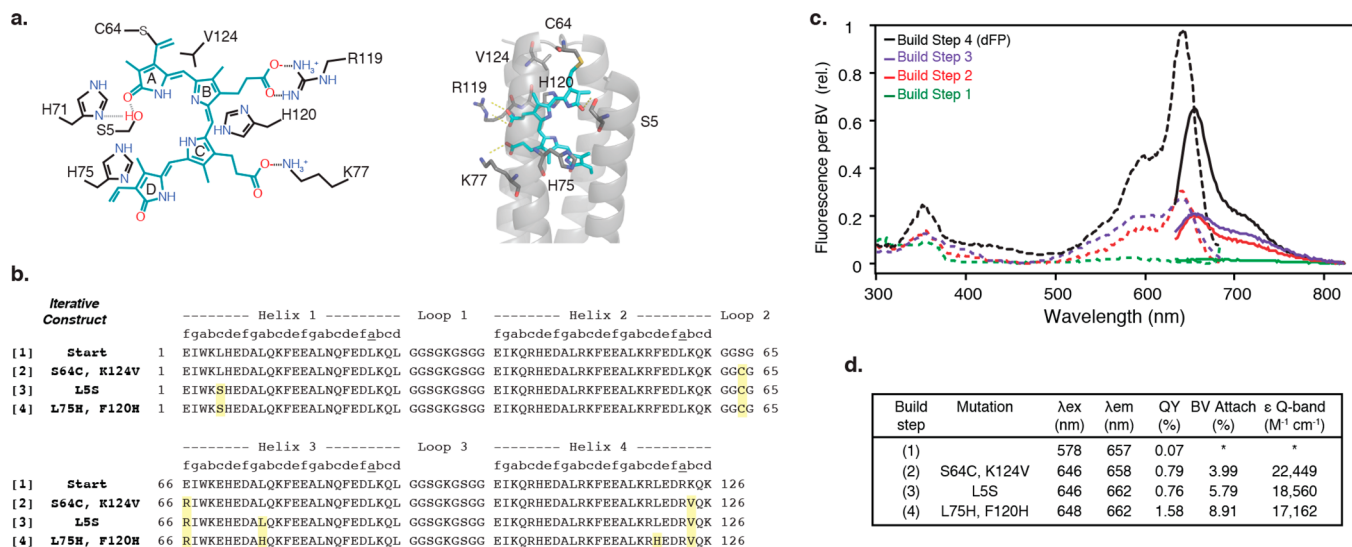


Figure 3. Rational engineering of a biliverdin-binding *de novo*-designed fluorescence-activating protein (dFP). (a) Homology-based contact schematic (left) for BV stabilization (black, side chains; green, BV) and Pymol visualization (right) of the BV-binding site in the Rosetta-modeled core. (b) Sequence alignment of the build series (yellow, mutated residues). The E66R mutation was introduced with the S64C mutation based on “CXR” motifs of natural bili-proteins but did not contribute to stabilization. (c) Excitation (dashed lines; $\lambda_{ex} > 715$ nm) and emission spectra (solid lines; $\lambda_{ex} = 600$ nm) of the stepwise construction. (d) Photophysical summary. Abbreviations: QY, relative quantum yield vs Cy5; ϵ , extinction coefficient. An asterisk indicates a value below the quantification limit. In panels b–d, build step 4 = dFP.

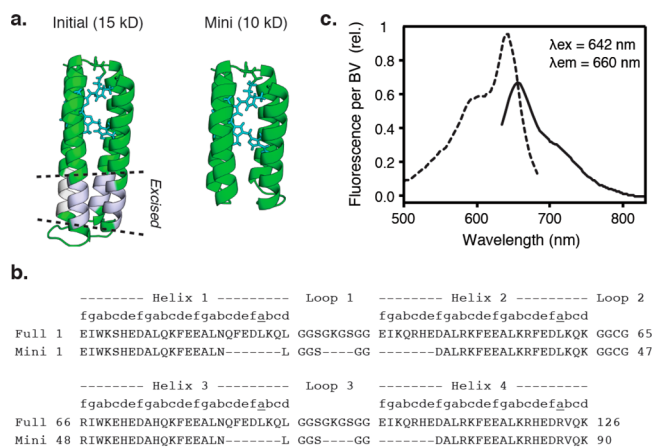


Figure 4. Miniature 10 kDa *de novo*-designed bili-protein (mini). (a) Pymol renders the BV-binding site in Rosetta-modeled full-length and mini dFPs. (b) Sequence alignment. (c) Mini excitation (dashed line; $\lambda_{ex} > 715$ nm) and emission (solid line; $\lambda_{ex} = 600$ nm) spectra. $\phi_F = 1.48\%$. $\epsilon = 16209$ $cm^{-1} M^{-1}$; 16.44% attachment efficiency.

predicts all BV pyrroles remain solvent-shielded. To preserve bundle stability, truncation began after the hydrophobic caps at the helical termini closest to the D-ring. Loops 1 and 3 were also shortened.

A stable 10 kDa mini was formed with helices that extend a half-heptad repeat beyond the furthest modeled D-ring contact residue. Shorter proteins terminating at the final hypothetical contact were unstable. dFP-mini had photophysical properties nearly identical to those of full-length dFP (Figure 4c). Likewise, the mini forms a thioether bond with the BV vinyl group (Figure S4), and the cofactor D-ring adopts a 15Z orientation (Figure S5).

The mini is approximately one-third the size of GFP (27 kDa) and Bph-derived FPs (30–35 kDa) and half that of a minimal domain engineered from a Bph (18 kDa).¹⁴ The facile truncation and relative compactness reflect the structural

simplicity of the *de novo*-designed scaffold. Other than three interhelical loops and hydrophobic caps, the mini lacks accessory structural elements beyond the cofactor-binding pocket itself.

Compact protein fluorochromes are advantageous in molecular sensors because they shorten Förster distances and limit potential interference with the activity of fusion partners. Because BV is endogenous to eukaryotes, this study is a valuable step toward fully genetically encoded and compact *de novo*-designed reporters, with the primary next steps being increasing the quantum yield and bilin uptake for robust performance.

The uptake here reflects an *in vitro* autocatalytic attachment efficiency, without an evolutionarily conserved bilin lyase domain (BLD)^{25,26} or accessory bilin lyase.²⁹ Analogous efficiencies of Bph-derived FPs (before separating apoprotein from holoprotein) are largely unreported. BV likely attaches to the *de novo*-designed scaffold by partitioning into the core and stabilizing within the binding pocket before thioether formation (as described for phytochromes^{25,30}), given that cysteine-to-serine mutations only partially decrease the level of holoprotein formation and that attachment efficiency trended with quantum yield. We anticipate that improved cofactor stabilization will enhance thioether formation and consequent fluorescence properties.

These enhancements may result from complementary directed evolution approaches and/or new computational design tools, including a recently reported “rotamer interaction field” (RIF) algorithm that decouples ligand-docking optimization from overall backbone optimization in Rosetta.¹⁸ This algorithm yielded a *de novo*-designed β -barrel¹⁸ that binds exogenously supplied DFHBI (chromophore of GFP) and fluoresces with a quantum yield ($\phi_F = 2\%$) similar to that described above.

To summarize, we rationally constructed compact *de novo*-designed proteins that covalently bound biliverdin and stabilized it in a fluorescent conformer. In keeping with the

tenets of synthetic biology and protein design, we built them from the bottom up from first principles rather than from the top down using natural protein starting points.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.biochem.8b01076](https://doi.org/10.1021/acs.biochem.8b01076).

Materials and Methods, cofactor rigidification and scaffold stabilization by holoprotein formation (Figure S1), and covalent attachment of BV (Figure S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: bchow@seas.upenn.edu. Telephone: (215) 898-5159.

ORCID

Molly M. Sheehan: 0000-0001-9114-5528

Michael S. Magaraci: 0000-0001-6800-0101

Ivan A. Kuznetsov: 0000-0003-0023-3405

Goutham Kodali: 0000-0001-6356-9016

Christopher C. Moser: 0000-0003-4814-8568

P. Leslie Dutton: 0000-0002-3063-3154

Brian Y. Chow: 0000-0001-6784-6404

Author Contributions

M.M.S., M.S.M., I.A.K., and J.A.M. contributed equally to this work. All authors contributed to experiment and protein design, data analysis, and manuscript preparation. M.M.S., M.S.M., I.A.K., J.A.M., G.K., and C.C.M. conducted protein characterization. I.A.K. and M.S.M. built the Rosetta models. C.C.M., P.L.D., and B.Y.C. coordinated research.

Funding

Research was supported by the National Institutes of Health (1R21DA040434, 1R21EY027562, and 1R01NS101106), the National Science Foundation (NSF) (CBET 126497 and MCB 1652003), and the Penn University Research Foundation (B.Y.C.), the U.S. Department of Energy (DESC0001035) (P.L.D.), and an NSF GRFP Fellowship (M.S.M.). I.A.K. was supported by the Paul and Daisy Soros Fellowship for New Americans.

Notes

The authors declare no competing financial interest. Plasmids will be made available at Addgene (plasmid nos. 120542 and 120543).

■ ACKNOWLEDGMENTS

The authors thank Kushol Gupta, Leland Mayne, and Kendrick Laboratories for technical assistance and Donald Bryant, J. Clark Lagarias, Katrina Forest, Nathan Rockwell, and Vladislav Verkhusha for helpful discussion.

■ REFERENCES

- (1) Huang, P.-S., Boyken, S. E., and Baker, D. (2016) The coming of age of de novo protein design. *Nature* 537, 320–327.
- (2) Dahiyat, B. I., and Mayo, S. L. (1997) De novo protein design: fully automated sequence selection. *Science* 278, 82–87.
- (3) Beasley, J. R., and Hecht, M. H. (1997) Protein design: the choice of de novo sequences. *J. Biol. Chem.* 272, 2031–2034.
- (4) Regan, L., and DeGrado, W. (1988) Characterization of a helical protein designed from first principles. *Science* 241, 976–978.

- (5) Farid, T. A., Kodali, G., Solomon, L. A., Lichtenstein, B. R., Sheehan, M. M., Fry, B. A., Bialas, C., Ennist, N. M., Siedlecki, J. A., Zhao, Z. Y., Stetz, M. A., Valentine, K. G., Anderson, J. L. R., Wand, A. J., Discher, B. M., Moser, C. C., and Dutton, P. L. (2013) Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* 9, 826–833.

- (6) Mancini, J. A., Sheehan, M., Kodali, G., Chow, B. Y., Bryant, D. A., Dutton, P. L., and Moser, C. C. (2018) De novo synthetic biliprotein design, assembly and excitation energy transfer. *J. R. Soc., Interface* 15, 20180021.

- (7) Sharp, R. E., Moser, C. C., Rabanal, F., and Dutton, P. L. (1998) Design, synthesis, and characterization of a photoactivatable flavocytochrome molecular maquette. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10465.

- (8) Huang, P. S., Oberdorfer, G., Xu, C., Pei, X. Y., Nannenga, B. L., Rogers, J. M., DiMaio, F., Gonen, T., Luisi, B., and Baker, D. (2014) High thermodynamic stability of parametrically designed helical bundles. *Science* 346, 481–485.

- (9) Murphy, G. S., Sathyamoorthy, B., Der, B. S., Machius, M. C., Pulavarti, S. V., Szyperki, T., and Kuhlman, B. (2015) Computational de novo design of a four-helix bundle protein—DND_4HB. *Protein Sci.* 24, 434–445.

- (10) Polizzi, N. F., Wu, Y., Lemmin, T., Maxwell, A. M., Zhang, S.-Q., Rawson, J., Beratan, D. N., Therien, M. J., and DeGrado, W. F. (2017) Strategy for designing hyperstable, non-natural protein-ligand complexes with sub-Å accuracy. *Nat. Chem.* 9, 1157–1164.

- (11) O'Neil, K. T., and DeGrado, W. F. (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* 250, 646–651.

- (12) Robertson, D. E., Farid, R. S., Moser, C. C., Urbauer, J. L., Mulholland, S. E., Pidikiti, R., Lear, J. D., Wand, A. J., DeGrado, W. F., and Dutton, P. L. (1994) Design and synthesis of multi-haem proteins. *Nature* 368, 425–432.

- (13) Shu, X., Royant, A., Lin, M. Z., Aguilera, T. A., Lev-Ram, V., Steinbach, P. A., and Tsien, R. Y. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* 324, 804–807.

- (14) Remyantsev, K. A., Shcherbakova, D. M., Zakharova, N. I., Emelyanov, A. V., Turoverov, K. K., and Verkhusha, V. V. (2016) Minimal domain of bacterial phytochrome required for chromophore binding and fluorescence. *Sci. Rep.* 5, 18348.

- (15) Shcherbakova, D. M., Baloban, M., and Verkhusha, V. V. (2015) Near-infrared fluorescent proteins engineered from bacterial phytochromes. *Curr. Opin. Chem. Biol.* 27, 52–63.

- (16) Bhattacharya, S., Auldridge, M. E., Lehtivuori, H., Ihalainen, J. A., and Forest, K. T. (2014) Origins of Fluorescence in Evolved Bacteriophytochromes. *J. Biol. Chem.* 289, 32144.

- (17) Meiler, J., and Baker, D. (2006) ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins: Struct., Funct., Genet.* 65, 538–548.

- (18) Dou, J., Vorobieva, A. A., Sheffler, W., Doyle, L. A., Park, H., Bick, M. J., Mao, B., Foight, G. W., Lee, M. Y., Gagnon, L. A., Carter, L., Sankaran, B., Ovchinnikov, S., Marcos, E., Huang, P.-S., Vaughan, J. C., Stoddard, B. L., and Baker, D. (2018) De novo design of a fluorescence-activating β -barrel. *Nature* 561, 485.

- (19) Lehtivuori, H., Bhattacharya, S., Angenent-Mari, N. M., Satyshur, K. A., and Forest, K. T. (2015) Removal of Chromophore-Proximal Polar Atoms Decreases Water Content and Increases Fluorescence in a Near Infrared Phytofluor. *Front. Mol. Biosci.* 2, 65.

- (20) Murphy, J. T., and Lagarias, J. C. (1997) The phytofluors: a new class of fluorescent protein probes. *Curr. Biol.* 7, 870–876.

- (21) Rodriguez, E. A., Tran, G. N., Gross, L. A., Crisp, J. L., Shu, X., Lin, J. Y., and Tsien, R. Y. (2016) A far-red fluorescent protein evolved from a cyanobacterial phycobiliprotein. *Nat. Methods* 13, 763–769.

- (22) Glazer, A. N., and Stryer, L. (1984) Phycofluor probes. *Trends Biochem. Sci.* 9, 423–427.

- (23) Kumagai, A., Ando, R., Miyatake, H., Greimel, P., Kobayashi, T., Hirabayashi, Y., Shimogori, T., and Miyawaki, A. (2013) A

bilirubin-inducible fluorescent protein from eel muscle. *Cell* 153, 1602–1611.

(24) Huang, S. S., Koder, R. L., Lewis, M., Wand, A. J., and Dutton, P. L. (2004) The HP-1 maquette: From an apoprotein structure to a structured hemoprotein designed to promote redox-coupled proton exchange. *Proc. Natl. Acad. Sci. U. S. A.* 101, 5536.

(25) Lamparter, T., Carrascal, M., Michael, N., Martinez, E., Rottwinkel, G., and Abian, J. (2004) The Biliverdin Chromophore Binds Covalently to a Conserved Cysteine Residue in the N-Terminus of *Agrobacterium* Phytochrome App1. *Biochemistry* 43, 3659–3669.

(26) Wu, S. H., and Lagarias, J. C. (2000) Defining the bilin lyase domain: lessons from the extended phytochrome superfamily. *Biochemistry* 39, 13487–13495.

(27) Ghosh, S., Yu, C.-L., Ferraro, D. J., Sudha, S., Pal, S. K., Schaefer, W. F., Gibson, D. T., and Ramaswamy, S. (2016) Blue protein with red fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* 113, 11513.

(28) Shang, L., Rockwell, N. C., Martin, S. S., and Lagarias, J. C. (2010) Biliverdin amides reveal roles for propionate side chains in bilin reductase recognition and in holophytochrome assembly and photoconversion. *Biochemistry* 49, 6070–6082.

(29) Kronfel, C. M., Kuzin, A. P., Forouhar, F., Biswas, A., Su, M., Lew, S., Seetharaman, J., Xiao, R., Everett, J. K., Ma, L.-C., Acton, T. B., Montelione, G. T., Hunt, J. F., Paul, C. E. C., Dragomani, T. M., Boutaghou, M. N., Cole, R. B., Riml, C., Alvey, R. M., Bryant, D. A., and Schluchter, W. M. (2013) Structural and Biochemical Characterization of the Bilin Lyase CpcS from *Thermosynechococcus elongatus*. *Biochemistry* 52, 8663–8676.

(30) Li, L., Murphy, J. T., and Lagarias, J. C. (1995) Continuous Fluorescence Assay of Phytochrome Assembly in Vitro. *Biochemistry* 34, 7923–7930.