Rational construction of compact de novo-designed biliverdin-binding proteins

Molly M. Sheehan,1,4 Michael S. Magaraci,1,4 Ivan A. Kuznetsov,1,4 Joshua A. Mancini,2,4 Goutham Kodali,2 Christopher C. Moser,2 P. Leslie Dutton,2 and Brian Y. Chow1,∗

1Department of Bioengineering & 2Department of Biochemistry and Biophysics, Johnson Research Foundation, University of Pennsylvania, Philadelphia, PA 19104. (∗) Equal contributions. (* Email: bchow@seas.upenn.edu. Tel: (215) 898-5159

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 tetrahedral bundles bind biliverdin IXa (BV) cofactor autcatalytically in vitro, similarly to photosensory proteins that bind BV (and related bilins, or linear tetrapyrroles) despite lacking sequence and structural homology to the natural counterparts. Upon identifying a suitable site for cofactor ligation 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 of the binding pocket. Holoprotein formation stabilized BV, resulting in increased far-red BV fluorescence. By removing segments extraneous to cofactor stabilization or bundle stability, the initial 15-kilodalton de novo-designed fluorescence-activating protein (“dFP”) was truncated without altering its optical properties, down to a miniature 10-kilodalton “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 scaffolds5,13. Self-assembling tetrahedral bundles4,10, created by binary patterning of hydrophobic and hydrophilic residues with high α-helical propensity11, comprise the best-established class of de novo-designed scaffolds. They provide stable frames for binding cofactors, as protein maquettes5,7,12 for rationally engineering artificial holoproteins in which the cofactor-interacting structure-function 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 propionates5.

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 tools13-16. Energy minimization modeling in Rosetta8,9,17,18 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 bilir-proteins were successfully forward-engineered as small as 10-kD molecular weight, or half that of a minimal fluorescent domain engineered from a Bph14.

RESULTS AND DISCUSSION

Rational design and construction strategy. Fluorescent proteins (FPs) have been evolved from Bph13,16, phycocyanin (Phy)19,20, allophycocyanins21,22 (AP), and fatty acid-binding muscle proteins23. These engineered proteins
are generally rigidified (i) to stabilize the cofactor in a fluorescent conformation, (ii) to limit solvent and oxygen access to the cofactor, and (iii) to prevent protein structural rearrangements intrinsic to their signaling roles.

Structural insights from Bph- and Phy-derived FPds\(^{15,16,19}\) led to a design strategy for stabilizing the bilin by hydrogen bonding to the BV propionates and A-ring, plus 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\(^{12-14}\) 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 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* due to autocatalytic bilin ligation to cysteine \(^6,23,24\). 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 a- and d-positions) were leucines to limit potential contributions to bilin stabilization by structured interactions within the core.

<table>
<thead>
<tr>
<th>Cysteine</th>
<th>Absorption</th>
<th>Fluorescence</th>
<th>Brightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorbent</td>
<td>Emission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC (μm)</td>
<td>Qmax (μm)</td>
<td>λEm</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>298</td>
<td>0.85 ± 0.11</td>
<td>656</td>
</tr>
<tr>
<td>16</td>
<td>380</td>
<td>0.29 ± 0.05</td>
<td>630</td>
</tr>
<tr>
<td>23</td>
<td>398</td>
<td>1.00 ± 0.10</td>
<td>657</td>
</tr>
<tr>
<td>26</td>
<td>398</td>
<td>0.95 ± 0.15</td>
<td>650</td>
</tr>
<tr>
<td>34</td>
<td>384</td>
<td>0.88 ± 0.07</td>
<td>654</td>
</tr>
<tr>
<td>41</td>
<td>380</td>
<td>0.31 ± 0.01</td>
<td>638</td>
</tr>
<tr>
<td>51</td>
<td>398</td>
<td>0.18 ± 0.02</td>
<td>624</td>
</tr>
<tr>
<td>55</td>
<td>390</td>
<td>0.29 ± 0.08</td>
<td>645</td>
</tr>
<tr>
<td>65</td>
<td>384</td>
<td>0.93 ± 0.20</td>
<td>650</td>
</tr>
<tr>
<td>67</td>
<td>390</td>
<td>0.27 ± 0.06</td>
<td>619</td>
</tr>
</tbody>
</table>

**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 ± s.e.). Fluorescence measured at fixed holoprotein concentrations (λex = 600 nm). Brightness calculated as absorbance × fluorescence. (AUC = area under 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 reported challenges in cofactor uptake into Bph-FPs\(^{15,16}\). Subsequent engineering proceeded faster with the loop-bound S64C maquette, the starting point of proteins hereon. 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 BV attachment efficiency, enhancing fluorescence quantum yield (φf), and sharpening absorbance Q-bands (**Figures 3 and S1**), with the latter two events indicative of bilin rigidification\(^{27}\). 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 introducing a serine (LSS) 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 pi-stack to the pyrroles and to provide hydrophobic core bulk that restricts protein movement and core water access (Build Step 4), which has stabilized tetrapyroles effectively in previous maquettes\(^6-9\). 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-kD monomer fluoresced modestly in the far-red spectrum (λex = 648 nm, λem = 662 nm, φf = 1.58%). The quantum yield is similar to that of sanderceyanin, a natural BV-binding fish pigment (φf = 1.6%)\(^{27}\), and is less than those of Bph- and AP-derived directed evolution products (φf ~ 7-18%)\(^{13,16,21}\). For simplicity, we hereon 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 (Tm-apo = 44.7°C, Tm-holo = 50.8°C (**Figure S2**)). Mass spectrometry and zinc acetate-staining of denaturing protein gels confirmed covalent bilin attachment (**Figure S3a-c**). Non-covalently 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, as evident by diminished uptake, Q-band absorbance, and fluorescence quantum yield (φf = 0.8%) (**Figure S4b**). BV was stripped from dFP(C64S) upon denaturation and column filtration, and thus, non-covalently bound to this mutant. Similarly, cysteine-containing dFP bound less mesobiliverdin (meso-BV), which differs from BV by its reduced vinyl sidechains (to ethyl), and was stripped from the holoprotein upon denaturation (**Figure S4c**).

Biliverdin of denatured dFP did not appreciably photoconvert upon stimulation (λ = 610+/− 5 nm or > 650 nm) based on the difference spectrum (**Figure S5**), suggesting that its
D-ring adopts a 15Z configuration as designed (Figures 1a and 3b), but the hydrophobic caps at the helical termini nearest to the D-ring. Loops 1 and 3 were also shortened.

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

(b-d) Build step 4 = dFP.

Figure 4. Miniature 10 kD de novo-designed bili-protein (mini). (a) Pymol renders of the BV binding site in Rosetta-modeled full-length and mini dFPs. (b) Sequence alignment. (c) Mini excitation (dashed, λem > 715 nm) and emission spectra (solid, λem = 600 nm). ϕE = 1.48%, ε = 16,209 cm⁻¹ M⁻¹, 16.44% attachment efficiency.

A stable 10-kD 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 nearly identical photophysical properties to full-length dFP (Figure 4c.). Likewise, the mini forms a thioether bond with the BV vinyl group (Figure 4d), and the cofactor D-ring adopts a 15Z orientation (Figure 4e).

The mini is approximately one-third the size of GFP (27 kD) and Bph-derived FP’s (30-35 kD), and half that of a minimal domain engineered from a Bph (18 kD)[24]. The facile truncation and relative compactness reflect the structural simplicity of the de novo-designed scaffold. Other than three inter-helical loops and hydrophobic caps, the mini lacks accessory structural elements beyond the cofactor-binding pocket itself.

Compact protein fluorochromes are advantageous in molecular sensors by shortening Förster distances and limiting potential interference with the activity of fusion partners. Since BV is endogenous to eukaryotes, this study is a valuable step toward fully genetically encoded and compact de novo-designed reporters, with primary next steps of increasing 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 FP’s (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, 26]), given that cysteine-to-serine mutations only partially reduce 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
. This algorithm begot a de novo-designed β-barrel that binds exogenously supplied DFHBI (chromophore of GFP) and fluoresces with a quantum yield ($\Phi_F = 2\%$) similar to here.

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, they were built from the bottom-up from first principles, rather than engineered from the top-down using natural protein starting points.

ASSOCIATED CONTENT

Materials and Methods. Refer to Supporting Information. Plasmids will be made available at Addgene.

Supporting Information. Materials and Methods. Figure S1: Cofactor rigidification and scaffold stabilization by holoprotein formation. Figure S2. Covalent attachment of BV.

AUTHOR INFORMATION

Corresponding Author
* Address: 210 South 33rd Street, Philadelphia, PA 19104. Tel: (215) 898-5159. Email: bchow@seas.upenn.edu

Author Contributions
All authors contributed to experiment and protein design, data analysis, and manuscript preparation. MMS, MSM, IAK, JAM, GK, and CCM conducted protein characterization. IAK and MSM built the Rosetta models. CCM, PLD, and BYC coordinated research.

Funding Sources
Research was supported by: [BYC] National Institutes of Health (1R21DA040434, 1R21EV027562, 1R01NS101106), National Science Foundation (CBET 126497, MCB 1652003), and the Penn University Research Foundation. [PLD] Department of Energy (DESC0001035). [MSM] NSF GRFP Fellowship. [IAK] was supported by the Paul and Daisy Soros Fellowship for New Americans.

ACKNOWLEDGMENT
We 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


Rational construction of compact de novo-designed biliverdin-binding proteins

Molly M. Sheehan,1,† Michael S. Magaraci,1,† Ivan A. Kuznetsov,1,† Joshua A. Mancini,2,† Goutham Kodali,2 Christopher C. Moser,2 P. Leslie Dutton,2 and Brian Y. Chow1,*