

A naturally monomeric infrared fluorescent protein for protein labeling *in vivo*

Dan Yu^{1,2}, Michelle A Baird^{3,4}, John R Allen^{3,4}, Elizabeth S Howe^{3,4}, Matthew P Klassen^{2,5,6}, Anna Reade², Kalpana Makhijani^{1,2}, Yuanquan Song^{2,5,6}, Songmei Liu², Zehra Murthy², Shao-Qing Zhang^{1,2,7}, Orion D Weiner², Thomas B Kornberg², Yuh-Nung Jan^{2,5,6}, Michael W Davidson^{3,4} & Xiaokun Shu^{1,2}

Infrared fluorescent proteins (IFPs) provide an additional color to GFP and its homologs in protein labeling. Drawing on structural analysis of the dimer interface, we identified a bacteriophytochrome in the sequence database that is monomeric in truncated form and engineered it into a naturally monomeric IFP (mIFP). We demonstrate that mIFP correctly labels proteins in live cells, *Drosophila* and zebrafish. It should be useful in molecular, cell and developmental biology.

GFP and its red homologs are powerful tools for cell and molecular biology^{1,2}. Recently, this fluorescent protein (FP) palette has been extended into the infrared region by introduction of bacteriophytochrome (BphP)-derived IFPs (for example, IFP1.4 and iRFP) that autocatalytically incorporate biliverdin (BV) as the chromophore^{3–7}. As a protein fusion tag, an FP should be monomeric so as not to perturb the stoichiometry of the protein of interest. However, most BphPs function as multimeric complexes^{8–10}; both IFP1.4 and iRFP are derived from dimeric and truncated BphPs (Δ BphP denotes truncations including only the PAS and GAF domains). Whereas iRFP is dimeric, in our previous characterization we reported IFP1.4 and its variant IFP2.0 (ref. 11) to be monomeric. However, we have found that IFP1.4 and IFP2.0 both tend to dimerize at high concentration, with dissociation constants of 7.8 μ M and 3.7 μ M, respectively (Supplementary Fig. 1). To develop a robust protein tag in the infrared spectrum, we decided to engineer a naturally monomeric IFP: mIFP.

We first identified a monomeric Δ BphP that, judging by the biological fitness of the bacteria expressing the full-length protein,

is unlikely to misfold or aggregate in cells¹². We hypothesized that some of the many BphP sequences in protein sequence databases might be monomeric in the Δ BphP form, which is likely to occur if the protein lacks strong hydrophobic interactions at the putative dimer interface. In contrast, in IFP1.4's parent, *Dr*BphP¹³, the dimer interface includes several residues (Fig. 1a): Leu311 appears to play a critical role, as the mutation L311K disrupts the dimer interface³. Analysis of ~40 BphP sequences from the NCBI database revealed *Br*BphP (from *Bradyrhizobium*) as a potential candidate, as the residue corresponding to Leu311 in *Dr*BphP is a polar threonine (Fig. 1b). Indeed, size-exclusion chromatography indicated that Δ *Br*BphP eluted later than dimeric Δ *Dr*BphP and at a time similar to that of the monomeric form of IFP1.4 (Supplementary Fig. 2), suggesting that Δ *Br*BphP is a monomer.

We engineered the nonfluorescent Δ *Br*BphP into a fluorescent mutant. In brief, we selected several residues (Asp199, Tyr168, Val178 and Asn258) surrounding BV for saturation mutagenesis, which was followed by DNA shuffling¹⁴ and random mutagenesis. The final fluorescent mutant mIFP absorbed maximally at 683 nm (Supplementary Fig. 3), with excitation and emission maxima of 683 and 704 nm, respectively (Fig. 1c), a quantum yield of 8% and an extinction coefficient of 82,000 M⁻¹ cm⁻¹ (Supplementary Table 1). We confirmed that mIFP was monomeric at high concentrations (17 and 34 μ M) (Fig. 1d and Supplementary Fig. 4). It contains 19 mutations (Supplementary Figs. 5 and 6), including 5 near the D-ring of BV that likely limit its rotation, contributing to the engineered fluorescence by increasing radiative decay of the excited state (Fig. 1e). Mutated residues in mIFP, IFP1.4 and iRFP do not overlap and thus might be targeted for further engineering (Supplementary Fig. 7). mIFP was stable in pH 4–10 (Fig. 1f). Its molecular brightness was similar to that of IFP1.4 and iRFP (Supplementary Table 1), and its cellular brightness was similar to that of iRFP and tenfold greater than that of IFP1.4 in live HeLa cells (Fig. 1g,h and Supplementary Fig. 8). mIFP was 6.3 times more photostable than IFP1.4 (Supplementary Fig. 9) but approximately one-fifth as photostable as iRFP in HEK293 cells (Supplementary Table 1). Photobleaching of mIFP was irreversible (Supplementary Fig. 10), suggesting no residual photoisomerization. mIFP was similar to IFP2.0 and iRFP in terms of maturation rate, BV binding kinetics and affinity (Supplementary Figs. 11–13).

To demonstrate mIFP as a protein tag for use in live-cell imaging, we constructed ~30 mIFP fusion proteins, targeting both the N and C termini with an appropriate-length linker (Online Methods). We successfully expressed and imaged the mIFP fusions in cultured cells without addition of the cofactor, which suggests that mIFP

¹Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California, USA. ²Cardiovascular Research Institute, University of California, San Francisco, San Francisco, California, USA. ³National High Magnetic Field Laboratory, Tallahassee, Florida, USA.

⁴Department of Biological Science, The Florida State University, Tallahassee, Florida, USA. ⁵Department of Physiology, University of California, San Francisco, San Francisco, California, USA. ⁶Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, California, USA. ⁷Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania, USA. Correspondence should be addressed to X.S. (xiaokun.shu@ucsf.edu).

RECEIVED 10 JANUARY; ACCEPTED 31 MARCH; PUBLISHED ONLINE 22 JUNE 2015; DOI:10.1038/NMETH.3447

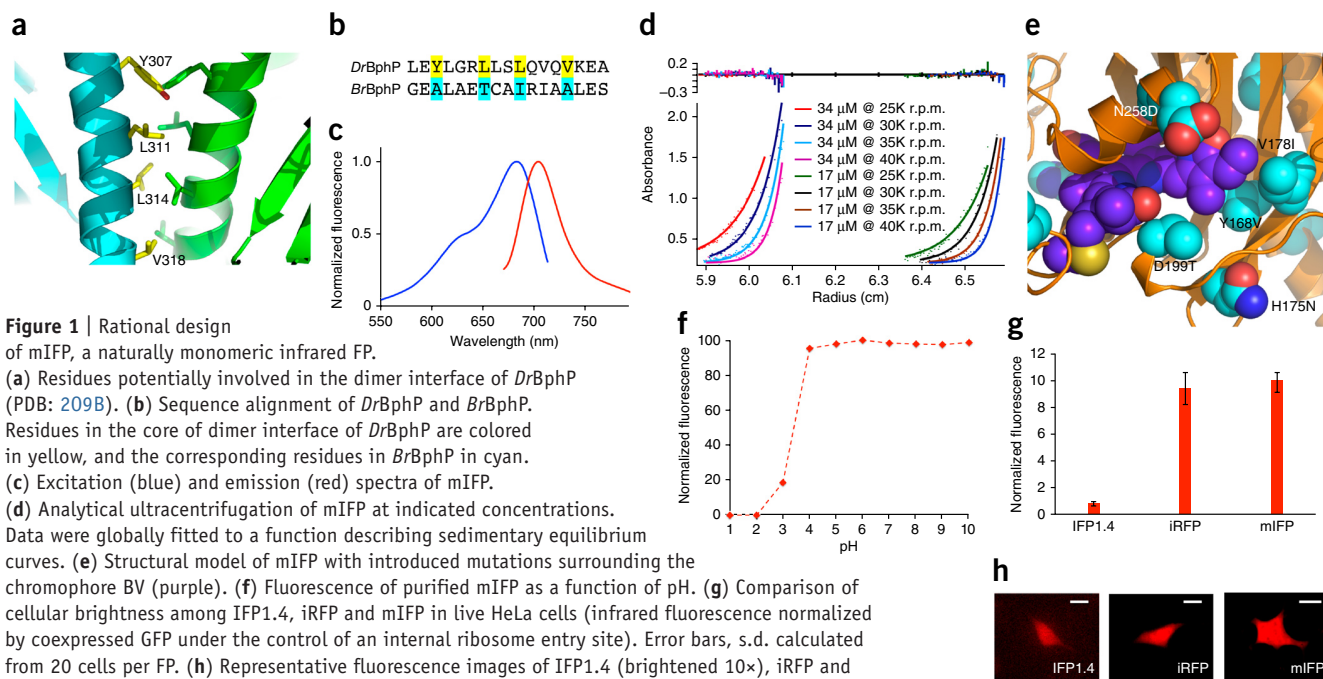


Figure 1 | Rational design

of mIFP, a naturally monomeric infrared FP.

(a) Residues potentially involved in the dimer interface of *DrBphP* (PDB: 209B). (b) Sequence alignment of *DrBphP* and *BrBphP*. Residues in the core of dimer interface of *DrBphP* are colored in yellow, and the corresponding residues in *BrBphP* in cyan.

(c) Excitation (blue) and emission (red) spectra of mIFP.

(d) Analytical ultracentrifugation of mIFP at indicated concentrations. Data were globally fitted to a function describing sedimentary equilibrium curves.

(e) Structural model of mIFP with introduced mutations surrounding the chromophore BV (purple). (f) Fluorescence of purified mIFP as a function of pH. (g) Comparison of cellular brightness among IFP1.4, iRFP and mIFP in live HeLa cells (infrared fluorescence normalized by coexpressed GFP under the control of an internal ribosome entry site). Error bars, s.d. calculated from 20 cells per FP. (h) Representative fluorescence images of IFP1.4 (brightened 10 \times), iRFP and mIFP in HeLa cells. 20 cells were examined for each FP. Scale bars, 10 μ m.

(c) Excitation (blue) and emission (red) spectra of mIFP.

(d) Analytical ultracentrifugation of mIFP at indicated concentrations. Data were globally fitted to a function describing sedimentary equilibrium curves.

(e) Structural model of mIFP with introduced mutations surrounding the chromophore BV (purple). (f) Fluorescence of purified mIFP as a function of pH. (g) Comparison of cellular brightness among IFP1.4, iRFP and mIFP in live HeLa cells (infrared fluorescence normalized by coexpressed GFP under the control of an internal ribosome entry site). Error bars, s.d. calculated from 20 cells per FP. (h) Representative fluorescence images of IFP1.4 (brightened 10 \times), iRFP and mIFP in HeLa cells. 20 cells were examined for each FP. Scale bars, 10 μ m.

(f) Fluorescence of purified mIFP as a function of pH. (g) Comparison of cellular brightness among IFP1.4, iRFP and mIFP in live HeLa cells (infrared fluorescence normalized by coexpressed GFP under the control of an internal ribosome entry site). Error bars, s.d. calculated from 20 cells per FP. (h) Representative fluorescence images of IFP1.4 (brightened 10 \times), iRFP and mIFP in HeLa cells. 20 cells were examined for each FP. Scale bars, 10 μ m.

(g) Comparison of cellular brightness among IFP1.4, iRFP and mIFP in live HeLa cells (infrared fluorescence normalized by coexpressed GFP under the control of an internal ribosome entry site). Error bars, s.d. calculated from 20 cells per FP. (h) Representative fluorescence images of IFP1.4 (brightened 10 \times), iRFP and mIFP in HeLa cells. 20 cells were examined for each FP. Scale bars, 10 μ m.

(h) Representative fluorescence images of IFP1.4 (brightened 10 \times), iRFP and mIFP in HeLa cells. 20 cells were examined for each FP. Scale bars, 10 μ m.

used endogenous BV. These fusions localized properly in live cells (Supplementary Figs. 14–16 and Supplementary Videos 1–6), including those that require a high degree of monomeric character, such as α -tubulin, connexin 43 and intermediate filaments. We observed all phases of mitosis in fusions of mIFP to human histones H1 and H2B. Additionally, mIFP was compatible with structured illumination microscopy (Supplementary Fig. 17).

To demonstrate mIFP as a protein tag in live animals, we created histone fusions and imaged them *in vivo* in fruit flies and zebrafish. For the flies (*Drosophila melanogaster*), we created upstream activating sequence (UAS)-mIFP-histone 3.3 (H3.3) T2A heme oxygenase-1 (HO1) transgenic line and subsequently crossed it with the *engrailed-GAL4* line to promote expression of mIFP-H3.3 in a segmental pattern. Here T2A is a ‘self-cleaving’ peptide widely used in coexpression of multiple genes¹⁵ and HO1 converts heme to BV¹⁶. Coexpression of this enzyme was necessary to overcome the insufficient levels of endogenous BV in *Drosophila*. Confocal imaging of the embryo detected bright nuclear fluorescence with the expected segmental pattern (Fig. 2a,b). For zebrafish (*Danio rerio*), we expressed mIFP-H2B with HO1 by mRNA injection at the one-cell stage. Imaging of the eye region at 30 hours post fertilization (h.p.f.) revealed bright nuclear fluorescence with the proper expression pattern (Fig. 2c).

To apply mIFP in multicolor labeling *in vivo*, we coexpressed mIFP-H3.3 T2A HO1 and CD8-GFP in neurons of *Drosophila*. CD8 is a transmembrane protein that labels the cell membrane (as is CD4). We observed infrared fluorescence in the nucleus and green fluorescence in the cell membrane in the brain region of embryos, results suggesting correct targeting of the fusion proteins (Fig. 2d–g). We further expressed mIFP T2A HO1 in the abdominal muscle of *Drosophila* larvae, together with a GFP fusion trap of the extracellular matrix protein Viking (collagen) and CD4-tdTomato in class IV dendritic arborization (DA) neurons (*ppk::CD4-tdTomato*). Fluorescence imaging revealed separation of the three fluorophores (Fig. 2h–j) and the expected structural organization of the labeled

cell types (Fig. 2k). Expression of CD4-mIFP T2A HO1 in class IV DA neurons in *Drosophila* larvae clearly and evenly labeled dendrites and axons (Fig. 2l,m), as expected from previous results with CD4-GFP¹⁷.

To compare the naturally monomeric mIFP to our previously engineered monomeric IFP2.0 (which was derived from a dimeric parent), we expressed CD4-mIFP T2A HO1 and CD4-IFP2.0 T2A HO1 in epithelial cells of *Drosophila* larvae. Whereas mIFP-CD4 correctly labeled the epithelial cell membrane, IFP2.0-CD4 formed aggregates and failed to label the plasma membrane (Supplementary Fig. 18a,b). This is consistent with *in vitro* data showing that IFP2.0 tends to dimerize at high concentrations.

To demonstrate the use of mIFP in cell labeling, we expressed mIFP in different tissues and compared mIFP to engineered monomeric FPs derived from oligomeric parents, including the popular red FP mCherry, orange FP tdTomato¹⁸ and red FP FusionRed¹⁹ (Supplementary Fig. 18). We found that mCherry formed punctate structures in the muscles and neurons of *Drosophila* first instar larvae. We observed these puncta in the body muscles of each of the more than 200 examined animals. Expression of FusionRed in the leg muscles of adult *Drosophila* revealed many rounded structures, which varied in diameter from 0.5 to 1.3 μ m; tdTomato formed punctate structures with elongated shape varying in length from 2 to 6 μ m with width of \sim 0.4 μ m. In contrast, mIFP expression was homogeneous in all these contexts, similarly to what we observe with GFP (Supplementary Fig. 19).

To examine potential toxicity of mIFP and HO1, we ubiquitously expressed them in *Drosophila* and conducted a viability assay. We did not find obvious toxicity differences between mIFP or mIFP T2A HO1 and GFP (Supplementary Fig. 20a). Furthermore, we did not observe any defects in the eye morphogenesis of *Drosophila* ubiquitously expressing either mIFP or mIFP T2A HO1, and *Drosophila* embryos expressing mIFP-H3.3 T2A HO1 and CD8-GFP in neurons displayed normal ventral nerve cord shortening without any obvious differences with respect to embryos

Figure 2 | Expression of mIFP fusions *in vivo*. (a,b) Fluorescence image of *Drosophila* embryo expressing UAS-mIFP-histone 3.3 T2A HO1 driven by *engrailed-GAL4*. (b) High-magnification view of the boxed area in a. (c) Fluorescence image of a zebrafish eye expressing mIFP-H2B and HO1. (d-g) Two-color fluorescence imaging of a *Drosophila* embryo expressing UAS-mIFP-histone 3.3 T2A HO1 and UAS-CD8-GFP, driven by *elav-GAL4*. (d) Whole embryo. (e-g) High-magnification views of the boxed area in d showing GFP expression (e), mIFP expression (f) and a merged image (g). (h-k) Three-color fluorescence imaging of *Drosophila* abdominal muscle expressing UAS-mIFP T2A HO1 driven by *Mef2R-GAL4*, class IV DA neurons expressing CD4-tdTomato (*ppk::CD4-tdTomato*) and extracellular matrix expressing GFP-Viking (collagen): tdTomato (in blue pseudocolor) (h), GFP (i), mIFP (j) and merged (k). (l,m) Class IV DA neurons expressing CD4-mIFP in *Drosophila* larvae. (m) High-magnification view of the boxed area in l. Scale bars: 50 μm (a,d,l), 10 μm (b,e-g) and 20 μm (c,h-k,m).

expressing CD8-GFP alone (Supplementary Videos 7 and 8). We note that coexpression with HO1 improved mIFP fluorescence by 30- to 40-fold in *Drosophila* muscle (Supplementary Fig. 20b,c). We also observed no obvious toxicity in zebrafish expressing myristoylated mIFP (*myr_mIFP*) and HO1 by RNA injection at the one-cell stage, evaluated by comparison to zebrafish expressing only GFP (Supplementary Fig. 20d).

To test whether coexpression of HO1 also improves iRFP fluorescence in zebrafish, we expressed myristoylated iRFP (*myr_iRFP*) with or without HO1 in the tail of embryo at 30 h.p.f. (Supplementary Fig. 21). We observed little iRFP fluorescence in the absence of HO1. Coexpression of HO1 substantially increased iRFP fluorescence, which was obvious in the cell membrane in the tail. Zebrafish expressing *myr_mIFP* and HO1 showed similar levels of infrared fluorescence in the tail muscle (Supplementary Fig. 21e,f), a finding consistent with our *in vitro* data indicating that mIFP and iRFP have comparable BV binding kinetics and affinity.

In summary, we have engineered a naturally monomeric mIFP that requires biliverdin as a cofactor and have demonstrated its use for protein labeling in living cells and *in vivo*. mIFP is a good template for the development of infrared fluorescent reporters, such as for visualizing cell signaling in live animals. For example, we have designed and recently published an infrared fluorogenic protease reporter based on mIFP that visualizes apoptotic pathways *in vivo*²⁰. Protein sequence databases contain thousands of BphPs. These proteins are promising starting points for rational design of future IFPs and related tools with desirable photophysical and photochemical properties.

METHODS

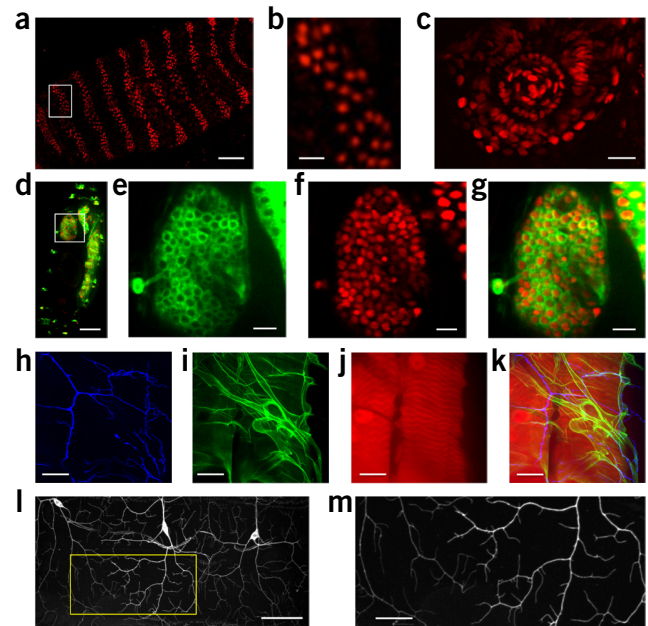
Methods and any associated references are available in the online version of the paper.

Accession codes. GenBank/EMBL/DDBJ: [KM285236](https://www.ncbi.nlm.nih.gov/nuclseq/KM285236).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by the Program for Breakthrough Biomedical Research (to X.S.); US National Institute of Health (NIH) GM030637 (to T.B.K.), GM084040 and GM096164 (to O.D.W.); and the Howard Hughes Medical Institute (to Y.-N.J.). S.-Q.Z. was supported by NIH GM054616 (to W.F. DeGrado) and US National Science Foundation DMR-1120901. We thank N. Joh for assistance in gel filtration chromatography; A. Royant for size-exclusion chromatography; W.F. DeGrado for providing access to the analytical ultracentrifuge;



L.D. Wilsbacher, A. Schepis and S.R. Coughlin for initial testing of mIFP in zebrafish; C.S. Craik for constructive comments; S. Woo at University of California, San Francisco (UCSF) for providing the pCS2+ vector; and S. Roy (UCSF) for providing the UAS-mCherry *Drosophila* line.

AUTHOR CONTRIBUTIONS

X.S. conceived the project. D.Y. and X.S. designed mIFP and the H2B fusion. K.M. and X.S. planned the *Drosophila* embryo imaging. M.W.D. planned the fusion constructs. T.B.K. planned the histone H3.3 fusion construct and the transgenic *Drosophila*. A.R. and O.D.W. planned the imaging of zebrafish. M.P.K., Y.S. and Y.-N.J. planned the imaging of epithelia, muscle and neurons in *Drosophila* larvae and adults. D.Y., M.A.B., J.R.A., E.S.H., M.P.K., A.R., K.M., Y.S., S.L., Z.M. and S.-Q.Z. performed the experiments. D.Y. and X.S. wrote the manuscript. All the authors contributed to the final draft.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Day, R.N. & Davidson, M.W. *Chem. Soc. Rev.* **38**, 2887–2921 (2009).
- Tsien, R.Y. *Angew. Chem. Int. Ed. Engl.* **48**, 5612–5626 (2009).
- Shu, X. *et al. Science* **324**, 804–807 (2009).
- Filonov, G.S. *et al. Nat. Biotechnol.* **29**, 757–761 (2011).
- Auldridge, M.E., Satyshur, K.A., Anstrom, D.M. & Forest, K.T. *J. Biol. Chem.* **287**, 7000–7009 (2012).
- Shcherbakova, D.M. & Verkhusha, V.V. *Nat. Methods* **10**, 751–754 (2013).
- Piatkevich, K.D., Subach, F.V. & Verkhusha, V.V. *Nat. Commun.* **4**, 2153–2162 (2013).
- Giraud, E. & Verméglio, A. *Photosynth. Res.* **97**, 141–153 (2008).
- Rockwell, N.C., Su, Y.S. & Lagarias, J.C. *Annu. Rev. Plant Biol.* **57**, 837–858 (2006).
- Karniol, B., Wagner, J.R., Walker, J.M. & Vierstra, R.D. *Biochem. J.* **392**, 103–116 (2005).
- Yu, D. *et al. Nat. Commun.* **5**, 3626–3632 (2014).
- Geiler-Samerotte, K.A. *et al. Proc. Natl. Acad. Sci. USA* **108**, 680–685 (2011).
- Wagner, J.R., Brunzelle, J.S., Forest, K.T. & Vierstra, R.D. *Nature* **438**, 325–331 (2005).
- Stemmer, W.P. *Nature* **370**, 389–391 (1994).
- Szymczak, A.L. *et al. Nat. Biotechnol.* **22**, 589–594 (2004).
- Cui, L. *et al. Biochem. Biophys. Res. Commun.* **377**, 1156–1161 (2008).
- Han, C., Jan, L.Y. & Jan, Y.-N. *Proc. Natl. Acad. Sci. USA* **108**, 9673–9678 (2011).
- Shaner, N.C. *et al. Nat. Biotechnol.* **22**, 1567–1572 (2004).
- Shemiakina, I.I. *et al. Nat. Commun.* **3**, 1204–1207 (2012).
- To, T.-L. *et al. Proc. Natl. Acad. Sci. USA* **112**, 3338–3343 (2015).

ONLINE METHODS

General methods and materials. A BrBphP gene was ordered from GenScript and cloned into a modified pBAD vector containing the heme oxygenase-1-encoding gene from cyanobacteria. All synthetic DNA oligonucleotides were purchased from Integrated DNA Technologies. Restriction enzymes were purchased from New England BioLabs. Site-specific saturation mutagenesis was performed using a QuickChange Mutagenesis Kit (Stratagene). Random mutagenesis was performed using a GeneMorph II Random Mutagenesis Kit (Stratagene). Libraries were expressed in *Escherichia coli* strain TOP10 (Invitrogen) and screened by imaging the agar plates with colonies using a BioSpectrum Imaging System (UVP). The brightest clone in each library was picked as a template for the next round of random mutagenesis. Cultured cells were not tested for the presence of *Mycoplasma*, as such contamination would not impact the conclusions made on the basis of our imaging results. The sequences of all primers are provided in **Supplementary Table 2**.

Protein purification and characterization. mIFP was expressed with a C-terminal polyhistidine tag in a pBAD expression vector (Invitrogen). Proteins were purified with the Ni-NTA purification system (Qiagen). Protein concentration was measured by the Pierce BCA method. Two different approaches were used to determine extinction coefficients. The first one was based on a comparison of absorbance values for the protein at the main peak (683 nm) with the absorbance value at the 391-nm peak, assuming the latter to have the extinction coefficient of the free BV, which is $39,900 \text{ M}^{-1} \text{ cm}^{-1}$. The second one was based on direct measurement of the protein concentrations with a BCA protein assay kit (Pierce) followed by the calculation of extinction coefficient using a Beer-Lambert-Bouguer equation. For determination of quantum yield, mIFP solution was prepared with the same absorbance as a solution of Alexa Fluor 647 at wavelength 630 nm (quantum yield = 0.33 in phosphate-buffered saline (PBS)). The absorbance of both solutions is around or below 0.05. Fluorescence from 650 to 800 nm was collected and used to calculate the quantum yield. pH titrations were performed using a buffer series (100 mM sodium acetate, 300 mM NaCl for pH 2.5–5.0 and 100 mM NaH_2PO_4 , 300 mM NaCl for pH 4.5–9.0).

To study protein maturation, we grew TOP10 bacterial cells at 37 °C overnight in a LB medium supplemented with ampicillin. The next morning, the cells were centrifuged, resuspended and cultured in LB medium with 0.002% arabinose, 0.001 mM IPTG, 100 μM ALA and 50 μM FeCl_3 for 1 h. The cells were washed and cultured in LB medium supplemented with 0.001 mM IPTG, 100 μM ALA and 50 μM FeCl_3 (no arabinose) at 37 °C. Fluorescence intensity of the cell suspension was measured every hour.

Gel-filtration chromatography was performed using a Superdex-200 HR 10/30 FPLC gel-filtration column (Amersham Biosciences). The column was equilibrated with sterile PBS in a cold room. 100 μl of purified protein at a concentration of 0.5 mg ml^{-1} in PBS were loaded in each column. Elution was performed in PBS at a flow rate of 0.5 ml min^{-1} for 45 min. The column effluent was monitored by absorbance at 280 and 630 nm. The gel-filtration protein standards thyroglobulin, BSA, azurin and aprotinin were also loaded under the same conditions to calibrate the column. The linear calibration curve representing the

logarithm of molecular mass as a function of the fraction number was used to calculate the molecular mass of mIFP.

Analytical ultracentrifugation of mIFP, IFP1.4 and IFP2.0 was carried out by equilibrium sedimentation performed at 25 °C using a Beckman XL-I analytical ultracentrifuge. mIFP solutions were prepared at 34 μM and 17 μM , and IFP1.4 and IFP2.0 solutions were prepared at 17 μM and 8.5 μM , each in a buffer of 50 mM Tris, pH 7.5, 300 mM NaCl, 10 mM imidazole. Centrifugation was conducted at speeds of 25K, 30K, 35K and 40K r.p.m. with an An-60 Ti rotor (**Fig. 1d**), and the radial gradient profiles were acquired by absorbance scans at 280 nm. Data were globally fitted to equilibrium sedimentation models of a single-species and/or monomer-dimer mixtures by a nonlinear least-squares method using Igor Pro (WaveMetrics).

Fusion plasmid construction. The mIFP mammalian expression vectors were constructed from C1 or N1 cloning vectors (Clontech-style). The mIFP cDNA was PCR amplified with a 5' primer encoding an AgeI site and a 3' primer encoding either a BspEI (C1) or NotI (N1) site, in reference to mIFP. To prepare the additional 14-amino-acid linker (GGGSGGGSGGGSSG) used for the C-terminal mIFP fusions, we used two primers: mIFP12-AgeI-C-f and mIFP12-BspEI-C-r (**Supplementary Table 2**). The PCR products were gel purified, digested and ligated into EGFP-C1 or EGFP-N1 cloning vectors, respectively, resulting in mIFP C1 and N1 cloning vectors.

To construct the mIFP C-terminal fusions (number of linker amino acids in parenthesis), we performed the following digests: human β -actin (30), NheI and BglII (cDNA source: Clontech; NM_001101.3); CAF1 (22), AgeI and BspEI (mouse chromatin assembly factor; cDNA source: A. Gunjan, Florida State University; NM_013733.3); human light chain clathrin (27), NheI and BglII (cDNA source: G. Patterson, NIH; NM_001834.2); human endosomes (26), NheI and BspEI (human RhoB GTPase; cDNA source: Clontech; NM_004040.2); human fibrillarin (19), AgeI and BspEI (cDNA source: Evrogen; NM_001436.3); H2B (10), BglII and NheI (human histone 2B, cDNA source: G. Patterson, NIH; NM_021058.3); human lamin A/C (30), NheI and BglII (cDNA source: D. Gilbert, Florida State University; NM_170707.2); human lasp1 (22) NheI and BglII (cDNA source: OriGene; NM_006148.3); human myotilin (26), AgeI and BspEI (cDNA source: OriGene; NM_006790.2); human Rab4a (19), BglII and BamHI (cDNA source: V. Allen, University of Manchester; NM_004578.3); rat sEpsin (30) NheI and BglII (cDNA source: OriGene; NM_019585.3); human α -tubulin (30), NheI and BglII (cDNA source: Clontech; NM_006082); human vinculin (35) NheI and EcoRI (cDNA source: OriGene; NM_003373.3).

To prepare the mIFP N-terminal fusions (number of linker amino acids in parenthesis), we performed the following digests: human calnexin (14), AgeI and NotI (cDNA source: OriGene; NM_001746.3); human CENP-B (22), BamHI and NotI (A. Khodjakov, Wadsworth Center; NM_001810.5); Cx43 (7), BamHI and NotI (rat α -1 connexin 43 cDNA source: M. Falk, Lehigh University; NM_001004099.1); human EB3 (7), BglII and BamHI (cDNA source: L. Cassimeris, Lehigh University; NM_012326.2); H1 (10), BamHI and NotI (mouse histone 1, cDNA source: G. Patterson, NIH; NM_008197.3); H2B (6), BamHI and NotI (human histone 2B, cDNA source: G. Patterson,



NIH; NM_021058.3); human keratin 18 (17), EcoRI and NotI (cDNA source: Open Biosystems; NM_199187.1); rat lysosomal membrane glycoprotein 1 (20), BamHI and NotI (LAMP1; cDNA source: G. Patterson, NIH; NM_012857.1); Lifeact (7), BamHI and NotI (cDNA source: Integrated DNA Technologies); human MAPTau (10), AgeI and NotI (cDNA source: OriGene; NM_016841.4); human nucleoporin 50 kDa (10), BamHI and NotI (NUP50; cDNA source: OriGene; NM_007172.3); human peroxisomal membrane protein (10), NotI and AgeI (PMP; cDNA source: OriGene; NM_018663.1); human translocase outer mitochondria membrane 20 (10), (TOMM-20; cDNA source: OriGene; NM_014765.2); human zyxin (6), BamHI and NotI (cDNA source: OriGene; NM_003461.4).

DNA for transfection was prepared using the Plasmid Maxi kit (Qiagen).

In summary, the protein of interest can be fused to either the N or C terminus of mIFP. But it is necessary to note that in contrast to N-terminal mIFP fusions, C-terminal fusions often require an extended amino acid linker (see above). On the basis of the crystal structure of an mIFP homolog, the N terminus (first 17 amino acids) is unstructured and forms a flexible loop, whereas the C terminus is highly structured, forming an α -helix. Limited separation of mIFP and the protein of interest may cause steric hindrance. Inadequate C-terminal linker lengths result in erroneous localization and expression patterns characterized by faint membrane localization and decreased fluorescence. However, optimal linker length must still be determined experimentally.

Deposition of mIFP fusion plasmid. The mIFP fusions are available at Addgene. For each fusion, the N- or C-terminal fusion and the linker amino acid length are indicated after the name of the targeted proteins, followed by the Addgene plasmid number: mIFP-N1-54620; mIFP-Calnexin-N-14-56214; mIFP-Endo-14-56219; mIFP-Golgi-7-56221; mIFP-LaminA-C-18-56226; mIFP-Nup50-N-10-56233; mIFP-PMP-N-10-56238; mIFP-Tubulin-C-18-56240; mIFP-Actin-C-18-56211; mIFP-Cx26-7-56216; mIFP-H2B-6-56223; mIFP-Lysosomes-20-56228; mIFP-MapTau-N-10-56230; mIFP-Vimentin-7-56242; mIFP-C-Src-7-56213; mIFP-EB3-7-56218; mIFP-ER-5-56220; mIFP-Keratin-17-56225; mIFP-MyosinIIC-N-18-56232; mIFP-PMP-C-10-56237; mIFP-Annexin-12-56212; mIFP-CytERM-N-17-56217; mIFP-H2B-C-10-56224; mIFP-MANNII-N-10-56229; mIFP-Mito-7-56231; mIFP-PDHA1-N-10-56236; mIFP-Zyxin-6-56243; mIFP-alpha-actinin-19-56210; mIFP-CENPB-N-22-56215; mIFP-H1-10-56222; mIFP-LC-myosin-N-7-56227; mIFP-Paxillin-22-56234; mIFP-Tomm20-N-10-56239; mIFP-VASP-5-56241; mIFP12-C1-54819; mIFP12-Annexin-12-56245; mIFP12-LASPI-C-10-56257; mIFP12-Caveolin-C-10-56247; mIFP12-H2B-C-10-56254; mIFP12-MyosinIIa-C-18-56259; mIFP12-Rab4a-7-56261; mIFP12-Actin-C-18-56244; mIFP12-Clathrin-15-56249; mIFP12-LaminA-C-18-56256; mIFP12-CD81-10-56248; mIFP12-Endo-14-56250; mIFP12-ILK-C-14-56255; mIFP12-sEspin-C-18-56262; mIFP12-CAF1-C-10-56246; mIFP12-FilaminA-C-14-56253; mIFP12-myopalladin-C-14-56258; mIFP12-myotilin-C-14-56260; mIFP12-ZO1-C-14-56265; mIFP12-Fibrillarin-7-56252; mIFP12-Farnesyl-5-56251; mIFP24-C1-54820; mIFP24-Caf1-C-10-56266; mIFP24-LaminA-C-18-56267.

Characterization and imaging in mammalian cells. To image mIFP and its fusions in mammalian cells, we transfected HeLa or HEK293 cells with mIFP fusions using the calcium phosphate transfection method, maintained them with regular growth medium (without exogenous BV) and imaged them 48 h after transfection on a Nikon Eclipse Ti inverted microscope equipped with a Yokogawa CSU-W1 confocal scanner unit (Andor), a digital complementary metal-oxide semiconductor (CMOS) camera ORCA-Flash4.0 (Hamamatsu) and an ASI MS-2000 XYZ automated stage (Applied Scientific Instrumentation). Laser inputs were provided by an Integrated Laser Engine (Spectral Applied Research) equipped with laser lines of 405 nm, 488 nm, 561 nm and 640 nm (Coherent). The confocal scanning unit was equipped with the following emission filters: 460/50 nm, 525/50 nm, 610/60 nm, 661/20 nm, 732/60 nm and 731/137 nm. Bright-field and DIC imaging was provided by a Lambda TLED LED transmitted light source (Sutter Instrument). The system was also equipped with a SOLA light engine (Lumencor) for wide-field fluorescence imaging. Images in this study were obtained with the following objectives: Nikon Plan Apo λ 20 \times air (numerical aperture (NA), 0.75), Nikon Apo λ S LWD 40 \times water (NA, 1.15), and Nikon Apo TIRF 60 \times oil (NA, 1.49). Image acquisition was controlled by the NIS-Elements Ar Microscope Imaging Software (Nikon). The infrared fluorescence signal of mIFP was excited with the 640-nm laser and collected through the 731/137-nm emitter at 60 \times . All images were processed and analyzed with ImageJ.

Photobleaching of mIFP, iRFP, IFP2.0 or IFP1.4 in live HeLa cells were performed at 60 \times . Cells were excited by the 640-nm laser, and images were taken every 10 s for 30 min and every 1 min afterward.

For studying the dependence of brightness of IFP-expressing cells on BV concentration, LN229 cells stably expressing mIFP, IFP1.4, IFP2.0 or iRFP (each coexpressed with GFP under IRES), grown on 60-mm dishes, were treated with various concentrations of BV for 2 h and collected by centrifugation. The IFP fluorescence was measured using the infinite M1000 plate reader (normalized by coexpressed GFP fluorescence) and was plotted against the BV concentration. s.d. was calculated from 3 independent measurements. To investigate BV binding kinetics, we treated live LN229 cells with 25 μ M BV. The green and infrared fluorescence intensities were monitored using the infinite M1000 plate reader over time. The infrared fluorescence intensity (normalized by GFP) was plotted against time. s.d. was calculated from 3 independent measurements. To compare the brightness of IFPs in cells by flow cytometer, we transiently transfected live HeLa cells with IFP1.4 IRES EGFP, IFP2.0 IRES EGFP, mIFP IRES EGFP or iRFP IRES EGFP. 48 h after transfection, cells were washed and suspended in 0.5 ml PBS and analyzed on a flow cytometer (FACS Aria III). EGFP fluorescence was detected in the FITC-A channel (488-nm laser, 505- to 535-nm emission filter). IFP fluorescence was detected in the Alexa 700-A channel (640-nm laser, 708- to 753-nm emission filter). Data were processed using FlowJo software.

Structure illumination microscopy (SIM) was performed on a Zeiss Elyra PS.1 super-resolution imaging microscope. Fluorescence was excited using a 642-nm diode laser and 3–5 rotations of the grid pattern, filtered using a 655-nm long-pass filter, and collected using a Zeiss Plan-APOCHROMAT 100 \times (NA, 1.46) oil-immersion objective. All SIM image processing was performed using the proprietary ZEN 2012 Black software (Zeiss).

Fluorescence imaging in *Drosophila*. Expression of UAS-mIFP-histone 3.3 was driven by *engrailed-GAL4*. Fluorescence images were taken with the Nikon Eclipse Ti inverted microscope. Transcripts encoding tdTomato, FusionRed, mIFP and mIFP-T2A-HO1 were subcloned into pJFRC81 and phiC31-integrated into the genome at position 75A10 [PBac{y+-attP-9A} VK00005]. UAS::mCherry and UAS::EGFP [P{UAS-2xEGFP}AH2] were previously generated. Expression of these UAS::xFPs was driven ubiquitously [P{tubP-GAL4}LL7, in neurons [P{GMR57C10-GAL4}attP2] or in muscles [P{GAL4-Mef2.R}R1]. For coincident imaging of fluorophores, recombinant flies expressing mIFP-T2A-HO1 in neurons [P{GMR57C10-GAL4}attP2] or muscles [P{GAL4-Mef2.R}R1] were intercrossed to either w[1118], UAS::mCherry, EGFP [P{UAS-2xEGFP}AH2], UAS::tdTomato, UAS::FusionRed. For the triple-labeling experiment, recombinant flies expressing mIFP T2A HO1 in muscles [P{GAL4-Mef2.R}R1] were intercrossed with flies expressing CD4-tdTomato in class IV DA neurons [ppk::CD4-tdTomato] and a GFP fusion trap of the extracellular matrix protein Viking [vkg-PT{G00205}]. Intact first instar larvae were acutely immobilized in water during imaging using a small thermoelectric cooler (Laird) maintained at 4 °C by a PID Relay (Watlow). Condensation arising owing to the temperature differential was mitigated by wiping a thin film of water across the coverslip. Third instar larvae were immobilized in 300 mM sucrose after we removed the distal ends of the cuticle and viscera. Adult female mid-legs were acutely removed and imaged *ex vivo*.

The larval genotypes we used for comparing CD4-mIFP and CD4-IFP2.0 in epithelial cells were w; ppk-CD4-tdGFP/repo-tdTomato; UAS-CD4-mIFP-T2A-HO1/A58-Gal4 and w; ppk-CD4-tdGFP/repo-tdTomato; UAS-IFP2.0-T2A-HO1/A58-Gal4. Embryos were collected for 2 h on grape juice agar plates treated with yeast and were aged at 25 °C. At the appropriate time, a single larva was mounted in 90% glycerol under coverslips sealed with grease.

Animals were imaged using the Nikon Eclipse Ti inverted microscope. GFP was imaged with the 488-nm laser and 525/50-nm emission filter at 40×, and acquisition time was 100 ms. tdTomato, mCherry and FusionRed were imaged with the 561-nm laser and 610/60-nm emission filter at 40×, and acquisition time was 300 ms. IFP2.0 and mIFP were imaged with the 640-nm laser and 731/137-nm emission filter at 40×, and acquisition time was

500 ms. Confocal z series were acquired using the 20× or 40× objectives and the digital CMOS camera ORCA-Flash4.0 (Hamamatsu) on the Nikon Eclipse Ti inverted microscope controlled by NIS-Elements Ar Microscope Imaging Software (Nikon).

Viability assay in *Drosophila*. Fluorophores were ubiquitously expressed at 25 °C using [P{tubP-GAL4}LL7 or P{w[+mC] = Act5C-GAL4}25FO1], and the adult progeny were examined for eye morphogenesis defects and reduced viability (% of animals carrying the fluorophore transgene relative to a balancer).

Imaging in zebrafish. Adult zebrafish, both TL and AB wild-type strains, were maintained under standard laboratory conditions. Expression plasmids pCS2-mIFP-H2B and pCS2-HO1 were created by PCR amplification of mIFP-H2B and HO1 ORFs, respectively, and then cloned into pCS2+. Capped mRNA was synthesized using the mMACHINE SP6 kit (Ambion). 100 pg of mIFP-H2B mRNA with 100 pg of HO1 mRNA were injected at the one-cell stage. Fluorescence and bright-field images were taken at 30 h.p.f. with a charge-coupled device (CCD) camera (Andor; Clara DR-2199) on the Nikon Eclipse Ti microscope. Manually dechorionated embryos were embedded in 1.5% low-melt agarose within glass-bottom Petri dishes (MatTek Corporation). Eye-specific images were taken with a 40×/1.3-NA Plan-Fluor oil objective. mIFP fluorescence was imaged with a 638-nm laser line and a red-shifted Cy5.5 filter. z stacks of 2.5-μm intervals were acquired for each magnification.

Toxicity assay in zebrafish. At the one- to two-cell stage, 0, 100, 150 or 200 pg of either Myr-mIFP and HO1 or *gfp* (control) mRNA per embryo were injected. Unfertilized embryos were removed on day 0, and phenotypes of each group were scored alongside uninjected control embryos from the same clutch on day 1 after manual dechoriation. The experiment was performed under constant blue light conditions (465 nm). Each construct had at least $n = 300$ embryos. Embryos were scored as follows: normal to unaffected embryos were considered to have a wild-type phenotype; embryos with the presence of a slightly curved tail and/or mild edema were considered mildly deformed; and embryos with smaller heads, major curves or a kink in the tail and/or severe edema were considered severely deformed.