Published as: *Nat Methods*. 2015 August ; 12(8): 763–765.

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 Reade**[2], **Kalpana Makhijani**[1,2], **Yuanquan Song**[2,5,6], **Songmei Liu**[2], **Zehra Murthy**[2], **Shao-Qing Zhang**[1,2,7], **Orion D. Weiner**[2], **Thomas B. Kornberg**[2], **Yuh-Nung Jan**[2,5,6], **Michael W. Davidson**[3,4], and **Xiaokun Shu**[1,2]

[1]Department of Pharmaceutical Chemistry, University of California – San Francisco, San Francisco, California, USA

[2]Cardiovascular Research Institute, University of California – San Francisco, San Francisco, California, USA

[3]National High Magnetic Field Laboratory, The Florida State University, Tallahassee, Florida, USA

[4]Department of Biological Science, The Florida State University, Tallahassee, Florida, USA

[5]Department of Physiology, University of California – San Francisco, San Francisco, California, USA

[6]Howard Hughes Medical Institute, University of California – San Francisco, San Francisco, California, USA

[7]Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19014, USA

Abstract

Infrared fluorescent proteins (IFPs) provide an additional color to GFP and its red homologs in protein labeling. Based on structural analysis of the dimer interface, a monomeric bateriophytochrome is identified from a sequence database, and is engineered into a naturally-monomeric IFP (mIFP). We demonstrate that mIFP correctly labels proteins in live *Drosophila* and zebrafish requiring no exogenous cofactor, and will thus be useful in molecular, cell and developmental biology.

Correspondence should be addressed to X.S. (xiaokun.shu@ucsf.edu)..

AUTHOR CONTRIBUTIONS

X.S. conceived the project. D.Y., X.S. designed mIFP and the histone 2B fusion. K.M., X.S. planned the *Drosophila* embryo imaging. M.W.D. planned the fusion constructs. T.B.K. planned histone H3.3 fusion construct and the transgenic *Drosophila*. A.R., O.D.W. planned the imaging of zebrafish. M.K., Y.S. and Y.N.J. planned the imaging of epithelia, muscle and neurons in *Drosophila* larvae and adult. D.Y., M.A.B, J.R.A, E.S.H., M.K., A.R., K.M., Y.S., S.L., Z.M., S.-Q.Z. performed the experiments. D.Y. and X.S. wrote the manuscript. All the authors contributed to the final draft.

Accession codes. Genebank: KM285236

COMPETING FINANCIAL INTERESTS

None declared.

GFP and its red homologs are powerful tools for cell and molecular biology[1,2]. Recently, this FP palette has been extended into infrared region by introduction of bacteriophytochrome (BphP)-derived IFPs (e.g. IFP1.4, iRFP) that autocatalytically incorporate biliverdin (BV) as the chromophore[3-7]. IFPs can potentially provide an additional color to GFP and its red homologs in protein labeling. As a protein fusion tag, an FP has to be monomeric so as not to perturb 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 (ΔBphP denoting truncations including only the PAS and GAF domains). Although IFP1.4 is engineered to be monomeric (iRFP is dimeric), here we found that IFP1.4 and IFP2.0 (a latest IFP1.4 mutant[11]) tended to dimerize at high concentration with dissociation constant 7.8 and 3.7 μM, 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 ΔBphP with little-to-no cytotoxicity based upon biological fitness[12]. Protein sequence databases such as NCBI (National Center for Biotechnology Information) contain many BphP sequences. We hypothesized that some might be monomeric for ΔBphP, most likely without strong hydrophobic interactions at dimer interface. In IFP1.4's parent *Dr*BphP[13], the dimer interface includes several residues (**Fig. 1a**): Leu311 appears to play a critical role as single mutation L311K disrupts the dimer interface[3]. Analysis of ~40 BphP sequences from NCBI revealed *Br*BphP (*Bradyrhizobium*) as a potential candidate, since the residue corresponding to Leu311 in *Dr*BphP is a polar threonine (**Fig. 1b**). Size exclusion chromatography (SEC) indicated that Δ*Br*BphP eluted later than dimeric Δ*Dr*BphP and at a time similar to the monomeric form of IFP1.4 (**Supplementary Fig. 2**), suggesting that Δ*Br*BphP is a monomer.

We next engineered the non-fluorescent Δ*Br*BphP into a fluorescent mutant. In brief, we selected several residues (Asp199, Tyr168, Val178, Asn258) surrounding BV for saturation mutagenesis, which was followed by DNA shuffling[14] and random mutagenesis. The final fluorescent mutant mIFP absorbs maximally at 683 nm (**Supplementary Fig. 3**), with excitation and emission maxima of 683 and 704 nm, respectively (Fig. 1c); quantum yield 8%; extinction coefficient 82,000 $M^{-1}cm^{-1}$ (**Supplementary Table 1**).

mIFP was confirmed to be monomeric at high concentrations (17 and 34 μM) (**Fig. 1d** and **Supplementary Fig.4**). It contained nineteen mutations (**Supplementary Figs.5** and **6**), including five mutations 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 did not overlap and thus might be targeted for further engineering (**Supplementary Fig. 7**). mIFP was stable in pH 4-10 (**Fig. 1f**). The molecular brightness of mIFP was similar to that of IFP1.4 and iRFP (**Supplementary Table 1**), and cellular brightness similar to iRFP and 10-fold greater than IFP1.4 in live HeLa cells (**Fig. 1g,h** and **Supplementary Fig. 8**). mIFP was 6.3-time more photostable than IFP1.4 (**Supplementary Fig. 9**), but ~5-time less photostable than iRFP in HEK293 cells (**Supplementary Table 1**). Photobleaching of mIFP was irreversible (**Supplementary**

Fig. 10), suggesting no residual photo-isomerization. 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, ~30 mIFP fusion proteins were constructed to the N and C termini with an appropriate-length linker (Online Methods). These fusions localized properly in live cells (**Supplementary Figs. 14–16**, **Supplementary Videos 1–6**), including those that require a high degree of monomeric character, such as α-tubulin, connexin 43 and intermediate filaments. All phases of mitosis were observed in fusions of human histone H1 and H2B. Additionally, mIFP was compatible with structured illumination microscopy (**Supplementary Fig. 17**).

To demonstrate mIFP as a protein tag in live animals, including *Drosophila* and zebrafish, we created histone fusions and imaged them *in vivo*. For *Drosophila*, we created a UAS-mIFP-histone 3.3 (H3.3) T2A heme oxygenase-1 (HO1) transgenic line and subsequently crossed with *engrailed*-GAL4 line to promote expression of mIFP-H3.3 in a segmental pattern. Here T2A is a “self cleaving” peptide widely used in co-expression of multiple genes[15]. HO1 converts heme to BV[16]. We co-expressed HO1 to overcome limitation of endogenous BV in *Drosophila* (see below). Confocal imaging of the embryo detected bright nuclear fluorescence with expected segmental pattern (**Fig. 2a,b**). For zebrafish, we expressed mIFP-H2B with HO1 by mRNA injection during the one-cell stage. Imaging of the eye region at 30 hpf (hours post fertilization) revealed bright nuclear fluorescence with proper expression pattern (**Fig. 2c**).

To demonstrate mIFP in multicolor protein and cellular labeling *in vivo*, we co-expressed mIFP-H3.3 T2A HO1 and CD8-GFP in neurons of *Drosophila*. CD8 is a transmembrane protein that labels cell membrane. Examination in the brain region of the embryo indicated infrared fluorescence in the nucleus and green fluorescence on the cell membrane, suggesting correct targeting of the fusion proteins (**Fig. 2d–g**). We then tested feasibility of mIFP in 3-color imaging. We expressed mIFP T2A HO1 in the abdomen muscle of *Drosophila* larvae, together with a GFP fusion trap of the extracellular matrix protein Viking (collagen), and CD4 (a transmembrane protein that labels cell membrane)-tdTomato in Class IV DA neurons (ppk::CD4-tdTomato). Fluorescence imaging revealed nice separation of 3 fluorophores’ signal (**Fig. 2h–j**) and expected structural organization of abdomen muscle, extracellular matrix and sensory neurons (**Fig. 2k**). Previously CD4 fused GFP was shown to label neuronal processes most efficiently[17], here expression of CD4-mIFP T2A HO1 in class IV DA neurons in *Drosophila* larvae clearly and evenly labeled dendrites and axons (**Fig. 2l, m**).

To demonstrate advantage of the naturally-monomeric mIFP, we first compared it to the engineered-monomeric IFP2.0 derived from a dimeric parent. Here we expressed CD4-mIFP T2A HO1 and CD4-IFP2.0 T2A HO1 in epithelial cells of *Drosophila* larvae. While mIFP-CD4 labeled epithelial cells’ membrane correctly, IFP2.0-CD4 formed aggregates and failed to label the plasma membrane (**Supplementary Fig. 18a,b**). This is consistent with the in vitro data that IFP2.0 tended to dimerize at high concentrations.

We then compared mIFP to other engineered-monomeric FPs derived from oligomeric parents, including the popular red FP mCherry, orange FP tdTomato[18], and FusionRed that was recently engineered to be a "pure" monomer, overcoming dimerization tendency of its parent mKate2 at high concentration[19]. We found that mCherry formed punctate structures in the muscles and neurons in the first instar larvae of *Drosophila*, whereas mIFP was homogenous in both tissues (**Supplementary Fig. 18c,d**). We examined mCherry fluorescence in the body muscles of more than 200 animals and all of them showed punctate structures. Expression of FusionRed in the leg muscle of adult *Drosophila* revealed many rounded structures, varying in size from 0.5 to 1.3 μm (**Supplementary Fig. 18e**). TdTomato also formed punctate structures with elongated shape varying in length from 2 to 6 μm with width ~0.4 μm (**Supplementary Fig. 18f**). In contrast, mIFP expression was homogenous in the leg muscle when co-expressed with either FusionRed or tdTomato (**Supplementary Fig. 18e,f**). We also imaged GFP expression in the muscles and neurons of *Drosophila* larvae and observed homogenous fluorescence (**Supplementary Fig. 19**), similar to mIFP.

To examine potential toxicity of mIFP and HO1 in animals, we conducted viability assay in *Drosophila* and did not find obvious difference for mIFP or mIFP T2A HO1 compared to GFP (**Supplementary Fig. 20a**). Furthermore, we did not observe any defects in the eye morphogenesis of *Drosophila* expressing either mIFP or mIFP T2A HO1, suggesting no obvious toxicity. On the other hand, co-expression of HO1 improved mIFP fluorescence by 30–40 folds in the *Drosophila* muscle (**Supplementary Fig. 20b,c**). Furthermore, *Drosophila* embryos expressing mIFP-H3.3 T2A HO1 and CD8-GFP displayed normal ventral nerve cord shortening, without obvious difference compared to embryos expressing CD8-GFP (**Supplementary Videos 7,8**). Thus, mIFP-histone fusion with HO1 does not seem to have observable effect on animal development, consistent with the above viability assay. We also conducted toxicity assay of zebrafish expressing myristoylated mIFP (myr_mIFP) and HO1 by RNA injection, with comparison to GFP (**Supplementary Fig. 20d**). Our data did not reveal any obvious toxicity.

To demonstrate that co-expression of HO1 in zebrafish not only improved mIFP fluorescence but also other BphP-derived IFPs, we imaged zebrafish expressing myr_iRFP without HO1, which revealed little iRFP fluorescence (**Supplementary Fig. 21a,b**). Co-expression of HO1 substantially increased iRFP fluorescence, which was obvious on cell membrane in the tail muscle (**Supplementary Fig. 21c,d**). Zebrafish expressing myr_mIFP and HO1 showed similar level of infrared fluorescence in the tail muscle (**Supplementary Fig. 21e,f**), consistent with the in vitro data that mIFP and iRFP have comparable BV binding kinetics and affinity.

In summary, we have engineered a naturally-monomeric mIFP and demonstrated that mIFP performs excellently in protein labeling *in vivo*. mIFP is also a good template in developing infrared fluorescent reporters, such as for visualizing cell signaling *in vivo*. For example, based on mIFP we have designed an infrared fluorogenic protease reporter that visualizes apoptotic pathways *in vivo*[20]. Because protein sequence databases contain thousands of BphPs, rational design of future IFPs and related tools with desirable photophysical and photochemical properties is promising.

Online Methods

General methods and materials

A BrBphP gene was ordered from GeneScript and cloned into a modified pBAD vector containing the heme oxygenase-1 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 E. 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 M^{-1} 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 compared to a solution of Alexa Fluor 647 at wavelength 630 nm (quantum yield = 0.33 in 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 NaH2PO4, 300 mM NaCl for pH 4.5–9.0).

To study protein maturation, TOP10 bacterial cells were grown 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 FeCl3 for 1 h. The cells were washed and cultured in LB medium supplemented with 0.001 mM IPTG, 100 μM ALA and 50 μM FeCl3 (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 phosphate buffered saline (PBS) in a cold room. 100 ul of purified protein at a concentration of 0.5mg/ml in PBS was loaded in each column. Elution was performed in PBS, at a flow rate of 0.5ml/min 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 were 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; IFP1.4 and IFP2.0 solutions at 17 μM and 8.5 μM; all in the buffer 50 mM Tris pH 7.5, 300 mM NaCl, 10 mM Imidazole. Centrifugation was conducted at speeds 25K, 30 K, 35K and 40K r.p.m. 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 monomerdimer mixtures by nonlinear least-squares method using IGOR Pro (Wavemetrics).

Fusion plasmid construction

The monomeric infrared fluorescent protein (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**) utilized for the C-terminal mIFP fusions, 2 primers were used: 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), the following digests were performed: human β-actin (30), NheI and BglII (cDNA source: Clontech, Mountain View, CA, USA; NM_001101.3); CAF1 (22), AgeI and BspEI (mouse chromatin assembly factor; cDNA source: A. Gunjan, Florida State University, Tallahassee, FL, USA; NM_013733.3); human light chain clathrin (27), NheI and BglII (cDNA source: G. Patterson, National Institutes of Health, Bethesda, MD, USA; 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, Moscow, Russia; 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, Manchester, UK; 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), the following digests were performed: human calnexin (14), AgeI and NotI (cDNA source: Origene; NM_001746.3); human CENP-B (22), BamHI and NotI (A. Khodjakov, Wadsworth Center, Albany, NY, USA; NM_001810.5); Cx43 (7), BamHI and NotI (rat α-1

connexin 43 cDNA source: M. Falk, Lehigh University, Bethlehem, PA, USA; 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 keratin18 (17), EcoRI and NotI (cDNA source: Open Biosystems, Huntsville, AL, USA; 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, Coralville, IA, USA); 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 N or C termini 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). Based upon the crystal structure of a mIFP homolog, the N-terminus (first 17 amino acids) is unstructured and forms a flexible loop, while the C-terminus is highly structured, forming an alpha 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.

Characterization and imaging in mammalian cells

To image mIFP and its fusions in mammalian cells, HeLa or HEK293 cells were transfected with mIFP fusions using calcium phosphate transfection method, maintained under regular growth media (without exogenous BV), and imaged 48 hrs after transfection on a Nikon Eclipse Ti inverted microscope equipped with a Yokogawa CSU-W1 confocal scanner unit (Andor), a digital CMOS camera ORCA-Flash4.0 (Hamamatsu) and a 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 Instructment). The system was also equipped with a SOLA light engine (Lumencor) for wildfield fluorescence imaging. Images in this study were obtained with the following objectives: Nikon Plan Apo λ 20X air (N.A. 0.75), Nikon Apo λS LWD 40X water (N.A. 1.15), and Nikon Apo TIRF 60X oil (N.A. 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 60X. All images were processed and analyzed with ImageJ.

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

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

Structure illumination microscopy (SIM) was performed on a Zeiss Elyra PS.1 superresolution 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× (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. Fluorescent 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 1st 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 due to the temperature differential was mitigated by wiping a thin film of water across the coverslip. Third instar larvae were immobilized in 300mM Sucrose after removing the distal ends of the cuticle and viscera. Adult female mid legs were acutely removed and imaged ex vivo.

The larvae genotype we used for comparing CD4-mIFP and CD4-IFP2.0 in epithelial cells were w; ppk-CD4-tdGFP/repo-tdTomato; UAS-CD4-mIFP-T2A-H01/A58-Gal4 and w; ppk-CD4-tdGFP/repo-tdTomato; UAS-IFP2.0-T2A-H01/A58-Gal4. Embryos were collected for 2 hours on yeasted grape juice agar plates and were aged at 25°C. At the appropriate time, a single larvae 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 40X, and acquisition time 100 ms. TdTomato, mCherry and FusionRed were imaged with the 561 nm laser and 610/60-nm emission filter at 40X, and acquisition time 300 ms. IFP2.0 and mIFP were imaged with the 640 nm laser and 731/137-nm emission filter at 40X, and acquisition time 500 ms. Confocal z-series were acquired using the 20X or 40X 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 25C 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 ORF's, respectively, and then cloned into pCS2+. Capped messenger RNA was synthesized using the mMESSAGE mMACHINE SP6 kit (Ambion). 100 pg of mIFP-H2B mRNA with 100 pg of HO1 mRNA were injected at the one-cell stage. Fluorescent and brightfield images were taken at 30 h.p.f. with a charge-coupled device 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 638nm laser line and a redshifted 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, 0pg, 100 pg, 150 pg, 200 pg of Myr-mIFP and HO1, or 0pg, 100pg, 150pg, 200pg, 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 dechorionation. 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; embryos with smaller heads, major curves or a kink in the tail and/or severe edema were considered severely deformed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by Program for Breakthrough Biomedical Research (to X.S.), 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 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, S. Woo for providing pCS2+ vector, S. Roy for providing UAS-mCherry *Drosophila* line.

REFERENCES

1. Day RN, Davidson MW. Chem Soc Rev. 2009; 38:2887–2921. [PubMed: 19771335]
2. Tsien RY. Angew Chem Int Ed Engl. 2009; 48:5612–5626. [PubMed: 19565590]
3. Shu X, et al. Science. 2009; 324:804–807. [PubMed: 19423828]
4. Filonov GS, et al. Nat Biotechnol. 2011; 29:757–761. [PubMed: 21765402]
5. Auldridge ME, Satyshur KA, Anstrom DM, Forest KT. J Biol Chem. 2012; 287:7000–7009. [PubMed: 22210774]
6. Shcherbakova DM, Verkhusha VV. Nat Methods. 2013; 10:751–754. [PubMed: 23770755]
7. Piatkevich KD, Subach FV, Verkhusha VV. Nature Communications. 2013; 4:2153–2162.
8. Giraud E, Verméglio A. Photosyn Res. 2008; 97:141–153. [PubMed: 18612842]
9. Rockwell N, Su Y, Lagarias J. Annu. Rev. Plant Biol. 2006; 57:837–858. [PubMed: 16669784]
10. Karniol B, Wagner JR, Walker JM, Vierstra RD. Biochem J. 2005; 392:103–116. [PubMed: 16004604]
11. Yu D, et al. Nature Communications. 2014; 5:3626–3632.
12. Geiler-Samerotte KA, et al. Proc Natl Acad Sci USA. 2011; 108:680–685. [PubMed: 21187411]
13. Wagner JR, Brunzelle JS, Forest KT, Vierstra RD. Nature. 2005; 438:325–331. [PubMed: 16292304]
14. Stemmer WP. Nature. 1994; 370:389–391. [PubMed: 8047147]
15. Szymczak AL, et al. Nat Biotechnol. 2004; 22:589–594. [PubMed: 15064769]
16. Cui L, et al. Biochem Biophys Res Commun. 2008; 377:1156–1161. [PubMed: 18983822]
17. Han C, Jan LY, Jan Y-N. Proceedings of the National Academy of Sciences. 2011; 108:9673–9678.
18. Shaner NC, et al. Nat Biotechnol. 2004; 22:1567–1572. [PubMed: 15558047]
19. Shemiakina II, et al. Nature Communications. 2012; 3:1204–7.
20. To T-L, et al. Proceedings of the National Academy of Sciences. 2015; 112:3338–3343.

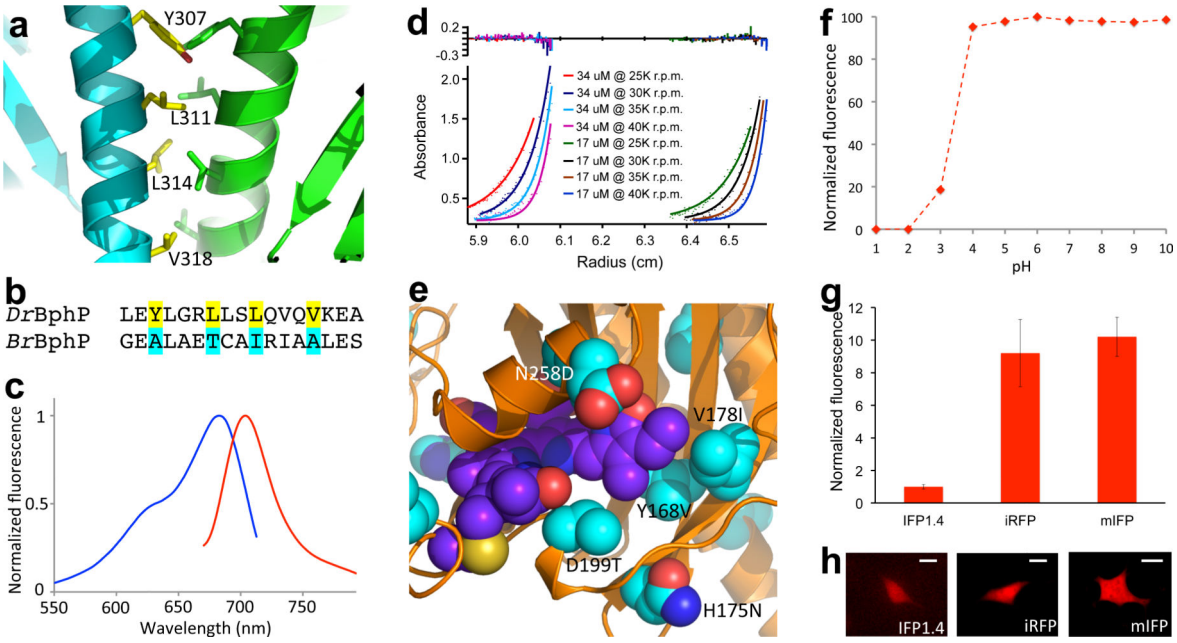

Figure 1. Rational design of a naturally-monomeric infrared fluorescent protein

(**a**) Residues potentially involved in the dimer interface of *Dr*BphP (PDB: 2O9B). (**b**) Sequence alignment of *Dr*BphP and *Br*BphP. Residues in the core of dimer interface of *Dr*BphP are colored in yellow, and the corresponding residues in *Br*BphP in cyan. (**c**) Excitation (blue) and emission (red) spectra of mIFP. (**d**) Analytical ultracentrifugation of mIFP at high concentrations (17 and 34 μM). Data were globally fitted to a function describing sedimentary equilibrium curves. Single-species fitting agrees well with the data over the entire range and yielded the molecular weight of mIFP 33.05 ± 0.98 kD, which is close to the expected molecular weight of a monomer (35.69 kD). (**e**) A structure model of mIFP with introduced mutations surrounding the chromophore BV (in purple). (**f**) Fluorescence of purified mIFP against pH. (**g**) Comparison of cellular brightness among IFP1.4, iRFP, and mIFP in live HeLa cells (infrared fluorescence normalized by co-expressed GFP under IRES). The standard deviation was calculated based on 20 cells for each IFP. (**h**) Representative fluorescence images of IFP1.4 (brightened 10 times), iRFP and mIFP in HeLa cells. 20 cells were examined for each IFP. Scale bar: 10 μm.

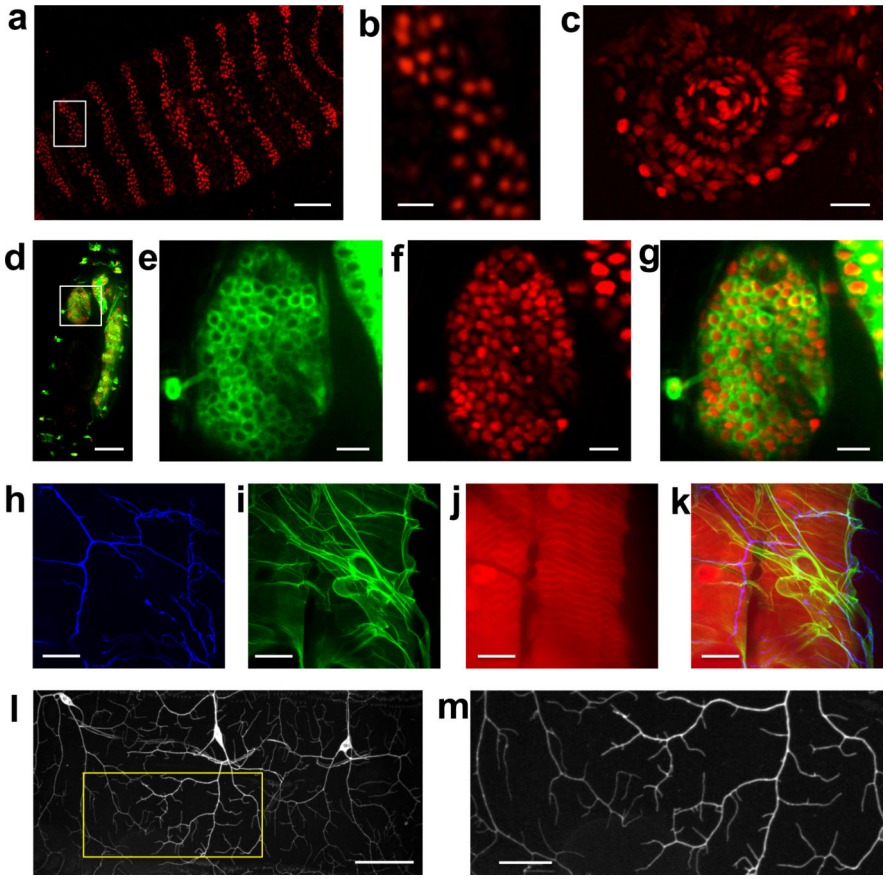

Figure 2. Expression of mIFP, histone or membrane fusions in model organisms

(**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 zebrafish eye expressing mIFP-H2B and HO1. (**d-g**) Two-color fluorescence imaging of *Drosophila* embryo expressing UAS-mIFP-histone 3.3 T2A HO1 and UAS-CD8-GFP, driven by *elav*-GAL4: (**d**) Whole embryo; (**e-f**) High magnification view of the boxed area in **d**. (**e**) GFP; (**f**) mIFP; (**g**) merged. (**h-k**) Three-color fluorescence imaging of *Drosophila* abdomen muscle expressing UAS-mIFP T2A HO1 driven by *Mef2R*-GAL4, Class IV DA neurons expressing CD4-tdTomato (ppk::CD4-tdTomato), extracellular matrix expressing GFP-viking (collagen): (**h**) tdTomato (in blue pseudo-color); (**i**) GFP; (**j**) mIFP; (**k**) merged. (**l,m**) Class IV DA neurons expressing CD4-mIFP in *Drosophila* larvae: (**m**) High magnification view of the boxed area in **l**. Scale bar, 50 μm (**a**); 10 μm (**b**); 20 μm (**c**); 50 μm (**d**); 10 μm (**e-f**); 20 μm (**h-k**); 50 μm (**l**); 20 μm (**m**).