

PROTEIN DESIGN

De novo design of orthogonal far-red, orange, and green fluorophore-binding proteins for multiplexed imaging

Long Tran^{1,2,†}, Steffen Klein^{3,†}, David Juergens^{2,4,5,6,†}, Shajesh Sharma^{2,7,†}, Justin Decarreau^{2,†}, Gyu Rie Lee^{2,4,8}, Yujia Wang^{2,4}, Wei Chen^{2,4}, Asim K. Bera², Alex Kang², Jon Woods², Emily Joyce², Dionne K. Vafeados², Nicole Roullier², Xinting Li², Bingxu Liu^{2,4}, Yang Bo^{2,4}, Edin Muratspahic^{2,4}, Tim A. Brown⁹, Jonathan B. Grimm⁹, Ronak Patel⁹, Luke D. Lavis⁹, Julia Mahamid^{3,10*}, Linna An^{2,4*}, David Baker^{2,4,11*}

Fluorescent proteins and small-molecule dyes offer complementary advantages for biological imaging: Proteins are amenable to genetic tagging, whereas dyes provide superior brightness and photostability. To combine these strengths, we used de novo protein design to generate small, nanomolar-affinity, high-selectivity binders (NovoTags) for three cell-permeable dyes spanning the visible spectrum. We show that the NovoTag fluorescent lifetimes can be tuned and demonstrate their application in lifetime- and wavelength-based multiplexed fluorescence imaging. We also designed a two-chain version (NovoSplit) that functions as a chemically induced dimerization system with fluorescent readout in living cells or as a minimally perturbing proximity probe in fixed cells. Our approach combines the advantages of fluorescent proteins and small-molecule dyes, thus expanding the toolkit for cellular imaging.

Biological imaging has been transformed by the ability to tag specific cellular components with genetically encoded fluorescent proteins. But fluorescent proteins have limited brightness and photostability, constraining the resolution, duration, and complexity of imaging experiments (1). By contrast, small-molecule dyes offer superior optical properties (2–4). Advanced synthetic fluorophores such as the Janelia Fluor (JF) dyes are notable for their quantum yields across a range of emission wavelengths, photostabilities, and cell permeabilities (2, 3). However, the use of JF dyes in cellular imaging requires a mechanism to target the proteins of interest. Current approaches primarily use fusion tags, including HaloTag (5) and SNAP-tag (6), which covalently bind modified versions of the dyes. Although effective, these tags are relatively large (20 to 35 kDa) and not dye specific; they bind a general ligand moiety that can be conjugated to a compatible dye (5, 6). This has complicated the development of orthogonal versions, thereby limiting multiplexing capabilities (7).

We reasoned that designing a set of small proteins, each selectively binding to a specific JF dye across the visible spectrum, could provide

a general solution to the multiplexed imaging problem. This, however, poses a challenging design task because JF dyes are structurally analogous rhodamine derivatives that lack known natural binders (8). To address this challenge, we leveraged recent advances in machine learning-based de novo protein design (9, 10) to generate genetically encodable, small JF dye-binding proteins with high affinity and specificity for single dyes and to explore the use of such designs for multiplexed fluorescence imaging.

Results

Design and evaluation of fluorophore-binding proteins

We selected three JF dyes, JF₄₉₄, JF₅₉₆, and JF₆₅₇ (Fig. 1, A to C), with well-separated excitation and emission spectra spanning the visible range (2) as design targets (Fig. 1D and fig. S1). These fluorophores share a common rhodamine-derived core structure but differ in specific chemical modifications that enhance brightness and tune their excitation and emission maxima (2). We reasoned that a design strategy that generates proteins with extensive complementarity to an input ligand structure should achieve specificity, as even small changes in the chemistry could introduce clashes or eliminate favorable contacts (9, 11).

We developed a machine learning-based approach for generating binders with internally repeating closed structures (pseudocycles) that encircle the target JF dyes. We used the generative deep learning method α RFdiffusion (10) (fig. S2) with symmetry conditioning to build protein structures around the JF₄₉₄, JF₅₉₆, and JF₆₅₇ dyes. We encoded the ligand structure using the α RFdiffusion motif template input and omitted symmetry operations in any ligand-ligand or ligand-protein blocks in the pair representation (see the materials and methods). We found that this method effectively generated monomeric pseudocyclic structures around the three selected dyes (Fig. 1, A to C). This RFdiffusion-based method has an advantage over our previous hallucination-based approach (9) because it generates pseudocyclic proteins directly around a target ligand, thus avoiding a second ligand-docking step. After obtaining the structures, we generated sequences using LigandMPNN with FastRelax cycling (12). Each design comprised a single-domain protein of 110 to 160 amino acids with a predicted cavity tailored to accommodate the targeted JF dye. We filtered the designs with Rosetta (13) for ligand binding and with AlphaFold 2 (AF2) (14) for protein sequence-structure consistency. A total of 4843, 5800, and 6032 binder variants were selected for JF₄₉₄, JF₅₉₆, and JF₆₅₇ binding, respectively. These sequences were synthesized on oligonucleotide microarrays, expressed in yeast for surface display, and screened by fluorescence-activated cell sorting (FACS) (fig. S3). Next-generation sequencing of enriched pools revealed multiple high-affinity binders: 56 designs for JF₄₉₄, 236 designs for JF₅₉₆, and 334 designs for JF₆₅₇ exhibited binding affinities of 5 μ M or lower on yeast.

For each dye, we selected a single binder design based on the FACS enrichment and single-clone flow cytometry signal intensity, for subsequent characterization (fig. S3). These were expressed in *Escherichia coli*, purified by affinity chromatography, and their binding affinities measured using fluorescence polarization (FP) assays. The selected designs for each target dye (NovoTag₄₉₄, NovoTag₅₉₆, and NovoTag₆₅₇ for JF₄₉₄, JF₅₉₆, and JF₆₅₇, respectively) have molecular weights of 15.1, 13.3, and 15.6 kDa and binding affinities (K_d) of 19.1, 1.5, and 2.0 nM, respectively (Fig. 1, A to C). We used stopped-flow fluorescence polarization to determine the binding kinetics of JF₄₉₄ to NovoTag₄₉₄, yielding a k_{on} of $1.8 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ and a k_{off} of 0.00336 s^{-1} (fig. S4), consistent with the measured K_d of 19.1 nM (Fig. 1A). Next, we assessed the specificity of the NovoTags to their JF targets using FP and found that they bound their respective targets with at least 1000-fold greater affinity than the alternative ligands (fig. S5). We further characterized the absorbance maximum (λ_{abs}), emission maximum (λ_{em}), extinction coefficient (ϵ), fluorescence quantum yield (Φ_f), and fluorescence lifetime (τ_f) of the free dyes and their conjugates (Fig. 1E). For both NovoTag₆₅₇ and NovoTag₅₉₆, we observed modest shifts in λ_{abs} and λ_{em} and larger ϵ ($\sim 15\%$) and Φ_f ($\sim 40\%$) values upon

¹Department of Chemical Engineering, University of Washington, Seattle, WA, USA. ²Institute for Protein Design, University of Washington, Seattle, WA, USA. ³Molecular Systems Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany. ⁴Department of Biochemistry, University of Washington, Seattle, WA, USA. ⁵Graduate Program in Molecular Engineering, University of Washington, Seattle, WA, USA. ⁶Department of Chemistry, Stanford University, Stanford, CA, USA. ⁷Department of Bioengineering, University of Washington, Seattle, WA, USA. ⁸Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. ⁹Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA. ¹⁰Cell Biology and Biophysics Unit, EMBL, Heidelberg, Germany. ¹¹Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. *Corresponding author. Email: julia.mahamid@embl.de (J.M.); la72@rice.edu (L.A.); dabaker@uw.edu (D.B.) †These authors contributed equally to this work. ‡These authors contributed equally to this work.

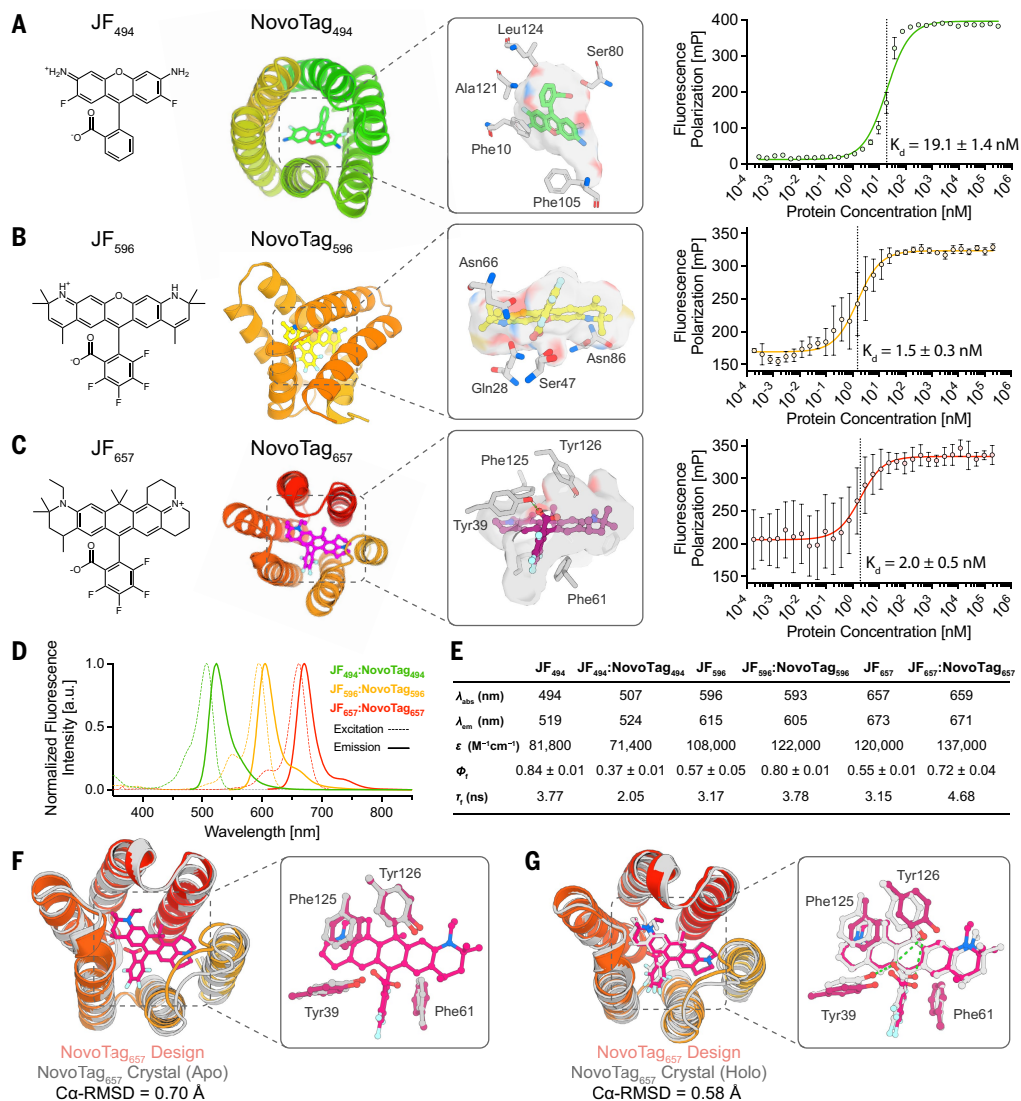


Fig. 1. Design and characterization of JF-binding proteins. (A to C) NovoTag designs for the binding of JF₄₉₄ (A), JF₅₉₆ (B), and JF₆₅₇ (C). For each JF dye–NovoTag pair, the chemical structure of the target dye (left), the model of the NovoTag design (middle) with the relevant dye-interacting residues highlighted (insets), and FP titration curves (right) are shown. The FP titrations were repeated three to eight times from independent serial dilutions, and data are presented as means ± SDs over all titrations. K_d values and SEs were determined from a standard binding isotherm model by nonlinear regression. (D) Excitation (dotted lines) and emission (solid lines) of JF dye–NovoTag pairs: JF₄₉₄:NovoTag₄₉₄ (green), JF₅₉₆:NovoTag₅₉₆ (orange), and JF₆₅₇:NovoTag₆₅₇ (red). (E) Photophysical characterization of free JF dyes and JF dye–NovoTag conjugates. (F and G) Superposition of the design model of NovoTag₆₅₇ (red) with the apo (F) and holo (G) crystal structure (gray). Green dashed lines indicate hydrogen bonds.

protein binding, resulting in a substantial increase in fluorescence emission intensity of the protein–dye complexes compared with the free dyes. By contrast, NovoTag₄₉₄ showed a decrease in both ϵ (~15%) and Φ_f (~50%). The measured τ_f values matched the trends in Φ_f with NovoTag₆₅₇ and NovoTag₅₉₆ increasing the fluorescence lifetime and NovoTag₄₉₄ decreasing it compared with the free dyes. Compared with the state-of-the-art synthetic fluorophore JF₆₄₆–HaloTag ligand bound to HaloTag7 ($\epsilon = 152,000 \text{ M}^{-1} \text{ cm}^{-1}$, $\Phi_f = 0.54$) (3) or the JF₆₅₇–HaloTag ligand bound to HaloTag7 ($\epsilon = 128,000 \text{ M}^{-1} \text{ cm}^{-1}$, $\Phi_f = 0.54$), JF₆₅₇ bound to NovoTag₆₅₇ exhibited 20 and 43% higher brightness, respectively ($\epsilon = 137,000 \text{ M}^{-1} \text{ cm}^{-1}$, $\Phi_f = 0.72$).

We characterized the biophysical properties of the NovoTag designs using circular dichroism (CD) and size-exclusion chromatography (SEC). CD spectra revealed predominantly α -helical secondary structures, consistent with our computational models, and thermal denaturation experiments indicated high thermostability (fig. S6). SEC traces showed that NovoTags are soluble and monomeric in solution.

We determined the crystal structures of NovoTag₆₅₇ in the apo and holo states at 1.44 and 2.37 Å resolution, respectively, which closely matched the design model over the protein backbone (Cα-RMSD–design–Apo = 0.70 Å, Cα-RMSD–design–Holo = 0.58 Å; Fig. 1, F and G, and tables S1 and S2). The aromatic side chains of Phe⁶¹ and Phe¹²⁵, designed to form hydrophobic contacts and π - π stacking interactions with the dye, as well as Tyr³⁹ and Tyr¹²⁶, which form hydrogen bonds with JF₆₅₇, were correctly positioned. All-by-all AlphaFold 3 (AF3) (15) prediction comparisons of NovoTags in complex with JF dyes yielded the highest-confidence predictions [highest interface-predicted template modeling (ipTM) and highest ligand predicted local distance difference test (pLDDT)] for the designed on-target pairs (fig. S7). These structures and our in silico results show how the high protein–small-molecule shape complementarity generated by our design methods can lead to specificity even among small-molecule analogs with high structure similarity (fig. S5).

NovoTags enable multiplexed fluorescence microscopy in cells

We next evaluated each NovoTag as a genetically encoded fluorescent tag for microscopy. We fused the NovoTags either to mScarlet (NovoTag₄₉₄) or enhanced green fluorescent protein (eGFP) (NovoTag₅₉₆ and NovoTag₆₅₇) and targeted each of the constructs to the outer mitochondrial membrane using a mitochondrial localization signal derived from Tom70 (16) (MitoTag) in HeLa cells. Cells were incubated with the corresponding JF dye (5 nM for 15 min). After extensive washing, live-cell fluorescence imaging showed that the JF dye signal colocalized with the mScarlet or eGFP signal on the mitochondrial membrane (fig. S8), demonstrating that all three NovoTags effectively bind and recruit the cell-permeable JF dyes to specific subcellular sites. The nanomolar affinities (Fig. 1, A to C) and slow k_{off} rate (fig. S4) of the NovoTags enabled live-cell imaging after washing, thereby minimizing nonspecific background, highlighting the benefits of high-affinity binders. To assess the impact of chemical fixation on NovoTag staining, MitoTag-NovoTag₆₅₇-eGFP was expressed in HeLa cells, and JF dye staining was performed either before or after fixation (fig. S9). In both cases, the JF₆₅₇ signal colocalized with eGFP, demonstrating that the NovoTag system is compatible with this common imaging workflow. Comparing the photostability of JF₆₅₇ bound to NovoTag₆₅₇ with that of the well-established, commercially available JFX₆₅₀-HaloTag ligand bound to HaloTag7 (17) showed similar rates of photobleaching in HeLa cells,

with 50% signal loss after 14 or 15 bleaching iterations, respectively (Fig. 2C, fig. S10, and movie S1).

We evaluated the in-cell specificity of the NovoTags by testing each combination of NovoTag (NovoTag₄₉₄, NovoTag₅₉₆, and NovoTag₆₅₇), JF dye (JF₄₉₄, JF₅₉₆, and JF₆₅₇), and excitation wavelength (473, 575, and 631 nm) in HeLa cells and quantifying the fluorescence intensities (Fig. 2D and fig. S11). Only the specific combination of a matching NovoTag, JF dye, and excitation wavelength yielded a substantial fluorescent signal, demonstrating high specificity in cells.

To explore the use of the NovoTags for multiplexed fluorescence imaging, we simultaneously targeted NovoTag₄₉₄ to the early endosomal membrane (2×FYVE tag) (18), NovoTag₅₉₆ to the inner nuclear membrane (emerin) (19, 20), and NovoTag₆₅₇ to the mitochondrial membrane (MitoTag) in HeLa cells. Cells were incubated with a dye cocktail consisting of 5 nM JF₄₉₄, 20 nM JF₅₉₆, and 5 nM JF₆₅₇ for 15 min. After three washes, live imaging revealed clear signal separation (fig. S12), demonstrating orthogonal and specific dye binding in cells and the multiplexing capability of this system.

Given the reported high photostability of JF dyes (21, 22) and the observed high in-cell specificity of our binders, we evaluated the NovoTags for multiplexed super-resolution fluorescence microscopy using stimulated emission depletion (STED) imaging on fixed (Fig. 2A

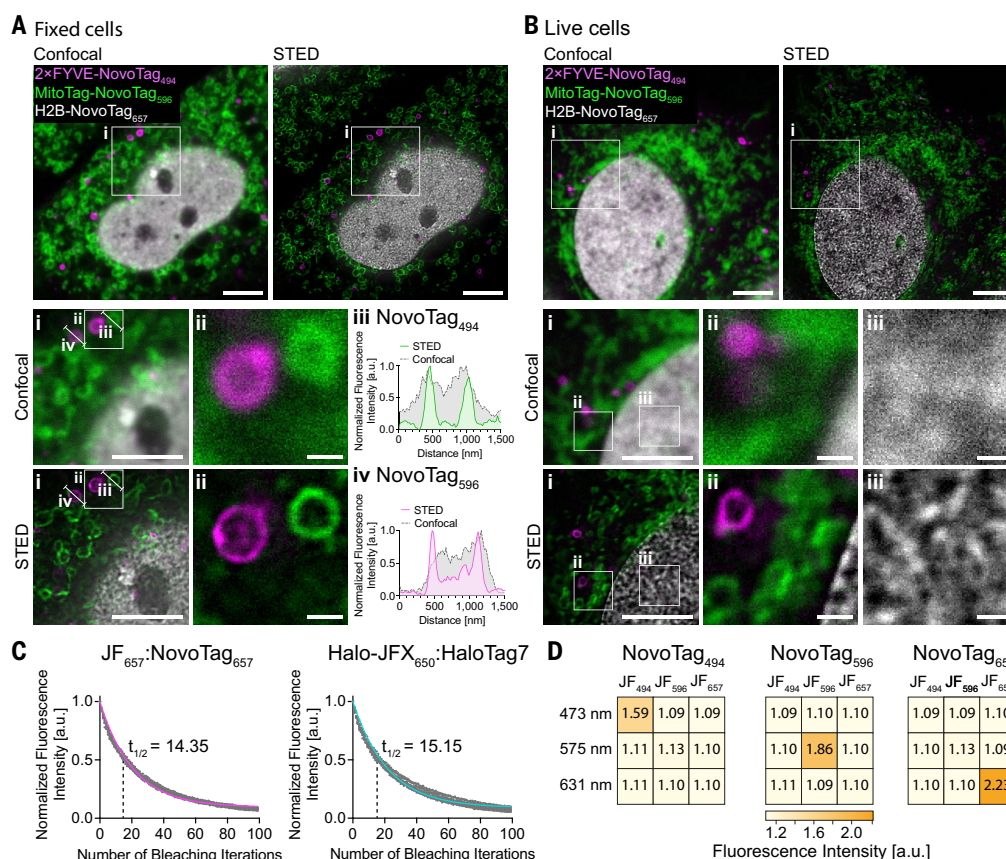


Fig. 2. Multiplexed super-resolution fluorescence microscopy in live and fixed cells using NovoTags. (A and B) Multiplexed fluorescence microscopy in fixed (A) and live (B) HeLa cells. NovoTag₄₉₄ labels endosomes (2×FYVE, magenta), NovoTag₅₉₆ labels mitochondria (MitoTag, green), and NovoTag₆₅₇ labels chromatin (H2B, white). Cells were stained with 50 nM concentrations of each dye (JF₄₉₄, JF₅₉₆, and JF₆₅₇) for 30 min at 37°C, followed by three washing steps. Top: confocal and STED images (presented side by side) were acquired on a Leica Stellaris 8 STED Falcon microscope. Bottom: magnified views of indicated areas and line plots for NovoTag₄₉₄ (iii) and NovoTag₅₉₆ (iv) for confocal and STED. (C) Photostability of JF dyes. HeLa cells expressing either NovoTag₆₅₇ or HaloTag7 localized to mitochondria (MitoTag) were fluorescently labeled with 50 nM of each dye (JF₆₅₇ and Halo-JFX₆₅₀) as above. Fluorescence intensity was measured after each bleaching iteration. For each sample, photobleaching of 10 individual cells was acquired in two independent experiments. All data points are shown (gray). Nonlinear regression (one-phase decay) of the fluorescence signal was fitted. (D) Specificity of the NovoTags. Each NovoTag (NovoTag₄₉₄, NovoTag₅₉₆, and NovoTag₆₅₇) was localized to mitochondria (MitoTag) in HeLa cells, labeled with one of each JF dye (JF₄₉₄, JF₅₉₆, or JF₆₅₇), and excited with one of three wavelengths (474, 575, or 631 nm). Fluorescence intensity measurements were obtained from 12 fields of view for each condition in three independent experiments. For each combination, the mean fluorescence intensity is plotted. Scale bars in (A), (B), and magnified views (i), 5 μm; and in magnified views (ii) and (iii), 500 nm. a.u., arbitrary units.

and fig. S13) and live (Fig. 2B) HeLa cells. NovoTag₄₉₄ was targeted to the early endosomal membrane (2×FYVE tag), NovoTag₅₉₆ to the outer mitochondrial membrane (MitoTag), and NovoTag₆₅₇ to chromatin (H2B) (23). Cells were incubated with a cocktail of all three JF dyes (50 nM each for 30 min). After extensive washing, STED imaging revealed clear signal separation (fig. S13) and enhanced resolution compared with confocal imaging (Fig. 2, A and B). The combination of the three described NovoTags thus enables multiplexed cellular imaging in live, fixed, and super-resolution modalities.

Tuning NovoTag₄₉₄ for multiplexed fluorescence lifetime imaging microscopy

Our observation that the NovoTags modulate the fluorescence lifetimes of the JF dyes upon binding (Fig. 1E) suggests that the photophysical properties can be fine-tuned by modifying the design of the binding

pockets. As a proof of principle, we sought to design binders for JF₄₉₄ with distinct, well-separated fluorescence lifetimes to enable multiplexed fluorescence lifetime imaging microscopy (FLIM) (24). We characterized the dye binding and fluorescence intensity of 36 redesigned variants of the NovoTag₄₉₄-binding pocket, incorporating more hydrogen bonds, fewer π - π stacking interactions, and an altered electrostatic environment (Fig. 3A), and selected two designs: NovoTag₄₉₄^S (short lifetime) and NovoTag₄₉₄^L (long lifetime) (fig. S14 and Fig. 3, B to D). The excitation and emission profiles were similar to those of the original design; NovoTag₄₉₄^L showed a slight blue shift (Fig. 3, B and C), NovoTag₄₉₄^S exhibited slightly reduced brightness (Fig. 3D), and both displayed a modest decrease in binding affinity (Fig. 3E). FLIM measurements in HeLa cells showed a mean fluorescence lifetime of 2.00 ns (SD = 0.09) for NovoTag₄₉₄, 1.27 ns (SD = 0.25) for NovoTag₄₉₄^S, and 2.78 ns (SD = 0.40) for NovoTag₄₉₄^L (Fig. 3F). These lifetime distributions were

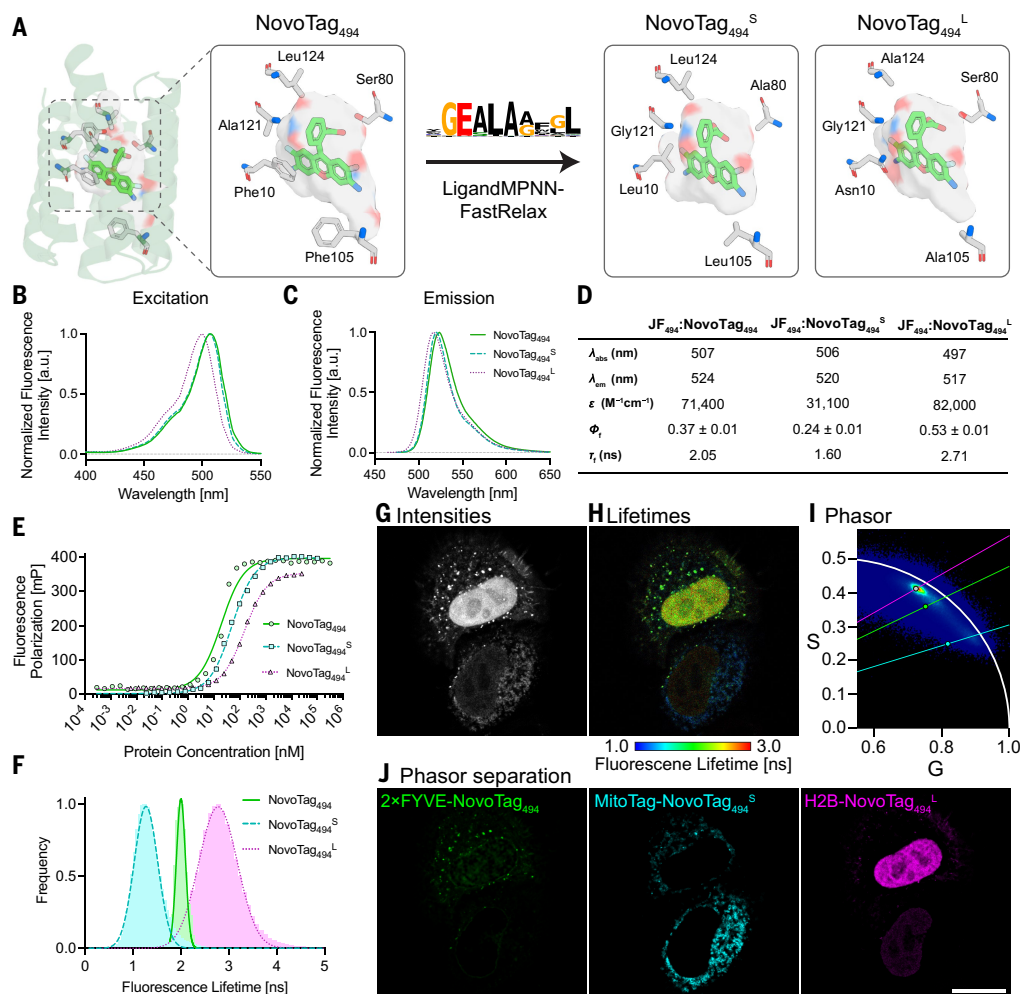


Fig. 3. Fluorescence lifetime-based multiplexed microscopy using NovoTags. (A) The NovoTag₄₉₄-binding site was subjected to sequence redesign using LigandMPNN-FastRelax, and two variants with different hydrogen-bonding networks, π - π stacking profiles, and electrostatics of the binding pocket were selected with shorter or longer fluorescence lifetimes (NovoTag₄₉₄^S and NovoTag₄₉₄^L). The sequence logo represents the relative frequency (letter height) of each amino acid in the redesigned residues. (B and C) Excitation (B) and emission (C) of JF₄₉₄ dye bound to NovoTag₄₉₄ and the two lifetime variants. (D) Photophysical characterization of the JF₄₉₄-NovoTag₄₉₄ variant conjugates. (E) Fluorescence polarization titration of NovoTag₄₉₄ and of the two lifetime variants. A nonlinear regression (standard binding isotherm model) was fitted and the K_d values were determined. NovoTag₄₉₄^S and NovoTag₄₉₄^L showed affinities to JF₄₉₄ of 49.8 and 150.8 nM, respectively. For each sample, three or four independent measurements were acquired. (F) Distribution of fluorescence lifetimes of each NovoTag₄₉₄ variant in cells. HeLa cells were transfected with one of the following plasmids: MitoTag-NovoTag₄₉₄^S, 2×FYVE-NovoTag₄₉₄, or H2B-NovoTag₄₉₄^L, chemically fixed and imaged in the presence of 10 nM JF₄₉₄. For each sample, fluorescence images of 11 to 16 cellular regions were acquired. The histograms of fluorescence lifetimes are plotted and fitted with a Gaussian distribution. (G to J) HeLa cells were cotransfected with MitoTag-NovoTag₄₉₄^S, 2×FYVE-NovoTag₄₉₄, and H2B-NovoTag₄₉₄^L, chemically fixed and imaged in the presence of 10 nM JF₄₉₄ with a confocal microscope (Leica Stellaris 8 Falcon). Shown are the total fluorescence intensity (G), the fluorescence lifetimes (H), and the phasor plot (I). Using phasor-based lifetime unmixing, the signals of the three labeled components were separated (J). Scale bar, 20 μ m.

well separated, with Cohen's d values of 4.5 (NovoTag₄₉₄^S:NovoTag₄₉₄^L), 3.9 (NovoTag₄₉₄^S:NovoTag₄₉₄), and 2.7 (NovoTag₄₉₄^L:NovoTag₄₉₄) (Fig. 3F).

We evaluated multiplexed FLIM with these designs (Fig. 3, G to J) by cotransfecting HeLa cells with MitoTag-NovoTag₄₉₄^S, 2×FYVE-NovoTag₄₉₄, and H2B-NovoTag₄₉₄^L, followed by fixation and staining with 10 nM JF₆₅₇. Using phasor-based fluorescence lifetime unmixing, the signals of the three labeled cellular components were separated. Signals showed specific signatures of mitochondria (MitoTag-NovoTag₄₉₄^S), endosomes (2×FYVE-NovoTag₄₉₄), and chromatin (H2B-NovoTag₄₉₄^L), with minimal cross-talk between lifetimes, demonstrating successful fine-tuning of our NovoTags for multiplexed FLIM-based cellular imaging.

Design of a covalent NovoTag

Covalent dye binding is of particular interest for applications such as pulse-chase experiments and single-molecule tracking (25–27). We thus investigated the design of dye-specific covalent NovoTags. Fluorinated dyes such as JF₆₅₇ can undergo nucleophilic aromatic substitution (S_NAr) on the fluorinated pendant phenyl ring (28). We surmised that placing a cysteine residue in NovoTag₆₅₇ within close proximity to its electrophilic moiety would facilitate the formation of a covalent thioether bond between the dye and the protein (fig. S16A).

We expressed and purified 47 such designs, incubated them with an excess of JF₆₅₇ (10 μM for 12 hours), and investigated the formation of a protein-dye covalent adduct using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). Twelve of the 47 designs showed covalent binding (figs. S15 and S16B), and we named the design with the highest fluorescence intensity in the SDS-PAGE gel NovoTag₆₅₇^{CV}. Liquid chromatography–mass spectrometry (LC-MS; fig. S16C) analysis revealed peaks corresponding to the labeled NovoTag₆₅₇^{CV}, confirming the formation of the covalent bond with the dye. LC-MS time-course analysis revealed 47.8% (SD = 1.6%) labeling after 3 hours of incubation, suggesting a half-time ($t_{1/2}$) of ~3 hours for S_NAr conjugation kinetics (under excess JF₆₅₇ conditions; fig. S16D). Incubation of live *E. coli* or HeLa cells expressing NovoTag₆₅₇^{CV} with JF₆₅₇ for up to 18 hours, followed by SDS-PAGE analysis of whole-cell lysates (fig. S16, E and F), showed that the covalent bond forms in cells. Live fluorescence imaging of the HeLa cells confirmed colocalization between MitoTag-NovoTag₆₅₇^{CV}-eGFP and JF₆₅₇ (fig. S17), demonstrating the feasibility of covalent labeling with the NovoTags for cellular imaging.

NovoSplit enables chemically induced dimerization

Chemically induced dimerization (CID) systems have been widely used to promote protein-protein interactions for applications ranging from signal transduction and gene expression to protein localization and degradation (29–32). We reasoned that designed split variants of the NovoTags could function as a CID system in which the fluorescent JF dye functions to induce dimerization within live cells, allowing for precise control of protein interactions and downstream processes, while simultaneously providing a direct fluorescent readout of the induced interactions.

We designed split versions of NovoTag₆₅₇, and 47 designs predicted to assemble in the presence of the dye were evaluated in a mammalian two-hybrid screen for minimal self-dimerization in the absence of the dye and inducible dimerization upon addition of JF₆₅₇ (fig. S18). The five designs exhibiting the highest dye-induced reporter activation (see the materials and methods) were further evaluated by fluorescence colocalization analysis in HeLa cells (fig. S19). The best of these split designs (fig. S19B) was derived from the top noncovalent NovoTag₆₅₇ design (Fig. 1C). To reduce residual self-dimerization in the absence of JF₆₅₇ (fig. S19B), we removed a salt bridge between Glu¹⁴ and Arg⁵³ (fig. S20A) by introducing an Arg⁵³→Ala substitution; this eliminated the noninduced dimerization, whereas induction of dimerization by JF₆₅₇ was not impaired (fig. S20, B and C). We named this split system

NovoSplit₆₅₇. The AF3 prediction of NovoSplit₆₅₇ aligned closely with the holo crystal structure of NovoTag₆₅₇ (Cα-RMSD = 0.38 Å) (Fig. 4B).

We evaluated NovoSplit₆₅₇ as a CID system in HeLa cells. One split fragment, NovoSplit₆₅₇^A, was targeted to the outer mitochondrial membrane using a MitoTag and additionally fluorescently labeled with mStayGold (33). The second split fragment, NovoSplit₆₅₇^B, was fluorescently labeled with mCherry (34) and expressed as a cytosolic protein (Fig. 4C). Incubation with JF₆₅₇ (200 nM for 1 hour) induced translocation of mCherry-NovoSplit₆₅₇^B from the cytosol to the mitochondria, with a Pearson's correlation coefficient (r) between mStayGold and mCherry of 0.74 (SD = 0.15). The JF₆₅₇ fluorescent signal colocalized with both mStayGold-NovoSplit₆₅₇^A and mCherry-NovoSplit₆₅₇^B (r = 0.78, SD = 0.06 and r = 0.88, SD = 0.05, respectively), demonstrating that the small molecule both induces dimerization and provides a fluorescent readout of the newly formed interaction (Fig. 4, D and E). Titration with JF₆₅₇ showed that concentrations as low as 50 nM induced dimerization, with higher dye concentrations yielding faster kinetics (fig. S21). Live-cell fluorescence imaging with 1 μM JF₆₅₇ showed a $t_{1/2}$ of 5.8 min (95% confidence interval: 5.0 to 6.7 min) (fig. S22 and movie S2). Evaluation of the dye-binding specificity of NovoSplit₆₅₇ showed that only JF₆₅₇ induced dimerization, whereas JF₄₉₄ and JF₅₉₆ did not (fig. S23).

NovoSplit enables native protein-protein proximity biosensing

We hypothesized that NovoSplit₆₅₇ could be additionally used to probe native protein-protein interactions, similar to bimolecular fluorescence complementation systems (35) such as split-GFP, if the two components of NovoSplit₆₅₇ were fused to two proteins of interest and the JF₆₅₇ dye was only introduced after cell fixation. We reasoned that if the proteins of interest were in close proximity within a cell, then small rearrangements after fixation could occur to enable JF₆₅₇ dye-facilitated dimerization of the split system, thereby providing a direct fluorescent readout of protein proximity. Conversely, if the two proteins of interest were far apart, then fixation would prevent the large-scale rearrangements required for assembly of the two halves of the split system and no fluorescence would be observed.

To evaluate protein-protein proximity biosensing using the NovoSplit system, we used a de novo–designed protein pair, LHD(A) and LHD(B), which forms stable heterodimers (36). LHD(A) was fused to NovoSplit₆₅₇^A-mScarlet and targeted to the outer mitochondrial membrane using a MitoTag. LHD(B) was fused to NovoSplit₆₅₇^B-mNeonGreen and expressed as a cytosolic protein. Because LHD(A) and LHD(B) spontaneously form heterodimers, we expected that both constructs would colocalize, bringing NovoSplit₆₅₇^A and NovoSplit₆₅₇^B into proximity (Fig. 4F). After chemical fixation and incubation with JF₆₅₇ (5 nM for 30 min), fluorescence microscopy showed colocalization of mScarlet and mNeonGreen (r = 0.93, SD = 0.06), indicating the formation of the LHD(A)–LHD(B) heterodimer, and a clear JF₆₅₇ fluorescence signal that colocalized with both fluorescent proteins (r = 0.74, SD = 0.09 and r = 0.79, SD = 0.08, respectively) (Fig. 4, G and J), demonstrating that the NovoSplit₆₅₇^A-NovoSplit₆₅₇^B-JF₆₅₇ ternary complex can assemble after fixation. As a control, we repeated the same experiment, expressing NovoSplit₆₅₇^B fused to mNeonGreen but omitting LHD(B) (Fig. 4H). Without the LHD(A)–LHD(B) pair to drive association, the NovoSplit₆₅₇^B-mNeonGreen was distributed throughout the cytosol, and the two constructs showed minimal colocalization (r = 0.31, SD = 0.12) and no localized JF₆₅₇ fluorescence (Fig. 4, I and J). These data confirm our conjecture that the addition of dye after fixation only leads to the reconstitution of the fluorescent ternary complex if the two halves of the split system are already in close proximity. Thus, by adding the dye after fixation, the NovoSplit system can be used to detect interactions between two proteins of interest without introducing spurious associations. The fluorescent signal provides a snapshot of the extent of association at the time of fixation rather than the cumulative amount of complex formed, as with some irreversible split fluorescent protein systems (37).

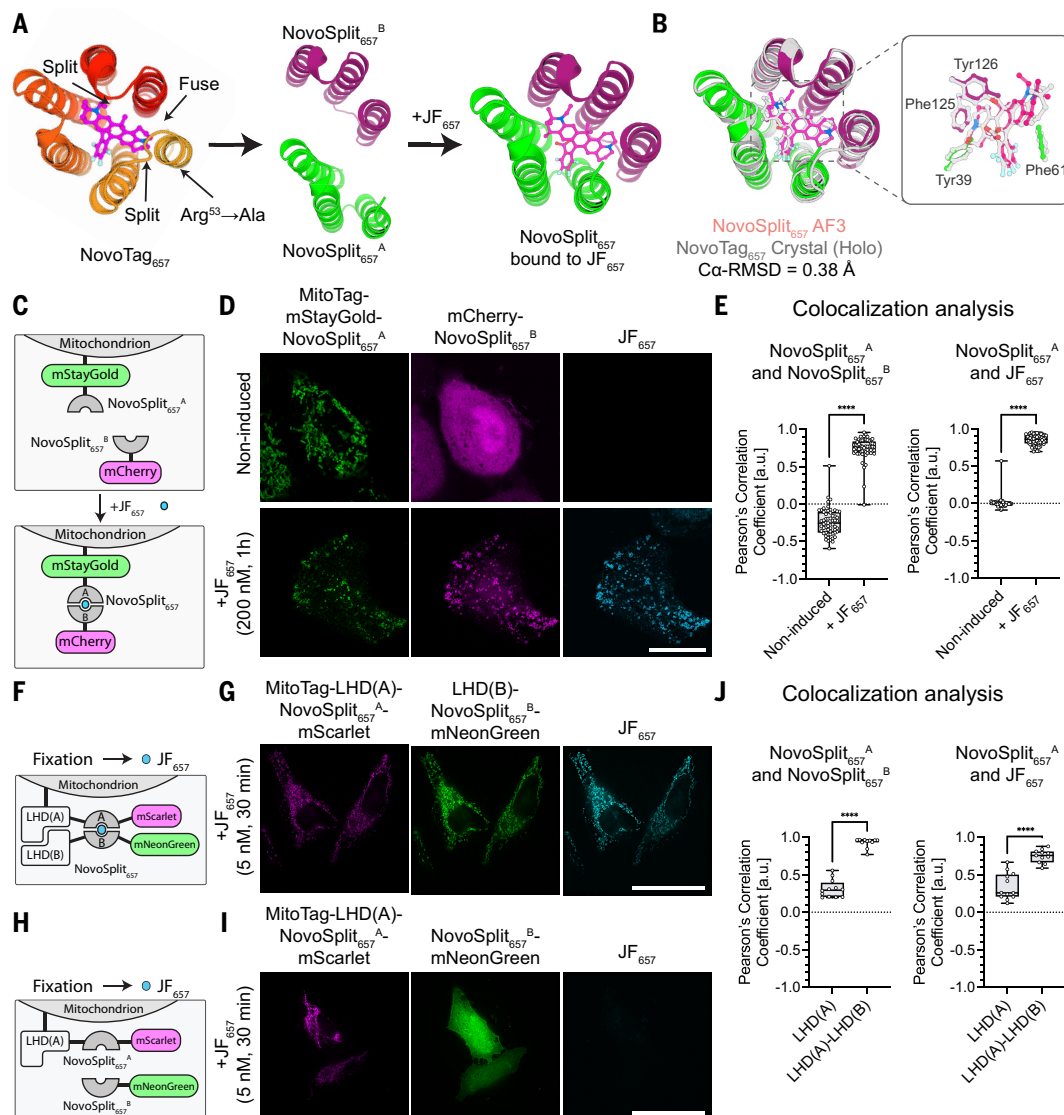


Fig. 4. Chemically induced dimerization and proximity biosensing using NovoSplit₆₅₇. (A) To generate NovoSplit₆₅₇, NovoTag₆₅₇ was split at two sites, and the original C and N termini were fused with a short linker. The fluorescent dye JF₆₅₇ induces dimerization of the two monomers NovoSplit₆₅₇^A and NovoSplit₆₅₇^B. To prevent noninduced dimerization, an Arg⁵³→Ala substitution was introduced into NovoSplit₆₅₇^A. (B) The NovoSplit₆₅₇ AF3 prediction (in color) superimposed with the crystal structure of NovoTag₆₅₇ (gray). (C to E) NovoSplit₆₅₇ as a CID system: HeLa cells were transfected with MitoTag-mStayGold-**NovoSplit₆₅₇^A** and mCherry-**NovoSplit₆₅₇^B** (C). Cells were incubated with 200 nM JF₆₅₇ for 1 hour, washed three times, chemically fixed, and imaged with a confocal microscope (Zeiss LSM 980 AiryScan) (D). Colocalization was analyzed by calculating the Pearson's correlation coefficient between NovoSplit₆₅₇^A and NovoSplit₆₅₇^B, as well as between NovoSplit₆₅₇^A and JF₆₅₇ (E). For each sample, 59 noninduced and 61 JF₆₅₇-induced cells were analyzed. (F to J) NovoSplit₆₅₇ as a proximity biosensor: HeLa cells were transfected with MitoTag-LHD(A)-**NovoSplit₆₅₇^A**-mScarlet and either LHD(B)-**NovoSplit₆₅₇^B**-mNeonGreen (F) or NovoSplit₆₅₇^B-mNeonGreen [as a negative control (H)]. Cells were chemically fixed and fluorescently labeled with 5 nM JF₆₅₇ for 30 min, followed by three washing steps. Fluorescence images were acquired on an OMX SR microscope [(G) and (I)]. For the experiment in (G), colocalization was analyzed by calculating the Pearson's correlation coefficient between NovoSplit₆₅₇^A and NovoSplit₆₅₇^B, as well as between NovoSplit₆₅₇^A and JF₆₅₇ (J). For each sample, 11 [LHD(A)-LHD(B)] and 12 [LHD(A)] cells were analyzed. All data are shown as box and whisker plots indicating the median (center of box), 25 and 75% quartiles (bounds of box), and minimum and maximum values (bars). Each data point represents one cell. Unpaired two-sided Welch's *t* tests were conducted. Statistical significance: *****P* < 0.0001. Scale bars: (D), 20 μm; (G) and (I), 40 μm.

Discussion

NovoTags combine the superior photophysical properties of small-molecule fluorophores with the genetic programmability of proteins. Because the designed binding sites are dye specific, NovoTags enable the direct and simultaneous use of multiple bright, photostable, non-modified JF dyes in the same cell for multiplex imaging. By contrast, the shared conjugation chemistry of HaloTag (5) or SNAP-tag (6) only allows use of one dye with each tag in a single experiment. In addition, the smaller size of NovoTags should reduce steric hindrance and potential perturbation to the tagged proteins. The tunability of

the NovoTag-binding sites further expands opportunities for advanced microscopy and approaches to modulate or measure protein interactions in cells: The covalent NovoTag₆₅₇^{CV} offers advantages in applications requiring permanent labeling, including pulse-chase experiments, single-molecule tracking (25–27), and in vivo imaging; NovoTags that modulate fluorescent lifetimes extend imaging beyond spectral unmixing, enabling multiplexed FLIM that is especially useful for live-cell imaging because reducing the number of spectrally distinct channels reduces photobleaching; NovoSplit converts dye binding into conditional protein dimerization with high

specificity and a direct fluorescent readout. With the fluorophore introduced after fixation, NovoSplit functions as a very sensitive (due to the brightness of the JF dyes) and low background native interaction proximity sensor.

The design strategy described here should be extendable to a wider range of synthetic fluorophores, including near-infrared dyes and fluorophores with specialized properties such as photoactivation or blinking (21, 38, 39). Given that advanced spectral unmixing can achieve spectral resolution for fluorescent dyes with emission maxima separated by 25 nm (40, 41), it should be possible to develop 10 or more resolvable NovoTag-dye combinations spanning the entire spectrum. The additional capability of separation along the orthogonal lifetime dimension, demonstrated here with the three lifetime-resolvable NovoTag₄₉₄ variants, could expand such a panel to $3 \times 10 = 30$ or more simultaneously resolvable probes. The ability to simultaneously resolve and track many different tagged proteins in the same cell or tissue could enable considerable advances in molecular cell biology. Together with recent related studies (42–44), our work outlines a general framework for integrating synthetic chemistry with protein engineering to advance biological imaging and related fields.

REFERENCES AND NOTES

- M. Hirano *et al.*, *Nat. Biotechnol.* **40**, 1132–1142 (2022).
- J. B. Grimm *et al.*, *J. Am. Chem. Soc.* **145**, 23000–23013 (2023).
- J. B. Grimm, L. D. Lavis, *Org. Lett.* **13**, 6354–6357 (2011).
- K. Yan *et al.*, *Nat. Commun.* **15**, 2593 (2024).
- G. V. Los *et al.*, *ACS Chem. Biol.* **3**, 373–382 (2008).
- A. Keppler *et al.*, *Nat. Biotechnol.* **21**, 86–89 (2003).
- M. Holtmannspötter *et al.*, *Angew. Chem. Int. Ed.* **62**, e202219050 (2023).
- E. A. Specht, E. Braselmann, A. E. Palmer, *Annu. Rev. Physiol.* **79**, 93–117 (2017).
- L. An *et al.*, *Science* **385**, 276–282 (2024).
- A. Lauko *et al.*, *Science* **388**, eadu2454 (2025).
- L. Lu *et al.*, *Science* **384**, 106–112 (2024).
- J. Dauparas *et al.*, *Nat. Methods* **22**, 717–723 (2025).
- A. Leaver-Fay *et al.*, in *Methods in Enzymology* (Elsevier, 2011), vol. 487, pp. 545–574; <https://linkinghub.elsevier.com/retrieve/pii/B9780123812704000196>.
- J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
- J. Abramson *et al.*, *Nature* **630**, 493–500 (2024).
- D. G. Millar, G. C. Shore, *J. Biol. Chem.* **268**, 18403–18406 (1993).
- J. B. Grimm *et al.*, *JACS Au* **1**, 690–696 (2021).
- D. J. Gillooly *et al.*, *EMBO J.* **19**, 4577–4588 (2000).
- A. Nagano *et al.*, *Nat. Genet.* **22**, 254–259 (1996).
- S. Manilal, T. M. Nguyen, C. A. Sewry, G. E. Morris, *Hum. Mol. Genet.* **5**, 801–808 (1996).
- J. B. Grimm *et al.*, *Nat. Methods* **17**, 815–821 (2020).
- J. B. Grimm *et al.*, *Nat. Methods* **14**, 987–994 (2017).
- T. Kanda, K. F. Sullivan, G. M. Wahl, *Curr. Biol.* **8**, 377–385 (1998).
- L. El Hajji *et al.*, *Adv. Sci. (Weinh.)* **11**, e2404354 (2024).
- A. Gautier *et al.*, *Chem. Biol.* **15**, 128–136 (2008).
- R. A. Merrill *et al.*, *J. Biol. Chem.* **294**, 16164–16171 (2019).
- T. Ha, C. Kaiser, S. Myong, B. Wu, J. Xiao, *Mol. Cell* **82**, 304–314 (2022).
- A. M. Spokoyny *et al.*, *J. Am. Chem. Soc.* **135**, 5946–5949 (2013).
- B. Z. Stanton, E. J. Chory, G. R. Crabtree, *Science* **359**, eaao5902 (2018).
- T. Inobe, N. Nukina, *J. Biosci. Bioeng.* **122**, 40–46 (2016).
- T. Miyamoto *et al.*, *Nat. Chem. Biol.* **8**, 465–470 (2012).
- S. Bottone *et al.*, *Nat. Methods* **20**, 1553–1562 (2023).
- R. Ando *et al.*, *Nat. Methods* **21**, 648–656 (2024).
- N. C. Shaner *et al.*, *Nat. Biotechnol.* **22**, 1567–1572 (2004).
- H. Ren *et al.*, *Biomolecules* **14**, 859 (2024).
- D. D. Sahtoe *et al.*, *Science* **375**, eabj7662 (2022).
- M. G. Romei, S. G. Boxer, *Annu. Rev. Biophys.* **48**, 19–44 (2019).
- J. B. Grimm *et al.*, *Nat. Methods* **13**, 985–988 (2016).
- K. L. Holland *et al.*, *Nat. Methods* **23**, 909–913 (2026).
- J. Seo *et al.*, *Nat. Commun.* **13**, 2475 (2022).
- T. Zimmermann, J. Marrison, K. Hogg, P. O'Toole, in *Confocal Microscopy: Methods and Protocols*, S. W. Paddock, Ed. (Springer, 2014), pp. 129–148; https://doi.org/10.1007/978-1-60761-847-8_5.
- Y. Chen *et al.*, De novo designed bright, hyperstable rhodamine binders for fluorescence microscopy. *bioRxiv* 661379 [Preprint] (2025); <https://doi.org/10.1101/2025.06.24.661379>.
- J. Kompa *et al.*, *J. Am. Chem. Soc.* **148**, 1419–1429 (2026).
- R. Petrenas *et al.*, Rapid assessment of size, shape, and chemical complementarity of ligands for computational protein design. *BioRxiv* 662286 [Preprint] (2025); <https://doi.org/10.1101/2025.06.30.662286>.
- L. Tran *et al.*, Code archive for: De novo design of orthogonal far-red, orange, and green fluorophore-binding proteins for multiplexed imaging, Zenodo (2026); <https://doi.org/10.5281/ZENODO.20068066>.
- L. Tran *et al.*, Raw data for: De novo design of orthogonal far-red, orange, and green fluorophore-binding proteins for multiplexed imaging, Zenodo (2026); <https://doi.org/10.5281/ZENODO.20554312>.

ACKNOWLEDGMENTS

We thank the EMBL Advanced Light Microscopy Facility, especially M. Fritsch, for support with fluorescence microscopy and D. Agard, H. Eisenach, S. Honda, T. Schlichthaerle, A. Swartz, J. Wilhelm, X. Wang, Y. Liu, H. Choi, G. Ahn, M. Abedi, S. Hanna, N. Zhu, T. J. Martinez, and A. Tkachuk for helpful discussions and comments during the preparation of this manuscript. Crystallographic diffraction data were collected at the Advanced Photon Source (APS) and the National Synchrotron Light Source II (NSLS-II). At the APS, this work used the Northeastern Collaborative Access Team (NE-CAT) beamlines, funded by the National Institute of General Medical Sciences (NIGMS) from the National Institutes of Health (NIH) (P30 GM124165), using the Eiger 16M detector on beamline 24-ID-E (NIH-ORIP HEI grant S100D021527). This research was performed on APS beam time under an award (<https://doi.org/10.46936/APS-192610/60015997>) from the APS, a US Department of Energy (DOE) Office of Science User Facility operated by Argonne National Laboratory under contract no. DE-AC02-06CH11357. At NSLS-II, this work used beamline 17-ID-2 (FMX), supported by the Center for BioMolecular Structure (CBMS). CBMS is primarily supported by the NIH/NIGMS through a Center Core P30 Grant (P30GM133893) and by the DOE Office of Biological and Environmental Research (KP1605010). NSLS-II is a US DOE Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract no. DE-SC0012704. This work was delivered in part as part of the MATCHMAKERS team supported by the Cancer Grand Challenges partnership funded by Cancer Research UK (CGCATF-2023/100008), the National Cancer Institute (OT2CA297288), and the Mark Foundation for Cancer Research. **Funding:** The research was supported by the Howard Hughes Medical Institute (GRO20267: C19 HHMI INITIATIVE - 66-0656 - 2021, to A.K., E.J., and N.R.; Janelia Research Campus to T.A.B., R.P., J.B.G., and L.D.L.); Merck (GRO24322: MERCK COLLABORATION - 66-8270 - 2021, to L.T. and D.V.); The Audacious Project at the Institute for Protein Design (PG117878: Audacious Hub, PG117866: Audacious Discretionary Sub, to D.B.); the European Commission through the ARISE program (Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant 945405 to S.K.); an Erwin Schrödinger Postdoctoral Fellowship (J-4663 to E.M.); the Chan Zuckerberg Initiative (grant 2023-332009 to J.M. and D.B.); and the Damon Runyon Cancer Research Foundation (grant DRG-2507-23 to W.C.). **Author contributions:** L.T.: project design and execution, manuscript preparation; S.K.: project design, pipeline design for split binder, cellular super-resolution imaging, figure and manuscript preparation; S.S.: library preparation; D.J.: computational scripts collaboration; J.D.: cellular imaging, analysis, and discussion, manuscript preparation; B.L.: discussion; Y.W.: pipeline design; A.B.: crystal structure solving; X.L.: LC-MS data curation; A.K.: xtal curation; D.K.V.: library preparation; N.R.: library preparation; Y.B.: experimental preparation; E.M.: experimental preparation; W.C.: experimental preparation; G.R.L.: design pipeline and experimental preparation; T.A.B.: stop-flow experiments; R.P.: fluorescence lifetime measurements; J.B.G.: chemical synthesis; L.D.L.: spectroscopy and project design; J.M.: project design, imaging supervision, and manuscript preparation; L.A.: project design, computational analysis, and manuscript preparation; D.B.: project design and manuscript preparation. L.T. and S.K. contributed equally as first authors, D.J., S.S., and J.D. contributed equally as second authors, and everyone agrees that their authorship order can be exchanged to benefit their own career development. **Competing interests:** US patent 12,344,594 and US patent application US20260001857 describing fluorophores and variant compositions (with inventors J.B.G. and L.D.L.) are assigned to HHMI. L.D.L. is a scientific cofounder and shareholder of Eikon Therapeutics. L.T., S.K., L.A., J.D., L.D.L., J.M., and D.B. have filed a provisional patent application 64/041,005 that incorporates discoveries described in this article. The authors declare no competing interests. **Data, code, and materials availability:** Crystal structures of NovoTag₆₅₇ are available at the world wide Protein Data Bank (wwPDB) with accession numbers pdb_00009PVL for the apo structure and pdb_000011MQ for the holo structure. Precommercial samples of JF494, JF596, and JF657 can be requested at <https://dyes.janelia.org>. Plasmids generated in this work (listed in table S3) are available at Addgene with accession numbers 255873, 255874, 255875, 255876, 255877, 255878, 255879, 255880, 255881, 255882, 255883, 255884, and 255885. The pseudocycle Cα RFdiffusion code is available on GitHub (https://github.com/baker-laboratory/CA_RFdiffusion); the design pipeline is available on GitHub (<https://github.com/steffen-klein/novotag-design>) and archived on Zenodo (46). **License information:** Copyright © 2026 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the Author Accepted Manuscript (AAM) of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.aeb0822](https://doi.org/10.1126/science.aeb0822)
Materials and Methods; Figs. S1 to S23; Tables S1 to S3; Movies S1 to S3; Data S1; MDAR Reproducibility Checklist; References (47–65)

Submitted 30 July 2025; resubmitted 9 March 2026; accepted 26 June 2026;
published online 16 July 2026

10.1126/science.aeb0822



De novo design of orthogonal far-red, orange, and green fluorophore-binding proteins for multiplexed imaging

Long Tran, Steffen Klein, David Juergens, Shajesh Sharma, Justin Decarreau, Gyu Rie Lee, Yujia Wang, Wei Chen, Asim K. Bera, Alex Kang, Jon Woods, Emily Joyce, Dionne K. Vafeados, Nicole Roullier, Xinting Li, Bingxu Liu, Yang Bo, Edin Muratspahi#, Tim A. Brown, Jonathan B. Grimm, Ronak Patel, Luke D. Lavis, Julia Mahamid, Linna An, and David Baker

Science **393** (6813), . DOI: 10.1126/science.aeb0822

Editor's summary

Fluorescent proteins permit genetic encoding of spatial markers in cells, but they are constrained in their photochemistry relative to small-molecule dyes. Tran *et al.* developed an innovative solution to this problem by designing high-affinity fluorescent dye-binding proteins, thus permitting genetic tagging by small-molecule dyes. The authors demonstrated various applications for super-resolution and multiplexed imaging, including a split-protein design that can induce or sense proximity between two encoded fusion protein targets. —Michael A. Funk

View the article online

<https://www.science.org/doi/10.1126/science.aeb0822>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works