A general method to optimize and functionalize red-shifted rhodamine dyes

Jonathan B. Grimm, Ariana N. Tkachuk, Liangqi Xie, Heejun Choi, Boaz Mohar, Natalie Falco, Kathy Schaefer, Ronak Patel, Qinsi Zheng, Zhe Liu, Jennifer Lippincott-Schwartz, Timothy A. Brown and Luke D. Lavis

Expanding the palette of fluorescent dyes is vital to push the frontier of biological imaging. Although rhodamine dyes remain the premier type of small-molecule fluorophore owing to their bioavailability and brightness, variants excited with far-red or near-infrared light suffer from poor performance due to their propensity to adopt a lipophilic, nonfluorescent form. We report a framework for rationalizing rhodamine behavior in biological environments and a general chemical modification for rhodamines that optimizes long-wavelength variants and enables facile functionalization with different chemical groups. This strategy yields red-shifted ‘Janelia Fluor’ (JF) dyes useful for biological imaging experiments in cells and in vivo.

The development of hybrid small-molecule-protein labeling strategies enable the use of chemical fluorophores in living cells and in vivo. Optimizing small-molecule dyes for these complex biological environments is important, as synthetic fluorophores are often brighter and more photostable than fluorescent proteins.1 We recently developed general methods to improve’ and fine-tune’ rhodamine fluorophores by incorporating four-membered azetidines into the structure, yielding the JF dyes. Although our existing tuning strategies allow optimization of blue- to red-excited rhodamines, we discovered these methods cannot be applied to analogs excited with far-red and near-infrared (NIR) light due to their propensity to adopt a colorless form. Here we describe a general rubric relating the chemical properties of rhodamine dyes to their performance in biological systems and report a new complementary tuning strategy that allows rational optimization of a broader palette of fluorophores. This general method also serves as a basis for facile functionalization, enabling the synthesis of new cell- and tissue-permeant rhodamine labels for biological imaging experiments. Our work substantially expands the JF palette and yields new fluorophores useful for biological imaging, including the green fluorescence protein- (GFP)-like JF479, the highly bioavailable far-red JF669, and the NIR-excited dyes JF711 and JF722.

Results
A general rubric to predict rhodamine performance. A key property of rhodamine dyes is an equilibrium between a lipophilic, colorless lactone and the polar, fluorescent zwitterion (Fig. 1a)). Based on our previous work2–4,7, we outlined a general rubric that directly correlates the lactone–zwitterion equilibrium constant (K_L–Z) to performance in biological environments (Fig. 1b). Dyes with high K_L–Z exist almost exclusively in the zwitterionic form, making them useful as environmentally insensitive biomolecule labels.5 Rhodamines with intermediate K_L–Z values exhibit improved cell and tissue permeability due to the modestly higher propensity of the molecule to adopt the lipophilic lactone form and rapidly traverse biological membranes6,7. Dyes exhibiting even lower K_L–Z values preferentially adopt the closed lactone form, which can be exploited to create ‘fluorogenic’ dyes7,9–11 as binding of ligands and stains to their cognate biomolecular targets typically shifts the equilibrium to the fluorescent form. This property also complicates in vivo use, however, due to problems with solubility and sequestration in membranes.

Finally, dyes with extremely low K_L–Z values exist completely in the nonfluorescent lactone form, rendering them effectively unusable in biological experiments.

We compared a series of JF rhodamines with different fluorophoric systems (1–8, Fig. 1c). Compounds 2, 4 and 5 were described previously and include the azetidine-containing rhodamine (2), which we termed ‘Janelia Fluor 549’ (JF549), and the carborhodamine 9,10 analogs 4 and 5 (JF480 and JF560, respectively; Fig. 1c). We expanded the wavelength range of the JF dyes by synthesizing new azetidine-containing analogs of known rhodamine structures using metatelenol of bis(2-bromoarenes) (Fig. 1c and Supplementary Note), which we previously established as a general method for rhodamine synthesis8. These included compounds based on classic dyes containing nitrogen9–17 (X=NCH3, 1) and sulfur18 (X=S, 3) atoms as well as recently described variants containing phosphinate19,20 (X=P(O)Ph, 6), phosphine oxide21,22 (X=P(O)Ph, 7) and sulfone23 (X=SO2, 8) moieties. We measured the absorption maximum (λ(abs), extinction coefficient at λ(abs), (ε), fluorescence emission maximum λ(em) and fluorescence quantum yield (Φ) of these dyes in aqueous buffer and the K_L–Z in a dioxane:water mixture (Table 1). Comparing K_L–Z and λ(abs) uncovered an inverse correlation (Fig. 1d), with the short-wavelength NCH3-containing JF480 (1) exhibiting a high K_L–Z of 4.33 and the NIR dyes containing P(O)Ph (7) and SO2 (8) showing a low K_L–Z of ≈10–4. This correlation likely stems from the electron-withdrawing character of the X substituents13 as recently demonstrated in a computational study of rhodamines containing O, C(CH3)2 and Si(CH3)2 moieties14.

Optimizing short-wavelength dyes. In previous work, we focused on tuning the K_L–Z lower to improve tissue permeability and create fluorogenic ligands based on short-wavelength dyes. This allowed us to estimate thresholds between different categories of dyes based on measured K_L–Z values (Fig. 1b,d)). One general strategy to decrease K_L–Z involved incorporation of 3,3-difluoroazetidines1. This modification also elicits a concomitant hypsochromic shift.

The development of hybrid small-molecule-protein labeling strategies enable the use of chemical fluorophores in living cells and in vivo. Optimizing small-molecule dyes for these complex biological environments is important, as synthetic fluorophores are often brighter and more photostable than fluorescent proteins.1 We recently developed general methods to improve’ and fine-tune’ rhodamine fluorophores by incorporating four-membered azetidines into the structure, yielding the JF dyes. Although our existing tuning strategies allow optimization of blue- to red-excited rhodamines, we discovered these methods cannot be applied to analogs excited with far-red and near-infrared (NIR) light due to their propensity to adopt a colorless form. Here we describe a general rubric relating the chemical properties of rhodamine dyes to their performance in biological systems and report a new complementary tuning strategy that allows rational optimization of a broader palette of fluorophores. This general method also serves as a basis for facile functionalization, enabling the synthesis of new cell- and tissue-permeant rhodamine labels for biological imaging experiments. Our work substantially expands the JF palette and yields new fluorophores useful for biological imaging, including the green fluorescence protein- (GFP)-like JF479, the highly bioavailable far-red JF669, and the NIR-excited dyes JF711 and JF722.

Results
A general rubric to predict rhodamine performance. A key property of rhodamine dyes is an equilibrium between a lipophilic, colorless lactone and the polar, fluorescent zwitterion (Fig. 1a)). Based on our previous work2–4,7, we outlined a general rubric that directly correlates the lactone–zwitterion equilibrium constant (K_L–Z) to performance in biological environments (Fig. 1b). Dyes with high K_L–Z exist almost exclusively in the zwitterionic form, making them useful as environmentally insensitive biomolecule labels5. Rhodamines with intermediate K_L–Z values exhibit improved cell and tissue permeability due to the modestly higher propensity of the molecule to adopt the lipophilic lactone form and rapidly traverse biological membranes6,7. Dyes exhibiting even lower K_L–Z values preferentially adopt the closed lactone form, which can be exploited to create ‘fluorogenic’ dyes7,9–11 as binding of ligands and stains to their cognate biomolecular targets typically shifts the equilibrium to the fluorescent form. This property also complicates in vivo use, however, due to problems with solubility and sequestration in membranes.

Finally, dyes with extremely low K_L–Z values exist completely in the nonfluorescent lactone form, rendering them effectively unusable in biological experiments.

We compared a series of JF rhodamines with different fluorophoric systems (1–8, Fig. 1c). Compounds 2, 4 and 5 were described previously and include the azetidine-containing rhodamine (2), which we termed ‘Janelia Fluor 549’ (JF549), and the carborhodamine 9,10 analogs 4 and 5 (JF480 and JF560, respectively; Fig. 1c). We expanded the wavelength range of the JF dyes by synthesizing new azetidine-containing analogs of known rhodamine structures using metatelenol of bis(2-bromoarenes) (Fig. 1c and Supplementary Note), which we previously established as a general method for rhodamine synthesis8. These included compounds based on classic dyes containing nitrogen9–17 (X=NCH3, 1) and sulfur18 (X=S, 3) atoms as well as recently described variants containing phosphinate19,20 (X=P(O)Ph, 6), phosphine oxide21,22 (X=P(O)Ph, 7) and sulfone23 (X=SO2, 8) moieties. We measured the absorption maximum (λ(abs), extinction coefficient at λ(abs), (ε), fluorescence emission maximum λ(em) and fluorescence quantum yield (Φ) of these dyes in aqueous buffer and the K_L–Z in a dioxane:water mixture (Table 1). Comparing K_L–Z and λ(abs) uncovered an inverse correlation (Fig. 1d), with the short-wavelength NCH3-containing JF480 (1) exhibiting a high K_L–Z of 4.33 and the NIR dyes containing P(O)Ph (7) and SO2 (8) showing a low K_L–Z of ≈10–4. This correlation likely stems from the electron-withdrawing character of the X substituents13 as recently demonstrated in a computational study of rhodamines containing O, C(CH3)2 and Si(CH3)2 moieties14.

Optimizing short-wavelength dyes. In previous work, we focused on tuning the K_L–Z lower to improve tissue permeability and create fluorogenic ligands based on short-wavelength dyes. This allowed us to estimate thresholds between different categories of dyes based on measured K_L–Z values (Fig. 1b,d)). One general strategy to decrease K_L–Z involved incorporation of 3,3-difluoroazetidines1. This modification also elicits a concomitant hypsochromic shift.
of ~25 nm, transforming JF549 (2, K_{L–Z} = 3.5) into JF325 (9, K_{L–Z} = 0.68, Fig. 1e and Table 1). The HaloTag ligand based on JF325 shows improved cell permeability relative to the JF480 derivative, which is consistent with our K_{L–Z} rubric (Fig. 1b,f) and is blood–brain–barrier permeant, making it useful for in vivo voltage imaging using the Voltron indicator. This strategy also transformed JF480 (4, K_{L–Z} = 0.091) into the fluorogenic JF503 (10, K_{L–Z} = 0.001, Fig. 1e and Table 1), again supporting our K_{L–Z}-based framework (Fig. 1b,f). We applied this same tuning approach to JF490 (1) to yield the fluorinated JF509 (11, Extended Data Fig. 1a–c). Given the high K_{L–Z} = 4.33 for 1, however, this modification only moderately decreased K_{L–Z} to 2.88 (Fig. 1e,f and Table 1) and the JF490–HaloTag ligand derivative (11HTL, Fig. 1g) exhibited similar cell permeability to our previously described rhodol-based JF90–HaloTag ligand (12HTL, Extended Data Fig. 1d–i). Although the molecular brightness of the parent JF490 (11, ε = 47,900 M^{-1} cm^{-1}, Φ = 0.62; Table 1) is lower than JF480 (ε = 83,000 M^{-1} cm^{-1}, Φ = 0.87)\(^4\), the shorter λ_{ex} is advantageous for multicolor imaging experiments. JF480–HaloTag ligand (11HTL) exhibits similar spectral properties to enhanced green GFP when excited with 488 nm light but no appreciable signal when excited with 532 nm light, allowing two-color imaging with JF503–cpSNAP-tag ligand (9STL, Fig. 1i and Extended Data Fig. 2a–e).\(^4\)

In contrast, the longer λ_{ex} of 12HTL results in unwanted excitation by 532 nm light, which makes spectral separation difficult when paired with 9STL (Fig. 1i and Extended Data Fig. 2b,f–h).

Another method to decrease K_{L–Z} involves direct fluorination on the xanthen system of rhodamine dyes. Using this strategy, we previously created JF532 (13, K_{L–Z} = 0.70; Fig. 1j and Table 1); the JF532–HaloTag ligand shows improved cell permeability compared to JF490 derivatives.\(^7\) This approach is complementary to incorporation of 3,3-difluoroazetidines and combining these modifications
yielded the fluorogenic JF526 (14, $K_{L-Z} = 0.005$; Fig. 1j and Table 1). The performance of both of these dyes is consistent with our $K_{L-Z}$ rubric (Fig. 1b,k) and we sought further validation by exploring a new SNAP-tag ligand based on JF 552. The SNAP-tag is typically inferior to the HaloTag in live-cell imaging experiments due to slower labeling kinetics25,26, higher nonspecific interactions of the SNAP-tag ligands27,28, and other factors29. We compared the performance of chloropyrimidine30 derivatives of JF549 and JF 552 (2STL and 13STL, respectively; Fig. 1l). The shift in $K_{L-Z}$ resulted in the JF552 compound (13STL) showing low nonspecific staining across a wider range of concentrations (Fig. 1m) and faster live-cell labeling (Extended Data Fig. 3a–d) compared to the JF 549 compound (2STL). Compounds 2STL and 13STL exhibited comparable brightness and photostability in single-particle tracking (SPT) experiments using SNAP-tag–histone H2B fusions (Extended Data Fig. 3e,f), but the 13STL compound showed significantly lower nonspecific cytosolic staining, matching the performance of the widely used JF549–HaloTag ligand (2HTL, Fig. 1n and Extended Data Fig. 3g–i)31.

| Dye | Name | X | Y | Z | NR₂ | $\lambda_{em}$ (nm) | ε (M⁻¹ cm⁻¹) | $\lambda_{em}$ (nm) | Φ | $K_{L-Z}$ |
|-----|------|---|---|---|-----|-------------------|-------------|-------------------|---|----------|
| 1   | JF502| H | H | H | - | 502               | 57,800      | 533               | 0.71 | 4.33     |
| 2   | JF549| H | H | H | - | 549               | 101,000     | 571               | 0.88 | 3.47     |
| 3   | JF570| H | H | H | - | 570               | 83,600      | 593               | 0.63 | 2.24     |
| 4   | JF608| H | H | H | - | 608               | 99,000      | 631               | 0.67 | 0.091    |
| 5   | JF646| H | H | H | - | 646               | 5,600       | 664               | 0.54 | 0.0014   |
| 6   | JF668| H | H | H | - | 668               | 26,700      | 687               | 0.34 | 0.0093   |
| 7   | -    | H | H | H | - | ≥704              | <200        | ≥723              | ND  | <0.0001  |
| 8   | -    | H | H | H | - | ND                | <100        | ND                | ND  | <0.0001  |
| 9   | JF525| H | H | H | - | 525               | 94,000      | 549               | 0.91 | 0.068    |
| 10  | JF585| H | H | H | - | 585               | 1,500       | 609               | 0.78 | ≥0.001    |
| 11  | JF679| H | H | H | - | 479               | 47,900      | 517               | 0.62 | 2.88     |
| 13  | JF552| F | H | H | - | 552               | 95,000      | 575               | 0.83 | 0.70     |
| 14  | JF526| F | H | H | - | 526               | 19,000      | 550               | 0.87 | 0.005    |
| 15  | JF669| H | F | H | - | 669               | 112,000     | 682               | 0.37 | 0.262    |
| 16  | JF690| H | F | H | - | 690               | 150,000     | 707               | 0.24 | 2.90     |
| 17  | JF722| H | F | H | - | 722               | 87,200      | 743               | 0.11 | 0.026    |
| 18  | JF724| H | F | H | - | 724               | 6,600       | 748               | 0.05 | ≥0.001   |
| 19  | JF771| H | F | H | - | 571               | 101,000     | 590               | 0.78 | 7.93     |
| 20  | JF593| H | F | H | - | 593               | 90,300      | 612               | 0.55 | 6.06     |
| 37  | JF599| H | F | H | - | 559               | 106,000     | 579               | 0.85 | 6.22     |
| 38  | JF711| H | F | H | - | 711               | 12,400      | 732               | 0.17 | ≥0.001   |

All properties were measured in 10 mM HEPES, pH 7.3 except for $K_{L-Z}$, which was determined in 1:1 dioxane:H2O. Properties for 2, 4, 5, 9, 10, 13 and 14 were taken from previous work3,4,7. ND, not determined due to low absorbance.
Improving and derivatizing long-wavelength dyes. Having conceived and validated our $K_{L-Z}$-based framework with the short-wavelength dyes, we then turned to the far-red and NIR rhodamines (5–8, Fig. 1c), where the $K_{L-Z}$ versus $\lambda_{abs}$ relationship reveals the need for alternative tuning strategy to increase $K_{L-Z}$ (Fig. 1d). This would improve the in vivo performance of Si-rhodamines such as JF569 and rescue the colorless P(O)Ph- and SO$_2$-containing dyes (7, 8). We previously showed that halogenation of the pendant phenyl ring system can substantially increase the $K_{L-Z}$ of Si-rhodamine dyes$^9$, presumably by lowering the $pK_{a}$ of the benzoic acid moiety; this substitution also elicits a bathochromic shift$^{12}$. For example, JF569 (5; $\lambda_{abs}/\lambda_{em} = 646/664$ nm) exhibits a $K_{L-Z} = 0.0014$ but the fluorinated analog, JF569 (15; $\lambda_{abs}/\lambda_{em} = 669/682$ nm), is higher with $K_{L-Z} = 0.262$ (Fig. 2a,b and Extended Data Fig. 4a,b). This shift in $K_{L-Z}$ manifests in a higher absorptivity in aqueous solution with 5 exhibiting $\epsilon = 5,600$ M$^{-1}$·cm$^{-1}$ but the fluorinated analog 15 showing $\epsilon = 112,000$ M$^{-1}$·cm$^{-1}$ (Fig. 2c and Table 1). We expected this reaction would be general and prepared the fluorinated PO$_2$H-, P(O)Ph-, and SO$_2$-containing rhodamines (16–18, Fig. 2a) by replacing phthalic anhydride with tetrafluorophthalic anhydride in our synthetic scheme (Extended Data Fig. 4a and Supplementary Note). This modification universally increased $K_{L-Z}$ and $\lambda_{abs}$, while eliciting an ~23 nm red shift in $\lambda_{em}$. In particular, the fluorinated phosgene oxime derivative 17 strongly absorbs visible light in aqueous solution ($\epsilon = 87,000$ M$^{-1}$·cm$^{-1}$; $\lambda_{abs} =722$ nm) compared to the parent compound 7 ($\epsilon < 200$ M$^{-1}$·cm$^{-1}$; $\lambda_{abs} = 704$ nm, Fig. 1c and Table 1). This trend was generalizable to oxygen- and sulfur-containing rhodamines based on 2 and 3 where the fluorine substituents on the pendant phenyl ring also increased $K_{L-Z}$ and $\lambda_{abs}$ (19, 20, Fig. 2d,e, Table 1 and Extended Data Fig. 4f,g). We then explored derivatives of these new far-red and NIR dyes. In addition to increasing $\lambda_{abs}$ and $K_{L-Z}$, the halogenated phenyl ring motif can also serve as an electrophile in nucleophilic aromatic substitution ($S_NAr$) reactions. Thiol nucleophiles have been used for decades to prepare conjugatable derivatives of fluorinated xanthene fluorophore derivatives$^{13-15}$, including some Alexa Fluor dyes$^{16}$. The reactivity of nucleophiles other than thiols was largely unexplored; however, we discovered that NCO, CN$^-$, NH$_2$, and NH$_2$OH could react with JF569 (15) to provide derivatives 21–24 (Fig. 2f). This reaction type was generalizable to other fluorinated rhodamines and regioselective at the 6 position (Supplementary Note).
Although beyond the scope of this report, we briefly investigated some of these derivatives, finding azide 21 was an excellent reactant in strain-promoted ‘click chemistry’ with cyclic alkynes \(^\text{25–26}\) to form triazole adducts \(^\text{27–28}\) (Extended Data Fig. 5a), validating the regiochemistry of the amine addition to form 22 using intermediates 21 and 29 (Extended Data Fig. 5b), and testing the reactivity of amine-containing ion-chelating groups 30 and 31 that generated new prototype far-red indicators for K\(^{+}\) and Zn\(^{2+}\) (\(^\text{32–33}\), Extended Data Fig. 5c–e).

**Synthesis of fluorescent labels for cellular imaging.** We then sought derivatives optimized for labeling strategies such as the HaloTag and SNAP-tag. Ligands based on 6-carboxyrhodamines are particularly attractive since compounds with this regiochemistry show superior labeling efficiency and lower toxicity\(^\text{27,30}\). Although 6-carboxy-4,5,7-trifluororhodamines were unknown, we were encouraged by the selectivity of thiol and amine addition (Fig. 2f). We therefore explored malonates and related carbon nucleophiles, all of which gave a regioselective reaction at the 6 position.

---

**Fig. 3 | Fine-tuning of NIR rhodamines.**

**a**, Structures of HaloTag ligands 15\(\text{HTL}\)–20\(\text{HTL}\), and high-magnification images of fixed U2OS cell nuclei expressing HaloTag-histone H2B and labeled with 15\(\text{HTL}\)–20\(\text{HTL}\) (100–200 nM, 30 min, 3x wash). Scale bars, 10 μm. The imaging experiments were duplicated with similar results. **b**, Absorption spectra of JF711–HaloTag ligand in the absence (18\(\text{HTL}\)) or presence (18\(\text{HTL}\)+HT) of excess HaloTag protein. **c**, Absorption spectra of JF711–HaloTag ligand in the absence (17\(\text{HTL}\)) or presence (17\(\text{HTL}\)+HT) of excess HaloTag protein. **d**, Structures of dyes 17 and 38. **e**, Plot of \(K_{L-Z}\) versus \(\lambda_{\text{abs}}\) showing decreased \(K_{L-Z}\) for dye 38. **f**, Structure of JF711–HaloTag ligand (38\(\text{HTL}\)). **g**, Absorption spectra of JF711–HaloTag ligand in the absence (38\(\text{HTL}\)) or presence (38\(\text{HTL}\)+HT) of excess HaloTag protein. **h**, Nuclear fluorescence versus time on addition of ligands 17\(\text{HTL}\) or 38\(\text{HTL}\) (100 nM) to fixed cells expressing HaloTag–histone H2B; error bars indicate s.e.m.; \(n=100\) nuclei from three fields of view except for \(t=30\) min with compound 17\(\text{HTL}\) where \(n=97\) nuclei. **i**, Confocal imaging experiment of fixed U2OS cells expressing HaloTag–tOMM20 labeled with 38\(\text{HTL}\) (1 μM, 30 min, 3x wash). Scale bar, 20 μm. This imaging experiment was duplicated with similar results.
(Supplementary Note). In particular, the addition of masked acyl cyanide reagent 34, an umpolung-type acyl anion equivalent, to JF669 resulted in intermediate 35, which could be deprotected to yield a reactive acyl cyanide suitable for direct conjugation with the HaloTag ligand amine (36) to form JF669–HaloTag ligand (15\text{HTR}, Fig. 2f and Extended Data Fig. 6a). As expected from the JF669, K_{\text{L-Z}} = 0.262 (Fig. 2b and Table 1), compound 15\text{HTR} was useful in cell biological experiments (Fig. 2g,h) and was also blood–brain-barrier permeant, labeling HaloTag-expressing neurons throughout the mouse brain after intravenous (i.v.) administration (Fig. 2i and Extended Data Fig. 6b). The fluorination also improves photo-stability, with 15\text{HTR} bruises slower compared to the parent JF669–HaloTag ligand (5\text{HTR}, Fig. 2) and Extended Data Fig. 6c). The JF669–SNAP-tag ligand (15\text{STL}, Fig. 2k) was an excellent live-cell label with low nonspecific staining (Fig. 2l and Extended Data Fig. 6e,d). This chemistry was generalizable across fluorinated rhodamines, allowing facile synthesis of HaloTag ligands 15\text{HTR}–20\text{HTR} from dyes 15–20 with λ_{\text{abs}} ranging from the yellow to NIR (Fig. 3a, Extended Data Fig. 6a,f and Supplementary Note). These compounds selectively labeled HaloTag fusions in cells (Fig. 3a) and, like the JF669–HaloTag ligand (15\text{HTR}, Fig. 2), the fluorinated dye ligands 19\text{HTR} and 20\text{HTR} showed higher photo-stability than their nonfluorinated congeners (2\text{HTR} and 3\text{HTR}, Extended Data Fig. 6c,g,h). We note this new late-stage, regioselective introduction of a carboxy group has distinct advantages over classic rhodamine syntheses that generate isomeric mixtures and will be useful for synthesizing derivatives beyond self-labeling tag ligands.

Fine-tuning of NIR labels. Finally, we sought to optimize NIR HaloTag ligands for biological imaging. The SO\text{2}-containing rhodamine JF21 (18) possessed a promising K_{\text{L-Z}} = 10^{-3} for creating fluorogenic compounds; the JF24–HaloTag ligand (18\text{HTR}, Fig. 3b) showed a 15-fold increase on reaction with HaloTag protein in vitro (Fig. 3b). Nevertheless, this dye was plagued with a low Φ = 0.05 (Table 1), making it suboptimal for imaging experiments. In contrast, the P(O)Ph-containing fluorophore JF22 (17) exhibits a larger Φ = 0.11 but also a relatively high K_{\text{L-Z}} = 0.026, as expected the JF22–HaloTag ligand (17\text{HTR}) was not fluorogenic (Fig. 3c). We investigated whether our two tuning strategies could work synergistically, using JF19 (19, Fig. 2d) as a proof-of-concept. We introduced a fluorine substituent on each azetidine ring to create JF39 (37, Extended Data Fig. 7a) and found this dye exhibits K_{\text{L-Z}} = 6.22, intermediate between JF19 (2, K_{\text{L-Z}} = 3.5) and JF17 (20, K_{\text{L-Z}} = 7.93; Table 1 and Extended Data Fig. 7b–d), demonstrating the compatibility of these strategies. The JF19–HaloTag ligand (37\text{HTR}) could be used in live-cell labeling (Extended Data Fig. 7e,f). We then applied this modification to JF29 synthesizing the 3-fluoroazetidinyl JF11 (38, Fig. 3d and Extended Data Fig. 7g). The fluorination in JF11 gave a further improvement in Φ = 0.17 (Table 1) and was predicted to yield fluorogenic ligands based on its K_{\text{L-Z}} = 10^{-3} (Fig. 3e). In line with our K_{\text{L-Z}} rubric, the JF11–HaloTag ligand 38\text{HTR} (Fig. 3f) showed a fivefold increase on binding HaloTag (Fig. 3g). We compared JF11 ligand 38\text{HTR} with the parent JF22–HaloTag ligand (17\text{HTR}) in cells. Consistent with the Φ values of the parent dyes (Table 1) and our K_{\text{L-Z}}-based framework (Figs. 1b and 3e) we found that the JF11 ligand 38\text{HTR} shows higher brightness in fixed cells (Fig. 3h,i), but the JF22–HaloTag ligand (17\text{HTR}) shows better loading kinetics in live cells (Extended Data Fig. 7h) along with modestly higher photo-stability (Extended Data Fig. 7i). Thus, JF11 derivatives could be most useful in experiments where high brightness and low background are crucial, and JF22-based compounds could be better-suited for live-cell applications.

Discussion

In summary, we developed a rubric to relate the performance of simple rhodamine dyes to a single parameter, K_{\text{L-Z}} (Fig. 1b), and discovered an inverse correlation between K_{\text{L-Z}} and λ_{\text{abs}} (Fig. 1d). We validated this rubric by using our established tuning strategies to decrease K_{\text{L-Z}} and λ_{\text{abs}}, resulting in the GFP-like JF49 (11, Fig. 1e–i) and an optimized SNAP-tag ligand based on JF39 (13, Fig. 1j–n). The NIR-excited dyes 7 and 8 posed a new challenge, with low K_{\text{L-Z}} values that rendered the compounds unusable in biological environments (Fig. 1d). We therefore established a complementary general method to increase both λ_{\text{abs}} and K_{\text{L-Z}} by incorporating fluorines on the pendant phenyl ring of rhodamine dyes (Fig. 2a,b) followed by facile, generalizable S\text{Ar} chemistry to install groups for bioconjugation (Fig. 2f). This strategy yielded the bioavailable JF669 (15, Fig. 2g–i) along with other new fluorophores (16–20, Fig. 3a) and could be combined with our previous tuning method to rationally design the fluorogenic NIR-excited JF11–HaloTag ligand (38\text{HTR}, Fig. 3f–i).

The development of the JF dyes has led to several methods to fine-tune rhodamine properties: decreasing both K_{\text{L-Z}} and λ_{\text{abs}} by incorporating substituted azetidines (Fig. 1e,f), shifting the K_{\text{L-Z}} lower without affecting λ_{\text{abs}} by fluorinating the core xanthene ring system (Fig. 1j,k)’ and—as described here—increasing K_{\text{L-Z}} and λ_{\text{abs}} through fluorination of the pendant phenyl ring (Fig. 2a,b). With these complementary methods established, we can now create JF dyes with specific combinations of K_{\text{L-Z}} and λ_{\text{abs}} across much of the visible spectrum and into the NIR (Fig. 3e). Together with our general rubric relating K_{\text{L-Z}} to cellular performance (Fig. 1b) and the new derivatization chemistry (Fig. 2f), these methods will allow the rational design and synthesis of new fluorophores, labels, stains and indicators in many colors for biological imaging experiments in cells or animals. Our future work will focus on the development of fine-tuned JF derivatives for specific biological applications, further improvements in the photophysics of the JF dyes, and expansion of these design principles and synthetic strategies beyond azetidine-containing rhodamines.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-020-0909-6.

Received: 20 December 2019; Accepted: 22 June 2020; Published online: 27 July 2020

References

1. Lavi, L. D. Chemistry is dead. Long live chemistry! Biochemistry 56, 5165–5170 (2017).
2. Abdelfattah, A. S. et al. Bright and photostable chemigenetic indicators for extended in vivo voltage imaging. Science 365, 699–704 (2019).
3. Grimm, J. B. et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. Nat. Methods 12, 244–250 (2015).
4. Grimm, J. B. et al. A general method to fine-tune fluorophores for live-cell and in vivo imaging. Nat. Methods 14, 987–994 (2017).
5. Hinckley, D. A. & Seybold, P. G. A spectroscopic/thermodynamic study of the rhodamine B lactone–zwitterion equilibrium. Spectrochim. Acta A 44, 1053–1059 (1988).
6. Grimm, J. B., Brown, T. A., Thachuk, A. N. & Lavi, L. D. General synthetic method for Si-fluorescens and Si-rhodamines. ACS Cent. Sci. 3, 975–985 (2017).
7. Zheng, Q. et al. Rational design of fluorogenic and spontaneously blinking labels for super-resolution imaging. ACS Cent. Sci. 5, 1602–1613 (2019).
8. Panchuk-Voloshina, N. et al. Alexa Dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. J. Histochem. Cytochem. 47, 1179–1188 (1999).
9. Lukinavičius, G. et al. A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. Nat. Chem. 5, 132–139 (2013).
10. Wang, L. et al. A general strategy to develop cell permeable and fluorogenic probes for multicell nanoscopy. Nat. Chem. Biol. 12, 165–172 (2020).
11. Chi, W., Qi, O., Lee, R., Xu, Z. & Liu, X. A unified push–pull model for understanding the ring-opening mechanism of rhodamine dyes. J. Phys. Chem. C. 124, 3793–3801 (2020).

12. Arden-Jacob, J., Frantzkeskos, J., Kemnitzer, N. U., Zilles, A. & Drexlage, K. H. New fluorescent markers for the red region. Spectrochim. Acta A. 57, 2271–2283 (2001).

13. Grimm, J. B. et al. Carbofluoresceins and carborhodamines as scaffolds for high-contrast fluorescent probes. ACS Chem. Biol. 8, 1303–1310 (2013).

14. Hansch, C., Leo, A. & Taft, R. W. A survey of Hammett substituent constants. Chem. Rev. 52, 11206–11209 (2011).

15. Ioffe, I. S. & Zelenin, K. N. Synthesis of azafluorescein. Zh. Obshch. Khim. 34, 2811 (1964).

16. Van Duuren, B. L., Goldschmidt, B. M. & Seltzman, H. H. The synthesis and biophysical characterizations, and comprehensive NMR analysis. Eur. J. Org. Chem. 21, 13344–13356 (2015).

17. Fischer, C. & Sparr, C. Direct transformation of esters into heterocyclic fluorophores. Angew. Chem. Int. Ed. 57, 2436–2440 (2018).

18. Hanaoka, K., Terai, T. & Nagano, T. Evolution of Group 14 rhodamines as platforms for near-infrared fluorescence probes utilizing photoinduced electron transfer. ACS Chem. Biol. 6, 600–608 (2011).

19. Zhou, X., Fang, Y., Lesiak, L. & Stains, C. I. A phosphinate-containing fluorophore capable of selectively inducing apoptosis in cancer cells. Chem. Bio. Chem. 20, 1712–1716 (2019).

20. Gorgeous, P. et al. A highly photostable near-infrared labeling agent based on a phospha-rhodamine for long-term and deep imaging. Angew. Chem. Int. Ed. 57, 10137–10141 (2018).

21. Mitronova, G. Y. et al. Functionalization of the meso-phenyl ring of rhodamine dyes through SNAr with sulfur nucleophiles: synthesis, biophysical characterizations, and comprehensive NMR analysis. Eur. J. Org. Chem. 2015, 337–349 (2015).

22. Emmel, R. S. et al. Labeling strategies matter for super-resolution microscopy: A comparison between HaloTags and SNAP-tags. Cell Chem. Biol. 26, 584–592.e586 (2019).

23. Correa, I. R. et al. Substrates for improved live-cell fluorescence labeling of SNAP-tag. Carr. Pharm. Des. 19, 5414–5420 (2013).

24. Liu, Z. et al. 3D imaging of Sox2 enhancer clusters in embryonic stem cells. eLife 3, e04236 (2014).

25. Liu, J. et al. Rational design and synthesis of a novel class of highly fluorescent rhodamine dyes that have strong absorption at long wavelengths. Tetrahedron Lett. 44, 4355–4359 (2003).

26. Stagge, F., Mitronova, G. Y., Belov, V. N., Wurm, C. A. & Jakobs, S. SNAP-CLIP and Halo-tag labelling of budding yeast cells. PLoS ONE 8, e78745–e78745 (2013).

27. Presman, D. M. et al. Quantifying transcription factor binding dynamics at the single-molecule level in live cells. Methods 123, 76–88 (2017).

28. Erdmann, R. S. et al. Labeling strategies matter for super-resolution microscopy. A comparison between HaloTags and SNAP-tags. Cell Chem. Biol. 26, 584–592.e586 (2019).

29. Presman, D. M. et al. Quantifying transcription factor binding dynamics at the single-molecule level in live cells. Methods 123, 76–88 (2017).

30. Dwight, S. J. & Levin, S. Scalable regioselective synthesis of rhodamine dyes. J. Org. Chem. 55, 7905–7908 (1996).

31. Patterson, D. M., Nazarova, L. A. & Prescher, J. A. Finding the right (bioorthogonal) chemistry. ACS Chem. Biol. 9, 592–605 (2014).

32. Shuen, Z., Lu, Z., Chhalbar, P. Y., O’Herron, P. & Kara, P. An artery-specific fluorescent dye for studying neurovascular coupling. Nat. Methods 9, 273–276 (2012).

33. Emmel, R. S. et al. Facile and general synthesis of photoactivatable xanthene dyes. Angew. Chem. Int. Ed. 50, 11206–11209 (2011).

34. Wysocki, L. M. et al. A highly photostable near-infrared labeling agent based on a phospha-rhodamine for long-term and deep imaging. Angew. Chem. Int. Ed. 57, 10137–10141 (2018).

35. Mitronova, G. Y. et al. Functionalization of the meso-phenyl ring of rhodamine dyes through SNAr with sulfur nucleophiles: synthesis, biophysical characterizations, and comprehensive NMR analysis. Eur. J. Org. Chem. 2015, 337–349 (2015).

36. Emmel, R. S. et al. Labeling strategies matter for super-resolution microscopy: A comparison between HaloTags and SNAP-tags. Cell Chem. Biol. 26, 584–592.e586 (2019).

37. Gretarsson, E. S. et al. Facile and general synthesis of photoactivatable xanthene dyes. Angew. Chem. Int. Ed. 50, 11206–11209 (2011).

38. Dwight, S. J. & Levin, S. Scalable regioselective synthesis of rhodamine dyes. J. Org. Chem. 55, 4515–4516 (1990).

39. Wysocki, L. M. et al. Facile and general synthesis of photoactivatable xanthene dyes. Angew. Chem. Int. Ed. 50, 11206–11209 (2011).

40. Dwight, S. J. & Levin, S. Scalable regioselective synthesis of rhodamine dyes. Org. Lett. 18, 5316–5319 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
two-photon excitation spectra were measured as previously described. Briefly, mixtures were incubated at 4 °C overnight. These HaloTag conjugate solutions were prepared as dilute solutions (< 0.1 μM final) into 10 mM HEPES, pH 7.3 buffer solution.

Fluorescence measurements were performed after the HaloTag conjugate solutions were diluted 5x (1 μM final [ligand]) into 10 mM HEPES, pH 7.3 buffer solution. Spectra are averages (n = 2). The spectra of GFP (Fig. 1b) were taken from FPbase (https://www.fpbase.org/protein/egfp/20).

Multiphoton spectroscopy of dyes and HaloTag conjugates. For compounds 1, 11 and the fluorescein control (Extended Data Fig. 1c) solutions of the free dyes (5 μM) were prepared in 10 mM HEPES buffer, pH 7.3. For other rhodamines (Extended Data Fig. 6f), spectra of the HaloTag conjugates were measured. As above, solutions of HaloTag ligands compounds 15HTL, 18HTL, and 38STL (5 μM) were prepared in 10 mM HEPES, pH 7.3 containing 0.1 mg/ml CHAPS. An aliquot of HaloTag protein (1.5 equiv., 7.5 μM final [HaloTag]) was added and the resulting mixture was incubated at 4 °C overnight. Fluorescence measurements were performed after the HaloTag conjugate solutions were diluted 5x (1 μM final [ligand]) into 10 mM HEPES, pH 7.3 buffer solution. Spectra are averages (n = 2). The spectra of GFP (Fig. 1b) were taken from FPbase (https://www.fpbase.org/protein/egfp/20).

Multiphoton spectroscopy of dyes and HaloTag conjugates. For compounds 1, 11 and the fluorescein control (Extended Data Fig. 1c) solutions of the free dyes (5 μM) were prepared in 10 mM HEPES buffer, pH 7.3. For other rhodamines (Extended Data Fig. 6f), spectra of the HaloTag conjugates were measured. As above, solutions of HaloTag ligands compounds 15HTL, 18HTL, and 38STL (5 μM) were prepared in 10 mM HEPES, pH 7.3 containing 0.1 mg/ml CHAPS. An aliquot of HaloTag protein (2 equiv., 10 μM final) was added and the resulting mixture was incubated at 4 °C overnight. These HaloTag conjugate solutions were diluted 5x (1 μM final [ligand]) into 10 mM HEPES, pH 7.3 and the two–photon excitation spectra were measured as previously described. Briefly, measurements were taken on an inverted microscope (IX81, Olympus) equipped with a x60, 1.2 numerical aperture (NA) water objective (Olympus). Dye–protein samples were excited with pulses from an 80 MHz Ti-Sapphire laser (Chameleon Ultra II, Coherent) for 710–1,080 nm and with an OPO (Chameleon Compact OPO, Coherent) for 1,000–3,000 nm. Fluorescence collected by the objective was passed through a dichroic filter (675DCXR, Omega) and a short pass filter (720SP Semrock) and detected by a fiber-coupled Avalanche Photodiode (SPCM_AQRH-14, PerkinElmer). All excitation spectra are corrected for the wavelength–dependent transmission of the dichroic and bandpass filters, and quantum efficiency of the detector. Spectra are averages (n = 2).

General cell culture and fluorescence microscopy. All cell lines underwent regular mycoplasma testing by the Janelia Cell Culture Facility. Unless otherwise noted, U2OS cells (ATCC) were cultured in Dulbecco’s modified Eagle medium (DMEM, phenol red-free; Life Technologies) supplemented with 10% v/v fetal bovine serum (FBS, Life Technologies), 1 mM GlutaMAX (Life Technologies) and maintained at 37 °C in a humidified 5% v/v CO2 environment. For confocal and widefield imaging of cell nuclei (Figs. 1a and 3a and Extended Data Figs. 1b, 2b–h and 7f), we used U2OS cells with an integrated HaloTag–histone H2B fusion protein expressing plasmid via the piggyBac transposition system unless otherwise noted. For confocal imaging of mitochondria (Fig. 3i), we used U2OS cells with an integrated TOMM20–HaloTag fusion protein expressing plasmid unless otherwise noted; TOMM20 is an outer mitochondrial membrane protein as part of a protein translocase complex.

Multiplexed imaging comparison JF20 and JF39. See Fig. 1a and Extended Data Fig. 2c. U2OS cells stably expressing HaloTag–histone H2B fusion protein were transiently transfected with plasmids encoding pSNAP–TOMM20 fusion protein as part of fluorescence (Lonza). Live cells were incubated with JF20–HaloTag ligand (1 μM, 300 nM) or JF39–HaloTag ligand (1 μM, 500 nM, Extended Data Fig. 1d) for 3 h followed by addition of JF17–cSNAP-tag ligand (9 μM, 100 nM, Extended Data Fig. 2a) and incubated for an additional 30 min. These cells were then washed three times in dye-free media and imaged (Fig. 1f) using tunable white light laser (WLL) excitation at 488 or 532 nm on a Leica SP8 Falcon confocal microscope with an HC PL-APo x63/1.2 NA objective. Zeiss LSM 800 confocal microscope with a Plan APO x20/0.8 NA M27 objective or Plan APO x63/1.2 NA oil DIC M27 objective; Zeiss LSM 880 with a C-APO x100/1.2 NA W Corr MCF2 M27 objective. The Leica and Zeiss LSM 800 confocal images were processed using Fiji. Unless otherwise noted, live cells were washed three times in dye-free media and fixed cells were washed in phosphate-buffered saline. We use the following shorthand in the figures: HT, HaloTag; ST, SNAP-tag; H2B, histone H2B; PDGFR, C-terminal transmembrane anchoring domain from platelet-derived growth factor receptor (PDGFR) fused to the HaloTag protein (HaloTag–PDGFR). Unless otherwise noted, cells were imaged on the following microscopes: Nikon Eclipse Ti with a Plan APO 100× objective; Leica SP8 Falcon confocal microscope with an HC PL-APo x63/1.2 NA water objective; Zeiss LSM 800 confocal microscope with a Plan APO x20/0.8 NA M27 objective or Plan APO x63/1.2 NA oil DIC M27 objective; Zeiss LSM 880 with a C-APO x100/1.2 NA W Corr MCF2 M27 objective. The Leica and Zeiss LSM 800 confocal images were processed using Fiji.

Flow cytometry loading experiments. See Fig. 1m and Extended Data Fig. 3a–d. This experiment used the mouse embryonic stem cell line JMR4, a gift from R. Tjian (Berkeley), derived from the C57BL/6N strain. The JMR4.N4 cells were authenticated by short tandem repeat DNA profiling and approved by the NIH 4D Nucleome project as a Tier2 cell line. Wild-type mouse embryonic stem cells or embryonic stem cells stably expressing SNAP-tag–histone H2B fusion protein were plated onto 96-well microplates precoated with 0.1% gelatin (Corning). Cells were washed three times for 10 min, trypsinized and loaded onto CytoFLEX 5 flow cytometer equipped with a plate loader (Beckman Coulter). The embryonic stem cell population was designated based on its forward light scatter (FLS) and side light scatter (SCL) characteristics (Extended Data Fig. 3a). The gating strategy to determine the nonfluorescent cell population used control embryonic stem cell samples not incubated with SNAP-tag ligands, plotting SSC versus fluorescence from the Y585-PE channel (phycoerythrin; 561 nm laser excitation), 585 nm with a 42 nm bandpass emission, Avalanche Photodiode detector gain setting, 130; SSC threshold Automatic settings were as follows: FSC Avalanche Photodiode detector gain setting, 12; FSC threshold Automatic. In some cases, not all concentrations or time points could be sampled during every run due to flow rate and instrument constraints. For experiments varying loading
concentration (Fig. 1m) replicates were as follows: experiments using 2HTL n = 7 except for [ligand] = 3 nM where n = 5; experiments using 13HTL n = 3. For experiments varying loading time (Extended Data Fig. 3d) replicates were as follows: experiments using 2HTL n = 4 except for t = 7.5 min where n = 2 and t = 210 min where n = 3.

**SPT experiments.** See Fig. 1n and Extended Data Fig. 3e–i. SPT experiments were performed in U2OS cells with an integrated SNAP-tag–histone H2B fusion protein expressing plasmid using a phosphate buffer saline (PBS) and immediately imaged using an integrated HaloTag–histone H2B fusion protein expressing plasmid via the piggyBac transposon system. SNAP-tag–histone H2B expressing cells were labeled with 2 nM of either JF5-cpSNAP-tag ligand (2HTL) or JF5-cpSNAP-tag ligand (13HTL). HaloTag–histone H2B fusion protein expressing cells were labeled with 2 nM of JF5-HaloTag–histone H2B (2HTL). Cells were incubated with dyes for 15 min at 37°C and then washed three times for 30 min each. SPT was performed at 0.5 μm/s (1 μm/frames) and 5 μm/frames and results were normalized to 0.5 μm/s per cell. Single molecules were localized and tracked by a MATLAB implementation of multiple target tracking and SLIMFast. For SPT brightness (photons per s) we used Extended Data Fig. 3e) n = 19,008 single-molecule events using 2HTL and n = 9,511 single-molecule events using 13HTL. For SPT track length (s) (Extended Data Fig. 3f) n = 10,822 single-molecule events using 2HTL and n = 9,387 single-molecule events using 13HTL. Trajectories were fitted into a two-state model: chromatin bound and diffusive (free) using diffusion coefficient (D) cut off D_{\text{chop}} (0.0005, 0.008) and D_{\text{none}} (0.15, 0.25). The fraction of chromatin-bound molecules per cell are plotted (Fig. 1n); n = 12 cells for experiments using 2HTL and n = 8 cells for experiments using 13HTL and 2HTL.

**Airyscan imaging experiments using 15HTL.** See Fig. 2g. U2OS cells were transiently transfected with HaloTag–Sec61β fusion protein expressing plasmid or HaloTag–TOMM20 fusion protein expressing plasmid using FuGENE HD (Promega) and maintained in DMEM containing 10% v/v FBS and penicillin–streptomycin–glutamine (15HTL) is an endoplasmic reticulum membrane protein translocator protein. Cells were incubated with 100 nM JF669–HaloTag ligand (15HTL) for 1 h with τ = 0.5 h with 38HTL where n = 86; t = 1 h with 38HTL where n = 94; t = 2 h with 38HTL where n = 96; t = 0.5 h with 17HTL where n = 94.

**Airyscan imaging experiments using 15HTL.** See Fig. 2l and Extended Data Fig. 6d.e. U2OS cells stably expressing SNAP-tag–histone H2A2Z and HaloTag–Sec61β fusion proteins were labeled with JF5-cpSNAP-tag ligand (15HTL, 30 nM) and JF5-HaloTag–histone H2B (2HTL, 30 nM) for 30 min, sustaining with Hoechst 33342 (1 μg/ml) (thermosilf). Cells were washed (10 min), imaged using the Zeiss LSM 880 platform under the Airyscan SR mode using a plan-apochromatic x63/1.4 NA objective. The Airyscan images were processed using the Zen software (Zeiss) and the fluorescence intensity line-scan (Extended Data Fig. 6e) was extracted using FIJI.

**Statistics and reproducibility.** For spectroscopy measurements reported n values for absorption spectra, extinction coefficient (ε) and quantum yield (Φ) represent measurements of different samples prepared from the same dye DMSO stock solution or HaloTag conjugate stock solution. For flow cytometry experiments, reported n denotes separate cell samples taken from different microplate wells. For SPT brightness and track-length experiments, n indicates separate events extracted by the multiple target tracking algorithm. For fraction chromatin bound from SPT experiments, n indicates the number of individual cells; one-way analysis of variance gave adjusted P = 0.0013 (**) for 2HTL versus 13HTL and adjusted P = 0.9963 (not significant) for 13HTL versus 2HTL (F2, 25) = 11.38. For photobleaching experiments, n indicates the number of separate cellular experiments where the intensity of the entire field of view was measured at the indicated time points. For cell loading experiments, n represents the number of intensity values from individual cells extracted from three fields of view at the indicated time points. For representative fluorescence microscopy and flow cytometry experiments, all procedures were duplicated at least once on a separate biological sample to ensure results were similar as indicated in the figure legends. Additional information can be found in the Nature Research Reporting Summary.

**Data availability.** The data that support the findings of this study are provided in the Source Data files or available from the corresponding author upon request. Source data are provided with this paper.

**References.**
41. Suzuki, K. et al. Reevaluation of absolute luminescence quantum yields of standard solutions with a spectrometer using a spectrally broad and a back-thinned CCD detector. Phys. Chem. Chem. Phys. 11, 9850–9860 (2009).
42. Lambert, T. J. fPhase: a community-editable fluorescent protein database. Nat. Methods 16, 277–278 (2019).
43. Mütze, J. et al. Excitation spectra and brightness optimization of two-photon excited probes. Biophys. J. 102, 934–944 (2012).
44. Akerboom, J. et al. Optimization of a GCaMP calcium indicator for neural activity imaging. J. Neurosci. 32, 13819–13840 (2012).
45. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012).
46. Serge, A., Bertrand, N., Rigoulet, H. & Marguet, D. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. Nat. Methods 5, 678–679 (2008).

**Acknowledgements.** We thank S. Sernson for initial discussions on unpalpable reagents; K. Svoboda for advice on in vivo labeling experiments; C. Deo and E. Schreiter for purified HaloTag.
protein and contributive discussions; B. Mensh and P. Kumar for a critical reading of the manuscript; R. Tjian (University of California, Berkeley) for the mouse embryonic stem cell line and the Janelia Cell and Tissue Culture, Anatomy and Histology and Vivarium teams for assistance with biological experiments. This work was supported by the Howard Hughes Medical Institute.

Author contributions
L.D.L. and J.B.G. conceived the project. J.B.G. contributed organic synthesis and one-photon spectroscopy measurements. A.N.T., L.X. and H.C. contributed cultured cell imaging experiments. B.M. contributed in vivo labeling and tissue imaging experiments. N.F. and Q.Z. contributed organic synthesis. L.X. and K.S. contributed flow cytometry experiments. R.P. contributed two-photon spectroscopy measurements. Z.L., J.L.-S. and T.A.B. directed the project. L.D.L. directed the project and wrote the paper with input from the other authors.

Competing interests
US Patent 9,933,417 and patent applications US 2019/0367736 and US 2019/0106573 describing azetidine-containing fluorophores and variant compositions (with inventors J.B.G. and L.D.L.) are assigned to HHMI.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41592-020-0909-6.
Supplementary information is available for this paper at https://doi.org/10.1038/s41592-020-0909-6.
Correspondence and requests for materials should be addressed to L.D.L.
Peer review information Rita Strack was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.
Reprints and permissions information is available at www.nature.com/reprints.
Extended Data Fig. 1 | Utility of JF479–HaloTag (11HTL) ligand in cellular imaging experiments. a–b, Fluorescence excitation (ex) and emission (em) spectra of 1 (a) and 11 (b). c, Two-photon absorption spectra of 1, 11, and reference dye fluorescein. d, Structure of JF503–HaloTag ligand (12HTL). e, Nuclear fluorescence vs. time upon addition of ligands 11HTL or 12HTL (200 nM) to live cells expressing HaloTag–histone H2B; error bars indicate SE; n = 100 nuclei. f–i, Confocal imaging experiments of fixed U2OS cells expressing either HaloTag histone–H2B fusion protein (f, h; nucleus) or HaloTag–PDGFR transmembrane domain (TMD) fusion protein (g, i; plasma membrane) labeled with JF479–HaloTag ligand (11HTL; 100 nM, 1 h, 2× wash; f, g) or JF503–HaloTag ligand (12HTL; 100 nM, 1 h, 2× wash; h, i); scale bars: 21 μm; these imaging experiments were duplicated with similar results.
Extended Data Fig. 2 | Comparison of JF479–HaloTag ligand (11HTL) and JF503–HaloTag ligand (12HTL) in two-color experiments with JF525–cpSNAP-tag ligand (9STL).

a, Structure of JF525–cpSNAP-tag ligand (9STL).
b, Fluorescence excitation spectra of JF479–HaloTag ligand (11HTL) or JF503–HaloTag ligand (12HTL) bound to HaloTag protein. Dashed lines highlight 488 nm or 532 nm excitation.
c–e, Enlarged confocal images and line-scans from Fig. 1i showing live U2OS cells expressing HaloTag–histone H2B labeled with 11HTL (500 nM, 3.5 h, 3× wash) and tOMM20–SNAP-tag labeled with 9STL (100 nM, 30 min, 3× wash) excited with 532 nm; this imaging experiment was duplicated with similar results.

c, Confocal image from Fig. 1i with blue line indicating line-scan position; d, Line-scan profile; e, Over-exposed image showing low nuclear signal.

f–h, Enlarged confocal images and line-scan from Fig. 1i showing U2OS cells expressing HaloTag–histone H2B labeled with 12HTL (500 nM, 3.5 h, 3× wash) and TOMM20–SNAP-tag labeled with 9STL (100 nM, 30 min, 3× wash) excited with 532 nm; this imaging experiment was duplicated with similar results.

f, Confocal image from Fig. 1i with green line indicating line-scan position; g, Line-scan profile; h, Over-exposed image showing high nuclear signal. Scale bars for all images: 10 μm.
Extended Data Fig. 3 | Utility of JF<sub>552</sub>–cpSNAP-tag ligand (13STL) in cellular imaging experiments. 

a, Representative flow cytometry plot showing forward light scatter (FSC) vs. side light scatter (SSC) demonstrating gating strategy to separate cells from debris; experiment was duplicated with similar results.
b, Representative flow cytometry plot showing SSC vs. fluorescence of 13STL–SNAP-tag measured using the Y585-PE channel (561 nm laser excitation, 585 nm with a 42 nm bandpass emission) to demonstrate gating strategy to separate fluorescent and nonfluorescent cells; experiment was duplicated with similar results. 
c, Representative flow cytometry plots showing the change in % fluorescent cells as a function of incubation time with 13STL (10 nM); top row: wild-type (WT) embryonic stem (ES) cells; bottom row: SNAP-tag–histone H2B (ST) expressing ES cells; experiment was duplicated with similar results. 
d, Plot of fluorescent mouse ES cells (%) vs. time determined by the flow cytometry experiment exemplified in c. WT cells or ST cells were incubated with 2STL or 13STL (10 nM) for different times and % fluorescent cells were measured; error bars show SE; experiments using 2STL n = 3; experiments using 13STL n = 4 except for t = 7.5 min where n = 2 and t = 210 min where n = 3.  
e, Violin plot of photon counts from a single-particle tracking (SPT) experiment using U2OS cells expressing SNAP-tag–histone H2B and labeled with 2STL or 13STL (2 nM, 30 min, 3x wash); lines indicate median and quartiles; n = 19008 single-molecule events for experiment using 2STL and n = 9511 single-molecule events for experiment using 13STL. 
f, Histogram of track lengths from SPT experiment using cells expressing SNAP-tag–histone H2B and labeled with 2STL or 13STL (2 nM, 30 min, 3x wash); n = 10822 single-molecule events for experiment using 2STL and n = 9387 single-molecule events for experiment using 13STL. 
g, Structure of JF<sub>549</sub>–Halotag ligand (2HTL).  
h, Image of individual SPT traces in live U2OS cells expressing SNAP-tag–histone H2B and labeled with 2STL (2 nM, 30 min, 3x wash); dashed line indicates outline of nucleus; arrows highlight nonspecific staining in cytosol; scale bar: 2 μm. 
i, 3D kymograph showing data from h detailing single-particle track position and length as a function of time. Diffusion coefficient values (D) are calculated from single-particle tracking data and color-coded; experiment was duplicated with similar results.
Extended Data Fig. 4 | Synthesis and spectral properties of 15–20. a, Synthesis of Janelia Fluor dyes 15–20. b–g, Fluorescence excitation (ex) and emission (em) spectra of JF669 (15; b), JF690 (16; c), JF722 (17; d), JF724 (18; e), JF571 (19; f), and JF593 (20; g).
Extended Data Fig. 5 | Derivatization of JF669. a, Reaction of azide 21 with strained alkynes 25 or 26 to form triazole adducts 27 or 28. b, Synthesis of amine 22 via reaction of JF669 (15) with NH₃, reduction of azide 21, or Curtius rearrangement starting from ester 29 showing consistent regioselectivity of SNAr reactions. c, Reaction of 15 with amine-containing chelator groups 30 and 31 to form far-red K⁺ indicator 32 and far-red Zn²⁺ indicator 33. d, Fluorescence emission spectra of 32 in the absence or presence of 100 mM K⁺. e, Fluorescence emission spectra of 30 in the absence or presence of 10 μM Zn²⁺.
Extended Data Fig. 6 | Synthesis and properties of new HaloTag and SNAP-tag ligands based on 15–20 and 37–38. a, Expanded synthetic scheme of HaloTag ligands 15HTL–20HTL and 37HTL–38HTL starting with nucleophilic aromatic substitution (SNAr) of 15–20 and 37–38 with masked acyl cyanide 34. b, Two-color montage image of fixed coronal slices with zoom-in regions from a mouse expressing HaloTag–gFP in neurons transduced by IV administration of the viral vector PHP-eB-Syn-HaloTag–gFP followed by IV administration of 15HTL (100 nmol), perfusion, and slicing; gFP signal in green and 15HTL signal in magenta; scale bar = 3 mm; experiment was duplicated with similar results.

c, Structures of JF669–Halotag ligand (15HTL), JF646–Halotag ligand (5HTL), JF593–Halotag ligand (20HTL), and JF570–Halotag ligand (3HTL).

d–e, Confocal image and line-scan from Fig. 2l showing live U2OS cells expressing Sec61β–Halotag labeled with 2HTL (30 nM, 30 min, 3x wash) and SNAP-tag–histone variant H2A.Z labeled with 15STL (30 nM, 30 min, 3x wash); costained with Hoechst 33342 (1 μM, 30 min, 3x wash); d, Confocal image with white line indicating line-scan position; e, Line-scan profile; scale bar: 5 μm; experiment was duplicated with similar results.

f, Two-photon absorption spectra of the HaloTag conjugates of HaloTag ligands 15HTL–20HTL and 37HTL–38HTL.

g, Plot of fluorescence from cells expressing HaloTag–H2B labeled with 2HTL (200 nM) or 19HTL (200 nM) over 30 bleach cycles; error bars indicate SE; n = 3 independent cellular samples.

h, Plot of fluorescence from fixed U2OS cells expressing HaloTag–H2B labeled with 3HTL (200 nM) or 20HTL (200 nM) over 30 bleach cycles; error bars indicate SE; n = 3 independent cellular samples.
Extended Data Fig. 7 | Further fine-tuning of JF571 (19) and JF722 (17). a, Structure of JF571 (19) and JF559 (37). b–c, Full plot of $K_{L-Z}$ vs. $\lambda_{abs}$ (b) and zoom-in (c) showing decreased $K_{L-Z}$ for dye 37. d, Fluorescence excitation (ex) and emission (em) spectra of JF559 (37). e, Structure of JF559–Halotag ligand (37HTL). f, Widefield imaging experiment of U2OS cells expressing HaloTag–histone H2B labeled with 37HTL (100 nM, 30 min, 3x wash); scale bar: 51 μm; experiment was duplicated with similar results. g, Fluorescence excitation (ex) and emission (em) spectra of JF711 (38). h, Nuclear fluorescence vs. time upon addition of ligands 17HTL (200 nM) or 38HTL (200 nM) to live cells expressing HaloTag–histone H2B; error bars indicate SE; $n = 100$ nuclei except for: $t = 0.5$ h with 38HTL where $n = 86$ nuclei; $t = 1$ h with 38HTL where $n = 94$ nuclei; $t = 2$ h with 38HTL where $n = 96$ nuclei; $t = 0.5$ h with 17HTL where $n = 94$ nuclei. i, Plot of fluorescence from cells expressing HaloTag–H2B labeled with 17HTL (200 nM) or 38HTL (200 nM) over 30 bleach cycles; error bars indicate SE; $n = 3$ independent cellular samples.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

☑ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

☑ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☑ The statistical test(s) used AND whether they are one- or two-sided

☑ Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☑ A description of all covariates tested

☑ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☑ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☑ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted

Give P values as exact values whenever suitable.

☑ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☑ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☑ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

The Shimadzu analytical HPLC instrument was controlled by Shimadzu LabSolutions 5.51 software. The Shimadzu preparative HPLC instruments were controlled by Shimadzu LabSolutions 5.42 SP4, 5.85, or 5.97 software. The Shimadzu LC/MS system was controlled by LabSolutions 5.42 SP4 software. The Agilent LC/MS system was controlled by Agilent LC/MSD ChemStation Rev. B04.02 [54] software. The Biotage Isolera flash chromatography system was controlled by Biotage OS 852M software. The Buchi Rotavapor was running Buchi firmware version 01.01.00.00. The NMR was controlled by IconNMR 4.7.7 Build 20 software. The Cary Model 100 spectrometer was controlled by CaryWinUV 4.20[468] software. The Cary Eclipse was controlled by Cary Eclipse Scan Application 1.1(132) software. The Quantaurus Quantum Yield spectrophotometer was controlled by PLQY software U6039-05 3.4.2. The Zeiss LSM 880 confocal microscope was controlled by Zen2.1 SP3 software.These cytometers were controlled by CytoFLEX CytExpert v.2.3 software.

Data analysis

NMR data were analyzed using Bruker TopSpin 3.2 or MestReNova 14.1.1-24571 software. Spectra and other graphical data were analyzed using GraphPad Prism 8 for macOS version 8.4.2 (464) software. Fluorescence microscopy images were processed and processed using the following software: Zen2.1 SP3 software from Zeiss, NIS-Elements AR 4.40.0, 64-bit (build 1084) software from Nikon, or Fiji Version:2.0.0-rc-69/1.52p; Build:269a69d59f. Flow cytometry data was analyzed using FlowJo v.10.6.1 software.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are provided in the Source Data files or available from the corresponding author upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size
All quantitative imaging experiments were performed on replicate samples or fields of view as reported in the Methods and Figure legends.

Data exclusions
No data were excluded from this study.

Replication
All imaging and flow cytometry data were repeated and showed similar results.

Randomization
We compared a relatively small number of compounds allowing the use of cell culture samples that were the same age, treated in the same way, and analyzed by a standard protocol. Randomization was not necessary for these experiments.

Blinding
We compared a relatively small number of compounds allowing the use of cell culture samples that were the same age, treated in the same way, and analyzed by a standard protocol. Blinding was not necessary for these experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants
- Clinical data

Methods

- Involved in the study

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)
U2OS cells were from ATCC. The mouse ES JM8.N4 cells, derived from the C57BL/6N mouse strain, were a gift from Robert Tjian’s lab (University of California, Berkeley).

Authentication
The U2OS cells were authenticated by ATCC and were not re-authenticated at Janelia. The JM8.N4 cells were authenticated by short tandem repeat DNA profiling and approved by the NIH 4D Nucleome project as a Tier2 cell line.

Mycoplasma contamination
The Janelia Cell Culture Facility regularly tests for mycoplasma contamination; these cell lines tested negative.

Commonly misidentified lines
No commonly misidentified cell lines were used in this study.
Animals and other organisms

Policy information about studies involving animals: ARRIVE guidelines recommended for reporting animal research

| Laboratory animals | Adult C57/BL6 male mice, 2-4 months old, were used to express a GFP-HaloTag fusion protein throughout the brain. |
|--------------------|-------------------------------------------------------------------------------------------------------------|
| Wild animals       | No wild animals were used in this study.                                                                     |
| Field-collected samples | No field-collected samples were used in this study.                                                          |
| Ethics oversight   | All experimental protocols were conducted according to the National Institutes of Health guidelines for animal research and were approved by the Institutional Animal Care and Use Committee at the Janelia Research Campus, HHMI. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation Wild type mouse embryonic stem (ES) cells or ES cells stably expressing the SNAP-tag–histone H2B were plated into onto flat-bottom 96-well plates precoated with 0.1% gelatin (Corning). Cells at 70% confluency were stained with JF549–cpSNAP-tag ligand (2STL) or JF552–cpSNAP-tag ligand (13STL) at different concentrations (3 nM, 10 nM, 30 nM, 100 nM, 300 nM) for 15 min or at different time points (15 min, 30 min, 60 min, 120 min, 210 min) at 10 nM concentration. Cells were washed 3× for 10 min, trypsinized, and loaded onto the flow cytometer equipped with a plate loader.

Instrument CytoFLEX S flow cytometer equipped with a plate loader (Beckman Coulter).

Software Data was collected with the CytoFLEX CytExpert v.2.3 operating software and analyzed with FlowJo v.10.6.1 software.

Cell population abundance The cells were not sorted. For the experimental replicates, the mouse embryonic stem cell line used for flow cytometry analysis had between 85–98% expression of SNAP-tag–histone H2B as determined against negative fluorescence gating.

Gating strategy The mouse ES cell population was designated based on its forward light scatter (FSC) and side light scatter (SSC) characteristics (Supplementary Fig. 3a). The "nonfluorescent" cell population was determined from control ES cell samples not incubated with SNAP-tag ligands, plotting SSC-H vs. PE fluorescence (phycoerythrin; 585 nm with a 42 nm bandpass; Supplementary Fig. 3b). Sample dilution and flow rate were adjusted to optimize event recordings for the 96-well microplate format. The settings were as follows: FSC avalanche photodiode (AP) detector gain setting = 12; SSC AP detector gain setting = 130; FSC threshold Automatic; PE channel AP gain setting = 1.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.