

SNAP-tag2 for faster and brighter protein labeling

Received: 9 September 2024

Accepted: 15 May 2025

Published online: 3 July 2025

Check for updates

Stefanie Kühn¹, Veselin Nasufovic¹, Jonas Wilhelm¹, Julian Kompa¹, Eline M. F. de Lange^{2,3}, Yin-Hsi Lin¹, Cornelia Egoldt¹, Jonas Fischer¹, Artem Lennoi¹, Mirosław Tarnawski⁴, Jochen Reinstein⁵, Rifka Vlijm⁶, Julien Hiblot^{1,7}✉ & Kai Johnsson^{1,6,7}✉

SNAP-tag is a powerful tool for labeling proteins with synthetic fluorophores in bioimaging. However, its utility in live-cell applications can be constrained by its relatively slow labeling kinetics and the limited cell permeability of its substrates. Here, we introduce improved labeling substrates and an engineered SNAP-tag for faster labeling in vitro and in live cells. SNAP-tag2 presents a second-order rate constant with rhodamine substrates that approaches $10^7 \text{ s}^{-1} \text{ M}^{-1}$, a 100-fold improvement over the corresponding SNAP-tag–substrate pairs. When labeled with highly fluorogenic dyes, SNAP-tag2 also shows a fivefold increase in fluorescence brightness relative to currently used SNAP-tag. The increased labeling kinetics and brightness of SNAP-tag2 translate into greatly improved performance in various live-cell (super-resolution) imaging applications.

Self-labeling protein tags (SLPs) can be specifically and covalently labeled with synthetic probes in vitro and in live cells^{1,2}. Popular examples of such tags are SNAP-tag^{3,4}, HaloTag7 (ref. 5) and, to a lesser extent, CLIP-tag⁶. Their main field of use has been in bioimaging, where these tags offer the opportunity to attach bright and photostable synthetic fluorophores⁷ to proteins of interest. The development of fluorogenic probes for protein labeling (that is, probes that only become highly fluorescent upon binding to the protein of interest) has further facilitated the use of SLPs in live-cell bioimaging as it decreases background signal from unbound dye^{8,9}.

HaloTag7 is at present the most popular SLP for live-cell imaging. HaloTag7 was engineered to react fast and efficiently with chloroalkane (CA)–rhodamine substrates¹⁰ (Supplementary Fig. 1). Rhodamines are well-suited synthetic fluorophores for live-cell imaging as they cover a broad spectral range, possess excellent spectroscopic properties¹¹, can be fluorogenic^{8,12} and are permeable in live cells^{13,14}. The fluorogenicity and permeability of rhodamines can probably be attributed to the reversible spirocyclization of the zwitterionic, fluorescent rhodamine

to a nonpolar and nonfluorescent spirolactone^{9,14,15}. For selected CA–rhodamines, labeling kinetics of HaloTag7 are close to the diffusion limit. The high reactivity of HaloTag7 toward CA–rhodamines is at least partially because of specific interactions of the protein surface with the rhodamine dye¹⁶. The combination of the high labeling velocity of HaloTag7 with CA–rhodamines, their relatively good permeability and the excellent performance of these dyes in various microscopy applications have made the approach a powerful tool in bioimaging¹⁷.

SNAP-tag undergoes a nucleophilic substitution reaction with *O*⁶-benzylguanine (BG) or chloropyrimidine (CP) derivatives, in which the benzyl group of the substrate is irreversibly transferred to an active site reactive cysteine residue (Fig. 1a). SNAP-tag is a highly engineered human *O*⁶-alkylguanine-DNA alkyltransferase (hAGT) that was optimized for reaction with BG derivatives^{3,18–20}. While BG derivatives show faster labeling kinetics with SNAP-tag than the corresponding CP derivatives¹⁶, the less polar CP derivatives generally perform better in live-cell imaging²¹. However, compared to HaloTag7, SNAP-tag labeling kinetics with rhodamine-based substrates are in general at least two

¹Department of Chemical Biology, Max Planck Institute for Medical Research, Heidelberg, Germany. ²Molecular Biophysics, Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands. ³Molecular Cell Biology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands. ⁴Protein Expression and Characterization Facility, Max Planck Institute for Medical Research, Heidelberg, Germany. ⁵Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, Heidelberg, Germany. ⁶Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. ⁷These authors contributed equally: Julien Hiblot, Kai Johnsson. ✉e-mail: julien.hiblot@mr.mpg.de; johnsson@mr.mpg.de

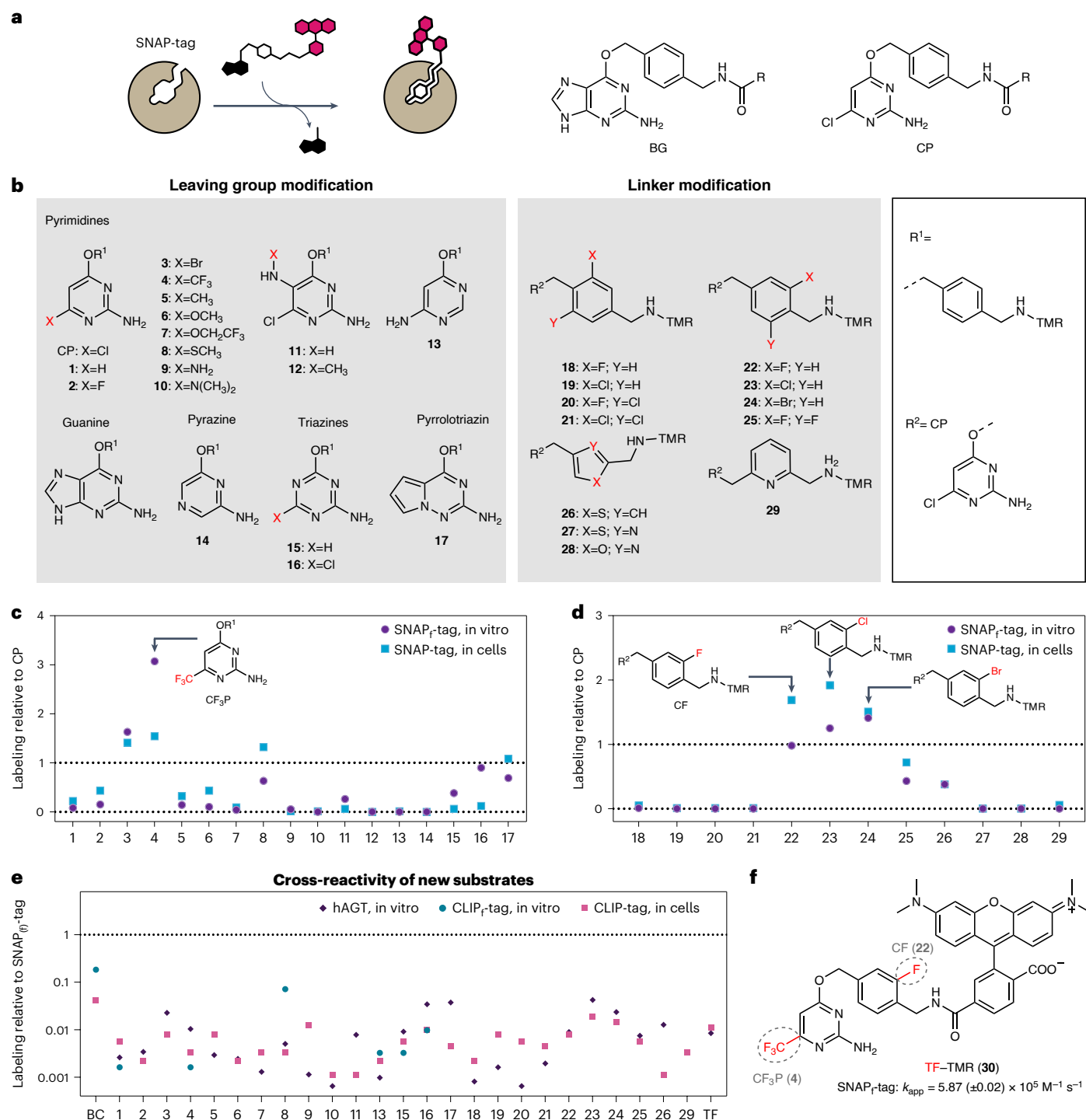


Fig. 1 | SNAP-tag substrate screening for more efficient labeling in vitro and in live mammalian cells. a, Scheme of SNAP-tag labeling reaction with fluorophore substrates. The chemical structures of SNAP-tag substrates BG and CP are shown on the right. R represents the functional moiety to be linked to SNAP-tag.

b, Chemical structures of modified SNAP-tag substrates divided into two groups: leaving group modification and linker modification. Modified leaving groups were linked to TMR over linker R¹. Modified linkers coupled to TMR were tested on the CP (R²) scaffold. **c, d**, Comparison of substrates 1–17 (**c**) and substrates 18–29 (**d**) relative to SNAP_F-tag labeling with CP–TMR regarding their in vitro labeling kinetics and performance in live-cell labeling. In vitro labeling kinetics of SNAP_F-tag with different substrates were measured recording FP traces over time. The k_{app} was calculated (Supplementary Table 1) and normalized to the

k_{app} of SNAP_F-tag with CP–TMR. Live-cell performance of developed substrates was tested by labeling of U2OS cells stably expressing a mEGFP–SNAP-tag fusion protein with TMR substrates at 100 nM for 2 h. Cells were washed and analyzed by flow cytometry. Fluorescence intensity ratios of TMR to mEGFP were calculated (Supplementary Table 1) and normalized to the ratio obtained for SNAP-tag with CP–TMR. Leaving group substrate 4 and linker substrates 22, 23 and 24 showed the most promising results relative to CP–TMR. **e**, Reactivity of selected SNAP-tag substrates with hAGT and CLIP_F-tag compared to SNAP_F-tag labeling with CP–TMR. Experiments were conducted as previously described. **f**, Chemical structure of TF–TMR (30) found through the combination of leaving group 4 and linker 22. The k_{app} of SNAP_F-tag labeling with TF–TMR is depicted below.

orders of magnitude slower¹⁶. Additionally, HaloTag7 shows higher fluorogenicity with rhodamines than SNAP-tag^{9,22}. This is presumably because of specific interactions between the fluorophore and the HaloTag7 surface, which promote an equilibrium shift from the nonfluorescent spirolactone to the fluorescent zwitterionic state²³. No such specific interactions between SNAP-tag and rhodamine-based fluorophores have been identified¹⁶.

The limited cell permeability of the SNAP-tag substrate and the reduced labeling kinetics and fluorogenicity of the SNAP-tag can impede its use in certain live-cell imaging applications. Here, we present SNAP-tag2, a highly engineered SNAP-tag mutant that rapidly reacts with optimized pyrimidine-based substrates for live-cell applications, and demonstrate the effectiveness of SNAP-tag2 and its substrates in a range of bioimaging applications.

Results

Improved SNAP-tag substrates

To improve the performance of SNAP-tag in live-cell imaging, we decided to first develop different substrate scaffolds for higher intrinsic reactivity in vitro and improved live-cell compatibility, followed by engineering of SNAP-tag for faster reaction kinetics and enhanced fluorescence brightness with these substrates. SNAP-tag substrates can be divided into two parts: the leaving group and the linker that connects the synthetic probe to the reactive cysteine (Fig. 1b). We synthesized 17 different heterocycles connected through a benzyl linker to tetramethylrhodamine (TMR) and measured their labeling kinetics in vitro with SNAP_F-tag, a single-point mutant of SNAP-tag with increased labeling kinetics^{16,24}, as well as their performance of labeling SNAP-tag in a cellular assay (Fig. 1c,d and Supplementary Table 1). The selection of heterocycles was based on previous observations that pyrimidines can be good substrates for SNAP-tag^{21,25} but also comprised pyrazines, triazines and a pyrrolotriazine. Of these compounds, the trifluoromethyl-substituted pyrimidine **4** (CF₃P) showed the best results (Fig. 1c), with a roughly 3-fold increase in labeling kinetics in vitro and 1.5-fold higher fluorescence labeling in live cells. Additionally, we investigated 12 different linkers connecting CP as the leaving group and TMR as the fluorescent probe. The fluoro and chloro derivatives **22** and **23** showed 2-fold increased fluorescence labeling in the cellular assay (Fig. 1d). However, the chloro analog **23** showed increased reactivity against parental hAGT (Fig. 1e). Therefore, to reduce the risk of unwanted background reactions, it was excluded. We furthermore computed the octanol–water partition coefficient (QP logP o/w), aqueous solubility (QP logS) and apparent MDCK permeability (QP PMDCK) of the various substrates as their nonfluorescent *N*-acetylated derivatives (Extended Data Fig. 1 and Supplementary Table 2). The substrates with the highest predicted permeability and logP o/w, which additionally showed fast labeling of SNAP_F-tag in vitro when coupled to TMR, also performed best in live-cell fluorescence labeling experiments (Extended Data Fig. 1). An exception was compound **7**, for which high permeability was predicted but low efficiency for labeling of SNAP-tag in live cells was observed (Fig. 1c). Combining the features of the best-performing substrates **4** (CF₃P) and **22** (CF) yielded trifluoromethyl fluorobenzyl pyrimidine **30** (TF), which, as a TMR derivative (Fig. 1f), showed an apparent second-order rate constant (k_{app}) of $5.87 (\pm 0.02) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for the labeling of SNAP_F-tag (Supplementary Fig. 2), corresponding to 3.9-fold faster labeling kinetics than its reaction with CP–TMR ($k_{app} = 1.51 (\pm 0.01) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$)¹⁶. Moreover, TF–TMR showed 1.7-fold higher fluorescence labeling than CP–TMR in live cells (Supplementary Fig. 3). We further demonstrated that the optimized substrates do not significantly influence cell viability, even at substrate concentrations higher than used in labeling (Extended Data Fig. 2).

Engineering of SNAP-tag protein

In addition to substrate optimization, we used protein engineering to increase the reactivity of SNAP-tag toward pyrimidine-based substrates

and its brightness when labeled with fluorogenic rhodamines such as MaP618 (ref. 9). Specifically, we used mutagenesis of selected residues, computational design and directed evolution to identify mutants with increased reactivity and fluorogenicity (Fig. 2 and Supplementary Fig. 4). The engineering was guided by the available X-ray structures of SNAP-tag in its apo (Protein Data Bank (PDB) 3KZY), benzylated (PDB 3L00), BG-bound (PDB 3KZZ)⁴ and TMR-labeled (PDB 6Y8P)¹⁶ states and by past engineering efforts^{3,18,19,24} (Fig. 2a). We modified S135 to arginine, assuming it might favorably interact with the carboxylate of rhodamine substrates. We also used the computational method PROSS²⁶ to identify amino acid substitutions that could increase the thermal stability of SNAP-tag (Fig. 2a). Additionally, the N-terminal domain of SNAP-tag comprises a long, unstructured region (residues 37–54; Fig. 2a–c), which might compromise the protein's stability. To address this, we used RosettaRemodel²⁷ to design a short eight-residue loop as a replacement for this unstructured region (Fig. 2b,c). For directed evolution experiments, protein libraries of two distinct regions in close proximity to the substrate-binding site were created by saturation mutagenesis (Fig. 2a): the β -strand proximal to the active site (residues 29–36) and the active site loop (residues 155–161). Lastly, a synthetic deep mutational scanning library (sdMSL)²⁸ was screened, in which each residue in SNAP-tag was individually replaced by all other 19 amino acids, excluding cysteines but including deletions. Deep mutational scanning allows systematically studying the impact of individual variations in the context of SNAP-tag and identifying the best amino acids at each position. Libraries were screened for enhanced labeling kinetics with the developed TMR substrates and/or increased fluorescence intensity of MaP618 derivatives (Supplementary Fig. 1) using yeast surface display (YSD) combined with fluorescence-activated cell sorting (Supplementary Fig. 4). These engineering efforts led to the identification of SNAP-tag2, which carries 11 substitutions and a replacement of 18 amino acids (residues 37–54) by a peptide of eight amino acids (Fig. 2b,c), without destabilizing the protein (melting temperature (T_m) of around 65 °C; Supplementary Fig. 5). The k_{app} for its reaction with TF–TMR was $8.22 (\pm 0.79) \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Table 1, Supplementary Fig. 6 and Supplementary Tables 3 and 4), which was about 100-fold faster than the reaction of SNAP-tag with CP–TMR (Fig. 2d) and close to that for the reaction of HaloTag7 with CA–TMR ($k_{app} = 1.88 (\pm 0.01) \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$)¹⁶. SNAP-tag2 also showed increased reaction kinetics with CP-based and CF-based substrates and reached its fastest labeling reaction with TF–CPY in vitro with a k_{app} of $1.04 (\pm 0.12) \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ (Table 1, Supplementary Figs. 7–9 and Supplementary Tables 3 and 4). Furthermore, SNAP-tag2 showed increased reactivity toward nonfluorescent probes (Fig. 2d, Supplementary Fig. 10 and Supplementary Table 5). The in vitro labeling kinetics of SNAP-tag2 with BG substrates were determined to be very similar to those of SNAP-tag with these substrates (Fig. 2d and Supplementary Figs. 10 and 11). The extinction coefficients and quantum yields (QYs) of SNAP-tag2 labeled with different rhodamines were similar to those of its predecessor, except for very fluorogenic MaP618 substrates (Supplementary Table 6). For these, SNAP-tag2 showed a significantly higher absorbance than SNAP_F-tag, (Fig. 2e), indicating that SNAP-tag2 might have a higher capacity to shift the dye toward its zwitterionic, fluorescent state.

Performance of SNAP-tag2 in live-cell imaging

We then tested the performance of SNAP-tag2 and its improved substrates in live-cell imaging applications. We first compared the labeling kinetics of SNAP-tag2 and SNAP_F-tag with different fluorescent probes (that is, TMR, carbopyronine (CPY) and silicon rhodamine (SiR)) in live U2OS cells. The fluorophores were coupled to CP, CF or TF (Supplementary Fig. 1) to identify the best substrate–fluorophore pair for each fluorophore individually to use in live-cell labeling. We used U2OS cells stably coexpressing nuclear-localized HaloTag7–SNAP-tag2 or HaloTag7–SNAP_F-tag fusions together with mTurquoise2, which was used for normalization of the protein expression level. Cells were

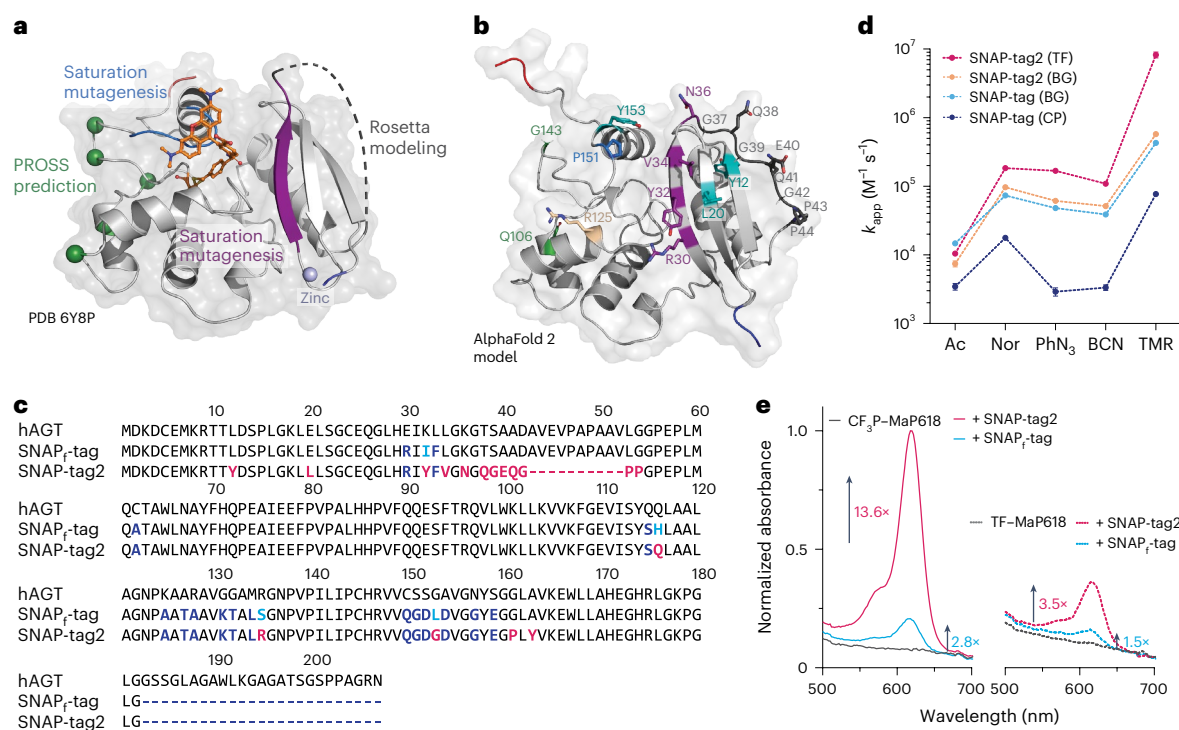


Fig. 2 | SNAP-tag engineering aiming for faster and brighter fluorescence labeling. **a**, Crystal structure of SNAP-tag labeled with TMR (PDB 6Y8P) with engineered regions highlighted. SNAP-tag is represented as a light-gray cartoon and the TMR ligand is represented as sticks. Saturation mutagenesis libraries were constructed on the C-terminal loop region (residues 155–161) and the β -strand proximal to the active site (residues 29–36), highlighted in blue and violet, respectively. Substitutions predicted by PROSS²⁶ to increase the protein thermal stability are highlighted as green spheres. The unstructured region in SNAP-tag (residues 37–55) highlighted as a gray dotted line was redesigned using RosettaRemodel²⁷. Termini are highlighted in blue (N terminus) and red (C terminus) and the coordinated zinc ion is illustrated as a light-blue sphere. **b**, AlphaFold 2 (ref. 36) model of SNAP-tag2. Introduced substitutions are highlighted as sticks and color-coded on the basis of the engineering rationale. Green, PROSS²⁶ prediction; marine blue and deep purple, saturation mutagenesis libraries; wheat, rational design; teal, sDMSL; dark gray, Rosetta-modeled loop. Termini are highlighted in blue (N terminus) and red (C terminus). **c**, Sequence alignment of hAGT, SNAP_f-tag and SNAP-tag2. Dark blue, common differences

of SNAP_f-tag and SNAP-tag2 versus hAGT; turquoise, unique differences SNAP_f-tag versus hAGT; pink, unique differences SNAP-tag2 versus hAGT and/or SNAP_f-tag. Amino acid deletions are presented as dotted lines. **d**, Comparison of labeling kinetics between SNAP-tag2 and SNAP-tag. The k_{app} of SNAP-tag2 with TF substrates demonstrate one to two orders of magnitude faster labeling kinetics compared to the reaction of parental SNAP-tag with CP substrates. Labeling kinetics of SNAP-tag2 with BG substrates remained unchanged compared to SNAP-tag labeling. CP and BG results for SNAP-tag labeling were taken from Wilhelm et al.¹⁶. Abbreviations: Ac, acetate; BCN, bicyclononyne; Nor, (1*S*,4*S*)-5-methylbicyclo[2.2.1]hept-2-ene (norbornene); PhN₃, phenylazide. **e**, Normalized absorbance spectra of fluorogenic CF₃P-MaP618 and TF-MaP618 substrates (15 μ M) in the presence or absence of SNAP-tag2/SNAP_f-tag protein (30 μ M). Fold changes are referred to the absorbance of the dye only. SNAP-tag2 shows approximately 4.9-fold and 2.3-fold increases in absorbance of fluorogenic substrates for CF₃P-MaP618 and TF-MaP618, respectively, compared to SNAP_f-tag.

Table 1 | The k_{app} of SNAP-tag2 labeling with different TMR and CPY substrates

| TMR substrate | k_{app} (\pm s.d.) ($M^{-1}s^{-1}$) | CPY substrate | k_{app} (\pm s.d.) ($M^{-1}s^{-1}$) |
|---------------|--|---------------|--|
| TF-TMR | $8.22 (\pm 0.79) \times 10^6$ | TF-CPY | $1.04 (\pm 0.12) \times 10^7$ |
| CF-TMR | $6.21 (\pm 3.65) \times 10^6$ | CF-CPY | $3.62 (\pm 0.73) \times 10^6$ |
| CP-TMR | $7.82 (\pm 2.78) \times 10^6$ | CP-CPY | $5.29 (\pm 1.62) \times 10^6$ |

Values represent the average k_{app} values calculated from experimental replicates (Supplementary Table 3).

incubated with the respective fluorescent probes and the fluorescence intensity changes were recorded as a function of time. For all tested fluorescent substrates, SNAP-tag2 showed significantly faster labeling than SNAP_f-tag with CP substrates (Fig. 3a–d and Supplementary Fig. 12). Among the various fluorophores, the differences between SNAP-tag2 and SNAP_f-tag were most pronounced for the labeling with SiR derivatives (Fig. 3c,d). The half-labeling time ($t_{1/2}$) for reaching saturation of the fluorescence signal was around 20 min for SNAP-tag2 labeled with CF-SiR, whereas SNAP_f-tag showed only relatively weak labeling with CP-SiR under the same experimental conditions, not reaching a plateau within 90 min (Fig. 3c,d). For all CPY substrates,

SNAP-tag2 showed around 4-fold faster labeling than SNAP_f-tag. The $t_{1/2}$ values for SNAP-tag2 labeling with different CPY substrates were similar ($t_{1/2} = 10$ –12 min), highlighting a small influence of the substrate core for live-cell labeling with this fluorophore. For TMR substrates, the differences between SNAP-tag2 and SNAP_f-tag were less pronounced (1.6–2.2-fold). This suggests that live-cell labeling with TMR substrates is limited by the entry of the substrate into the cell.

Similar results were obtained in a flow cytometry-based assay for the labeling of mEGFP-SNAP-tag2 and mEGFP-SNAP_f-tag fusion proteins in U2OS cells. All tested substrates led to a more efficient labeling of SNAP-tag2 than SNAP_f-tag, with the most pronounced differences observed for SiR and MaP618 substrates (Extended Data Fig. 3). BG substrates, however, showed reduced labeling in live cells in general (Extended Data Fig. 3).

Moreover, we compared the fluorescence brightness of labeled SNAP-tag2 in live cells to that of SNAP_f-tag and HaloTag7. Previously described U2OS cells stably coexpressing nuclear-localized HaloTag7-SNAP-tag2 or HaloTag7-SNAP_f-tag fusions together with mTurquoise2 were labeled overnight with CA-TMR, CA-MaP618, TF-TMR or TF-MaP618 to achieve full labeling (Fig. 3e). Quantification of the SLP fluorescence signal normalized to mTurquoise2 expression

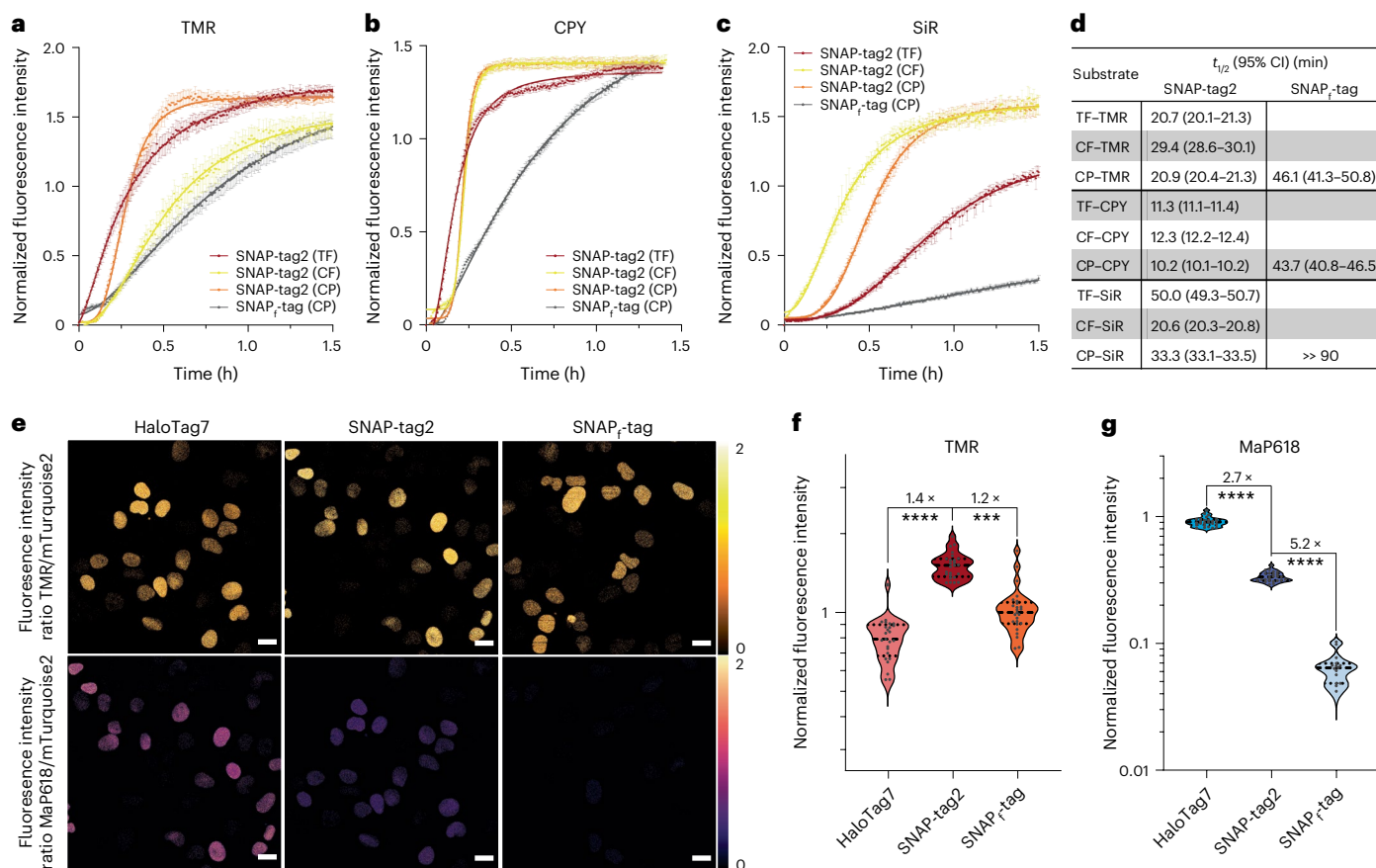


Fig. 3 | Labeling kinetics and fluorescence brightness for SNAP-tag2 in live cells. Experiments were conducted in live U2OS cells stably coexpressing HaloTag7–SNAP_f-tag or HaloTag7–SNAP-tag2 together with mTurquoise2 (expression marker) in the nucleus. **a–c**, Kinetic traces of SNAP-tag2 and SNAP_f-tag fluorescence labeling in live cells with TMR (**a**), CPY (**b**) and SiR (**c**) substrates. U2OS cells were labeled with TF–fluorophore, CF–fluorophore and CP–fluorophore substrates (50 nM for TMR and CPY, 100 nM for SiR) and the labeling reaction was followed by confocal fluorescence microscopy. Fluorescence intensity changes of the substrates were normalized to the mTurquoise2 fluorescence over time. Data were fitted to a sigmoidal curve (n (TMR) \geq 66 cells, n (CPY) \geq 105 cells, n (SiR) \geq 27 cells). Data are presented as the mean values \pm 95% CI. Representative results from a biological duplicate are shown (replicate in Supplementary Fig. 12). **d**, The $t_{1/2}$ of SNAP-tag2 and SNAP_f-tag

calculated from in-cell kinetic measurements described in **a–c**. Values represent the mean values \pm 95% CI from a biological duplicate. **e**, Comparison of HaloTag7, SNAP-tag2 and SNAP_f-tag fluorescence brightness in confocal fluorescence microscopy with nonfluorogenic TMR and highly fluorogenic MaP618 substrates. U2OS cells were labeled with SLP substrates CA–TMR, CA–MaP618, TF–TMR and TF–MaP618 (100 nM overnight) and washed before imaging. Ratiometric projections are presented corresponding to fluorescence intensities of label to mTurquoise2 using orange-hot (TMR) and mpl-magma (MaP618) lookup tables. Scale bar, 20 μ m. **f,g**, Quantitative analysis of single cells shown in **e** represented as violin plots. Numbers represent fold changes between the different SLPs ($n \geq$ 15 cells). Statistical analysis was conducted using a two-tailed unpaired *t*-test with Welch's correction. **** $P < 0.0001$ and *** $P = 0.0003$.

showed smaller differences among all three SLPs for TMR labeling (Fig. 3f and Extended Data Fig. 4c). Similar results were obtained for SiR substrates (Extended Data Fig. 4a,b,f) and CPY substrates (Extended Data Fig. 4d), for which differences in the fluorescence intensities among all SLPs, including HaloTag7, remained below twofold. For labeling with TF–MaP618, SNAP-tag2 showed a 5.2-fold higher fluorescence intensity than SNAP_f-tag in cells (Fig. 3g and Extended Data Fig. 4e), which is in line with the measured absorbance difference in vitro (Fig. 2e). However, the brightness of MaP618-labeled SNAP-tag2 remained 2.7-fold dimmer than that of MaP618-labeled HaloTag7. Thus, SNAP-tag2 probably shifts the equilibrium of spirocyclization of very closed rhodamine dyes more toward the zwitterionic form than SNAP_f-tag but less than HaloTag7.

We further tested the performance of SNAP-tag2 in super-resolution stimulated emission–depletion (STED)²⁹ microscopy. HeLa cells stably coexpressing HaloTag7, SNAP-tag2 or SNAP_f-tag fused to the Cox8a presequence (mitochondrial matrix localization) and mEGFP as an expression marker (no specific localization) were labeled with their respective CA–SiR or TF–SiR substrates (Fig. 4a).

SNAP-tag2 showed much stronger fluorescence labeling inside mitochondria, both in confocal laser scanning microscopy (CLSM) and STED microscopy than observed for SNAP_f-tag under the same experimental conditions, even though a cell with stronger mEGFP fluorescence (\triangleq higher expression level) was chosen for SNAP_f-tag. Detection of the dim SNAP_f-tag labeling required a significant increase in laser power (Extended Data Fig. 5). The observed fluorescence signal for SNAP-tag2 was comparable to the signal obtained for the corresponding HaloTag7 labeling at similar expression levels. Similar differences in fluorescence intensities were observed for SNAP-tag2 and SNAP_f-tag labeling of intermediate filaments (vimentin (Vim) fusion) (Fig. 4b). Because of the increased photon count, SNAP-tag2 showed improved STED resolution, as highlighted by the reduced full-width at half-maximum (FWHM) of Vim filaments (108 ± 23 nm) compared to SNAP_f-tag (146 ± 7.0 nm). We also tested the STED performance of SNAP-tag2 with other fusion proteins such as CalR–KDEL (endoplasmic reticulum lumen; Fig. 4c) and TOMM20 (outer mitochondrial membrane; Fig. 4d) and observed bright and specific labeling of SNAP-tag2 with CF–SiR. Furthermore, we performed dual-color live-cell STED imaging in U2OS cells expressing

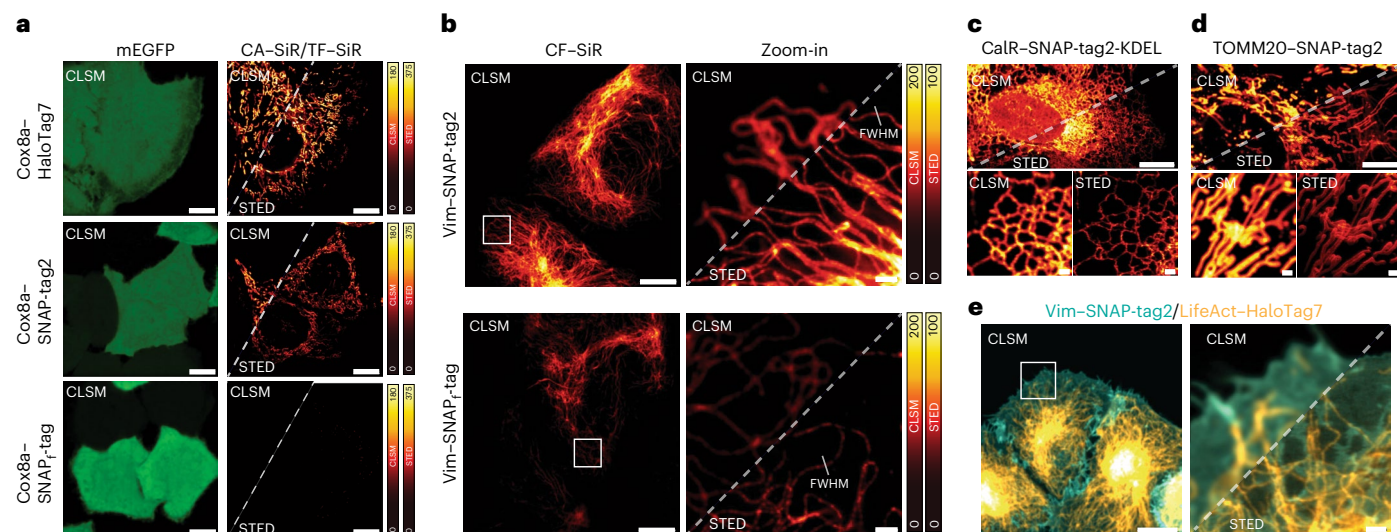


Fig. 4 | SNAP-tag2 performance in live-cell super-resolution microscopy.

a, Comparison of HaloTag7, SNAP-tag2 and SNAP_r-tag performance in CLSM and STED microscopy. HeLa cells stably coexpressing HaloTag7, SNAP-tag2 or SNAP_r-tag in the mitochondria (Cox8a localization sequence) together with mEGFP (no specific localization) were labeled with CA-SiR or TF-SiR (100 nM) for 1 h and washed afterward. SNAP-tag2 and HaloTag7 show comparable performance in STED imaging, while SNAP_r-tag shows insufficient signal under the same imaging conditions. This experiment was performed three times with similar results. Scale bars, 10 μm. Lookup tables: green (mEGFP) and red-hot (SiR). **b, c**, CLSM and STED images of U2OS cells stably expressing Vim-SNAP-tag2 (**b**) or Vim-SNAP_r-tag (**c**) labeled with CF-SiR (100 nM) for 1 h. Cells were washed before imaging. White squares in the overview images (left) highlight the area chosen for magnification and STED imaging (right). The experiment was repeated twice with similar results. The FWHM of single intermediate filament

fibers highlighted in **b** was determined to be 108 (±23) nm and 146 (±7.0) nm, underlining the higher resolution achieved with SNAP-tag2 compared to SNAP_r-tag (15 individual filaments from three individual images per condition). Scale bars, 5 μm (overview) and 1 μm (magnification). **c, d**, CLSM and STED images of U2OS cells stably expressing SNAP-tag2 (**c**) in the lumen of the endoplasmic reticulum (CaR-KDEL) and (**d**) on the outer mitochondrial membrane (TOMM20) labeled with CF-SiR (100 nM) for 1 h, demonstrating the versatility of using SNAP-tag2 in different cellular compartments. Cells were washed before imaging. Scale bars, 5 μm (overview) and 1 μm (magnification). **e**, Dual-color CLSM and STED images of U2OS cells expressing Vim-SNAP-tag2 and LifeAct-HaloTag7. SNAP-tag2 was labeled with CF-SiR (100 nM) and HaloTag7 was labeled with CA-MaP618 (100 nM) for 1 h; cells were washed afterward. Scale bars, 5 μm (overview) and 1 μm (magnification). Lookup tables: cyan (SNAP-tag2-SiR) and orange-hot (HaloTag7-MaP618).

Vim-SNAP-tag2 and LifeAct-HaloTag7 (F-actin) using CF-SiR and CA-MaP618, respectively (Fig. 4e). These results demonstrate the superiority of SNAP-tag2 over previous SNAP-tag versions in live-cell imaging experiments.

We then investigated whether the improved SNAP-tag2 system translates to more efficient fluorescence labeling in yeast. Chemical labeling in yeast is challenging, presumably because of the presence of its cell wall, an additional permeability barrier and the expression of various multidrug efflux pumps³⁰. The peroxisomal membrane protein Pex3 has a key role in peroxisomal biogenesis, inheritance and degradation in yeast³¹. To study the precise localization, dynamics and distribution of Pex3 and its binding partners, live-cell STED nanoscopy is a suitable technique because it enables the quantification of the exact number and size of peroxisomes in live yeast³². Thus, we expressed either Pex3-SNAP-tag2 or Pex3-SNAP_r-tag in the methylotrophic yeast species *Hansenula polymorpha* and tested different SiR substrates for labeling of peroxisomes in live cells (Fig. 5a). *H. polymorpha* expressing Pex3-SNAP-tag2 showed homogenous and highly efficient labeling of peroxisomes, whereas the fluorescence signal was much weaker for yeast cells expressing Pex3-SNAP_r-tag (Fig. 5a,b). The correct localization of Pex3-SNAP-tag2 was furthermore confirmed through STED microscopy (Fig. 5c and Supplementary Fig. 13a,b). Similar results were obtained for labeling with MaP555 (ref. 9) substrates, for which SNAP-tag2 worked best in combination with CP-MaP555 (Extended Data Fig. 6 and Supplementary Fig. 13c,d). These results demonstrate that SNAP-tag2 enables more efficient labeling in *H. polymorpha* yeast than its predecessor.

Discussion

In this work we used a combination of substrate optimization and protein engineering to develop SNAP-tag2, an improved version of the

SNAP-tag labeling system. SNAP-tag2 reacts faster with an improved set of pyrimidine-based substrates. An increased labeling speed is observed for a large variety of probes, including rhodamines as well as different reactive biorthogonal groups. The k_{app} for the reaction of SNAP-tag2 with the substrate TF-TMR in vitro approaches $10^7 \text{ M}^{-1} \text{ s}^{-1}$, which is ~100-fold faster than the labeling reaction of SNAP-tag with CP-TMR and is close to that of HaloTag7 with CA-TMR. SNAP-tag2 remained a very stable protein with a melting temperature of ~65 °C, which might facilitate further engineering attempts such as circular permutations and the generation of split versions of SNAP-tag2. The creation of such cpSNAP-tag2 and split-SNAP-tag2 systems could enable the development of biosensors and recorders for studying biological activities in cells and/or in vivo³³. Its size of 172 aa is slightly below that of its predecessor and substantially below that of HaloTag7 with 297 aa. SNAP-tag2 displaces the spirocyclization equilibrium of bound rhodamines more toward the open form compared to SNAP_r-tag, a property for which we directly screened in the directed evolution experiments. We assume that a possible interaction of SNAP-tag2 with the zwitterionic, fluorescent configuration of rhodamines results in higher brightness of SNAP-tag2 when labeled with the very fluorogenic dye MaP618. Nevertheless, SNAP-tag2 displaces the spirocyclization equilibrium to a lower degree than HaloTag7. Since very closed rhodamine derivatives tend to show reduced labeling performance in cellular experiments¹⁵ and in vivo³⁴, we believe that this is not a major limitation of SNAP-tag2 relative to HaloTag7. Owing to the lack of structural information on labeled SNAP-tag2, the role of individual amino acid substitutions in tuning the interactions of SNAP-tag2 with rhodamines remains unclear.

Most importantly, the increased reactivity of SNAP-tag2 toward its improved substrates translates into a more efficient fluorescence labeling in live cells. This observation was most pronounced for fluorescence labeling with the far-red SiR substrates, which enabled efficient

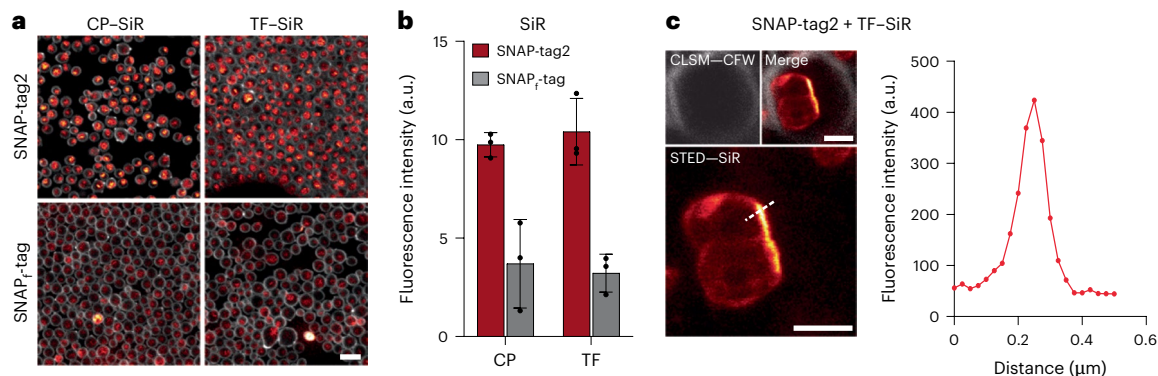


Fig. 5 | Comparison of SNAP-tag2 and SNAP-tag labeling of yeast peroxisomes.

a, CLSM images of *H. polymorpha* yeast cells expressing Pex3-SNAP-tag2 or Pex3-SNAP-tag fusion proteins labeled with different SiR substrates. Yeast cells were labeled with CP-SiR or TF-SiR (250 nM) for 18 h and the cell wall was stained with Calcofluor white (CFW; 25 $\mu\text{g ml}^{-1}$) for 15 min. Cells were washed before imaging. Scale bar, 10 μm . **b**, Bar plot representing the quantitative analysis of SNAP-tag2 and SNAP-tag labeling with SiR substrates in yeast. Experiments were conducted in biological triplicates ($n = 125$ cells for each replicate) and the mean

fluorescence intensity of SiR substrates was calculated (Supplementary Fig. 14). Error bars represent the s.d. SNAP-tag2 greatly outperforms SNAP-tag in labeling of live yeast peroxisomes. **c**, Bottom, STED image of Pex3-SNAP-tag2 labeled with TF-SiR. Top, CLSM image of CFW-stained cell wall and merge of both channels. Scale bar, 1 μm . Right, line profile of labeled peroxisomes in STED imaging (highlighted as dashed line in the image). SNAP-tag2 with TF-SiR is suitable to perform live-cell STED microscopy in yeast.

fluorescence labeling at substrate concentrations that only resulted in weak labeling of SNAP-tag. The relatively weak SiR labeling observed for SNAP-tag is in line with a previous report on differences of SNAP-tag and HaloTag7 labeling with SiR substrates²². In live-cell imaging, the different fluorophores showed some preferences among the different SNAP-tag substrates. For SiR, CF-SiR showed faster labeling in U2OS cells than TF-SiR and CP-SiR, even though all reached the same brightness after longer incubation periods. The labeling of SNAP-tag2 in live cells with TMR substrates and CPY substrates showed less dependency on the nature of the substrate (TF, CF and CP). The k_{app} values of CPY substrates for labeling SNAP-tag2 in vitro are all $>10^6\text{--}10^7 \text{ M}^{-1} \text{ s}^{-1}$ with TF-CPY being ~2-fold faster in comparison to CF-CPY and CP-CPY. The fact that the $t_{1/2}$ values of all CPY substrates for labeling in cells are rather similar suggests that the lower reactivity of CP-CPY and CF-CPY is offset by higher intracellular concentration of these substrates. This might be because of higher permeability and/or reduced binding to other cellular components. Thus, for the preparation of any new probes for live-cell applications with SNAP-tag2 or for new applications of the here-described probes, we recommend to first test the CF-based probes. The greater performance of SNAP-tag2 for fluorescence labeling in live cells also extends to super-resolution microscopy imaging techniques such as STED microscopy and to other cell types known to be refractory to chemical labeling such as yeast cells. As SNAP-tag2 together with its improved substrates is superior to the previously used SNAP-tag versions in every application we tested, we expect that it will also excel in in vivo applications. Lastly, implementation of the superior SNAP-tag2 system to already existing SNAP-tag-based biosensors such as Sniffts³⁵ should enhance their performance.

In summary, the improvements introduced in SNAP-tag2 are expected to advance the utility of this already widely adopted tool for live-cell imaging and other applications in life sciences.

Online content

Any methods, additional references, Nature Portfolio 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/s41589-025-01942-z>.

References

- Hinner, M. J. & Johnsson, K. How to obtain labeled proteins and what to do with them. *Curr. Opin. Biotechnol.* **21**, 766–776 (2010).
- Xue, L. et al. Imaging and manipulating proteins in live cells through covalent labeling. *Nat. Chem. Biol.* **11**, 917–923 (2015).
- Keppeler, A. et al. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2003).
- Mollwitz, B. et al. Directed evolution of the suicide protein O⁶-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* **51**, 986–994 (2012).
- Los, G. V. et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).
- Gautier, A. et al. An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15**, 128–136 (2008).
- Grimm, J. B. & Lavis, L. D. Caveat fluorophore: an insiders' guide to small-molecule fluorescent labels. *Nat. Methods* **19**, 149–158 (2022).
- Lukinavicius, G. et al. A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat. Chem.* **5**, 132–139 (2013).
- Wang, L. et al. A general strategy to develop cell permeable and fluorogenic probes for multicolour nanoscopy. *Nat. Chem.* **12**, 165–172 (2020).
- Encell, L. P. et al. Development of a dehalogenase-based protein fusion tag capable of rapid, selective and covalent attachment to customizable ligands. *Curr. Chem. Genomics* **6**, 55–71 (2012).
- Lavis, L. D. & Raines, R. T. Bright ideas for chemical biology. *ACS Chem. Biol.* **3**, 142–155 (2008).
- Lukinavicius, G. et al. Fluorogenic probes for live-cell imaging of the cytoskeleton. *Nat. Methods* **11**, 731–733 (2014).
- Grimm, J. B. et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods* **12**, 244–250 (2015).
- Lavis, L. D. Teaching old dyes new tricks: biological probes built from fluoresceins and rhodamines. *Annu. Rev. Biochem.* **86**, 825–843 (2017).
- Lardon, N. et al. Systematic tuning of rhodamine spirocyclization for super-resolution microscopy. *J. Am. Chem. Soc.* **143**, 14592–14600 (2021).
- Wilhelm, J. et al. Kinetic and structural characterization of the self-labeling protein tags HaloTag7, SNAP-tag, and CLIP-tag. *Biochemistry* **60**, 2560–2575 (2021).

17. Lavis, L. D. Chemistry is dead. Long live chemistry! *Biochemistry* **56**, 5165–5170 (2017).
18. Juillerat, A. et al. Directed evolution of O⁶-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol.* **10**, 313–317 (2003).
19. Gronemeyer, T. et al. Directed evolution of O⁶-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Eng. Des. Sel.* **19**, 309–316 (2006).
20. Juillerat, A. et al. Engineering substrate specificity of O⁶-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells. *ChemBioChem* **6**, 1263–1269 (2005).
21. Correa, I. R. Jr. et al. Substrates for improved live-cell fluorescence labeling of SNAP-tag. *Curr. Pharm. Des.* **19**, 5414–5420 (2013).
22. 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 (2019).
23. Frei, M. S. et al. Engineered HaloTag variants for fluorescence lifetime multiplexing. *Nat. Methods* **19**, 65–70 (2022).
24. Sun, X. et al. Development of SNAP-Tag fluorogenic probes for wash-free fluorescence imaging. *ChemBioChem* **12**, 2217–2226 (2011).
25. Kindermann, M. & Schwab, M. Pyrimidines reacting with O⁶-alkylguanine-DNA alkyltransferase fusion protein and method for detecting protein. US patent 8178314B2 (2012).
26. Goldenzweig, A. et al. Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell* **63**, 337–346 (2016).
27. Huang, P. S. et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS ONE* **6**, e24109 (2011).
28. Wei, H. & Li, X. Deep mutational scanning: a versatile tool in systematically mapping genotypes to phenotypes. *Front. Genet.* **14**, 1087267 (2023).
29. Hell, S. W. Nanoscopy with focused light (Nobel lecture). *Angew. Chem. Int. Ed. Engl.* **54**, 8054–8066 (2015).
30. McMurray, M. A. & Thorner, J. Septin stability and recycling during dynamic structural transitions in cell division and development. *Curr. Biol.* **18**, 1203–1208 (2008).
31. Jansen, R. L. M. & van der Klei, I. J. The peroxisome biogenesis factors Pex3 and Pex19: multitasking proteins with disputed functions. *FEBS Lett.* **593**, 457–474 (2019).
32. de Lange, E. M. F. et al. STED super-resolution microscopy unveils the dynamics of Atg30 on yeast Pex3-labeled peroxisomes. *iScience* **27**, 110481 (2024).
33. Huppertz, M. C. et al. Recording physiological history of cells with chemical labeling. *Science* **383**, 890–897 (2024).
34. Grimm, J. B. et al. A general method to optimize and functionalize red-shifted rhodamine dyes. *Nat. Methods* **17**, 815–821 (2020).
35. Sallin, O. et al. Semisynthetic biosensors for mapping cellular concentrations of nicotinamide adenine dinucleotides. *eLife* **7**, e32638 (2018).
36. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

Methods

Chemical synthesis and general methods

Detailed procedures for the synthesis and chemical characterization of all compounds are given in the Supplementary Methods. Buffer, media and reagent compositions can be found in Supplementary Table 13.

Computed chemical properties of SNAP-tag substrates

Molecular structures were prepared using the Schrödinger Maestro 12.3 interface to ensure suitability for QikProp calculations. The preparation included adding hydrogens, assigning bond orders and generating three-dimensional geometries. QikProp calculations were conducted using the Schrödinger suite. For each structure, key pharmacokinetic and physicochemical properties were calculated using default settings and according to the official protocol (<https://www.schrodinger.com/platform/products/qikprop/>). Calculated parameters were exported and are shown in Supplementary Table 2.

Plasmids for bacterial and mammalian protein expression

A pET51b(+) vector (Novagen) was used for protein production in *Escherichia coli*. Proteins were N-terminally Strep-tagged and C-terminally 10xHis-tagged for affinity purification. pcDNAs/FRT or pcDNAs/FRT/TO vectors (Thermo Fisher Scientific) were used for protein expression in mammalian cells by transient transfection or for stable cell line generation. Cloning was performed with Gibson assembly³⁷ or point mutations were introduced using the Q5-site-directed mutagenesis kit (New England Biolabs) according to the manufacturer protocol. Plasmid sequences were verified by Sanger sequencing (Eurofins Genomics or Microsynth) and plasmids were stored at -20°C .

Recombinant protein production and purification

pET51b(+) plasmids were transformed in *E. coli* BL21(DE3)-pLysS (Novagen) cells. Luria–Bertani (LB) ampicillin expression cultures were grown at 37°C to an optical density at 600 nm (OD_{600}) of 0.6–0.8 and protein expression was subsequently induced at 16°C by addition of IPTG (0.5 mM). After overnight expression, cells were harvested by centrifugation (4,000g, 4°C , 15 min), resuspended in His-tag extraction buffer (20–30 ml) and lysed by sonication (50% duty, 70% power, 7 min) on wet ice. All proteins were purified by immobilized metal ion affinity chromatography using either a gravity column or an ÄktaPure fast protein liquid chromatography (FPLC) instrument (Cytiva) equipped with a HisTrap FF crude column (Cytiva). Buffer was exchanged to activity buffer (50 mM HEPES and 50 mM NaCl, pH 7.3) either using Zeba spin desalting columns (7-kDa molecular weight cutoff (MWCO), 0.5 ml; Thermo Fisher Scientific) or HiPrep 26/10 desalting columns (Cytiva) on an FPLC system. Purified proteins were concentrated using Amicon Ultra centrifugal filters (10-kDa MWCO). The purity and correct size of the proteins were assessed by SDS–PAGE and high-resolution mass spectrometry (HRMS). Purified proteins were aliquoted, flash-frozen in liquid nitrogen and stored at -80°C .

Plasmids for YSD

For protein expression on the yeast surface, either a pCTcon2 vector (Addgene, 73152) with SNAP-tag proteins fused in between Aga2p–HA-tag (N-terminal) and MYC-tag (C-terminal) or a pJYDNg vector (Addgene, 162452) with a C-terminal fusion to Aga2p–HA-tag–MYC-tag–eUnaG2 was used. Epitope HA-tags and MYC-tags were used for immunostaining and the fluorescent reporter protein eUnaG2 was labeled with bilirubin for monitoring the protein expression level.

DNA library preparation for YSD

Libraries were created either using pCTcon2–SNAP-tag1.1 (SNAP–E30R–S135R–L153G) through one-pot saturation mutagenesis³⁸ or using pJYDNg–SNAP-tag1.5 (SNAP–E30R–I32Y–L34V–K36N–⁵⁰GHPEPQ–S135R–L153G–G161P) combining site-directed saturation mutagenesis with assembly PCR³⁹.

Library generation was initially performed by obtaining circular library plasmids from *E. coli*, which were then used for transformation into EBY100 yeast cells for YSD. For libraries generated through one-pot saturation mutagenesis, a mixture of degenerate primers was used, in which one or four NNKs covered the target regions for randomization (amino acids H29–K36 and 156–161; Supplementary Tables 15–17). Library DNA was transformed into *E. coli* cells and cells were recovered in 2 ml of SOC medium (New England Biolabs) for 1 h at 37°C . Serial dilutions were spread on selective LB agar plates containing kanamycin ($50\ \mu\text{g}\ \text{ml}^{-1}$) to determine the library size. Plates were incubated at 37°C overnight and isolated plasmids from ten single-picked colonies were sequenced to control the library diversity. Cells transformed with the libraries were grown in 3 ml of LB kanamycin medium at 37°C overnight and plasmids were extracted using a QIAprep Spin Miniprep kit (Qiagen).

Libraries were later generated as linear DNA fragments used for homologous recombination (HR) in yeast. Libraries on the active site loop of pJYDNg–SNAP-tag1.5 were generated using degenerated primers with three NNKs at different positions throughout the targeted region (156–GGYEGP–161). The N-terminal and C-terminal parts of SNAP-tag1.5 were PCR-amplified and used for assembly PCR in a mixture of primers carrying randomized positions surrounded by overlapping regions with the N-terminal and C-terminal fragments (Supplementary Table 18). This assembly PCR enabled the reconstitution of full-length SNAP-tag1.5 library genes, whose extremities overlapped with the plasmid by 50–80 bp to enable HR in yeast. The protocol for assembly PCR was adjusted from Routh et al.³⁹. Library inserts were purified using a QIAquick PCR purification kit (Qiagen).

An sDMSL designed on the SNAP-tag1.5 scaffold was purchased from Twist Bioscience. The library consisted of single-point variants, having each amino acid along the SNAP-tag1.5 sequence substituted to all other 19 amino acids, excluding cysteines and including deletions. The sDMSL was designed with a 50-bp overlap of the flanking regions to the pJYDNg expression vector for direct HR into yeast cells.

Electrocompetent yeast cell preparation for YSD

Single colonies of *Saccharomyces cerevisiae* strain EBY100 (American Type Culture Collection) spread on a YPD agar plate were used for inoculation of YPD liquid cultures. YPD expression cultures were grown at 250 rpm and 30°C until an OD_{600} of 1.3–1.5. Afterward, Tris–DTT buffer (800 μl) and Tris–lithium acetate buffer (2 ml) were added and the culture was incubated at 250 rpm and 30°C for 15 min. Cells were harvested by centrifugation (2,500g, 4°C , 3 min) and washed with ice-cold electroporation buffer (25 ml), followed by cell harvesting (2,500g, 4°C , 3 min). Cells were resuspended in electroporation buffer to a total volume of 300 μl . Cells were aliquoted (50 μl) and directly used for transformation by electroporation.

Yeast cell transformation and HR for YSD

For circular plasmid DNA transformation, DNA (1–3 μg) was diluted in double-distilled (dd) H_2O (<10 μl) and kept on ice. For HR, amplified vector (1 μg) and insert DNA (2–3 μg) were mixed to a final volume of 100 μl . The DNA was concentrated using alcohol precipitation (100 μl of isopropanol 100 μl , 10 μl of 3 M sodium acetate and 0.5 $\mu\text{g}\ \mu\text{l}^{-1}$ glycogen (Thermo Fisher Scientific), incubated at -20°C overnight). The mixture was centrifuged (10,000g, 20 min) and the DNA pellet was washed with 70% ethanol (200 μl) followed by centrifugation (10,000g, 20 min). The supernatant was removed and the pellet was air-dried. Electrocompetent EBY100 yeast cells (50 μl) were used to resuspend the DNA and transformation was conducted by electroporation. Electroporation cuvettes (GenePulser cuvette, 0.2-cm electrode gap; Bio-Rad Laboratories) were prechilled on ice. The cell–DNA mix was transferred to the prechilled electroporation cuvette and electroporation was conducted using a Bio-Rad GenePulser Xcell device (0.54 kV, 25 μF , infinite resistance with exponential decay). After electroporation, prewarmed

YPDS medium (1 ml) was immediately added to the cuvette and the cell suspension was transferred to culture tubes (15 ml). The cuvette was washed with 1 ml of YPDS medium and added to the culture. The culture was incubated at 250 rpm and 30 °C for 1 h. Yeast cells were collected by centrifugation (2,500g, 5 min) and resuspended in SDCAA selective medium (1 ml). Serial dilutions were plated on SDCAA agar plates to determine the transformation efficiency. Plates were incubated at 30 °C for 2–3 days. The remaining cell suspension was used to inoculate SDCAA culture (100 ml) and grown at 30 °C for 2 days. The culture was directly used for YSD screening and aliquots were frozen in a 1:1 mixture of culture and 50% glycerol in Tris buffer (pH 8).

Protein expression on the yeast surface

SDCAA medium (5 ml final volume) was inoculated either using a single colony from a SDCAA plate or by addition of fully grown yeast liquid culture (0.5 ml) after transformation. The culture was incubated at 250 rpm and 30 °C overnight. This preculture (0.5 ml) was used to inoculate 4.5 ml of SGCAA medium for protein expression on the yeast surface. Protein expression was conducted for at least 20 h at 250 rpm and 30 °C.

Protein labeling on yeast surface for FACS

For fluorescence labeling, 10^7 cells were obtained by centrifugation (14,000g, 1 min). The concentration of the yeast culture expressing the protein(s) of interest on the surface was determined by accounting that an $OD_{600} = 1$ corresponds to about 10^7 cells per ml (ref. 40). For antibody-based expression staining of yeast cells transformed with pTcon2 plasmids, yeast cells were resuspended in 1:10-diluted primary mouse anti-MYC antibody (OP10, EMD Millipore) in PBS (50 μ l) and incubated on a rotating wheel at 4 °C for 1 h. The cells were pelleted (14,000g, 1 min) and washed twice with PBS (125 μ l) with centrifugation in between (14,000g, 1 min). The cell pellet was resuspended in 1:50-diluted secondary goat anti-mouse–Alexa647 antibody (A-21236, Invitrogen, Thermo Fisher Scientific) in PBS (50 μ l) and incubated on a rotating wheel at 4 °C for 1 h. The cells were washed twice with PBS (125 μ l) before labeling with SNAP-tag substrates.

For protein libraries encoded by the pJYDNg expression vector, expression control was monitored over labeling of eUnaG2 with bilirubin. Yeast cells were resuspended in PBS (50 μ l) containing bilirubin (10 μ M) and BSA (1 mg ml⁻¹), gently vortexed and incubated on ice for 10 min. The cells were pelleted and washed twice with PBS (150 μ l) interspaced by centrifugation steps (14,000g, 1 min) before SNAP-tag labeling. Cells expressing SNAP-tag variants were labeled in PBS (50 μ l) using different substrates (CF₃P–TMR, CF₃P–MaP618, TF–TMR or TF–MaP618) with varying concentrations (10–500 nM) and incubation times (10–60 min) to adjust the screening stringency. Labeling was performed at room temperature on a rotating wheel. Cells were washed with PBS (125 μ l), resuspended in 1 ml of PBS and filtered through 5-ml round-bottom polystyrene test tubes with cell strainer snap caps (352235, Falcon) for FACS.

Protein library screening using FACS

Labeled yeast libraries were analyzed and subsequently sorted on a BD FACSMelody cell sorter (BD Biosciences) using the appropriate filter settings for each fluorophore (Supplementary Table 7). For yeast cell sorting, a 100- μ m sorting nozzle and a 1.5 neutral density filter was used. Photomultiplier tube (PMT) detector voltages were optimized on the basis of single-labeled and double-labeled samples and negative control samples. The range for positive signal was set to be approximately 10^3 – 10^4 , while negative and background staining was set to $\leq 10^2$. Gating strategies are depicted in Supplementary Fig. 15. Cells were gated for live and single-cell events on the basis of their size in the forward scatter (FCS) and side scatter (SSC). The single-cell population was then gated for the double-positive labeling signal and the top 0.5–1% of double-labeled cells (10^4 cells) were sorted in bulk.

After sorting, yeast cells were grown in SDCAA medium (5 ml) supplemented with penicillin–streptomycin (Gibco; 5,000 U per ml, 1:100 dilution) for 2 days. Plasmids were isolated from fully grown cultures (1 ml) using a Zymoprep yeast plasmid miniprep II kit (Zymo Research) according to the manufacturer's protocol and the remaining culture was propagated to start the next screening round.

Computational redesign of the unstructured loop in SNAP-tag

To redesign the unstructured loop (residues 37–54), the TMR-labeled SNAP-tag structure (PDB 6Y8P)¹⁶ was used as input. The covalent benzyl-TMR ligand was parameterized and a Rosetta constraint file was set up to fix the covalent bond between the ligand the protein analogous to published procedures³³. Point substitutions E30R, I32Y, L34V, K36N, S135R, L153G and G161P were modeled and the structure was minimized using a RosettaScripts⁴¹ protocol. The top-scoring structure was used as input for the loop design using the RosettaRemodel²⁷ application. Blueprint files for 7-aa, 8-aa and 9-aa loops were set up, allowing any amino acids except cysteine in the designed regions. For energy minimization, the Rosetta FastRelax method was applied using the 'ref2015' Rosetta score function and the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm with Armijo inexact line search conditions ('lbfgs_armijo_nonmonotone'). Example scripts can be found in the Supplementary Information. The top five scoring sequences for each loop length were tested experimentally.

Next-generation sequencing

Randomized regions from isolated plasmids were amplified with next-generation sequencing (NGS) primers (Supplementary Table 19) using standard PCR conditions and purified (Qiagen PCR purification kit). NGS was performed using adaptor ligation Illumina sequencing at Eurofins Genomics. The NGS package for each sample included 10 million total reads comprising of 2 \times 150-bp paired-end reads (5 million read pairs). NGS amplicon sizes varied from 150 to 300 bp and were pooled in equivalent quantities. The NGS primers' unique adaptor sequences enabled later reattribution of sequences to sorted libraries. NGS samples comprised 2 μ g of DNA in 100 μ l of ddH₂O.

Analysis of NGS data

Illumina sequencing results were first split into separate files on the basis of their experiment-specific and selection-round-specific sequence barcodes using 'je demultiplex' (ref. 42). Forward and reverse reads were combined into a single file for each barcode. Primer and adaptor sequences were trimmed to match the translation frame of the template sequence using cutadapt⁴³. Trimmed reads were then aligned to the template sequence with bowtie2 (ref. 44). The resulting SAM file was converted to a BAM file using SAMtools⁴⁵. The alignments were then analyzed with a custom R script to calculate amino acid frequencies for randomized positions (for saturation mutagenesis libraries) or point mutation frequencies (for sDMSL). Reads containing deletions, insertions relative to the template sequence or ambiguous bases were removed. The remaining reads were translated into amino acid sequences. For saturation mutagenesis libraries, amino acid frequencies at the positions of interest were calculated by dividing the number of reads featuring a specific amino acid at a particular position by the total number of reads. For DMSLs, mutation frequencies were calculated by dividing the number of reads featuring a specific mutation by the number of reads that cover the corresponding position. The script used for NGS data analysis can be found in the Supplementary Information.

Protein thermal stability analysis

Thermal stabilities of final protein variants were measured on a Prometheus NT48 nanoscale differential scanning fluorimeter (NanoDSF). Protein samples were prepared in activity buffer (0.8 mg ml⁻¹) and changes in tryptophan fluorescence at 330 nm were followed over a temperature range of 20–95 °C with a temperature increase

of 1 °C min⁻¹. Measurements were performed in technical triplicates. The inflection point of the first derivative corresponds to the proteins' melting temperature.

In vitro labeling kinetics measured using a microplate reader fluorescence polarization assay

Labeling kinetics were measured in black nonbinding flat-bottom 96-well plates (200- μ l final reaction volume) or in black nonbinding low-volume 384-well plates (20- μ l final reaction volume) (Corning) on a microplate reader (Spark 20M, Tecan) by recording fluorescence polarization (FP) traces at 37 °C over time. All measurements were performed in technical triplicates in FP buffer. Labeling reactions were started by either addition of substrate to the protein using a multichannel pipette or using the injector module of the plate reader for fast reaction conditions. Filter settings were chosen according to the fluorophore substrate (Supplementary Table 8). The G factors were calculated using a buffer only (blank) and free fluorophore substrate (reference) control. The optical gain, an amplification factor for the photomultiplier tube, was adjusted to 50%. Baselines were determined by recording FP kinetic traces of the free fluorophore substrates.

For screening purposes, labeling kinetics were recorded using 50 nM protein and 20 nM fluorophore substrate. For fast labeling kinetics, FP kinetics were conducted with 10 nM protein and 2 nM substrate in 384-well plates. A one-phase association (Eq. (2)) was fitted to the data using GraphPad Prism and k_{app} was calculated using Eq. (1).

$$Y = Y_0 + (Y_{max} - Y_0) \times (1 - e^{-kx}) \quad (1)$$

where Y is the FP (mP), Y_{max} is the FP plateau (mP), Y_0 is the y-intercept (mP), k is the labeling rate constant (s⁻¹) and x is time (s).

$$k_{app} = \frac{k}{[\text{protein}]} \quad (2)$$

For more accurate determination of k_{app} , a global fit approach to kinetic models was used as published in Wilhelm et al.¹⁶. In brief, kinetics were measured at fixed fluorophore substrate concentration (20 nM) and various protein concentrations (0 and 4.60–900 nM). Fast kinetics were measured at fluorophore substrate concentrations of 2 nM and protein concentrations of 0 and 0.78–50 nM. Kinetic data were preprocessed using a custom R script^{46,47}; subsequently, kinetic model 1 (Eq. (3)), described by the differential equations (Eqs. (4)–(6)), was fitted to the data using DynaFit⁴⁸. The FP baseline value and protein concentration were fixed parameters, while the fluorophore substrate concentration was adjustable to account for potential quantification errors. Including limiting protein concentration into the experimental conditions enabled accurate concentration fitting by impacting the final FP value (decreasing plateau). The s.d. and confidence interval (CI) were obtained using a Monte Carlo simulation⁴⁹ with standard settings ($n = 1,000$, 5% worst fits excluded). Exemplified DynaFit scripts can be found in the Supplementary Information. Representative graphics were generated using a custom R script (Supplementary Fig. 2).



where P is the SNAP-tag2 protein, S is the fluorophore substrate and PS is the fluorescently labeled SNAP-tag2.

$$\frac{d[P]}{dt} = -k_{app}[P][S] \quad (4)$$

$$\frac{d[S]}{dt} = -k_{app}[P][S] \quad (5)$$

$$\frac{d[PS]}{dt} = k_{app}[P][S] \quad (6)$$

Labeling kinetics of nonfluorescent substrates were determined in a competition assay as described in Wilhelm et al.¹⁶. In brief, competition kinetics were measured by mixing varying concentrations of nonfluorescent substrates (0 and 0.1–1 μ M) and BG–Alexa488 (100 nM) followed by addition of protein (200 nM) using an electronic 96-channel pipettor (Integra Bioscience). FP changes were recorded over time. Kinetic data were preprocessed using a custom R script^{46,47}; subsequently, a competition kinetic model (Eq. (7)), described by the differential equations (Eqs. (8)–(12)), was fitted to the data using DynaFit⁴⁸. The s.d. and CI were obtained using a Monte Carlo simulation⁴⁹ with standard settings ($n = 1,000$, 5% worst fits excluded). Exemplified DynaFit scripts can be found in the Supplementary Information. Representative graphics were generated using a custom R script (Supplementary Fig. 10).



where P is the SNAP-tag2 protein, S is the nonfluorescent substrate, PS is the nonfluorescently labeled SNAP-tag2 protein, I is the fluorescent competitor substrate and PI is the fluorescently labeled SNAP-tag2 protein.

$$\frac{d[P]}{dt} = -k(S)_{app}[P][S] - k(I)_{app}[P][I] \quad (8)$$

$$\frac{d[S]}{dt} = -k(S)_{app}[P][S] \quad (9)$$

$$\frac{d[I]}{dt} = -k(I)_{app}[P][I] \quad (10)$$

$$\frac{d[PS]}{dt} = k(S)_{app}[P][S] \quad (11)$$

$$\frac{d[PI]}{dt} = k(I)_{app}[P][I] \quad (12)$$

In vitro labeling kinetics of SNAP-tag2 using a stopped-flow device

Labeling kinetics of SNAP-tag2 with TF–TMR, CF–TMR, CP–TMR, TF–CPY, CF–CPY and CP–CPY substrates were measured by recording fluorescence anisotropy changes over time using a BioLogic SFM-400 stopped-flow instrument (BioLogic Science Instruments) in a single-mixing configuration at 37 °C as described in Wilhelm et al.¹⁶. The final substrate concentration was fixed to 0.5 or 1 μ M and the protein concentration was varied from 0.25 to 3.13 μ M in activity buffer (50 mM HEPES and 50 mM NaCl, pH 7.3) supplemented with 1 mM DTT and 0.1 mg ml⁻¹ BSA. The anisotropy of the free substrate was measured to obtain a baseline. The sampling time was set to 1 ms for a total recording time of 2 s or varied from 200 μ s for 0.2 s followed by 2 ms to a total duration of 3 s. For each condition, 15 technical replicates were recorded. Kinetic data were preprocessed using a custom R script^{46,47}, in which replicate values were averaged, pretrigger timer points were removed and times were adjusted to start at zero. Kinetic model 2 (Eq. (13)), described by differential equations (Eqs. (14)–(17)) was fitted to the data using DynaFit⁴⁸ as described in Wilhelm et al.¹⁶. The delay time, anisotropy baseline value and protein concentration were fixed parameters, while the fluorophore substrate concentration was adjustable to account for potential quantification errors. Including limiting protein concentration into the experimental conditions enabled accurate concentration fitting by impacting the final anisotropy value (decreasing plateau). The s.d. and CI were obtained using a

Monte Carlo simulation⁴⁹ with standard settings ($n = 1,000$, 5% worst fits excluded). Exemplified DynaFit scripts can be found in the Supplementary Information. Representative graphics were generated using a custom-made R script (Supplementary Figs. 6–9 and 11).



where P is the SNAP-tag2 protein, S is the fluorophore substrate, PS^* is the substrate-bound protein complex and PS is the fluorescently labeled SNAP-tag2.

$$\frac{d[P]}{dt} = -k_1[P][S] + k_{-1}[PS^*] \quad (14)$$

$$\frac{d[S]}{dt} = -k_1[P][S] + k_{-1}[PS^*] \quad (15)$$

$$\frac{d[PS^*]}{dt} = k_1[P][S] - k_{-1}[PS^*] - k_2[PS^*] \quad (16)$$

$$\frac{d[PS]}{dt} = k_2[PS^*] \quad (17)$$

The individual kinetic parameters for the dissociation constant (K_d) defining the protein affinity toward the substrate and the k_{app} defining the overall labeling speed were derived from Eqs. (18) and (19), respectively.

$$K_d = \frac{k_{-1}}{k_1} \quad (18)$$

$$k_{app} = k_1 \frac{k_2}{(k_2 + k_{-1})} \quad (19)$$

Multiple replicates of the experiment were conducted for TF-TMR (four replicates), CF-TMR (eight replicates) and CP-TMR (three replicates) and their corresponding CPY analogs (two replicates each). Results of the individual kinetic measurements are depicted in Supplementary Table 3. Replicates were averaged and the s.d. was calculated (Table 1 and Supplementary Table 4).

Photophysical properties of labeled SNAP-tag2 and SNAP_F-tag

For determination of the extinction coefficients (ϵ) of labeled SNAP-tag proteins, recombinant SNAP-tag2 and SNAP_F-tag (30 μ M) were labeled with fluorophore substrates (7.5 μ M) in FP buffer overnight. For measurements, a serial dilution of the labeling reaction was performed, targeting final substrate concentrations of 7.5–2.22 μ M. Absorbance spectra were recorded on a V-770 spectrophotometer (Jasco) in a quartz cuvette (3 mm). All measurements were performed in technical triplicates and absorbance values were baseline-corrected by setting the absorbance at 800 nm to zero using SpectraGryph version 1.2. For all fluorophores, the maximal absorbance value at each concentration was fitted to a linear function (Eq. (20)) and extinction coefficients were calculated from the slope b (Eq. (21)).

$$y(x) = a + bx \quad (20)$$

$$\epsilon = \frac{b}{0.3} \quad (21)$$

Absolute QYs of fluorescently labeled SNAP-tag2 and SNAP_F-tag were measured on a Quanta-QY spectrometer (model C11347,

Hamamatsu). Measurements were carried out with diluted samples from the same reaction solution used for ϵ measurements ($A = 0.03$ – 0.08).

Mammalian cell culture

All media were filtered before using sterile Nalgene Rapid-Flow filter units (0.22- μ m pore size, 500 ml; Thermo Fisher Scientific). All experimental steps were performed under sterile conditions. Mammalian U2OS Flp In T-REX cells⁵⁰ (U2OS cells) or HeLa Kyoto Flp In cells³³ (HeLa cells) were cultured in cell growth medium (DMEM + GlutaMAX supplemented with phenol red, high glucose (4.5 g l⁻¹), pyruvate (Gibco Life Technologies, 31966-021) and 10% FBS (Gibco Life Technologies)) in either T-25 or T-75 culture flasks (Sarstedt). The cells were cultivated in a humidified incubator at 37 °C and 5% CO₂ and passaged when reaching 80–90% confluency. Cell handling involved washing with PBS (1 \times , pH 7.4; Gibco Life Technologies, 10010-015), detachment (trypsin 0.05% + EDTA; TrypLE Express, Gibco Life Technologies, 12604-013), resuspension in DMEM + 10% FBS and transfer into fresh flasks. For imaging experiments, cells were handled in DMEM without phenol red, referred to as imaging medium (Gibco Life Technologies, 31053-028) supplemented with GlutaMAX (Gibco Life Technologies, 35050-038), pyruvate (Gibco Life Technologies, 11360-070) and 10% FBS (Gibco Life Technologies). For seeding cells at desired densities, cells were counted using a fluidlab R-300 cell counter (Anvajo) with Acella 100 chambers (20- μ l sample volume; Anvajo). Cells were seeded in 96-well tissue culture plates (TPP, 92096) or 8-well or 96-well microplates with a glass bottom (Ibidi, 80827 and 89627, respectively) in a total volume of 100–150 μ l of medium.

Generation of mammalian stable cell lines

Stable cell lines were generated using the Flp In T-REX system (Thermo Fisher Scientific) with U2OS or HeLa cells. Cells were grown to ~80% confluency in T-25 flasks. Lipofectamine 3000 (8 μ l; Thermo Fisher Scientific, L3000008) was diluted in Opti-MEM I reduced serum medium + GlutaMAX (200 μ l; Thermo Fisher Scientific, 51985026) in a 1.5-ml reaction tube and thoroughly mixed. In a separate tube, the plasmid DNA encoding the gene of interest on a pcDNA/FRT or pcDNA/FRT/TO vector (440 ng) and pOG44 (3,560 ng; encoding the Flp In recombinase) were diluted in Opti-MEM I reduced serum medium + GlutaMAX (200 μ l). P3000 (8.0 μ l) was added to the DNA mix. After thorough mixing, the DNA-P3000 mix incubated at room temperature for 10 min. Subsequently, the DNA-P3000 mix was combined with the diluted Lipofectamine 3000, mixed and left to incubate at room temperature for an additional 15 min. The resulting DNA-P3000-Lipofectamine 3000 mix was added to the T-25 flask and incubated overnight. The following day, the medium was exchanged with cell growth medium supplemented with 100 μ g ml⁻¹ hygromycin B (Gibco Life Technologies), selecting cells for plasmid integration. Selection was performed for 48–72 h. Survived cells were recovered in cell growth medium to reach confluency. Cells were sorted using FACS for high expression levels according to a positive fluorescent protein signal (mEGFP, FITC filter) or SLP labeling (SiR, APC filter). Generated stable cell lines can be found in Supplementary Table 14.

rAAV production and transduction

For dual-color STED imaging, recombinant adeno-associated viruses (rAAVs) were generated as described previously⁵¹. In brief, the AAV plasmid containing the CAG-Lifect-HT7 expression cassette flanked by AAV2 packaging signals (inverted terminal repeats) were cotransfected with plasmids pRV1 (AAV2 Rep and Cap sequences), pH21 (AAV1 Rep and Cap sequences) and pFD6 (Adenovirus helper plasmid) using Lipofectamine 3000 into HEK293 cells. Then, 5 days after transfection, the cells were collected and lysed using TNT extraction buffer (20 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100 and 10 mM MgCl₂) for 10 min. The cell debris was removed by centrifugation (3,000g, 5 min, 4 °C),

the cell supernatant was treated with Benzonase (4.5 μ l, 30 min, 37 °C) and the rAAVs were purified by FPLC using HiTrap AVB Sepharose HP columns (Cytiva, 28411211). Purified rAAVs were concentrated using Amicon Ultra 15-ml (100-kDa MWCO) centrifugal filters (Merck) with buffer exchange to PBS pH 7.3.

U2OS F1p In T-REx cells stably expressing Vim-SNAP-tag2 were transduced with rAAVs 18 h before STED imaging by adding 0.5 μ l of purified rAAVs ($\sim 10^9$ – 10^{10} rAAV particles) into 200 μ l of imaging medium (DMEM GlutaMAX and 10% FBS, phenol red free; Gibco).

Live-cell labeling performance of SNAP-tag proteins determined with flow cytometry

Flow cytometry was conducted on a BD Fortessa X-20 cell analyzer (BD Biosciences) equipped with a high-throughput screening (HTS) module for 96-well plates. Cells were seeded on transparent 96-well cell culture plates and treated according to desired experiment in a reaction volume of 100 μ l. After treatment, labeling reaction was stopped by addition of recombinant SNAP-tag2 (2 μ M, 100 μ l, 10-min incubation), cells were washed twice with PBS (150 μ l, 10-min incubation), trypsinized (50 μ l of trypsin, 10-min incubation) and resuspended in FACS buffer (2% FBS in PBS) to a final volume of 200 μ l. The cell suspension was transferred to nonbinding U-bottom 96-well plates (Falcon) and analyzed on a flow cytometer using the HTS module. Flow rates were set to 3 μ l s^{-1} with three mixing steps and a subsequent washing volume of 400 μ l. Cells were gated for live cells (SSC-A/FSC-A) and single cells (FSC-H/FSC-A) as exemplified in Supplementary Fig. 16. Excitation laser and emission filter settings were adjusted according to the fluorophore properties (Supplementary Table 9). PMT settings were adjusted to positive and negative controls. Data were analyzed using FlowJo software (BD Biosciences).

Substrate screening for SNAP-tag labeling in live mammalian cells

U2OS F1p In T-REx (Thermo Fisher Scientific) cells stably expressing mEGFP-SNAP-tag or mEGFP-CLIP-tag fusion proteins were seeded into 96-well cell culture plates (10,000 cells per well) 1 day before the experiment. The cells were then incubated with substrates 1–30, CP or BC (100 nM) for 2 h at 37 °C. All substrates were tested in technical triplicates. Cells were washed twice with cell growth medium for 150 min incubation at 37 °C and once with sterile PBS pH 7.4 before detachment with trypsin (50 μ l, 10 min, 37 °C). Cells were resuspended in FACS buffer to a final volume of 200 μ l. The cell suspension was transferred to nonbinding U-bottom 96-well plates (Falcon) and analyzed on a flow cytometer using the HTS module, as previously described. Cells were gated for live-cell (SSC-A/FSC-A) and single-cell events (FSC-H/FSC-A) and for mEGFP expression signal (FITC channel) $\geq 10^3$. Gated cells were analyzed for their SLP labeling signal (TMR, PE channel) over expression level (mEGFP) and the medians of the ratios were derived using FlowJo software (BD Biosciences).

Cell viability assay

U2OS cells were seeded on transparent 96-well cell culture plates 1 day before the experiment. Cells were incubated with TF-fluorophore, CF-fluorophore and CA-fluorophore substrates (1 μ M) or DMSO (1% (v/v)) or remained untreated for 1 h at 37 °C. The medium was collected into nonbinding U-bottom 96-well plates (Falcon) and detached cells were harvested by centrifugation (3,000g, 5 min). The supernatant was removed. Additionally, adherent cells were detached with trypsin (30 μ l, 10 min, 37 °C), resuspended in FACS buffer (2% FBS in PBS) to a final volume of 100 μ l and added to the same wells to collect all dead and live cells. SYTOX blue dead cell stain (100 μ l, 2 μ M; Thermo Fisher Scientific) was added to the cells to a final concentration of 1 μ M and cells were subsequently analyzed by flow cytometry (10,000 cells per well; laser: 405 nm, BP filter: 450/50). The experiment was performed in technical triplicates. Data were analyzed using FlowJo software (BD Biosciences).

Confocal fluorescence microscopy of SNAP-tag2 in live mammalian cells

Confocal fluorescence microscopy of U2OS cells stably expressing HaloTag7-P30-SNAP-tag2/SNAP_F-tag-NLS-P2A-NLS-mTurquoise2 was conducted on a Stellaris 5 microscope (Leica Microsystems). Live-cell imaging was performed in 8-well or 96-well glass bottom dishes (Ibidi) at 37 °C and 5% CO₂ in a humidified chamber. Images were acquired with a $\times 40$ (HC PL APO CS2 $\times 40/1.10$) water objective, a scan speed of 600 Hz, line average of 2, optical zoom of 1–1.3, image size 1,024 \times 1,024 μ m and 12-bit pixel depth. Z-stacks were recorded with a step size of 2–5 μ m. HaloTag7 and SNAP_F-tag/SNAP-tag2 were labeled with their respective TF-fluorophore or CA-fluorophore substrates at 100 nM overnight. Cells were washed twice with PBS before imaging in imaging medium. Laser settings are described in Supplementary Table 10. Image analysis was performed using Fiji software⁵². For visualization purposes, maximum intensity projections of z-stacked images were calculated. For quantification, intensities of multiplane images were summed and regions of interest (ROIs) were defined manually. Mean fluorescence intensities of the ROI's fluorescence channels were calculated (multi-ROI measurement) and ratios calculated from fluorescent labels to respective expression signals were compared. Ratio-metric projections were generated using the BRET-analyzer plugin⁵³.

In-cell kinetic measurements of SNAP-tag2 and SNAP_F-tag

In-cell kinetic measurements of SNAP-tag2 and SNAP_F-tag labeling with TF-fluorophore, CF-fluorophore and CP-fluorophore substrates were performed in U2OS cells stably expressing HaloTag7-P30-SNAP-tag2/SNAP_F-tag-NLS-P2A-NLS-mTurquoise2 on a Stellaris 5 microscope as previously described. Fluorophore substrates were added at a final concentration of 50 nM (for TMR and CPY) or 100 nM (for SiR) to the cells and images with three z-stacks were acquired every 30 s over time. For quantification, fluorescence intensities of multiplane images were summed and the ratios of fluorophore substrate signals normalized to mTurquoise2 expression signals were analyzed using CellProfiler⁵⁴. Additional image acquisition parameters were as follows: $\times 20$ water objective (HC PL APO CS2 $\times 20/0.75$ IMM), $\times 1.28$ optical zoom, image size 1,024 \times 1,024 μ m, scan speed of 400 Hz and line average of 2. Laser settings are described in Supplementary Table 11. Kinetic data were analyzed by fitting a sigmoidal curve (Eq. (22)) in GraphPad Prism 10.2.3.

$$Y = Y_0 + \frac{(Y_{\max} - Y_0)}{\left(1 + \left(\frac{t_{1/2}}{x}\right)^H\right)} \quad (22)$$

where Y is the fluorescence intensity ratio (signal/mTurquoise2), Y_{\max} is the plateau of the fluorescence intensity ratio, Y_0 is the y-intercept, H is the Hill slope and x is time (min).

The experiment was conducted in biological duplicates and the average $t_{1/2}$ of both replicates was calculated. Respective CIs were propagated by Eqs. (23–26) as follows:

1. Calculate the mean $t_{1/2}$ of both replicates:

$$\overline{t_{1/2}} = \frac{(t_{1/2})_1 + (t_{1/2})_2}{2} \quad (23)$$

2. Compute the standard error (SE) for each $t_{1/2}$ of a replicate:

$$SE_{(t_{1/2})_1} = \frac{CI_{high} - CI_{low}}{2 \times 1.96}$$

$$SE_{(t_{1/2})_2} = \frac{CI_{high} - CI_{low}}{2 \times 1.96} \quad (24)$$

3. Calculate the s.e.m.:

$$SE_{\overline{(t_{1/2})}} = \sqrt{\frac{SE^2_{(t_{1/2})_1} + SE^2_{(t_{1/2})_2}}{2}} \quad (25)$$

4. Compute the 95% CI for the average $t_{1/2}$:

$$CI_{(t_{1/2})} = \overline{t_{1/2}} \pm 1.96 \times SE_{(t_{1/2})} \quad (26)$$

where $\overline{t_{1/2}}$ is the $t_{1/2}$ of the duplicate, CI_{high} is the higher 95% CI and CI_{low} is the lower 95% CI.

Live-cell STED microscopy in mammalian cells

HeLa cells stably expressing Cox8a–SLP fusion proteins together with mEGFP were labeled with respective CA–SiR or TF–SiR (100 nM) substrates for 1 h and washed twice with imaging medium before microscopy. U2OS cells stably expressing Vim–SNAP_F-tag, Vim–SNAP-tag2, TOMM20–SNAP-tag2 or CalR–SNAP-tag2-KDEL or coexpressing Vim–SNAP-tag2 and LifeAct–HaloTag7 (rAAV) were labeled with CF–SiR and CA–MaP618 (100 nM), respectively, for 1 h and washed twice with imaging medium before microscopy. Live-cell STED nanoscopy was performed using an Abberior STED Expert Line 595/775/RESOLFT QUAD scanning microscope (Abberior Instruments). SiR was excited using a 640-nm excitation line and depleted using a 775-nm STED line. For dual-color imaging, MaP618 was excited with a 561-nm excitation line and depleted with the same STED line. The microscope was equipped with a UPlanSApo $\times 100/1.4$ oil immersion objective lens (Olympus) and the fluorescence signal was detected using avalanche photodiodes. The pinhole was set to 0.9 Airy units, a gating of 0.75–8 ns was applied and dwell times of 7–10 μ s and a pixel size of 25 nm were used. For STED images, each line was scanned 4–8 times and the signal was accumulated. Imaging settings are further detailed in Supplementary Table 12. Lookup tables used for image representation in Fig. 4a–e were as follows: green, mEGFP; red-hot, SiR; cyan, SNAP-tag2–SiR; orange-hot, HaloTag7–MaP618. For image representation, a Gaussian blur filter with a 2σ radius was applied.

The FWHM of single-intermediate filaments (Vim–SNAP_F-tag/SNAP-tag2 fusion) was determined by extracting fluorescence intensity profiles perpendicular to single filaments using Fiji software (ImageJ 1.53t)⁵². Mean filament diameters were calculated from 15 individual fibrils and from three individual images ($n = 3$) by fitting a Gaussian function (Eq. (27)), yielding the FWHM from Eq. (28).

$$y = y_0 + \frac{A}{\omega \sqrt{\frac{\pi}{2}}} e^{-2 \frac{(x-x_c)^2}{\omega^2}} \quad (27)$$

where y is the normalized fluorescence intensity, x is the x coordinate (μ m), y_0 is the offset, A is the area, ω is the width and x_c is the center.

$$FWHM = \omega \sqrt{2 \ln(2)} \quad (28)$$

where ω is the Gauss width.

STED microscopy images in Fig. 4a are snapshots of three independent experiments (eight images in total) with similar outcome. Additionally, five individual experiments (12 images in total) with labeling substrate concentrations ranging from 25 to 250 nM and incubation times ranging from 1 h to overnight labeling in two different cell types (U2OS and HeLa cells) yielded similar brightness differences. Images shown in Fig. 4b are snapshots of two individual experiments (seven images in total). FWHM quantification was conducted for one individual experiment (15 filaments, three images). STED images in Fig. 4c–e were acquired in one individual experiment.

Strain and growth conditions of *H. polymorpha* yeast cells

H. polymorpha ku80 yeast cells were grown at 37 °C with shaking (200 rpm). For fluorescence microscopy studies, the cultures were grown in mineral medium (MM)⁵⁵ containing 30 μ g ml⁻¹ leucine, 1 \times vitamin solution, 0.25% ammonium sulfate and 0.5% methanol.

Construction of *H. polymorpha* strains expressing Pex3–SNAP-tag2/SNAP_F-tag

The SNAP_F-tag DNA sequence was synthesized by Genscript and SNAP-tag2 was PCR-amplified from pET51b-SNAP-tag2 using primers Fw_BglII_SNAP-tag2 (GATGCTAGATCTGATAAAGACTGCCAAATGAAACG) and Rv_Sall_SNAP-tag2 (CTCTAGAGTCGACCTAGC-CCAGCCTGGTTTACCC). Both SNAP-tag inserts were cloned into the pUC57 plasmid vector. pHIPZ_Pex3_mGFP⁵⁶ and pUC57_SNAP_F-tag/SNAP-tag2 were restricted using Sall and BglII according to the supplier protocol (Thermo Fisher Scientific) and ligated to generate pHIPZ_Pex3_SNAP_F-tag or pHIPZ_Pex3_SNAP-tag2 plasmids for peroxisome targeting in yeast. Plasmids were transformed into *E. coli* DH5 α cells selected on LB with ampicillin. Positive colonies were checked using XhoI restriction. The plasmids were linearized with EcoRI or MluI and transformed into *H. polymorpha* yku80 cells by electroporation as described in Faber et al.⁵⁷ Positive integrations in yeast were tested with colony PCR using primers Ppex3_Fw (CCTGTTGCGCAAGATATAG) and Seq_SNAP_Rv (CGTAGGAAGGCTGGATGTC). Plasmids were confirmed by LightRun sequencing (Eurofins Genomics) and compared using Clone Manager 9 (Scientific and Educational Software). Correct yeast integration was checked by colony PCR with DreamTaq enzyme (Thermo Fisher Scientific) of zymolyase-treated cells and imaged using the Gel Doc XR+ System (Bio-Rad).

Live-cell labeling of yeast peroxisomes

H. polymorpha cells were grown until stationary phase in MM supplemented with 0.5% glucose at 37 °C shaking (200 rpm). The culture was diluted to OD₆₆₀ = 0.1 and grown until OD₆₆₀ > 1.6. This culture was again diluted to OD₆₆₀ = 0.1 in MM and methanol and the cells were simultaneously stained using TF–SiR, CP–SiR, BG–SiR, TF–MaP555, CP–MaP555 or BG–MaP555 at a final concentration of 250 nM for 18 h at 37 °C shaking, reaching a final OD₆₆₀ of ~3.0. To stain the yeast cell walls, cells were incubated with 25 μ g ml⁻¹ Calcofluor white (Fluorescent Brightner 28; Sigma-Aldrich, 910090) for 15 min at 37 °C shaking. Cells were washed three times with PBS and immediately used for live-cell imaging in MM and methanol.

Live-cell STED microscopy of yeast peroxisomes

STED nanoscopy was conducted on a commercial microscope (Abberior Instruments) equipped with a STED depletion laser (775 nm), four excitation lasers (640, 561, 488 and 405 nm), a CoolLED pE-2 excitation system and a $\times 100$ oil immersion objective (Olympus UPLSAPO/1.40). Confocal overview images (80 \times 80 μ m) were taken using a pixel size of 180 nm, dwell time of 20 μ s and 6% laser power using the 561-nm excitation laser with 570–680-nm detection for MaP555 substrates or 0.5% laser power using the 640-nm excitation laser with 650–730-nm detection for SiR substrates. STED images were taken with a pixel size of 25 nm, dwell time of 20 μ s and 30% laser power using the 561-nm excitation laser with 570–680-nm detection for MaP555 substrates or 2% laser power using the 640-nm excitation laser with 650–730-nm detection for SiR substrates. A depletion laser power of 50% or 10% at 775 nm was applied for 561-nm or 640-nm excitation, respectively. Calcofluor white was imaged using 3% laser power with the 405-nm excitation laser and 410–480-nm detection. Image acquisition was carried out using Imspector software (version 16.3, Abberior Instruments). Analysis was performed using Fiji software (ImageJ 2.14.014)⁵². To compare the fluorescence intensities of SNAP_F-tag and SNAP-tag2 labeled with different fluorophore substrates, the mean and maximum fluorescence intensity of 125 cells (rectangles of 3 \times 3 μ m as the ROI) were measured in biological triplicates. Line profile measurements were performed for one cell per condition.

Statistics and reproducibility

All in vitro measurements were performed at least in technical triplicates. Flow cytometry-based experiments were performed at least in

technical duplicates. Live-cell kinetic measurements were performed as biological duplicates (that is, on two different days 1 week apart). STED images aiming to compare differences in labeling brightness or resolution in mammalian cells were repeated at least two times to ensure similar outcomes. Quantification of fluorescence intensities in yeast cells were performed in biological triplicates. Unless otherwise stated, STED images originated from single experiments. Statistical significance of a sample group over a reference group was performed by performing a two-tailed unpaired *t*-test with Welch's correction using GraphPad Prism (version 10.2.3). Sample sizes are indicated in the captions of corresponding figures. The following notations apply for all statistical analyses: NS (not significant), $P \geq 0.05$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Plasmids of interest from the study are available on Addgene. Addgene accession codes can be found in Supplementary Table 14. The data supporting the findings of this study are available within the paper and its Supplementary Information. Data are available from the corresponding authors upon request. Source data are provided with this paper.

Code availability

All code and example scripts used in this study are available within the manuscript and its Supplementary Information.

References

- Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
- Wrenbeck, E. E. et al. Plasmid-based one-pot saturation mutagenesis. *Nat. Methods* **13**, 928–930 (2016).
- Routh, S., Acharyya, A. & Dhar, R. A two-step PCR assembly for construction of gene variants across large mutational distances. *Biol. Methods Protoc.* **6**, bpab007 (2021).
- Chao, G. et al. Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1**, 755–768 (2006).
- Fleishman, S. J. et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS ONE* **6**, e20161 (2011).
- Girardot, C. et al. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics* **17**, 419 (2016).
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
- Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
- Wickham, H. et al. Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
- R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2010).
- Kuzmic, P. Program DYNAPFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273 (1996).
- Straume, M. & Johnson, M. L. Monte Carlo method for determining complete confidence probability distributions of estimated model parameters. *Methods Enzymol.* **210**, 117–129 (1992).
- Malecki, M. J. et al. Leukemia-associated mutations within the NOTCH1 heterodimerization domain fall into at least two distinct mechanistic classes. *Mol. Cell. Biol.* **26**, 4642–4651 (2006).
- Zolotukhin, S. et al. Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. *Methods* **28**, 158–167 (2002).
- Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
- Chastagnier, Y. et al. Image processing for bioluminescence resonance energy transfer measurement—BRET-Analyzer. *Front. Comput. Neurosci.* **11**, 118 (2017).
- Stirling, D. R. et al. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021).
- van Dijken, J. P., Otto, R. & Harder, W. Growth of *Hansenula polymorpha* in a methanol-limited chemostat. Physiological responses due to the involvement of methanol oxidase as a key enzyme in methanol metabolism. *Arch. Microbiol.* **111**, 137–144 (1976).
- Wu, H. et al. Peroxisome development in yeast is associated with the formation of Pex3-dependent peroxisome–vacuole contact sites. *Biochim. Biophys. Acta Mol. Cell. Res.* **1866**, 349–359 (2019).
- Faber, K. N. et al. Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. *Curr. Genet* **25**, 305–310 (1994).

Acknowledgements

This work was supported by the Max Planck Society, the École Polytechnique Fédérale de Lausanne (EPFL) and the Deutsche Forschungsgemeinschaft (German Research Foundation; TRR 186). S.K., J.W. and J.K. acknowledge support from the Heidelberg Biosciences International Graduate School. J.W. was supported by the Max Planck School Matter to Life. Y.-H.L. was supported by the EPFL Doctoral Program in Biotechnology and Bioengineering. We thank A. Bergner and B. Réssy for providing reagents and materials. We thank B. Koch, J. Kress and P. Breuer for help with stable cell line generation. We acknowledge the mass spectrometry facility (S. Fabritz, T. Rudi and J. Kling) for its support. We thank all members of the K.J. lab for critical discussions.

Author contributions

S.K., V.N., J.H. and K.J. designed and interpreted the experiments. V.N., A.L. and S.K. designed and synthesized the compounds used in this paper. S.K., J.W., Y.-H.L., C.E. and J.H. performed the protein engineering. V.N. and J.F. performed the substrate screening and photophysical characterization. V.N. computed the chemical substrate properties. M.T. performed the NanoDSF measurements. S.K. and J.R. designed, performed and analyzed the in vitro kinetic studies. S.K. performed the in vitro characterization, confocal microscopy and flow cytometry experiments of mammalian cells. J.K. performed the STED imaging in mammalian cells. E.M.F.d.L. and R.V. designed and performed the experiments in yeast. S.K. and K.J. wrote the manuscript with input from all authors.

Funding

Open access funding provided by Max Planck Society.

Competing interests

V.N., S.K., J.H. and K.J. are listed as inventors on a patent application (EP 24172286) on improved SNAP-tag substrates filed by the Max Planck Society. The remaining authors declare no competing interests.

Additional information

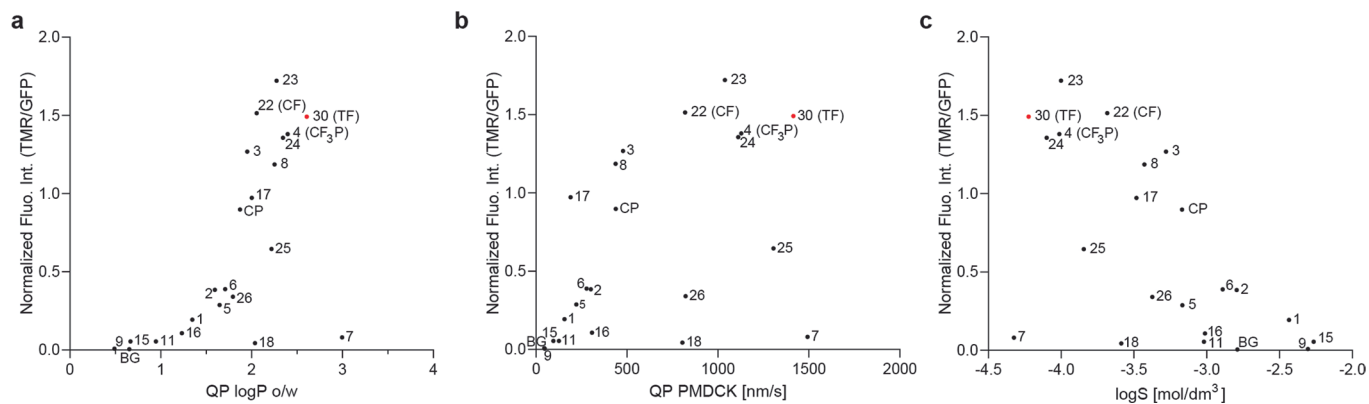
Extended data is available for this paper at <https://doi.org/10.1038/s41589-025-01942-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-025-01942-z>.

Correspondence and requests for materials should be addressed to Julien Hiblot or Kai Johnsson.

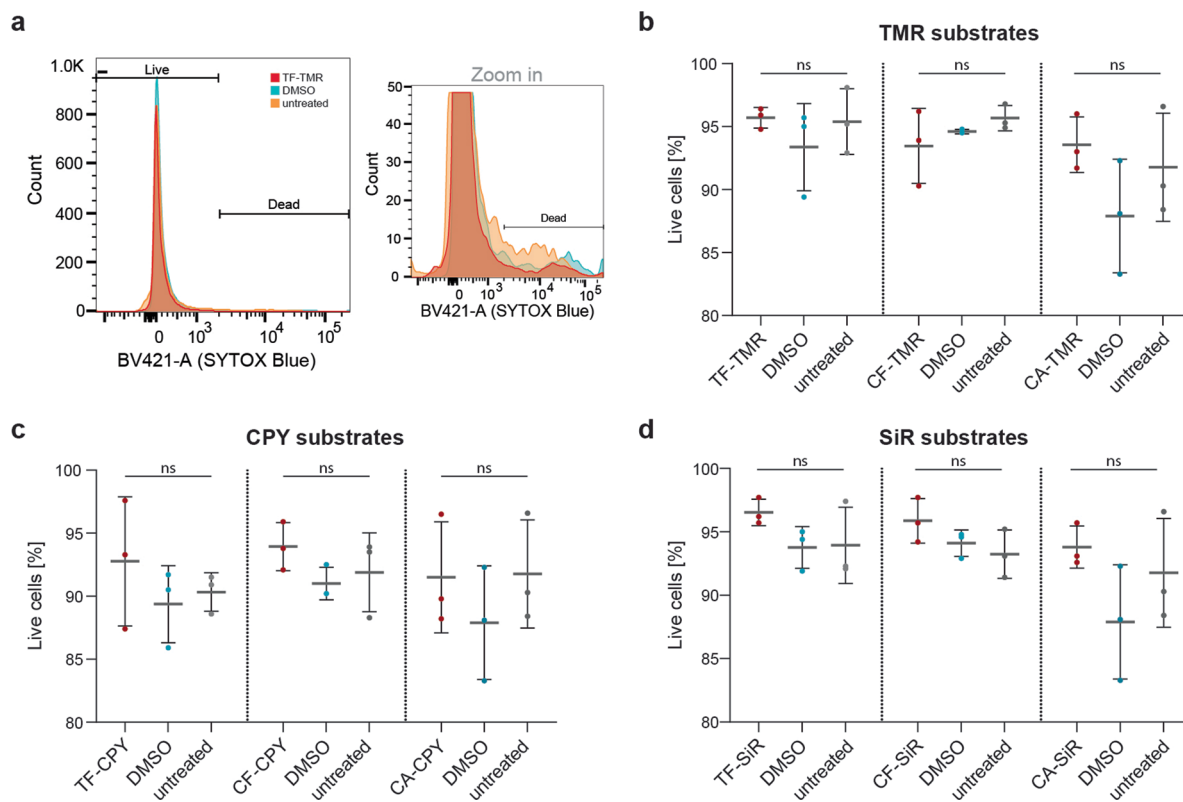
Peer review information *Nature Chemical Biology* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.



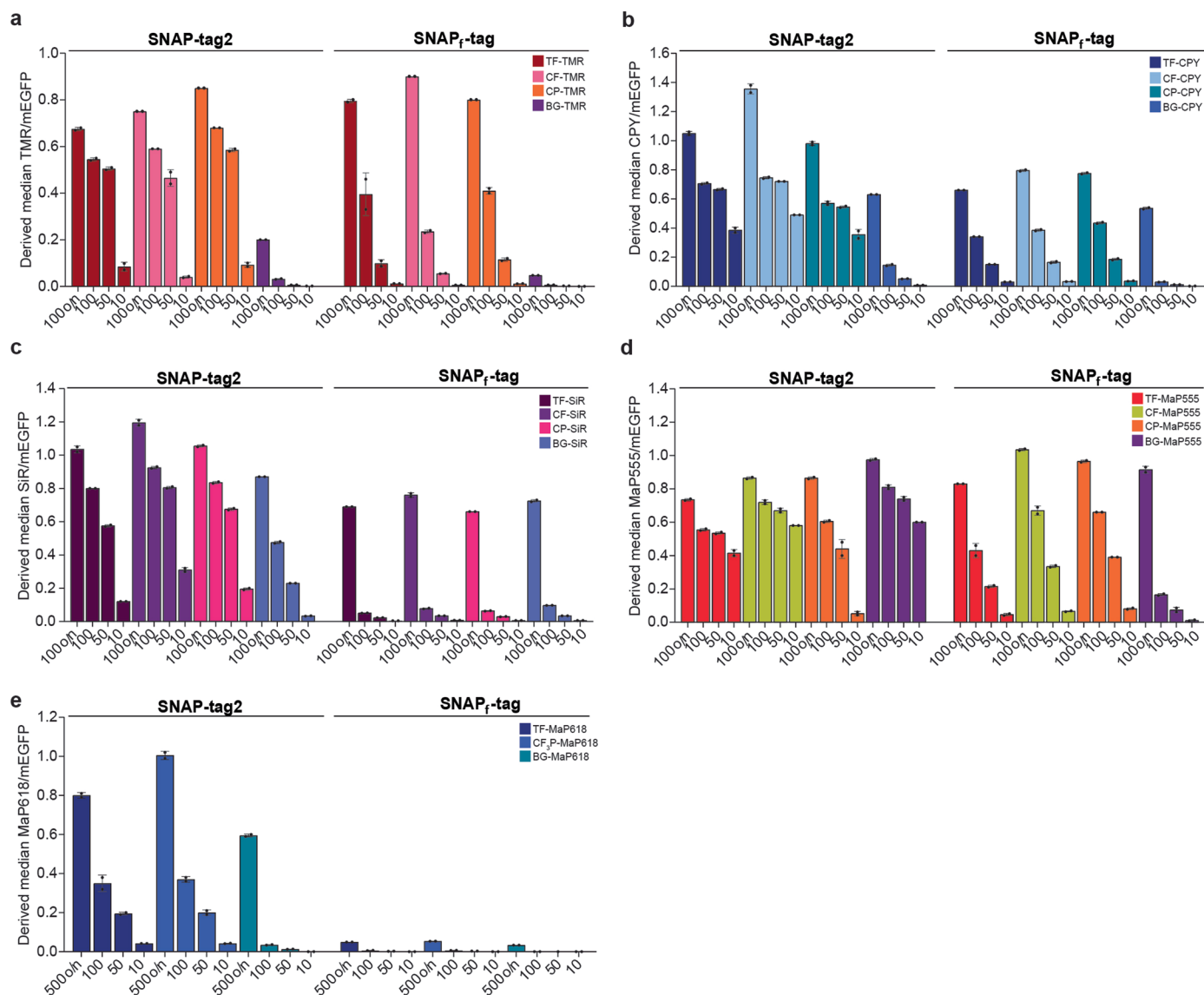
Extended Data Fig. 1 | Correlation between pharmacokinetic properties and live cell performance of selected substrates. QikProp (QP) Descriptors were calculated for non-fluorescent acetylated substrates using Schrödinger Maestro 12.3 software, regarding their: **a**, octanol/water partition coefficient (QP logP

o/w), **b**, apparent MDCK cell permeability (QP PMDCK) and **c**, aqueous solubility (logS) and correlated to the experimentally determined performance of their corresponding TMR-derivatives in live cells. Computed parameters of the acetylated substrates can be found in Supplementary Table 2.



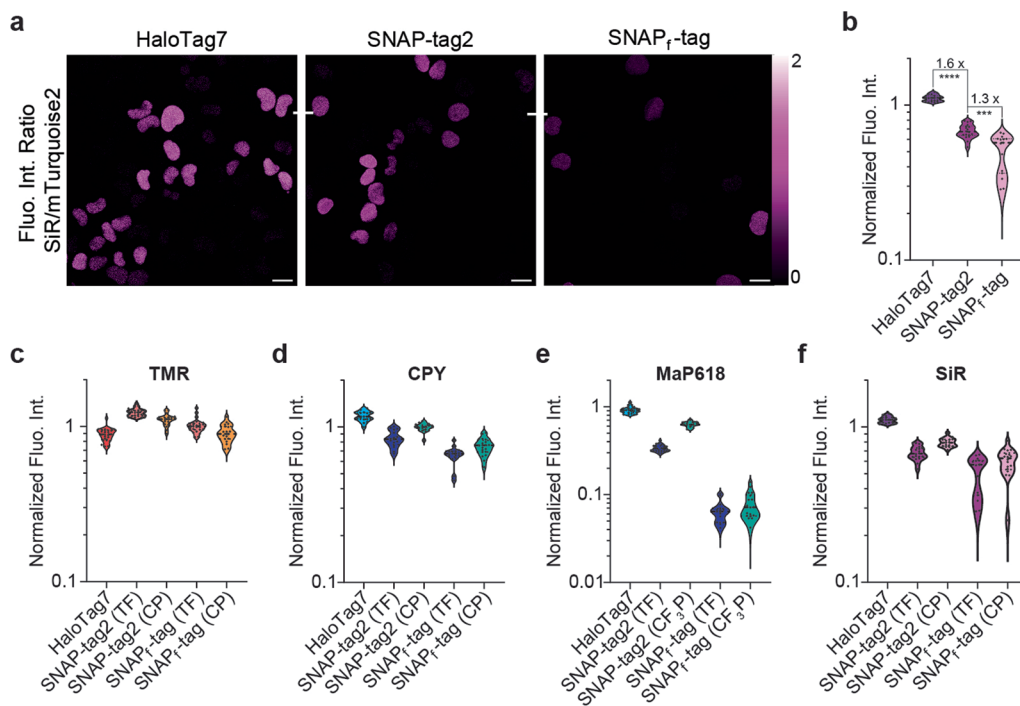
Extended Data Fig. 2 | Statistical analysis of cell viability using flow cytometry. U2OS cells were incubated either with TF-/CF-/CA-substrates [1 μ M], DMSO (1% v/v) or remained untreated. After 1 h incubation at 37 $^{\circ}$ C, both dead cells (supernatant) and live cells (detached with trypsin) were collected. SYTOX Blue dead cell stain [1 μ M] was added and the cells were subsequently analyzed by flow cytometry. Experiment was conducted in technical triplicates. **a**, Flow cytometry histograms of cells stained with SYTOX Blue, showing the gating strategy for live and dead cell events. **b–d** Dot plots showing the percentage of live cell events for

b TMR-substrates, **c** CPY-substrates and **d** SiR-substrates in comparison to DMSO and untreated control samples. Each dot represents one replicate. The mean percentage of live cells from each triplicate is indicated as a black horizontal line. Error bars represent the s.d. The differences in cell viability of treated and untreated cells were not statistically significant (two-tailed unpaired t-test with Welch's correction, $p > 0.05$, ns). Novel TF- and CF-substrates do not influence cell viability.



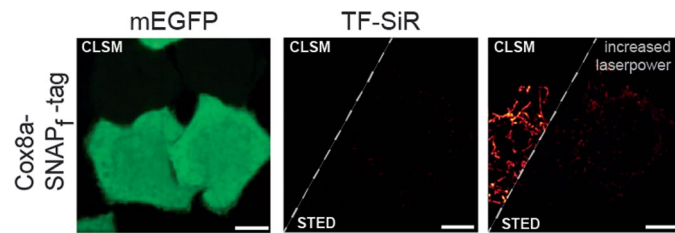
Extended Data Fig. 3 | Comparison of SNAP-tag2 und SNAP_f-tag for labeling with different fluorescent substrates in live cells. U2OS cell stably expressing mEGFP-SNAP-tag2 or mEGFP-SNAP_f-tag were incubated with different **a** TMR, **b** CPY, **c** SiR, **d** MaP555 or **e** MaP618 substrates at [100 nM] or [500 nM] overnight or at [100 nM, 50 nM and 10 nM] for 1 h. Cells were washed and analyzed by flow

cytometry (recorded $n = 10000$ cells). Analysis was done using FlowJo by gating for single cells and double positive labeling signal (as depicted in Supplementary Fig. 16) and the median of the fluorescent label/mEGFP was derived. Data are represented as median values \pm mean s.d. of technical duplicates. SNAP-tag2 showed for most of the substrates a higher labeling ratio.



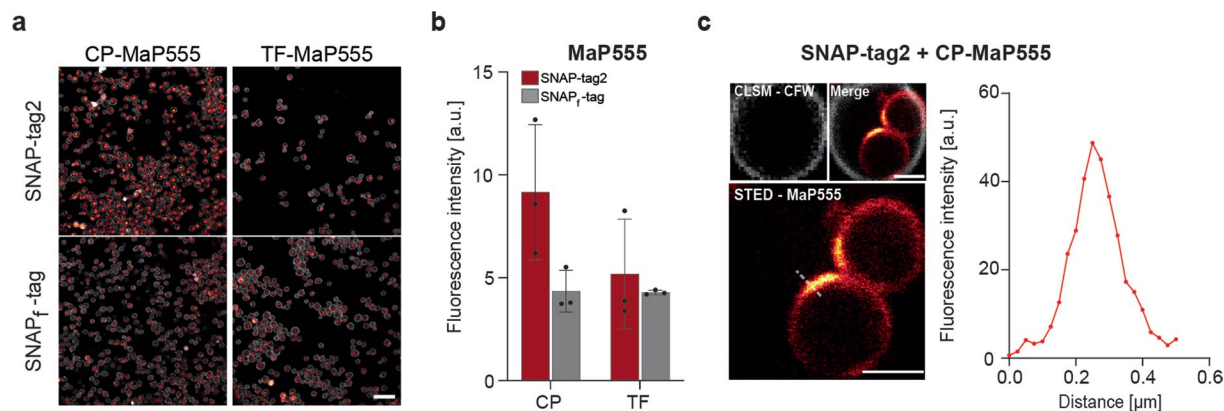
Extended Data Fig. 4 | Comparison of HaloTag7, SNAP-tag2 and SNAP₁-tag performances in confocal fluorescence microscopy. Experiments were conducted in live U2OS cells stably coexpressing HaloTag7-SNAP₁-tag or HaloTag7-SNAP-tag2 together with mTurquoise2 (expression marker) in the nucleus. **a**, Comparison of HaloTag7, SNAP-tag2 and SNAP₁-tag performances in confocal fluorescence microscopy with SiR substrates. U2OS cells were labeled with SLP-respective CA-/TF-SiR substrates at [100 nM] overnight and washed

afterward. Ratiometric projections are presented corresponding to SiR label/mTurquoise2 on a magenta-hot look-up table. Scale bar: 20 μm . **b**, Violin plots representing the quantitative analysis of single cells shown in **a**. Numbers represent fold-changes between the different SLPs ($n \geq 15$ cells, two-tailed unpaired t-test with Welch's correction: **** $P < 0.0001$, *** $P = 0.0003$). **c-f**, Violin plots representing the quantitative analysis of different fluorescent substrates/mTurquoise2 for **c** TMR, **d** CPY, **e**, MaP618 and **f** SiR ($n \geq 13$ cells).

**Extended Data Fig. 5 | SNAP_F-tag performance in CLSM and STED microscopy.**

HeLa cells stably co-expressing SNAP_F-tag in the mitochondria (Cox8a localization sequence) together with mEGFP (no specific localization) were labeled with TF-SiR (100 nM) for 1 h and washed afterward. SNAP_F-tag showed

insufficient signal under the same imaging conditions (0.5% laser power) used for SNAP-tag2 and HaloTag7 and required the use of an increase laser power (5%) to see a weak labeling signal. Scale bars: 10 μm. LUTs: green (mEGFP), red-hot (SiR).



Extended Data Fig. 6 | SNAP-tag2 and SNAP_r-tag labeling of live yeast peroxisomes. **a**, CLSM images of *H. polymorpha* yeast cells expressing Pex3-SNAP-tag2 or Pex3-SNAP_r-tag fusion proteins labeled with different MaP555 substrates. Yeast cells were labeled with CP- or TF-MaP555 (250 nM) for 18 h and the cell wall was stained with Calcofluor White (CFW; 25 $\mu\text{g}/\text{mL}$) for 15 min. Cells were washed prior to imaging. Scale bar: 10 μm . **b**, Bar plot representing the quantitative analysis of SNAP-tag2 and SNAP_r-tag labeling with MaP555 substrates in yeast. Experiments were conducted in biological triplicates ($n = 125$ cells for each replicate) and the mean fluorescence intensity of MaP555

substrates was calculated (Supplementary Fig. 14). Error bars represent the s.d. Combination of SNAP-tag2 with CP-MaP555 show the best results for labeling of yeast peroxisomes. **c**, Bottom, STED image of Pex3-SNAP-tag2 labeled with CP-MaP555. Top, CLSM image of CFW stained cell wall and merge of both channels. LUTs: red-hot (MaP555) and gray (CFW). Scale bar: 1 μm . Right, line profile of labeled peroxisomes in STED (highlighted as white dashed line in the image). SNAP-tag2 with CP-MaP555 is well suited to perform live cell STED in yeast.

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) 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

Plate reader: Tecan Sparkcontrol Method Editor Version 2.2
 Stopped-flow device: Biologic SFM-400
 Confocal Microscopy: Leica Stellaris 5
 STED Microscopy: Abberior STED Expert Line 595/775/RESOLFT QUAD scanning microscope
 FACS: BD FACSMelody™ Cell Sorter
 Flow cytometry: BD Fortessa™ X-20 Cell Analyzer
 NMR: BRUKER Advance III HD 400 NMR spectrometer equipped with a CryoProbe™
 Thermal stability: Prometheus NT48 nanoscale differential scanning fluorimeter
 Photophysical properties: Quantaurs-QY spectrometer (model C11347, Hamamatsu), V-770 Spectrophotometer (Jasco)
 Chemical property calculation: Schroedinger Maestro 12.3

Data analysis

General data analysis: GraphPad Prism (version 10.2.3), DynaFit 4 (version 4.11.050), SpectraGryph (version 1.2), R Studio (version 4.3.1),
 Molecular Biology: NEBaseChanger (version 1.3.3, [nebasechanger.neb.com](#)), Tm Calculator (version 1.15.0, [tcalculator.neb.com](#)), Geneious Prime (version 2023.1.2)
 Image analysis: ImageJ 1.54f, Leica LAS X 3.5.7.23225 (Confocal), CellProfiler (version 4.2.6), Inspector (version 16.3, Abberior Instruments)
 Chemistry: MestReNova (version 14.1.0-24037), ChemDraw (version 20.1.1)
 Crystal structures: PyMOL version 2.1.1.

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Plasmids encoding certain SNAP-tag2 constructs have been deposited on Addgene. Addgene numbers can be found in Supplementary Table 14. Correspondence and requests for materials should be addressed to Kai Johnsson. The data supporting the main findings of this study are available within the article and its Supplementary Information and are available from the corresponding author upon request. Source data are provided with this manuscript. The crystal structures of SNAP-tag in its apo (PDB: 3KZY), benzylated (PDB: 3L00), BG-bound (PDB: 3KZZ) and TMR-labeled (PDB: 6Y8P) states were available from the pdb.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

| | |
|-----------------------------|--|
| Reporting on sex and gender | Sex and gender was not necessary for the presented experiments, as the data is not relevant to a clinical trial. |
| Population characteristics | Population characteristics was not necessary for the presented experiments, as the data is not relevant to a clinical trial. |
| Recruitment | Recruitment was not necessary for the presented experiments, as the data is not relevant to a clinical trial. |
| Ethics oversight | Ethics oversight was not necessary for the presented experiments, as the data is not relevant to a clinical trial. |

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

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](https://www.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 | Samples were acquired until a clear trend was evident. |
| Data exclusions | No data was excluded. |
| Replication | In vitro measurements were performed in technical triplicates or as indicated. Microscopy experiments were performed on one to three independent sample preparations as stated in the method section. All replicates were successful. |
| Randomization | Randomization is not relevant to the study, as all experiments were performed with cell lines. No experiments involved allocation of different samples, organism, or participants into experimental groups. |
| Blinding | Blinding is not relevant to the study because no experiments involved allocation of different samples, organisms, or participants into experimental groups. |

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

Methods

| n/a | Involved in the study |
|-------------------------------------|---|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

| n/a | Involved in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used

primary Anti-c-Myc (Ab-1) Mouse mAb (9E10) (#OP10, EMD Millipore, Merck), secondary goat anti-mouse-Alexa647 antibody (#A-21236, Invitrogen, ThermoFisher Scientific)

Validation

Anti-c-Myc (Ab-1) Mouse mAb (9E10) is validated for use in FC, Frozen Sections, Immunoblotting, IF, IP, Chromatin IP, Paraffin Sections for the detection of c-Myc (Ab-1) as stated by the manufacturer (https://www.merckmillipore.com/TH/en/product/Anti-c-Myc-Ab-1-Mouse-mAb-9E10,EMD_BIO-OP10?ReferrerURL=https%3A%2F%2Fwww.google.com%2F#documentation).

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

U-2 OS Flp-In T-REx cell lines (Molecular and Cellular Biology 2006, 26 (12), 4642-4651) - from Blacklow lab. Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. HeLa Kyoto Flp-In cells (kind gift of Dr. Amparo Andres-Pons (EMBL, Heidelberg); Science 2024, 383, 890-897)

Authentication

Cell lines were not further authenticated.

Mycoplasma contamination

Cell lines have been tested by PCR and were negative.

Commonly misidentified lines
(See [ICLAC](#) register)

Not applicable as no commonly misidentified cell lines were used.

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

Protein labeling on yeast surface for fluorescence-activated cell sorting (FACS). For fluorescent labeling, 10^7 cells were harvested by centrifugation ($14'000\times g$, 1 min.). For antibody-based expression staining of yeast cells transformed with pCTcon2 plasmids, yeast cells were resuspended in 1:10-diluted primary mouse anti-cMyc antibody (#OP10, EMD Millipore, Merck) in PBS (50 μ L) and incubated on a rotating wheel at 4 °C for 1 h. The cells were pelleted ($14'000\times g$, 1 min.) and washed twice with PBS (125 μ L) with centrifugation in-between ($14'000\times g$, 1 min.). The cell pellet was resuspended in 1:50-diluted secondary goat anti-mouse-Alexa647 antibody (#A-21236, Invitrogen, ThermoFisher Scientific) in PBS (50 μ L) and incubated on a rotating wheel at 4 °C for 1 h. The cells were washed twice with PBS (125 μ L) prior to labeling with SNAP-tag substrates. For protein libraries encoded by the pYDNg expression vector, expression control was monitored by labeling of eUnaG2 with bilirubin. Yeast cells were resuspended in PBS (50 μ L) containing bilirubin [10 μ M] and BSA [1 mg/mL], gently vortexed and incubated on ice for 10 min. The cells were pelleted and washed twice with PBS (150 μ L) interspersed by centrifugation steps ($14'000\times g$, 1 min.) prior to SNAP-tag labeling. Cells expressing SNAP-tag variants were labeled in PBS (50 μ L) using different substrates (CF3P-TMR/-MaP618, TF-TMR/-MaP618) with varying concentrations [10-500nM] and incubation times (10-60 min) in order to adjust the screening stringency. Labeling was performed at r.t. on a rotating wheel. Cells were washed with PBS (125 μ L), resuspended in 1 mL of PBS and filtered through 5 mL round bottom polystyrene test tubes with cell strainer snap caps (#352235, Falcon®) for FACS.

Live cell labeling performance of SNAP-tag proteins determined with flowcytometry. Cells were seeded on transparent 96-well cell culture plates and treated according to desired experiment in a reaction volume of 100 μ L. After treatment, labeling reaction was stopped by addition of recombinant SNAP-tag2 (2 μ M, 100 μ L, 10 min. incubation), cells were washed twice with PBS (150 μ L, 10 min. incubation), trypsinized (50 μ L trypsin, 10 min. incubation) and resuspended in FACS buffer (2 %

FBS in PBS) to a final volume of 200 μ L. The cell suspension was transferred to non-binding u-bottom 96-well plates (Falcon) and analyzed on the flow cytometer using the HTS module.

Substrate screening for SNAP-tag labeling in live mammalian cells. U2OS Flp In™ T-REx™ (Thermo Fisher) cells stably expressing mEGFP-SNAP-tag or mEGFP-CLIP-tag fusion proteins were seeded into 96-well cell culture plates (10'000 cells/well) the day prior to experiment. The cells were then incubated with substrates 1 – 30, CP or BC [100 nM] for 2 h at 37 °C. All substrates were tested in technical triplicates. Cells were washed twice with cell growth medium for 15 min incubation at 37°C and 1x with sterile PBS (pH = 7.4) prior to detachment with trypsin (50 μ L, 10 min., 37 °C). Cells were resuspended in FACS buffer to a final volume of 200 μ L. The cell suspension was transferred to non-binding u-bottom 96-well plates (Falcon) and analyzed on the flow cytometer using the HTS module.

Cell viability assay. U2OS cells were seeded on transparent 96-well cell culture plates one day prior to experiment. Cells were incubated with TF-, CF- and CA-fluorophore substrates [1 μ M], DMSO [1 % (v/v)] or remained untreated for 1 h at 37 °C. The medium was collected into non-binding u-bottom 96-well plates (Falcon) and detached cells were harvested by centrifugation (3000 x g, 5 min). The supernatant was removed. Additionally, adherent cells were detached with trypsin (30 μ L, 10 min., 37 °C), resuspended in FACS buffer (2 % FBS in PBS) to a final volume of 100 μ L and added to the same wells to collect all dead and live cells. SYTOX Blue dead cell stain (100 μ L, 2 μ M; ThermoFisher) was added to the cells to a final concentration of 1 μ M and cells were subsequently analyzed by flow cytometry (10'000 cells/well; laser: 405 nm, BP filter: 450/50).

Instrument

BD Fortessa™ X-20 Cell Analyzer, BD FACSMelody™ Cell Sorter

Software

FlowJo (version 10.10.0)

Cell population abundance

For yeast: the final sorted population was 0.5-3% of all cell events
For mammalian cells: the final sorted population was up to 10-50% of all cell events.

Gating strategy

Hierarchical gating of live cells (SSC-A/FSC-A), single cells (FSC-H/FSC-A) and self-labeling protein labeling (label channel/ expression channel). Gating strategy is described in the method section and exemplified in Supplementary Figures 15 and 16.

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