Bright and stable monomeric fluorescent proteins derived from StayGold

Kiryl Piatkevich

kiryl.piatkevich@gmail.com

Westlake University https://orcid.org/0000-0002-7777-9468

Gleb Lesnov

National Research Center “Kurchatov Institute”

Hanbin Zhang

Westlake University https://orcid.org/0000-0002-0427-0070

Oksana Subach

National Research Center “Kurchatov Institute”

Anna Vlaskina

National Research Center “Kurchatov Institute”

Valeriya Samygina

National Research Center “Kurchatov Institute”

Alena Nikolaeva

National Research Center “Kurchatov Institute”

Azat Gabdulkhakov

Institute of Protein Research (Russian Academy of Sciences) https://orcid.org/0000-0003-1016-5936

Wenming Qin

Shanghai Advanced Research Institute CAS

Valentin Borshchevskiy

Moscow Institute of Physics and Technology (MIPT) https://orcid.org/0000-0003-4398-9712

Maxim Perfilov

Institute of Bioorganic Chemistry https://orcid.org/0000-0002-8987-9562

Alexander Mishin

Planta https://orcid.org/0000-0002-4935-7030

Fedor Subach

Kurchatov Institute https://orcid.org/0000-0003-2720-7821

Brief Communication

Keywords:

Posted Date: July 27th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3188559/v1

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Methods on February 26th, 2024.
See the published version at https://doi.org/10.1038/s41592-024-02203-y.

Abstract

High brightness and photostability of StayGold make it a particularly attractive probe for long-term live cell imaging. However, its dimeric nature precludes its application as a fluorescent tag for proteins. Here, we report the development and X-ray structures of a monomeric variant of StayGold (mStayGold), which preserves the beneficial properties of its precursor while serving as a tag for structural proteins and membranes. We compare mStayGold to other state-of-art GFPs and utilize it for super-resolution long-term live cell imaging.

Full Text

Since cloning the first GFP in 1992, protein engineers have been putting a lot of effort and thought into improving the brightness and photostability of FPs, which are crucial factors for many applications in fluorescence microscopy. Serendipitously, recently discovered GFP StayGold (a point mutant of wild-type GFP from *C. uchidae*) exhibits exceptional photostability of an order of magnitude higher than that for any currently available FP while characterized by high molecular brightness[1]. However, StayGold forms an obligate dimer that limits its applicability as a fluorescent tag for protein and membrane labeling. To address this issue, we employed a directed molecular evolution approach to generate a monomeric version of StayGold, referred to as mStayGold, which retains the beneficial properties of its precursor.

To facilitate screening for monomeric StayGold variants in *E. coli*, we decided to take advantage of the finding that the DNA-binding domain of the AraC protein ($AraC_{DNA}$), when not connected to a dimerization domain was unable to activate detectable transcription from the wild-type pBAD promoter, but $AraC_{DNA}$ protein in fusion with dimeric protein activated pBAD promoter in bacteria[2]. To test this possibility, we constructed an expression vector, which had fusion of FP of interest with $AraC_{DNA}$ domain under the control of the rhamnose promoter and a reporter gene, mTagBFP, under the control of the arabinose promoter (Fig. 1a). According to our model, dimerization of $AraC_{DNA}$ promoted by a dimeric FP of interest drives a higher expression level of reporter gene compared to monomeric FP (Fig. 1b). Indeed, validation of the expression vector using FPLC revealed that more monomeric StayGold variants induced lower mTagBFP expression compared to more dimeric StayGold mutants (Fig. 1c,d and Supplementary Fig. 1). Correspondingly, during each round of random mutagenesis we selected the brightest green colonies possessing dimmer blue fluorescence confirming their oligomeric state with FPLC. After eight rounds of directed evolution, we selected a bright monomeric variant, which was further modified with the N- and C-termini from EGFP protein to ensure its stability in mammalian cells, resulting in a GFP named mStayGold (Supplementary Fig. 2 and 3). FPLC demonstrated that mStayGold was ~99% monomeric at a high concentration (56 μM), which was higher than monomeric fraction for mNeonGreen (95% monomeric). At the same time, parental StayGold did not exhibit a monomeric fraction (Fig. 1e). To validate the monomeric behavior of mStayGold in mammalian cells, we expressed its fusion with CytERM in live HeLa cells and evaluated the fraction of transfected cells without visible OSER whorl structures. For comparison, we used parental StayGold and StayGold-E138D, reported in another manuscript

currently under review elsewhere in Nature Portfolio journal. StayGold and StayGold-E138D demonstrated only modest OSER scores of 59.0% and 70.9%, respectively. In comparison, mStayGold scored 88.6%, almost reaching the threshold of monomericity in OSER assay[3] (Fig. 1f). These results established mStayGold as a monomeric protein.

To understand the role of the introduced mutations, we analyzed the deposited X-ray structure of the StayGold dimer (PDP ID 7Y40) and solved X-ray structures of mStayGold with 1.45, 1.8, and 1.6 Å resolution at pH 6.5, pH 4.6 and 8.5, respectively (Fig. 2a,b,c). Two mutations (S55T and H141Q) were internal to the β-can, and six others were external (Supplementary Fig. 3 and 4). Among external mutations, Q140P mutation can potentially disrupt the dimeric interface of StayGold observed in X-ray structure since Q140 residues from both subunits form symmetrical contacts with N137 and T153 residues from opposite subunits (Supplementary Fig. 4). Analysis of FPLC chromatogram (Fig. 1e) and mutations introduced during StayGold evolution (Supplementary Fig. 1) revealed that on 4th round single C165Y mutation in StayGold resulted in efficient monomerization of the protein. Probably, bulky tyrosine residue in position 165 sterically prevented the efficient formation of the dimer by unfavorable interactions with e.g., R144. Analysis of mStayGold structure revealed its overall monomeric form and chromophore surrounding at different pH values of 4.6, 6.5, and 8.5 (Figure 2a-c and Supplementary Table 2).

Having validated mStayGold as a monomeric protein, we further characterized its spectral and biochemical properties (Fig. 2d-i, Supplementary Fig. 5, and Supplementary Table 1). The introduced mutations did not alter absorbance and fluorescence spectral profiles, however, mStayGold had 1.4-fold lower molecular brightness than StayGold though it was 1.2-fold brighter than mNeonGreen (Supplementary Fig. 5 and Supplementary Table 1). Rapidly maturing FPs are beneficial for imaging gene expression dynamics and possess higher brightness in growing cells than slowly maturing FPs[4]. mStayGold exhibited rapid maturation at 37^{o}C with a maturation half-time of 5 min, which was 2.6- and 3.4-fold faster than that for StayGold and mNeonGreen, respectively (Fig. 2d and Supplementary Table 1), making mStayGold one of the fastest maturing FPs[4]. Furthermore, mStayGold fluorescences appeared to be highly chemically stable. For example, compared to other monomeric GFPs, mStayGold was characterized by high pH stability with pK_a value of 4.26 (Fig. 2e and Supplementary Table 1). The high pH stability of mStayGold prompted us to check its chemical stability by extended incubation with 6M Gnd (guanidinium hydrochloride). After incubation of mStayGold in 6M Gdn for 24 h it increased its fluorescence by 18% (Fig. 2f); under the same conditions hfYFP, one of the most chemically stable FPs[5], increased its fluorescence by 1%, StayGold lost 2% of its fluorescence, and mNeoGreen completely lost its fluorescence.

Next, it was essential to verify that the introduced mutation did not have a detrimental effect on mStayGold brightness and photostability in mammalian cells. For this, we co-expressed StayGold, StayGold-E138D, and mStayGold with red FP FusionRed via P2A self-cleavage peptide and used red fluorescence for expression level normalization. The intracellular brightness of the mStayGold in live HEK

cells was ~11- and 5-fold higher than that of StayGold and its point mutant StayGold-E138D, respectively (Fig. 2g). PFA-fixation did not affect mStayGold brightness, in turn fluorescence of StayGold and StayGold-E138D dropped by 1.4- and 2.2-fold, respectively (Fig. 2g). We confirmed that such a dramatic difference in brightness was not due to different codon usage in the mStayGold and StayGold genes, cell type or expression vector and that mStayGold brightness was comparable to that of mNeonGreen in live HeLa cells (Supplementary Fig. 6 and Supplementary Table 1). We suggested that the reduced brightness of StayGold might be due to poor maturation in C-terminal fusions (as evident from the crystal structure, the C-terminus is positioned right against the other unit in the dimer, potentially causing sterical hindrance). Sterical hindrance may also explain why monomerizing E138D mutation increased the intracellular brightness of StayGold by almost 2-fold. For photobleaching experiments, we used 35-fold higher illumination power (i.e., 68 mW/mm^2), which we typically use for live cell imaging, in order to observe a difference in photobleaching rate within shorter periods (typical light power intensity did not photobleach proteins on tens of minutes timescale). In live HEK cells, after 15 min of continuous illumination mStayGold lost 47% of its initial brightness while StyGold and StayGold-E138D lost about 20% of fluorescence (Fig. 2h). PFA-fixation had a dramatic effect on photostability, making photobleaching of tested FPs faster and also inverting photostability order: mStayGold appears to be 2.5-fold more photostable compared to StayGold and StayGold-E138D (Fig. 2i). Similarly to live cells, mStayGold in microdroplets in oils photobleached 1.6-fold faster than StayGold and was 3.3-fold more photostable than mNeonGreen (Supplementary Fig. 5 and Supplementary Table 1).

Encouraged by the high performance of mStayGold in live mammalian cells, we sought to explore its applicability for long-term super-resolution imaging of highly dynamic subcellular structures in live cells. As imaging modalities, we chose to test spinning disk super-resolution by optical pixel reassignment and structure illumination microscopy (SIM), which are currently one of the most popular super-resolution imaging techniques for live cells. First, we confirmed that mStayGold possessed correct localization in demanding fusions with structural proteins, including vimentin, mitochondrial presequence of human cytochrome c oxidase subunit VIII, H2B, keratin, α-tubulin, β-actin (Fig. 3a-g). Using super-resolution confocal microscopy, we could image α-tubulin and β-actin dynamics for up to 60 min without photobleaching (Fig. 3g, h, Supplementary Video 1,2). Furthermore, mStayGold allowed tracking fast dynamics of EB3 protein with a high sampling rate under continuous illumination for 2.5 min without any photobleaching (Supplementary Video 3). Using SIM imaging, we achieved ~65 nm later resolution by imaging actin filaments for over 15 min without photobleaching (Figure 3i,j,k, Supplementary Video 4).

As a complementary super-resolution imaging modality, we used Bleaching/Blinking Assisted Localization Microscopy (BaLM)[6], which provides all benefits of other localization microscopy techniques, such as STORM and PALM, without the need for photoactivatable/photoswitchable fluorophores. Under the illumination of 0.5 kW/cm^2 at 488 nm light vimentin-mStayGold blinked on a single molecule level allowing the reconstruction of the subdiffraction image of vimentin fibers in live HeLa cells using the BaLM technique (Figure 3l-o). Compared to other blinking proteins (e.g., mNeonGreen, mAvicFP1), mStayGold required significantly lower illumination density[7], making it less

phototoxic for live cell imaging. We compared the imaging of vimentin with mStayGold at various frame rates. While all tested frame rates allowed long-term imaging (Supplementary Fig. 7a), 20 fps demonstrated the highest localization precision (Supplementary Fig. 7b). With decorrelation analysis[8], we estimated spatial resolution of 57 nm for reconstructed images and 311 nm for the standard deviation of all widefield frames recorded at 20 fps. Also, we calculated that with 20 fps, more than 50% of mStayGold molecules were detected on consequent frames for more than 100 ms in the nearest vicinity (100 nm) of their initial location (Supplementary Fig. 7d).

In conclusion, the mStayGold protein is a bright and photostable monomeric protein with excellent behavior in fusions, which enables long-term visualization of the cytoskeleton of the cells with sub-diffraction resolution in live cells. We believe that mStayGold will be a valuable probe for cell biologists and the super-resolution imaging community. In addition, the reported crystal structures may prompt further development of mStayGold spectral derivatives and facilitate its utilization for biosensor development.

Online Methods

Molecular cloning and mutagenesis

The StayGold gene was synthesized using polymerase chain reaction (PCR) with overlapping primers listed in Supplementary Table 3 and cloned as a BglII/EcoRI fragment in the pBAD/HisB or pBAD/HisB-Sumo vectors. For PCR amplification, we used a C1000 Touch Thermal Cycler (Bio-Rad, USA).

Random mutagenesis of StayGold gene was performed using PCR in the presence of Mn^{2+} ions in the conditions to achieve 2–3 random mutations per 1000 bp according to the Diversify PCR Random Mutagenesis Kit User Manual (Clontech, USA). The mutagenized StayGold PCR fragment was further cloned in pWA21cBP-mKate2-mTagBFP vector at PstI/BglII restriction sites to swap mKate2 gene and get StayGold in frame with $AraC_{DNA}$ gene and the library was transformed in BW25113 bacterial cells. Next, the obtained bacterial library was spread on LB/agar Petri dishes supplemented with 100 μg/ml ampicillin, 0.02% arabinose and 0.02% rhamnose and after incubation for 24 h at 37°C and 2-4 h at room temperature about twenty thousand of bacterial colonies were imaged using fluorescent stereomicroscope Leica M205FA (Leica, Germany) in blue (405/40 nm excitation and 450/40 nm emission) and green channels (480/40 nm excitation and 535/40 nm emission). About 60-70 the brightest green and dimmer blue fluorescent colonies were analyzed further on bacterial streaks. After each round the ten brightest green/dimmer blue fluorescent variants were further purified from 5 ml of LB medium supplemented with 100 μg/ml ampicillin and 0.004% rhamnose using Ni-NTA resin and their oligomeric state was assessed using FPLC chromatography.

Mammalian plasmid construction

In order to construct the pAAV-*CAG*-NES-mStayGold-P2A-mCherry, pAAV-*CAG*-NES-StayGold-P2A-mCherry, and pAAV-*CAG*-NES-mNeonGreen-P2A-mCherry plasmids, the mStayGold, StayGold, and mNeonGreen

genes were PCR amplified as BglII-EcoRI fragment and swapped with the NCaMP7 gene in the pAAV-*CAG*-NES-NCaMP7-P2A-mCherry vector.

In order to construct the pAAV-*CAG*-dMito-mStayGold plasmid, the mStayGold gene was PCR amplified as XhoI-EcoRI fragment and swapped with the mCherry gene in the pAAV-*CAG*-*d*Mito-mCherry vector.

In order to construct the pAAV-*CAG*-H2B-mStayGold plasmid, the mStayGold gene was PCR amplified as BglII-HindIII fragment and swapped with the B-GECO1 gene in the pAAV-*CAG*-H2B-B-GECO1 vector.

In order to construct the pLU-Vimentin-mStayGold plasmid, the mStayGold gene was PCR amplified as BamHI-BsrGI fragment using PCR with overlapping fragments to delete BsrGI restriction site and swapped with the NeonOxIrr gene in the pAAV-*CAG*-Vimentin-NeonOxIrr vector[9].

To construct the pEMTB-mStayGold plasmid, the mStayGold gene was PCR amplified as BamHI-NotI fragment and swapped with the mNeonGreen gene in the pEMTB-mNeonGreen vector (Addgene Plasmid #137802).

In order to construct the pmStayGold-Keratin plasmid, the mStayGold gene was PCR amplified as KpnI-NotI fragment and swapped with the mKate2 gene in the pmKate2-Keratin vector (Evrogene, Moscow, Russia).

In Piatkevich's group, synthetic DNA oligonucleotides used for cloning were synthesized by Tsingke Biotechnology Co., Ltd. or Zhejiang Youkang Biological Technology Co., Ltd., China. PrimeStar Max master mix (Takara, Japan) was used for high-fidelity PCR amplifications. Restriction endonucleases were purchased from New England BioLabs (USA) and used according to the manufacturer's protocols. DNA ligations were performed using OK Clon DNA Ligation Kit II from Accurate Biotechnology Hunan Co., Ltd, Changsha, China. The ligation products were chemically transformed into the TOP10 E. coli strain (Biomed, China) and cultured according to the standard protocols. Sequencing of bacterial colonies and purified plasmids were performed using Sanger sequencing (Zhejiang Youkang Biological Technology Co., Ltd., China). Small-scale isolation of plasmid DNA was performed with commercially available Mini-Prep kits (Tiangen, China); large-scale DNA plasmid purification was done with Midi-Prep kits (Qiagen, Germany). The gene of StayGold was *de novo* synthesized to substitute phiLOV3 in pAAV-CAG-phiLOV3-P2A-FusionRed plasmid by Tsingke Biotechnology Co., Ltd., China based on the DNA sequences reported on Genbank (https://www.ncbi.nlm.nih.gov/nuccore/2204333803). The gene of mStayGold was *de novo* synthesized by Synbiob Gene Technology Co., Ltd., China. To clone pAAV-CAG- mStayGold-P2A-FusionRed plasmid, the mStayGold's DNA were PCR amplified with KpnI/AgeI flanking sites and swapped with the StayGold gene in the pAAV-CAG- StayGold -P2A-FusionRed. To clone pAAV-CAG-StayGold(E138D)-P2A-FusionRed plasmid, overlap PCR with site mutation of E138D was performed with KpnI/AgeI flanking sites and swapped with the StayGold gene in the pAAV-CAG-StayGold-P2A-FusionRed. To construct CytERM (cytoplasmic end of an endoplasmic reticulum signal anchor membrane protein) fusions, StayGold, StayGold(E138D) and mStayGold were PCR amplified with AgeI/NotI flanking sites and swapped with the mScarlet gene in the pCytERM-mScarlet-N1 (Addgene plasmid #85066). To

construct plasmids for expression of structural protein fusions, mStayGold was PCR amplified and swapped with the corresponding FP genes in pActin-Electra1 (Addgene #184941), pEB3-mScarlet-I (Addgene plasmid #98826) and pTubulin-Electra1 (Addgene #184929) plasmids. For comparison on SIM, StayGold were PCR amplified and swapped with the Electra1 gene in pTubulin-Electra1 (Addgene #184929) plasmid.

The mammalian plasmids generated in the course of this study are available from the WeKwikGene plasmid repository at Westlake Laboratory, China (https://wekwikgene.wllsb.edu.cn/).

Protein purification and characterization

For protein expression, the genes of proteins were PCR amplified as BglII/EcoRI fragments and inserted into the pBAD/HisB (Invitrogen, USA) or pBAD/HisB-Sumo vectors at the BglII/EcoRII restriction sites, and the generated plasmids were transformed into BW25113 bacteria. The bacterial cultures were grown in 200 mL of LB medium supplemented with 0.004% arabinose and 100 μg/ml ampicillin overnight at 37°C and 180 rpm. The cultures were then centrifuged at 4648 g for 10 min. The cell pellets were resuspended in PBS buffer supplemented with 10 mM Imidazole, 300 mM NaCl, lysozyme (100 μg/ml final concentration), lysed by sonication on ice, centrifuged at 20000 rpm for 10 min, 4 °C and further purified using Ni-NTA resin (Qiagen, USA) and dialyzed for 12–16 h against PBS buffer.

For preparative mStayGold protein purification from 2.6 L of LB medium, the BW25113 bacterial cells expressing the HisB-small ubiquitin-like modifier (SUMO)-mStayGold protein were centrifuged for 20 min at 4648 g and 4 °C using the Avanti J-E centrifuge (Beckman Coulter, USA), cells were disrupted by sonication and the cell extract was centrifuged for 30 min at 28,000× *g,* 4 °C. The supernatant with protein was further purified using a 5 mL Ni-NTA Superflow column (Qiagen, Hilden, Germany). Next, His-SUMO-tag was cleaved using SUMO protease and the cleavage mix was applied to a Ni-NTA Superflow column (Qiagen, EU). The concentrated protein was further purified using a 1 mL ResourceQ column (GE Healthcare, Sweden) and finally concentrated until a 10 mg/mL concentration.

Extinction coefficients were determined by alkaline denaturation method by addition of the 2M NaOH water solution to protein solution in PBS buffer till 1M final concentration and assuming that GFP-like chromophore has extinction coefficient equal to 44,000 $M^{-1}cm^{-1}$ in 1 M NaOH[10]. The absorption spectra were registered using the NanoDrop 2000c spectrophotometer (Thermo Scientific, USA).

Quantum yields of purified proteins mStayGold and mNeonGreen were determined in the PBS buffer by the tangent of the slope of the dependence of integral fluorescence values in the range of 480-700 nm on absorbance at 470 nm. StayGold protein with a quantum yield of 0.93 was used as a standard. Excitation and emission spectra were recorded using a CM2203 spectrofluorometer (Solar, Belarus).

Photostabilities of purified proteins (50 μM concentration) in PBS buffer were determined in microdroplets in mineral oil. 10 ul of oil was mixed with 1.5 ul of protein solution and placed between coverslips. Protein droplets were photobleached under continuous wide-field illumination using Zeiss Axio

Imager Z2 microscope (Zeiss, Germany) equipped with a X-Cite 200DC XCT200 200 W mercury arc lamp (Lumen Dynamics, Canada), a 63x 1.4 NA oil immersion objective lens (PlanApo, Zeiss, Germany), a 470/40BP excitation filter, a FT 495 beam splitter, and 525/50BP emission filter. Light power density of 64 mW/cm^2 was measured at the rear focal plane of the objective lens. No corrections were applied to the acquired photobleaching curves.

For pH titrations, 5 µl of proteins (till 50 nM final concentration) in PBS buffer were added to 200 ul of buffer (30 mM citric acid, 30 mM borax, and 30 mM NaCl) with pH values (adjusted with HCl or NaOH) ranging from 2 to 11 in 0.5 pH units interval in a 96-well black clear bottom plate (Thermo Scientific, USA). After proteins incubation for 20 min at room temperature, the fluorescence values were measured using a ModulusTM II Microplate Reader (TurnerBiosystems, USA) equipped with fluorescence optical kit ex 490/em 510-570.

For comparison of stabilities of proteins in 6M guanidinium hydrochloride (GdnHCl), the 5 ul of proteins (till 50 nM final concentration) in PBS buffer were added to 200 ul of 30 mM HEPES, pH 7.80 buffer supplemented with 6M GdnHCl. After incubation of proteins at room temperature for 24 hours, the green fluorescence was registered using a ModulusTM II Microplate Reader (TurnerBiosystems, USA) equipped with fluorescence optical kit ex 490/em 510-570.

The oligomeric state of the proteins was characterized using ÄKTA prime plus and ÄKTA explorer 100 systems (GE Healthcare, Sweden) and a Superdex 75 10/30 GL column (GE Healthcare, Chicago, IL, USA) equilibrated with 20 mM Tris-HCl, pH 7.5, 200 mM NaCl.

Characterization in cultured mammalian cells

HEK293FT (Invitrogen) and HeLa (ATCC CCL-2) cells were authenticated by the manufacturer using STR profiling, reauthenticated in our lab by inspecting stereotypical morphological features under widefield microscope and tested negative for mycoplasma contamination to their standard levels of stringency. Authentication by morphology was performed every time before transient transfection. HEK293FT and HeLa cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin (PS), and seeded in 24-glass bottom well plate (P24-0-N Cellvis, USA) or glass bottom dish (MatTek, USA) after Matrigel (356235, BD Biosciences, USA) coating. Cells were transfected at 80-90% confluency using Hieff Trans Liposomal Transfection Reagent (Yeasen Biotechnology, 40802ES02) according to manufacturer's protocol, and imaged 24-48 h post transfection. Cell were imaged using a Nikon Ti2-E widefield fluorescence microscope equipped with Spectra III Light Engine (LumenCore, USA), the ORCA-Flash 4.0 V3 sCMOS camera (Hamamatsu, Japan), 10x NA0.45 and 20x NA0.75 objective lenses (Nikon, Japan) controlled by NIS-Elements AR 5.21.00 (Nikon Japan) or using a laser spinning-disk Andor XDi Technology Revolution multi-point confocal system (Andor Technology, Belfast, UK).

To measure the brightness of mStayGold in live cell, HEK cells were transfected with the P2A coexpression plasmids and were imaged in FITC (excitation 475/28 nm from SpectraIII LumenCor;

emission 594/40 nm) and TRITC (excitation 555/28 nm from SpectraIII LumenCor; emission 535/46 nm) channels. To obtain statistically significant datasets, we performed 2 independent transfections and ROIs were determined using auto-detect function of NIS Elements software limiting the ROI area to 50 μm^2 as a minimal size of HEK cells. The mean fluorescence intensity in FITC and TRITC channels for ROIs was extracted and FITC-to-TRITC were calculated after background subtraction for each channel, which were used for the comparison of intracellular brightness under corresponding imaging conditions. The data were excluded from the analysis if cells were out of focus.

To measure photostability in live HEK cells, HEK cells were transfected as described before and imaged under continuous FITC wide-field excitation. A 20x NA0.75 objective lens (Nikon, Japan) and SpectraIII LumenCor were used and set at 100% power. The illumination power at the focal spot was 68 mW/cm^2. The photobleaching curves were calculated for each cell individually and reported as the mean photobleaching curve for each protein (averaged from all individual curves). Cells that detached or died during photobleaching experiments were excluded from data analysis.

To evaluate the photostability under live-cell confocal imaging conditions, HeLa cells were transfected with pmStayGold-Tubulin-N1 in glass-bottomed 24-well plates and imaged for one hour after 24 h transfection using CSU-W1 SoRa imaging setup of Nikon Spinning Disk Field Scanning Confocal System with 488 nm excitation using a × objective lens (power at object plane mW).

To compare the brightness of mStayGold with other proteins after PFA fixation, HEK cells were transfected as described above, and imaged using fluorescence wide-field microscope 36-48 h post-transfection. Cells were washed with PBS twice and fixed with 4% PFA (15714, Electron Microscopy Sciences, USA) in PBS at room temperature for 10 min. Fixed cells were gently washed with PBS twice and imaged under identical imaging settings for each protein. Image analysis was performed as described above for live cells. To obtain statistically significant datasets, 2 independent transfections were performed and ROIs were determined using the auto-detect function of NIS Elements software.

To quantify monomeric state of mStayGold in mammalian cells, OSER assay (ref) was utilized as described before. Briefly, HeLa (ATCC CCL-2) cells were cultured and transfected with the corresponding plasmids as described for HEK cells and imaged 12-18 h post-transfection using FITC channel. For each protein we performed at least two independent transfections. To obtain the images of many cells in each well, large image mode and automated image stitching were used with 10% overlap in NIS elements software. Cells were analyzed as described previously, positive cells selected for analysis had overall similar fluorescence brightness, and cells that were significantly brighter were excluded (indications of unhealthy or highly stressed cells). Cells with non-spherical nuclei, ER sheet architectures, or condensed nuclei were also excluded from the assay.

Super-resolution BALM imaging of the cytoskeleton of cultured mammalian cells

Immediately before imaging cell medium was changed to minimal essential medium (MEM, Sigma-Aldrich) supplemented with 20 mM HEPES. Single-molecule localization super-resolution imaging of

living cells was performed using Nanoimager S (ONI, UK) microscope, equipped with Olympus UPlanSApo x100 NA 1.40 oil immersion lens, 488 nm laser, 560 nm on-camera beam splitter and Scope8 sCMOS camera. Imaging was performed using 475 W/cm^2 488 nm laser with 10-50 ms exposure time (100-20 frames per second).

Image acquisition and super-resolution reconstruction was performed using NimOS 1.18.3.15066 (ONI, UK). Image reconstruction was performed using default parameters. Data analysis was performed using FiJi ImageJ 1.53f51 (ref.[11]) and custom Python 3.9 scripts. For blinking duration calculation the "Tracking" tool of NimOS was used. Spatial resolution was calculated using decorrelation analysis[8] with default parameters. As a widefield image, the standard deviation of images stack was used.

Super-resolution HIS-SIM imaging of Tubulin and Actin in HeLa cells

The procedure for imaging was performed by following the previous report[12]. Briefly, cells were seeded in 3 cm^2 dishes with glass bottom and maintained at 37℃ and 5% CO_2 incubator. After 24 hours of transfection, the dishes were chambered in a humidified live cell imaging workstation for live SIM imaging. Super-resolution imaging of Actin and Tubulin structures were performed using commercialized HIS-SIM (High Intelligent and Sensitive Structured Illumination Microscope) provided by Guangzhou Computational Super-resolution Biotech Co., Ltd. Images were acquired using a 100×/1.5 NA oil immersion objective (Olympus) using 60% 488 nm laser with 10ms exposure time. To further improve the resolution and contrast in reconstructed images, sparse deconvolution was used by the previous report[13]. The software of HIS-SIM, IMAGER, was used for data collection and export. Data analysis was performed using Fiji 2.9.01/1.53t ImageJ.

Protein crystallization

An initial crystallization screening of mStayGold was performed with a robotic crystallization system (Rigaku, Woodlands, TX, USA) and commercially available 96-well crystallization screens (Hampton Research, Aliso Viejo, CA, USA and Anatrace, Maumee, OH, USA) at 15 °C using the sitting drop vapor diffusion method. The protein concentration was 8.3 mg/mL in the following buffer: 50 mM Tris, 150 mM NaCl pH 7.5. The initial conditions were optimized by the hanging-drop vapor-diffusion method in 24-4 Intelli plates. The crystals were obtained within several days under the following conditions: 28% PEG3350, 0.2 M lithium sulfate, 0.1 M Na Acetate pH4.6; 30%PEG8000, 0.1M NH4 Acetate, 0.1M Na Cocadilate pH 6.5, and PEG 3350 27%, 0.2M Lithium Sulfate, 0.1M Tris-Cl pH 8.5.

Data collection, structure solution and refinement

Crystals were briefly soaked in a cryosolution containing precipitant supplemented with 20% Glycerol (Hampton Research, Aliso Viejo, CA, USA) immediately prior to diffraction data collection and flash-frozen in liquid nitrogen. The X-ray data were collected from a single crystal (pH 6.5) at 100 K at beamline BL19U1 of the National Facility for Protein Science Shanghai at Shanghai Synchrotron (Shanghai, China). The X-ray data for other crystals (pH 4.6 and 8.5) were collected at

Regaku XtalLAB Synergy-S laboratory system (The Woodlands, Texas, USA) . The data were indexed, integrated, and scaled using the XDS program[14] (Supplementary Table 2). Structure of mStayGlold crystallized at pH 6.5 was solved using MOLREP program[15] with a subunit A of *pdb* 8BXT as a model. For other structures mStayGold-pH6.5 was used as a model. Refinement was carried out using the REFMAC5 program of the CCP4 suite[16]. The visual inspection of electron density maps and the manual rebuilding of the model were carried out using the COOT interactive graphics program[17]. The hydrogen atoms in fixed positions were introduced during the refinement. In the final model, an asymmetric unit of every structure contained two independent copies of the protein with chromophores and solvent molecules.

Data analysis and statistics

To estimate the significance of the difference between two values, we used the Mann–Whitney Rank Sum Test and provided *p*-values calculated for the two-tailed hypothesis. We considered the difference as significant if the *p* value was < 0.05.

Declarations

Data availability

All the data generated and analyzed in the study will be deposited online before publication.

Code availability

Custom code used in this study for BaLM imaging reconstruction is available at https://github.com/Perfus/mStayGold_SR.

Acknowledgements

We are grateful to Aisylu N. Asanova, Tatyana P. Kuzmicheva, and Anastasiya N. Romantsova for their help with mammalian plasmids cloning. We thank Guanzhou CSR Biotech Co. Ltd assistance with live-cell imaging and image analysis using HIS-SIM. This research was partially funded by an grant of the National Research Center Kurchatov Institute (development of mStayGold and characterization of mStayGold in vitro) and thematic plan of NRC Kurchatov Institute (crystallization and X-ray analysis); by the Ministry of Science and Higher Education of the Russian Federation for the development of the Kurchatov Center for Genome Research 075-15-2019-1659 (protein purification); by start-up funding from the Foundation of Westlake University, Westlake Laboratory of Life Sciences and Biomedicine, National Natural Science Foundation of China grants 32050410298 and 32171093 to K.D.P (characterization of mStayGold in mammalian cells and super-resolution imaging). The work was also supported by the Resource Centers department of the National Research Center Kurchatov Institute (imaging of bacteria). The authors acknowledge the BL19U1 beamline of Natinal Facility for Protein Science Shanghai at Shanghai Synchrotron.

Author contributions

G.D.L., O.M.S., and F.V.S. developed mStayGold and characterized it in vitro. A.V.V., and V.R.S., purified mStayGold. H.Z. and K.D.P. characterized mStayGold in mammalian cells and performed live cell imaging. M.M.P. and A.S.M. performed BALM microscopy. A.G.,V.B. and performed X-ray data collection. A.Yu.N. performed crystallization. V.R.S. perform crystallization, structure solution, refinement and analysis. K.D.P., F.V.S, H.Z., M.M.P., and A.S.M. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

Authors declare no competing interests.

References

1. Hirano, M. *et al.* A highly photostable and bright green fluorescent protein. *Nat. Biotechnol. 2022 407* **40**, 1132–1142 (2022).
2. Bustos, S. A. & Schleif, R. F. Functional domains of the AraC protein. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5638–5642 (1993).
3. Cranfill, P. J. *et al.* Quantitative assessment of fluorescent proteins. *Nat. Methods* **13**, 557–562 (2016).
4. Balleza, E., Kim, J. M. & Cluzel, P. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat. Methods 2017 151* **15**, 47–51 (2017).
5. Campbell, B. C., Paez-Segala, M. G., Looger, L. L., Petsko, G. A. & Liu, C. F. Chemically stable fluorescent proteins for advanced microscopy. *Nat. Methods 2022* 1–10 (2022). doi:10.1038/s41592-022-01660-7
6. Burnette, D. T., Sengupta, P., Dai, Y., Lippincott-Schwartz, J. & Kachar, B. Bleaching/blinking assisted localization microscopy for superresolution imaging using standard fluorescent molecules. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 21081–21086 (2011).
7. Gavrikov, A. S., Baranov, M. S. & Mishin, A. S. Live-cell nanoscopy with spontaneous blinking of conventional green fluorescent proteins. *Biochem. Biophys. Res. Commun.* **522**, 852–854 (2020).
8. Descloux, A., Grußmayer, K. S. & Radenovic, A. Parameter-free image resolution estimation based on decorrelation analysis. *Nat. Methods 2019 169* **16**, 918–924 (2019).
9. Subach, O. M. *et al.* Slowly reducible genetically encoded green fluorescent indicator for in vivo and ex vivo visualization of hydrogen peroxide. *Int. J. Mol. Sci.* **20**, (2019).
10. Subach, F. V. *et al.* Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods 2009 62* **6**, 153–159 (2009).
11. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods 2012 97* **9**, 676–682 (2012).

12. Huang, X. *et al.* Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy. *Nat. Biotechnol. 2018 365* **36**, 451–459 (2018).
13. Zhao, W. *et al.* Sparse deconvolution improves the resolution of live-cell super-resolution fluorescence microscopy. *Nat. Biotechnol. 2021 404* **40**, 606–617 (2021).
14. Kabsch, W. XDS. *urn:issn:0907-4449* **66**, 125–132 (2010).
15. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *urn:issn:0021-8898* **30**, 1022–1025 (1997).
16. Collaborative Computational Project, N. 4 & IUCr. The CCP4 suite: programs for protein crystallography. *urn:issn:0907-4449* **50**, 760–763 (1994).
17. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *urn:issn:0907-4449* **60**, 2126–2132 (2004).

Figures

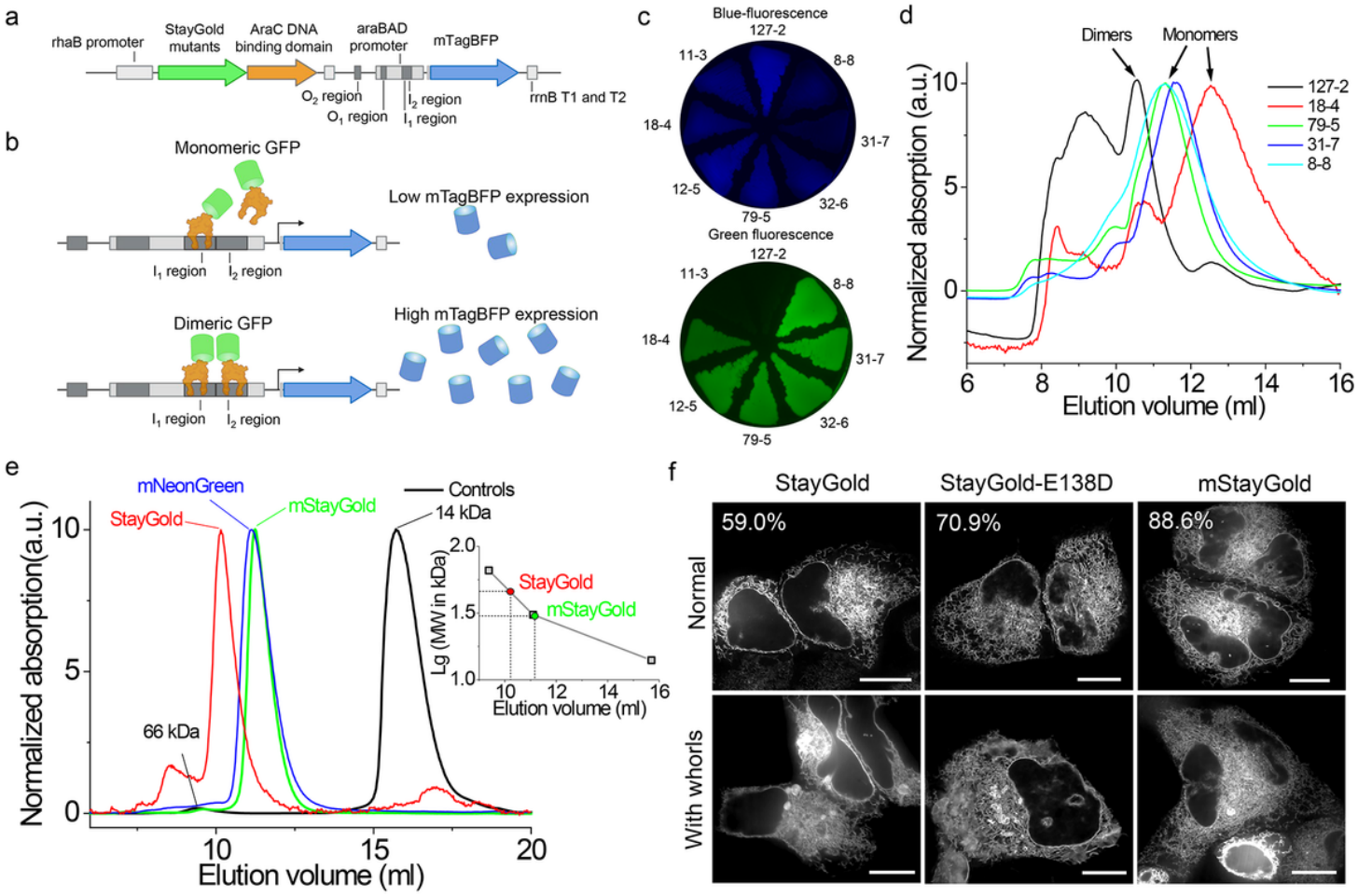

Figure 1

Monomerization of the StayGold protein. (a) Linear map of expression cassettes in the pWA21cBP vector design for screening of monomeric StayGold variants. (b) Proposed model for reporter gene expression regulation using the engineered vector shown in a. (c) Fluorescence images of bacterial streaks expressing the selected mutants using the pWA21cBP vector in blue and green channels. (d) FPLC chromatograms for the selected variants shown in e. (e) FPLC chromatograms of the mStayGold, mNeonGreen, and original StayGold proteins. Insert shows calibration curve for molecular mass calculation. (f) Representative images of live HeLa cells expressing CytERM fusions of StayGold, StayGold-E138D, and mStayGold (n=449, 396, and 671 cells from two independent transfections each, respectively) Upper row, representative normal cells; lower row, representative cells with whorls, the percentage indicates the fraction of normal cells. Scale bars, 10 μm.

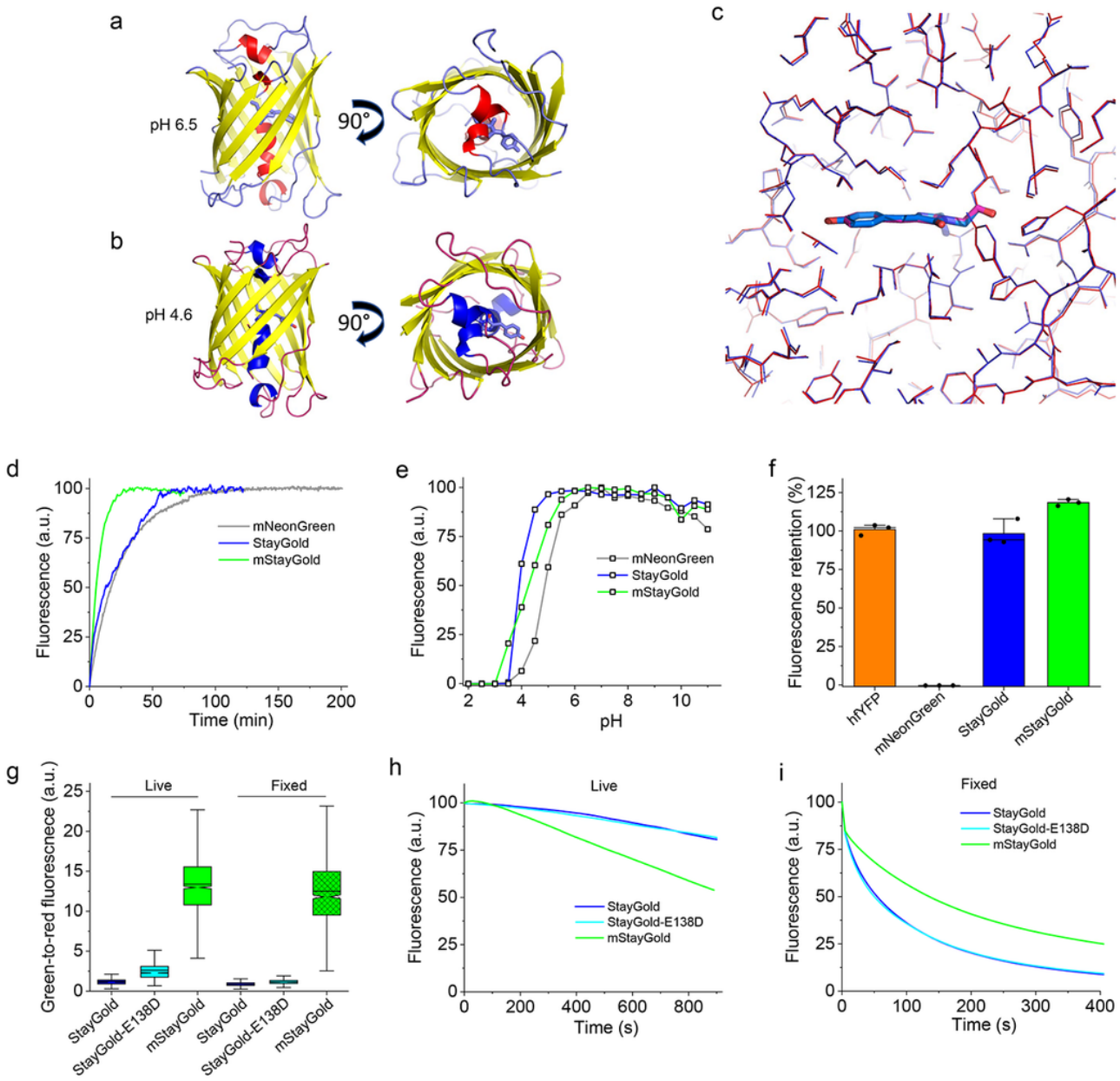

Figure 2

Structure and biochemical properties of purified mStayGold and its behavior in mammalian cells. (a,b) Overall 3D structure of mStayGold at 6.5 (a) and pH 6.5 (b). (c) Superposition of mStayGold structures at pH 4.6 (red) and 6.5 (blue) (d) Fluorescence maturation kinetics in solution at 37^{o}C. (e) pH stability (n = 3 technical replicates each). (f) Fluorescence retention in 6M GdnHCl solution after 24 h incubation (n = 3 technical replicates each). (g) Normalized intracellular brightness in live and PFA-fixed HEK cells in standard GFP filter (n = 2058, 3613, 2293, 1099, 540, 1198 cells for StayGold, StayGold-E138D, and mStayGold for live and fixed conditions, respectively, from two independent transfections each). Box plots with notches: narrow part of notch, median; top and bottom of the notch, 95% confidence interval for the median; top and bottom horizontal lines, 25% and 75% percentiles for the data; whiskers extend 1.5 × the interquartile range from the 25th and 75th percentiles; horizontal line, mean; outliers not shown but included in all calculations and available in the source datasets. (h Normalized photobleaching curves of StayGold, StayGold-E138D, and mStayGold in live HEK cells under continuous wide-field illumination (n=36, 49, and 37 cells, respectively, from two independent transfections each). (i) Normalized photobleaching curves of StayGold, StayGold-E138D, and mStayGold in PFA-fixed HEK cells under continuous wide-field illumination (n=42, 26, and 27 cells, respectively, from two independent transfections each).

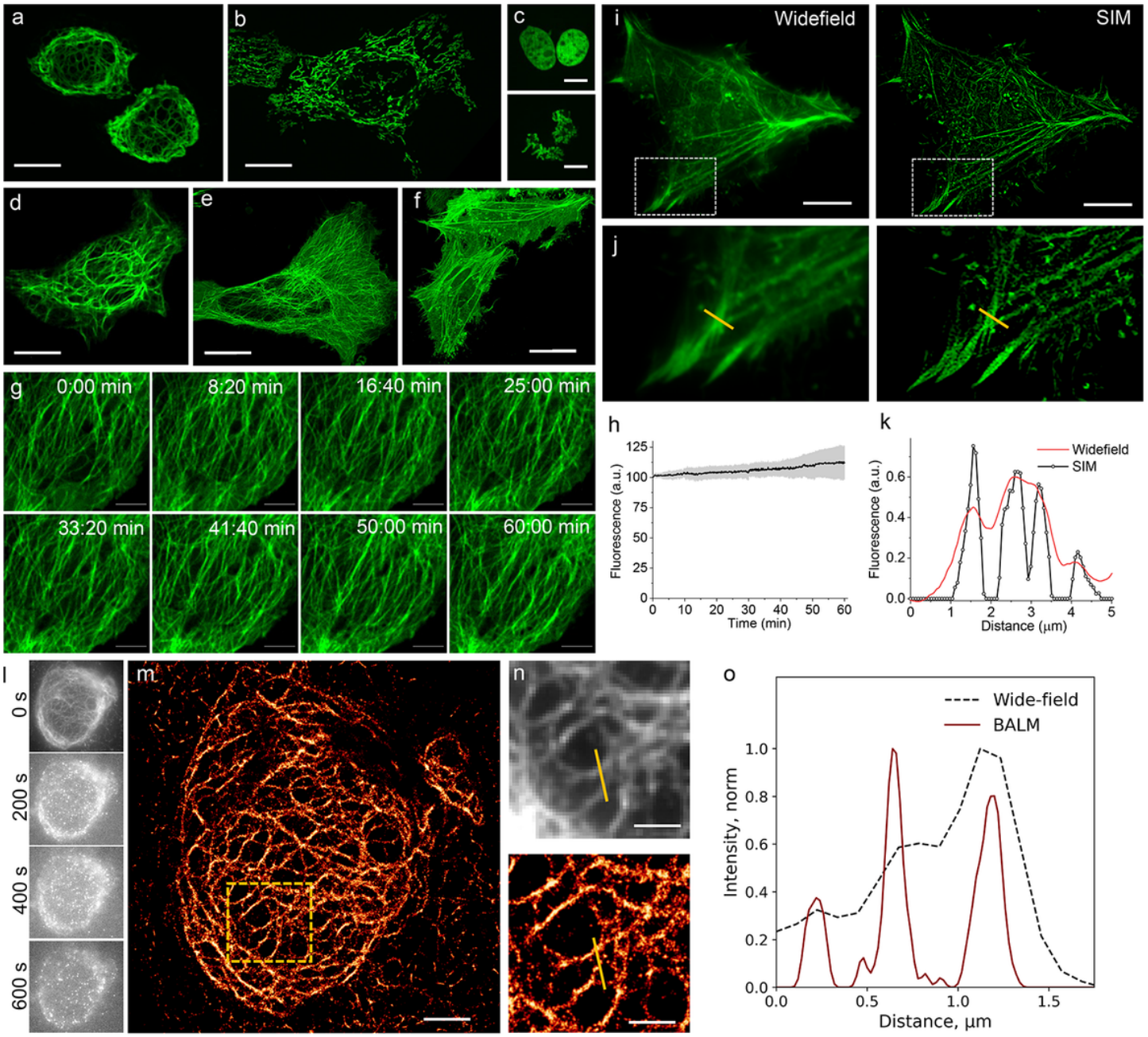

Figure 3

mStayGold enables long-term super-resolution imaging of live HeLa cells. (a-f) Confocal images of HeLa cells expressing mStayGold fusions with (a) vimentin, (b) mitochondrial presequence of human cytochrome c oxidase subunit VIII, (c) H2B, (d) keratin, (e) α-tubulin, (f) β-actin (n= 10, 9, 12, 4, 55, 23 cells from 4, 3, 2, 2, 4, and 2 independent transfections, respectively). Scale bars, 10 μm. (g) Long-term confocal super-resolution imaging of tubulin dynamics using CSU-W1 SoRa imaging setup (n = 15 cells from 4 independent transfections; imaging frequencies 0.2 Hz). Scale bars, 1 μm. (h) Photobleaching curve of tubulin microtubules during the experiment shown in g (n = 15 cells from 4 independent transfections; black line mean; shaded area, standard deviation). (i) Representative images of HeLa cell expressing mStayGold-actin under widefield (left) and SIM (right) microscopy (n=4 cells from 2

independent transfection). Scale bars, 10 μm. (j) Magnified views of boxed regions in i. (k) Quantification of line cuts in j. (l) Individual wide-field frames at indicated time points during BALM imaging of live HeLa cell expressing vimentin-mStayGold (n= 2 cells from 1 transfection; 20 frames per second). (m) The super-resolution reconstruction from 1-5000 frames for imaging shown in l. Scale bar is 10 μm. (n) Magnified views of the box region in m of standard deviation image of widefield (upper) and super-resolution (lower) imaging. (o) Quantification of line profiles in n demonstrating 57 nm of lateral resolution.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- mStayGoldSupplementaryInformationv1.docx
- SupplementaryVideo1.mp4
- SupplementaryVideo2.mp4
- SupplementaryVideo3.mp4
- SupplementaryVideo4.mp4