nature biotechnology

Brief Communication

https://doi.org/10.1038/s41587-023-02018-w

A monomeric StayGold fluorescent protein

Received: 12 March 2023

Accepted: 5 October 2023

Published online: 11 December 2023

Check for updates

Esther Ivorra-Molla[1,5], Dipayan Akhuli[1,2,5], Martin B. L. McAndrew[3,4], William Scott[1], Lokesh Kumar[1], Saravanan Palani[2], Masanori Mishima[1] ✉, Allister Crow[3] ✉ & Mohan K. Balasubramanian[1] ✉

StayGold is an exceptionally bright and stable fluorescent protein that is highly resistant to photobleaching. Despite favorable fluorescence properties, use of StayGold as a fluorescent tag is limited because it forms a natural dimer. Here we report the 1.6 Å structure of StayGold and generate a derivative, mStayGold, that retains the brightness and photostability of the original protein while being fully monomeric.

Fluorescent proteins, starting from the *Aequorea victoria* green fluorescent protein (GFP) to its variants and other fluorescent proteins, have advanced the study of biological processes across scales, from single molecules to whole-tissue behavior. The recently developed fluorescent protein StayGold (an engineered variant of a GFP from *Cytaeis uchidae*) is of special interest because of its brightness and exceptional resistance to photobleaching, with potential uses from single particle to volumetric imaging[1].

We determined the structure of StayGold using X-ray crystallography. Crystals of StayGold belong to space group $P6_1$ and diffract to 1.6 Å resolution. The structure of the StayGold dimer is shown in Fig. 1a with data collection and refinement statistics in Supplementary Table 1. The underpinning electron density is of high quality throughout (Supplementary Fig. 1). Each StayGold monomer is composed of an 11-stranded β-barrel that is almost identical to the well-characterized GFP of *A. victoria*[2]. The StayGold fluorophore is located at the center of the barrel and is formed from residues Gly57, Tyr58 and Gly59 (Fig. 1b). A notable feature of StayGold is the presence of a chloride ion immediately beside the fluorophore. The chloride ion is located within the same plane as the fluorophore and is held in place by electrostatic interactions with Lys61 and Lys192 on one side and Arg86 on the other. The chloride ion interacts with the carbonyl oxygen of the fluorophore and is close to the hydrogens on both the C_β and C_δ atoms of what was originally Tyr58. An equivalent chloride ion is also found in mNeonGreen[3]. Further experiments will be needed to understand the role of chloride in the brightness of StayGold.

Several residues on the inside of the StayGold protein barrel make hydrogen bonds with the fluorophore. These residues are Tyr84, His102, Asn137 and Glu211. The sidechains of Thr54, Val139, Leu154, Val152 and Lys192 are also positioned close to the fluorophore and could perhaps be targeted by mutagenesis campaigns seeking to generate new StayGold variants.

A current limitation of StayGold is that it forms a natural homodimer. Many researchers are wary of using StayGold as a fluorescent fusion protein because unintended dimerization of StayGold-tagged molecules might lead to experimental artifacts. Dimerization-induced artifacts are well documented in the fluorescent protein literature[4,5]. One solution to dimerization is to use tandem pairs of StayGold monomers in protein fusions[1]. However, tandem fusions do not necessarily preclude intermolecular dimerization, and the size of the tandem StayGold tag (~50 kDa) is prohibitively large. Furthermore, community investment in the future development of StayGold proteins for use in applications such as protein–protein Förster/fluorescence resonance energy transfer (FRET), multicolor imaging, split-fluorescent proteins or as biosensors (all mature technologies in conventional fluorescent protein systems) would be greatly accelerated if a monomeric form of StayGold were available. We therefore used our crystal structure to engineer a monomeric form of StayGold.

Using the computer program PISA[6], we initially identified two interfaces in the crystal structure that could conceivably represent the dimer interface. Neither interface is the same as the dimer interface found in the original GFP, and both are predicted by PISA to be only marginally stable. To experimentally distinguish the biologically relevant dimeric interface of StayGold from crystal contacts, we performed size-exclusion chromatography (SEC) experiments using StayGold variants with engineered single amino acid substitutions located at each of the potential interfaces (Fig. 2a and Supplementary Table 2). As expected from the crystal structure, the T195K variant (equivalent to A206K in *A. victoria* GFP[4]) does not monomerize the StayGold protein, confirming that the StayGold dimer is distinct. We also ruled out a potential interface formed by residues E24, E31 and K35 because alanine substitutions at these sites did not affect the protein's oligomeric state.

[1]Centre for Mechanochemical Cell Biology and Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry, UK. [2]Department of Biochemistry, Indian Institute of Science, Bangalore, India. [3]School of Life Sciences, University of Warwick, Coventry, UK. [4]Warwick Medical School, University of Warwick, Coventry, UK. [5]These authors contributed equally: Esther Ivorra-Molla, Dipayan Akhuli.
✉e-mail: m.mishima@warwick.ac.uk; allister.crow@warwick.ac.uk; m.k.balasubramanian@warwick.ac.uk

https://doi.org/10.1038/s41587-023-02018-w

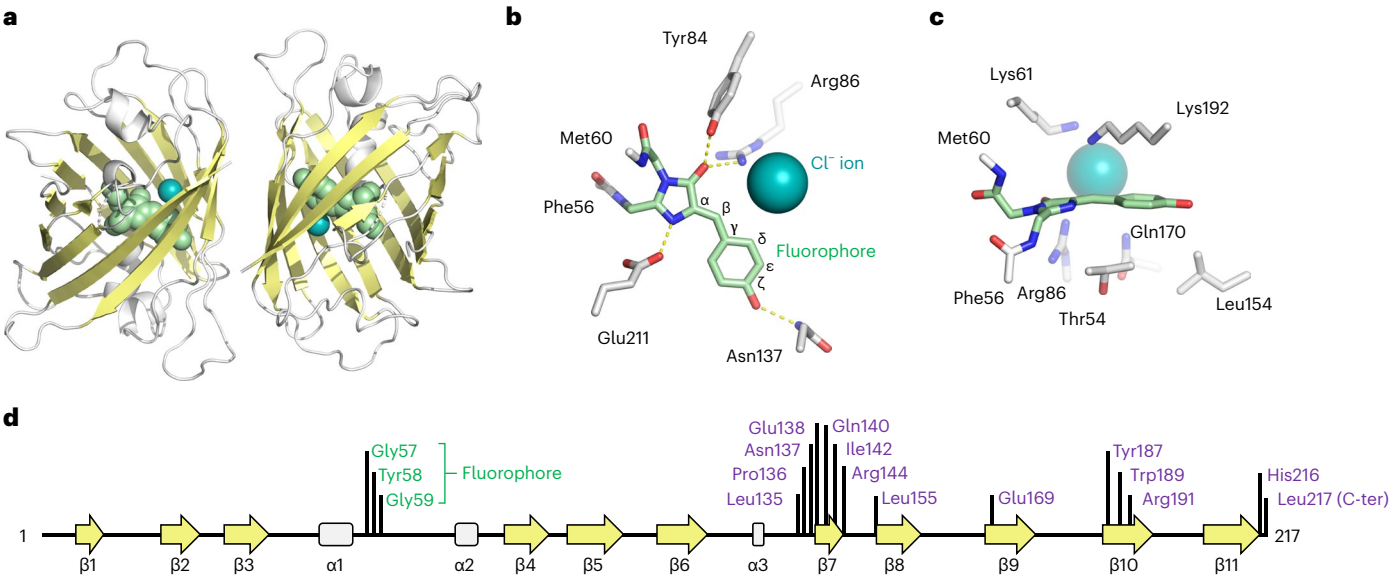

Fig. 1 | Structure of the StayGold fluorescent protein at 1.6 Å resolution. **a**, Overall structure of the StayGold dimer. **b**, Top–down view of the StayGold fluorophore and chloride ion. **c**, Side-on view of the StayGold fluorophore highlighting residues above and below the plane of the fluorophore. **d**, Secondary structure arrangement in the StayGold peptide. Locations of residues forming the fluorophore are shown in green and residues forming the dimer interface are purple.

Conversely, mutations in the remaining interface did disrupt oligomerization; both I142A and L155A have partly monomeric fractions, while E138A, E138D and Y187A variants are truly monomeric (Fig. 2b).

Integrating the peak areas in our size-exclusion chromatography data shows that the E138D mutant is more than 99% monomer, while the original StayGold is ~98% dimer (see Supplementary Table 2 for an analysis of all 11 variants). Size-exclusion measurements at higher protein concentrations show no evidence of concentration-dependent dimerization for the E138D and Y187A variants at 150 µM (Supplementary Fig. 2). Comparison with protein molecular weight standards (including superfolder GFP (sfGFP) variants and eGFP (enhanced GFP)) suggests an apparent molecular weight of 20 kDa for mStayGold and 40 kDa for StayGold consistent with monomer (26 kDa) and dimer (52 kDa), respectively. Our mutagenesis experiments therefore confirm the arrangement depicted in Fig. 2a as the functionally relevant dimer in StayGold. We note that similar interfaces have been observed for KillerRed[7] (a dimer) and the red fluorescent protein of Discosoma DsRed[8] (a tetramer) among others.

The monomerizing effect of the E138D mutation can be understood from the StayGold structure; shortening of the negatively charged side chain using a Glu-to-Asp substitution forces the carboxyl groups at the tip of each side chain into proximity causing repulsion between monomers (Fig. 2c). Mutation of Glu138 also disrupts stabilizing hydrogen bonding interactions with His216 and Arg191 that are colocated at the dimer interface (Fig. 2c). Indeed, Arg191 forms an important interprotein salt-bridge with the StayGold C-terminus—which may explain why both E138D and E138A mutations are effective in disrupting the dimer.

Having established a monomeric form of StayGold, we next compared its fluorescence properties to the original dimer (Supplementary Figs. 3–7 and Supplementary Table 3). We find that StayGold and mStayGold have near-identical excitation (Ex.) and emission (Em.) maxima (Ex., 497 nm and Em., 504 nm) and similar extinction coefficients (165,000 M^{-1} cm^{-1} for StayGold and 145,000 M^{-1} cm^{-1} for E138D mStayGold). Quantum yields for StayGold and mStayGold are similar to one another (0.91 and 0.87) and high in comparison to sfGFP (0.63). Both proteins are stable over a broad pH range with an apparent pK_a of 4.0 (StayGold) and 4.6 (mStayGold) likely corresponding to protonation of the fluorophore. The half pH unit difference in pK_a was the same for both E138D and Y187A variants suggesting the shift is due to the impact of monomerization itself rather than the nature of the amino acid substitutions used to produce monomers. No difference in the rates of chromophore maturation was observed between StayGold and mStayGold (Supplementary Fig. 8). We therefore find that the monomerizing mutations presented here do not adversely affect the fluorescence properties of mStayGold.

We then assessed the photostability of mStayGold in live and fixed cells. We first fused StayGold, mStayGold or a monomeric GFP to the myosin light chain in *Schizosaccharomyces pombe* and imaged the cytokinetic actinomyosin ring using live cell spinning disc confocal microscopy (Supplementary Video). At full laser intensity, the GFP fusions were quickly photobleached, losing ~90% of their initial emission intensity after 30 s. By contrast, both StayGold and mStayGold lost only ~40% brightness in the same period (Fig. 2d,e; further repeats shown in Supplementary Fig. 9). In a second imaging experiment, we observed the cytoskeleton of fixed human retinal pigment epithelial-1 (RPE-1) cells expressing the actin-binding protein tropomyosin[9] (a naturally dimeric coiled-coil protein[10]) fused to either mNeonGreen[11], StayGold[1] or mStayGold (Supplementary Fig. 10). The use of fixed cells confirmed that the long-lived brightness of StayGold and mStayGold was intrinsic to these proteins and not due to newly folded fluorescent molecules replenishing those lost to photobleaching. Fixed cells also eliminate the issue of protein degradation affecting bleaching measurements. Finally, we monitored photobleaching for freely expressed StayGold and mStayGold in immortalized human retinal cells (RPE-1) and human embryonic kidney (HEK293T) cells (Supplementary Fig. 11a,b). StayGold and mStayGold maintained their fluorescence substantially longer than mNeonGreen. No difference in cellular brightness was observed between StayGold and the monomeric E138D variant, although their brightness was lower than mNeonGreen (Supplementary Fig. 11c). The apparent cellular brightness can be affected by various factors such as rates of transcription, translation, folding, chromophore maturation and protein degradation. Further adaptation might benefit the use of StayGold in applications where the highest expression level of the free form is preferable—such as a cell-fate tracer.

https://doi.org/10.1038/s41587-023-02018-w

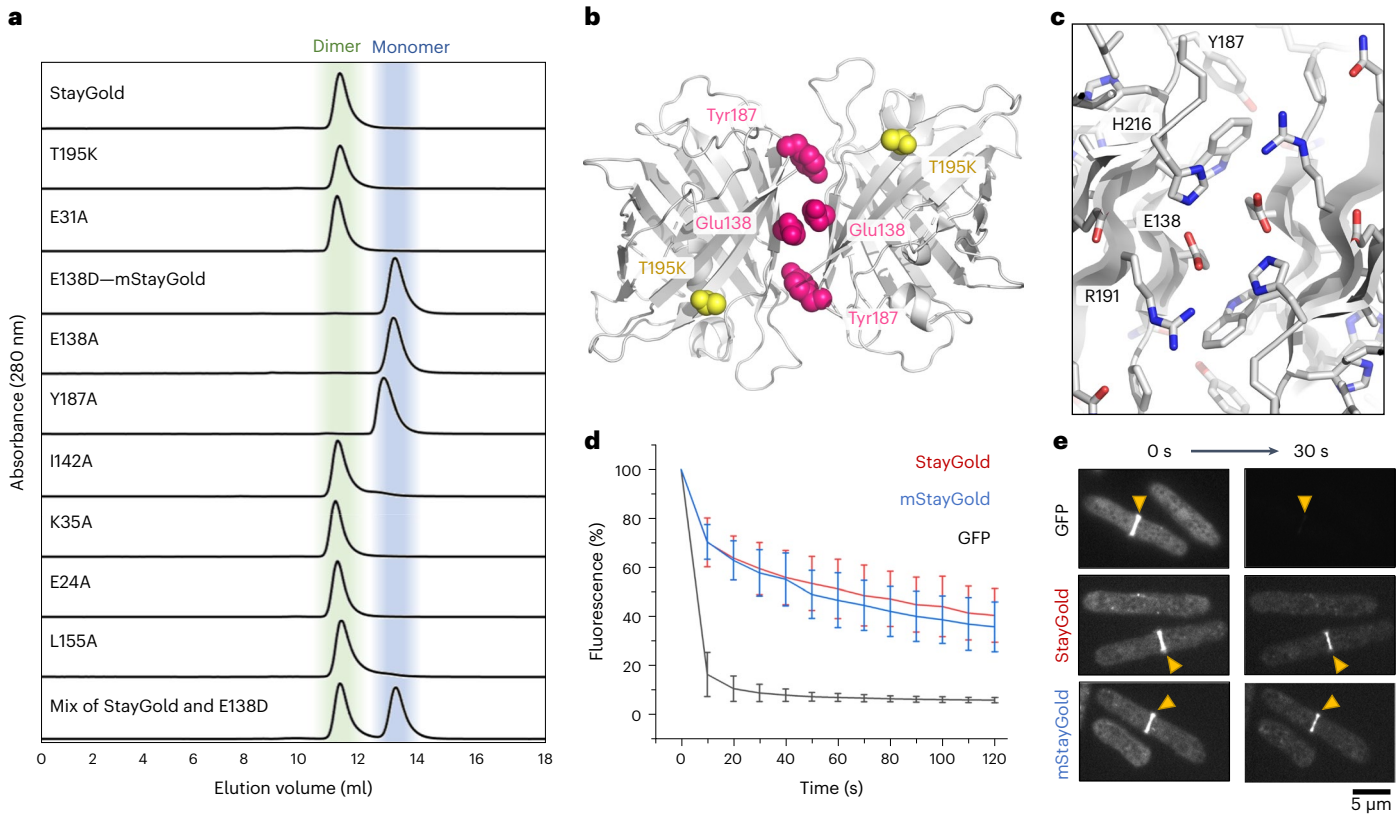

Fig. 2 | Rational design of a monomeric StayGold derivative for use in biological research. a, Size-exclusion chromatography showing the oligomeric state of StayGold and its variants. A mixture of StayGold and mStayGold (that is, E138D) is used to demonstrate their separability. **b**, Location of two monomerizing mutations (E138D and Y187A) and nonmonomerizing (T195K) mutation in the context of the StayGold dimer. **c**, Close-up view of the StayGold dimer interface. **d**, Photostability of StayGold and mStayGold relative to GFP. Fluorescence is given as a percentage of the initial brightness of the *S. pombe* cytokinetic ring for cells expressing the indicated fluorescent protein fusion to the myosin light chain. Error bars indicate s.d. ($n = 20$). **e**, Representative images showing photobleaching for *S. pombe* cytokinetic ring using GFP, StayGold or mStayGold protein fusions. Yellow triangles indicate the cytokinetic ring.

To probe photostability in vitro, we made photobleaching measurements using purified fluorescent proteins immobilized in polyacrylamide (Supplementary Fig. 12). These experiments exclude the effects of differential protein expression (or turnover) allowing us to focus on the photostability of the proteins themselves. Consistent with the data obtained in cells, we find that monomeric StayGold variants last substantially longer under laser illumination than conventional fluorescent proteins such as sfGFP. For example, we measured a half-life of ~70 s for mStayGold E138D compared to 11 s for sfGFP (Supplementary Fig. 12a). Superior photostability of StayGold and mStayGold was also observed under widefield illumination with a metal halide source (Supplementary Fig. 12b).

While testing StayGold for an imaging application in budding yeast, we identified fusions of the original StayGold to the septin collar protein (Shs1) that resulted in frequent mis-localization events and visible aggregation (Supplementary Fig. 13). Hypothesizing that the Shs1–StayGold aggregates were caused by promiscuous dimerization of the StayGold tag, we introduced the E138D mutation. The resultant mStayGold-Shs1 fusions behave as expected, labeling the collar of budding yeast without excess aggregates. These data lend further in vivo support for the monomerizing function of the E138D mutation in StayGold. Introduction of the E138D mutation may be a useful strategy to improve existing fusion protein constructs made by early adopters of StayGold.

Finally, to establish monomerization in vivo, we performed the organized smooth endoplasmic reticulum (OSER) assay[12] for StayGold and mStayGold E138D alongside dTomato (an established dimer) and mTurquoise (a known monomer; Supplementary Fig. 14). The data show that mStayGold E138D is monomeric, with an OSER score comparable to that of the mTurquoise control, while the original StayGold is a weak dimer, with an OSER score that is lower than mStayGold but higher than dTomato. The apparent fragility of the StayGold dimer in the OSER assay is consistent with our observation that the interface can easily be broken with single amino acid substitutions.

In conclusion, we have solved the crystal structure of StayGold and identified E138D as a mutation that renders the protein monomeric without loss of brightness or photostability. We name the E138D variant mStayGold. We anticipate that both the monomerizing mutations found in mStayGold, and the StayGold crystal structure itself will be useful tools for future development of bright fluorescent tags, FRET pairs, split-fluorescent proteins and biosensors.

Note: During the revision of our manuscript, two preprints describing distinct monomeric forms of StayGold were submitted to Research Square[13,14].

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/s41587-023-02018-w.

References

1. Hirano, M. et al. A highly photostable and bright green fluorescent protein. *Nat. Biotechnol.* **40**, 1132–1142 (2022).

2. Yang, F., Moss, L. G. & Phillips, G. N. The molecular structure of green fluorescent protein. *Nat. Biotechnol.* **14**, 1246–1251 (1996).
3. Clavel, D. et al. Structural analysis of the bright monomeric yellow-green fluorescent protein mNeonGreen obtained by directed evolution. *Acta Crystallogr. D Struct. Biol.* **72**, 1298–1307 (2016).
4. Zacharias, D. A., Violin, J. D., Newton, A. C. & Tsien, R. Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916 (2002).
5. Cranfill, P. J. et al. Quantitative assessment of fluorescent proteins. *Nat. Methods* **13**, 557–562 (2016).
6. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
7. Pletnev, S. et al. Structural basis for phototoxicity of the genetically encoded photosensitizer KillerRed. *J. Biol. Chem.* **284**, 32028–32039 (2009).
8. Wall, M. A., Socolich, M. & Ranganathan, R. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat. Struct. Biol.* **7**, 1133–1138 (2000).
9. Hatano, T. et al. mNG-tagged fusion proteins and nanobodies to visualize tropomyosins in yeast and mammalian cells. *J. Cell Sci.* **135**, jcs260288 (2022).
10. Brown, J. H. et al. Structure of the mid-region of tropomyosin: bending and binding sites for actin. *Proc. Natl Acad. Sci. USA* **102**, 18878–18883 (2005).
11. Shaner, N. C. et al. A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat. Methods* **10**, 407–409 (2013).
12. Costantini, L. M., Fossati, M., Francolini, M. & Snapp, E. L. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions: fluorescent protein oligomerization assay. *Traffic* **13**, 643–649 (2012).
13. Ando, R. et al. StayGold variants for molecular fusion and membrane targeting applications. Preprint at *Research Square* https://doi.org/10.21203/rs.3.rs-2941917/v1 (2023).
14. Piatkevich, K. et al. Bright and stable monomeric fluorescent proteins derived from StayGold. Preprint at *Research Square* https://doi.org/10.21203/rs.3.rs-3188559/v1 (2023).

Methods

Cloning, expression and purification of fluorescent proteins

StayGold and mStayGold were each cloned into pET-MCN vectors and transformed into BL21(DE3) *Escherichia coli* strain. In total, 2 l cultures at OD ~0.6 were induced overnight at 18 °C with 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranoside). Samples were purified with Ni-NTA (nickel-nitrilotriacetic acid) beads in wash buffer containing 50 mM phosphate buffer, 500 mM NaCl and 30 mM imidazole at pH 7.6. Proteins were eluted with 50 mM phosphate buffer, 500 mM NaCl and 500 mM imidazole at pH 7.6. Fluorescent proteins were then exchanged into a 20 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.2 and 150 mM NaCl using a PD10 desalting column. sfGFP and its variants were produced similarly with mutations confirmed by sequencing.

Size-exclusion chromatography

Size-exclusion chromatography was performed using an ÄKTA Pure FPLC with a Superdex 75 Increase 10/300 GL column. Proteins were concentrated to ~2 mg ml^{-1} (76 µM) and loaded onto the column via a 100 µl loop. The flow rate was 0.8 ml min^{-1}, and the buffer used was 150 mM NaCl and 20 mM HEPES (pH 7.2). Chromatograms were baseline-corrected using the mean absorbance value measured in the first 5 ml following injection. Monomer and dimer percentages were calculated after integrating the peaks corresponding to the monomer and dimer fractions. A higher concentration run using StayGold and mStayGold was performed using 4 mg ml^{-1} (154 µM) protein. Additional runs using a mixture of bovine serum albumin (66 kDa) and hen egg white lysozyme (14 kDa), sGFP (27 kDa) or BIORAD molecular weight standards (thyroglobulin, 670 kDa; γ-globulin, 158 kDa; ovalbumin, 44 kDa; myoglobin, 17 kDa and vitamin B_{12}, 1.35 kDa) were used for estimation of apparent molecular weights.

Crystallization

StayGold was concentrated to 12 mg ml^{-1} in a buffer composed of 150 mM NaCl and 20 mM HEPES (pH 7.5) before setting up crystallization screens in MRC 2-drop plates using a Formulatrix NT8 robot. Crystallization used the sitting drop vapor diffusion method with 1 µl drops composed of either 2:1 or 1:2 ratios of protein solution and crystallization reagent. Crystals of StayGold grew in 1–2 d with many similar hits across the screens. The crystals used for structure determination were obtained using the SG1 HT96 Eco Screen (Molecular Dimensions) condition H4 (0.2 M sodium acetate, 25% PEG (Polyethylene glycol) 3350 and 0.1 M Bis–Tris (pH 6.5)). Crystals formed as long needles that could be broken apart for single crystal data collection.

Structure determination using x-ray crystallography

Drops containing StayGold crystals were supplemented with a solution of 70% crystallization reagent (taken from the reservoir of the plate) and 30% glycerol, collected in litholoops and flash-frozen in liquid nitrogen. X-ray diffraction experiments were conducted at the Diamond Synchrotron using remote collection at beamline I04. Diffraction images were integrated using Dials[15] via Diamond's auto-processing pipeline. Subsequent processing was performed using tools from the CCP4 suite[16]. Space group determination and scaling were conducted with Aimless[17]. Phases were calculated by molecular replacement using a monomer from 5WJ2 (ref. 18) as the search probe. The molecular replacement solution (containing two molecules per asymmetric unit) was found using Phaser[19] software. After an initial round of refinement in Refmac[20], a new map was calculated and density modification (including non-crystallographic symmetry averaging, solvent flattening and histogram matching) was performed with Parrot[21]. A new model was then built into the density-modified map using Buccaneer[22], and the model was completed using iterative rounds of model building with Coot[23] and refinement using Refmac[20]. As refinement neared completion, water was added and the model was completed by modeling in the fluorophore and chloride ion. Later rounds of model building and refinement integrated validation tools including COOT[23], Procheck[24] and Rampage[25]. Protein–protein contacts were analyzed using PISA[6] and molecular images generated using PyMOL[26]. The StayGold structure has been deposited in the protein databank[27] with accession code 8BXT. X-ray data and refinement statistics are given in Supplementary Table 1.

Fluorescence excitation and emission spectra

Fluorescence spectra were obtained using a Cary Eclipse fluorescence spectrophotometer using 5 nm excitation and emission slits. Datapoints were collected at 1 nm intervals using a scan rate of 600 nm min^{-1}.

Microscopy and in vivo photostability

S. pombe strains were generated by chromosomal integration of the plasmids pDUAL-p^{adh11}-rlc1-GFP (pLK43), pDUAL-p^{adh11}-rlc1-40aa-StayGold (pLK114) and pDUAL-p^{adh11}-rlc1-40aa-StayGold (E138D, pLK126) at leu1-32 locus in MBY102. The plasmid, pLK43 was generated by Gibson cloning *rlc1* and fluorescent protein fragments into the pDUAL plasmid[28]. The plasmid pLK114 was generated by swapping GFP with a 40 amino acid linker (LEGSGQGPGSGQGSGPGSGQGSGPGQGSGPGQGSGPGQG) and StayGold fragment with the help of NEBuilder HiFi DNA Assembly Master Mix (NEB, E2621L). The *rlc1*-40aa-StayGold(E138) variant was generated using site-directed mutagenesis. The cells were grown at 30 °C in YES medium.

The mammalian cell line, RPE-1 cells were transfected with one of the following six expression plasmids: (1) mNeonGreen-Tropomyosin2 (*pCMV*-mNeonGreen-40aaLinker-*TPM2.2*)[9], (2) Tropomyosin2 StayGold dimer (*pCMV*-StayGold-40aaLinker-*TPM2.2*), (3) Tropomyosin2 StayGold monomer (*pCMV*-StayGold (E138D)-40aaLinker-*TPM2.2*) or (4–6) the equivalent versions lacking the tropomyosin sequences (which would express the unfused fluorescent proteins). The Tropomyosin2 StayGold plasmids were constructed by swapping mNeonGreen for StayGold using the NEBuilder HiFi DNA Assembly Master Mix (NEB, E2621L). The Tropomyosin2-mStayGold variant was generated by incorporating the E138D mutation using site-directed mutagenesis.

Immortalized (hTERT, human telomerase reverse transcriptase) diploid human RPE-1 cells (ATCC, CRL-4000) were cultured in Dulbecco's modified Eagle's medium/Nutrient Mixture F-12 Ham with 15 mM HEPES and sodium bicarbonate (Sigma-Aldrich, D6421) supplemented with 6 mM L-glutamine (Gibco, 25030-081), 10% FBS (Sigma-Aldrich, F7524), 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (Gibco, 15140-122) at 37 °C under 5% CO_2.

For transfection, RPE-1 cells were grown on ibiTreat 2 Well µ-Sildes (Ibidi, 80286) to 50% confluency. Cells were then transfected using Lipofectamine 2000 (Invitrogen, 11668-019) and Opti-MEM reduced-serum media (Gibco, 31985-062) according to the manufacturer's instructions. Transfected RPE-1 cells were fixed in 4% paraformaldehyde/PBS and sealed with Vectashield (Vector, H-1000) 20 h post-transfection for imaging.

For measurement of the practical brightness of free StayGold, mStayGold and mNeonGreen expressed in mammalian cells, the RPE-1 cells cultured on a µ-Slide 8 Well high (Ibidi, 80806) were imaged 30 h after transfection using a DeltaVision microscope system (Applied Precision) equipped with a CoolSNAP HQ2 camera (Photometrics). The images with the GFP filter set and 32% illumination were acquired by softWoRx software (v5.5.1) using a ×10 UPlanSApo lens (Olympus; numerical aperture (NA) 0.4, air) with 0.25 s exposure in a z-stack of 10 images with 1.5 µm spacing. Fluorescent cells were detected by a custom ImageJ/Fiji script using trainable Weka segmentation[29]. The mean fluorescence intensities of each cell above the background level were statistically analyzed by R version 4.3.1 (https://www.r-project.org/).

For comparison of photostability, images were acquired with a spinning disk confocal microscope (Andor Revolution XD imaging system, equipped with ×100 oil immersion 1.45 NA Nikon Plan Apo

Lambda, and a confocal unit Yokogawa CSU-X1, Andor iXon Ultra 888 EMCCD camera and Andor IQ acquisition software). The cells were imaged with a 488 nm laser at 100% power (~7.4 W cm^{-2}) at 0.5-s intervals. Fluorescence intensities were measured for the Rlc1 GFP, Rlc1 StayGold dimer and Rlc1 StayGold monomer at the cytokinetic ring (for *S. pombe*) and the entire cell (for RPE-1 cells) for each slice using the image processing software Fiji (https://imagej.net/software/fiji/). The datapoints were normalized to the first intensity measurement.

In vitro photostability

Photostability of pure proteins was measured using the method described in ref. 5. Photobleaching was quantified by embedding the fluorescent proteins into a polyacrylamide gel, which was polymerized between a 1.5 coverslip (Menzel, 12312128) and a glass slide. Other 1.5 coverslips were used as spacers. The final composition of the gel was 20% acrylamide (mono:bis 29:1), 150 mM NaCl and 20 mM Tris (pH 7.6). For laser illumination, the samples were observed using a CellR TIRFM system (Olympus) equipped with an iXon+ DU897 electron-multiplying CCD camera (Andor). The samples were illuminated with a 488 nm laser beam at a 0° angle of incidence. The power of light detected at a ×100 objective lens outlet was 2 mW, corresponding to illumination at 29 W cm^{-2} within a circular area with a radius of 46.5 µm. Similar experiments were performed under widefield illumination using a metal halide source.

Extinction coefficient determination

Absorbance was measured with a Cary 50 Conc UV Visible Spectrophotometer. Molar extinction coefficients (ε) were determined by standardizing the absorbance at the chromophore absorption maxima (488 nm for sfGFP or 496 nm for StayGold and its derivatives) either with the protein concentration[5] or with the chromophore concentration[1,30]. For the protein concentration method, the theoretical molar extinction coefficient at 280 nm ($\varepsilon_{280\,\mathrm{theory}}$) was calculated (https://web.expasy.org/protparam/) based on the protein sequence (minus one tyrosine). The molar extinction coefficient per protein ($\varepsilon_{\mathrm{protein}}$) was calculated by multiplying $\varepsilon_{280\,\mathrm{theory}}$ with the ratio of the measured absorbance values at 280 nm (A_{280}) and the fluorophore absorption maxima (488 nm or 496 nm, A_{peak}).

$$\varepsilon_{\mathrm{protein}} = \varepsilon_{280\mathrm{theory}} \times A_{\mathrm{peak}}/A_{280}$$

The $A_{\mathrm{peak}}/A_{280}$ ratio was determined as the gradient of a A_{peak} versus A_{280} plot.

For the denatured-chromophore method, the absorbance at 447 nm after denaturation with 0.1 M NaOH (5 min for sfGFP and 1.5 min for StayGold and its derivatives, A_{NaOH}) was measured as well as the absorbance at the absorption maxima under native conditions (488 nm for sfGFP or 496 nm for StayGold and its derivatives, A_{neutral}). The chromophore-based molar extinction coefficient ($\varepsilon_{\mathrm{chromophore}}$) was then determined by multiplying the ratio of these measurements to the extinction coefficient of the isolated chromophore at 447 nm ($\varepsilon_{\mathrm{isolated\ chromophore}}$ = 44,000 M^{-1} cm^{-1}) as described previously[1].

$$\varepsilon_{\mathrm{chromophore}} = \varepsilon_{\mathrm{isolatedchromophore}} \times A_{\mathrm{neutral}}/A_{\mathrm{NaOH}}$$

Quantum yield measurements

The quantum yield of each fluorescent protein was determined by comparing the ratio between the absorbance at the peak excitation and the integrated fluorescence emission, with that of fluorescein in 0.1 M NaOH (ref. 5). We used a quantum yield of 0.79 for fluorescein[31,32]. The fluorescence/absorbance ratios were each determined using the gradient from a plot of integrated fluorescence versus absorbance with at least 4 points with an absorbance below 0.1.

Measurement of the pH sensitivity and apparent pK_a

Absorption measurements were made at the peak absorbance wavelength after diluting purified proteins into preprepared buffers containing either 150 mM NaCl, 50 mM NaH_2PO_4 and 50 mM citrate (pH 2.5–7.5) or 150 mM NaCl, 50 mM Tris, 50 mM glycine (pH 8–11). Absorption measurements were normalized to the value at pH 7, and the apparent pK_a was estimated from the plot of normalized absorption versus pH.

Chromophore maturation

Oxidative maturation of the chromophore of StayGold was assessed by preparing the cell lysate containing StayGold expressed under a hypoxic condition and monitoring the increase of the absorbance at 496 nm by the matured chromophore under an aerobic condition[18,33]. The BL21(DE3) cells transformed with a pET-based vector encoding StayGold or a nonfluorescent protein as a control were cultured in 25 ml LB (Lysogeny Broth) with kanamycin to the log phase (A600 ~0.5) and induced for the recombinant protein expression with 0.2 mM IPTG for 4 h at 22 °C with good aeration (this improves expression under hypoxia in the following step). The cells were collected by centrifugation and resuspended with 12 ml of degassed LB with IPTG and kanamycin. The aliquots were transferred to cryotubes (~2.5 ml; Greiner). The tubes were filled to the brim, capped expelling the air and incubated at 22 °C with rotation at 200 r.p.m. for 16–30 h. Cells from a cryotube were lysed with 400 µl BugBuster Protein Extraction Reagent (Millipore) and centrifuged at 20,817*g* for 5 min. The 200 µl supernatant was incubated at room temperature (24 °C) in a disposal cuvette (UVette, Eppendorf) and the absorbance spectra were acquired in a time course. The gradual increase of the turbidity was observed independently of fluorescent protein, due to the increasing scattering, which is inversely proportional to the fourth power of wavelength (the Rayleigh scattering). The fraction due to scattering in the absorbance at the peak wavelength (496 nm) was estimated by extrapolating the absorbance from 525 nm to 575 nm, where the absorbance by the chromophore is minimum, and subtracted from the raw read at the peak. The peak absorbance after the correction for scattering, $f(t)$, was plotted against time, t, and fitted with $f(t) = \beta_0 - \beta_1 \exp(-kt)$ to determine the rate of maturation, k, by nonlinear least squares fitting using the nls() function in R version 4.3.1 (https://www.r-project.org/).

OSER assays

OSER assays were performed as per the method discussed in ref. 12. Hela cells were transfected with one of the following four plasmids: CytERM-dTomato (Addgene, 98834), CytERM-mTurquoise2 (Addgene, 98833), pcDNA3-CytERM-StayGold or pcDNA3-CytERM-StayGold(E138D). The two StayGold plasmids were constructed by cloning G-blocks (Integrated DNA Technologies) into pcDNA3 using NEBuilder HiFi DNA Assembly Master Mix. For imaging, 20 h post-transfection, cell media was replaced with Leibovitz's L-15 Medium (No Phenol Red). Images were acquired using Andor IQ3 software at either 69 nm per pixel with a Nikon Apo ×60/1.40 oil immersion objective lens or at 80 nm per pixel with a Nikon Plan Fluor ×40/1.30 oil immersion objective lens. The fluorophores were excited by laser lines at a wavelength of 405 nm (mTurquoise2), 488 nm (StayGold) or 561 nm (dTomato) as appropriate. Fixed cell images were acquired at room temperature, while live cell images were acquired at 37 °C. Z-stacks were obtained for each cell, followed by Z-projection in Fiji, and manual quantification was performed to count the number of cells with and without whorls.

Reporting summary

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

Data availability

The structure of the StayGold dimer has been deposited at the protein databank (8BXT). Source data for biochemical and spectrometric experiments (Fig. 2a and Supplementary Figs. 2–8) and numerical data

from the microscopy-based photokinetic or photobiological studies (Fig. 2d and Supplementary Figs. 9–14) are available for download. Raw images (>650 GB in total) from the microscopy-based studies (Fig. 2d and Supplementary Figs. 9–14) will be shared upon request.

References

15. Winter, G. et al. DIALS as a toolkit. *Protein Sci.* **31**, 232–250 (2022).
16. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
17. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
18. Campbell, B. C., Petsko, G. A. & Liu, C. F. Crystal structure of green fluorescent protein clover and design of clover-based redox sensors. *Structure.* **26**, 225–237 (2018).
19. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
20. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
21. Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 470–478 (2010).
22. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
23. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
25. Lovell, S. C. et al. Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins* **50**, 437–450 (2003).
26. Schrodinger, LLC. The PyMOL molecular graphics system, version 2.0.
27. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
28. Matsuyama, A. et al. pDUAL, a multipurpose, multicopy vector capable of chromosomal integration in fission yeast. *Yeast* **21**, 1289–1305 (2004).
29. Arganda-Carreras, I. et al. Trainable Weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017).
30. Chalfie, M. & Kain, S. R. (eds.). *Methods of Biochemical Analysis*, pp. 39–65 (John Wiley & Sons, 2005). https://doi.org/10.1002/0471739499.ch3
31. Umberger, J. Q. & LaMer, V. K. The kinetics of diffusion controlled molecular and ionic reactions in solution as determined by measurements of the quenching of fluorescence. *J. Am. Chem. Soc.* **67**, 1099–1109 (1945).
32. Kurian, A. et al. Effect of pH on quantum yield of fluorescein using dual beam thermal lens technique. *J. Opt.* **31**, 29–35 (2002).
33. Campbell, R. E. et al. A monomeric red fluorescent protein. *Proc. Natl Acad. Sci. USA* **99**, 7877–7882 (2002).

Acknowledgements

We thank the staff at the Diamond Synchrotron for the use of X-ray diffraction facilities—staff at the Computing and Advanced Microscopy Unit (CAMDU) for their support on microscopy imaging; C. Smith and R. Cross (University of Warwick, UK) for access to fluorimetry equipment; the Medical Research Council for funding Ph.D. work by M.B.L.M. (MR/N014294/1); the University of Warwick A*STAR Research Attachment Programme for funding Ph.D. work by E.I.M. M.K.B. is funded by a Wellcome Trust Senior Investigator Award (WT101885MA). S.P. is funded by DBT-Wellcome Trust India Alliance (IA/I/21/1/505633) and Science and Engineering Research Board (SERB-SRG/2021/001600). A.C. is funded by the BBSRC (BB/V017101/1).

Author contributions

M.K.B. led cell biology experiments and protein production. A.C. led the structural biology and size-exclusion chromatography experiments. M.M. led photophysical characterization experiments. E.I.-M. and D.A. performed cloning, protein purification and crystallization. E.I.-M., D.A. and A.C. conducted the structure determination. M.B.L.M. performed and analyzed SEC experiments. E.I.-M. and M.M. measured and analyzed photophysical data. L.K. performed imaging experiments in *S. pombe*. D.A. and S.P. performed imaging experiments in *Saccharomyces cerevisiae*. E.I.-M., W.S. and M.M. performed imaging experiments in mammalian cells. A.C., M.K.B. and M.M. wrote the manuscript with contributions from all authors. A.C. and M.K.B. conceived the study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-023-02018-w.

Correspondence and requests for materials should be addressed to Masanori Mishima, Allister Crow or Mohan K. Balasubramanian.

Peer review information *Nature Biotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Mohan K. Balasubramanian, Allister Crow, Masanori Mishima

Last updated by author(s): Sep 26, 2023

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	The GDA suite (for x-ray data collection - maintained by Diamond syncrotron). Andor iQ3, Micro-Manager 1.4.22, Olympus xcellence rt 2.0, Cary WinFLR, Cary WinUV
Data analysis	Pymol 2.5.4, CCP4 8.0.006, Refmac 5.8.0352, Aimless 0.7.8, Phaser 2.8.3, Parrot - 1.0.7, Buccaneer - 1.6.12, Rampage - 1.00.0, Microsoft Excel 2023, R version 4.3.1 (2023-06-16), OriginPro 2023a, Fiji 2.9.0/ImageJ 1.53t

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

The structure of the StayGold dimer has been deposited at the protein databank (8BXT). Source data for biochemical and spectrometric experiments (Figures 2a, S2, S3, S4, S5, S6, S7, and S8) and numerical data from the microscopy-based photokinetic or photobiological studies (Figures 2d, S9, S10, S11, S12, S13, and S14) are available for download as a zip file. Due to the large size of the dataset, raw images from the microscopy-based studies (Figures 2d, S9, S10, S11, S12, S13, and S14) will be shared on request.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

Life sciences study design

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

Sample size	No explicit sample size calculation was performed. We followed community norms. For time-lapse imaging of photobleaching (Figures 2d, S9, S10, S11, S12), the time required for photobleaching was a major factor that limits the sample size for the practically achievable experimental design. In biochemical and spectrometric experiments (Figures S2, S3, S4, S5, S6, S7, and S8) each experiment was repeated at least three times. Size exclusion chromatography was performed in duplicate.
Data exclusions	No data were excluded.
Replication	The photobleaching microscopy in Figures 2d+S9 were independently repeated 4 times. For photobleaching in Figures S10 and S11, the biochemical and spectrometric experiments (Figures S2, S3, S4, S5, S6, S7b, and S8) and visual scoring (Figures S13 and S14), three independent experiments were performed and the consolidated data were presented. Measurement of the sensitivity to pH>7 (Figure S7a) was performed once. The in vitro photobleaching with laser (Figure S12a) was repeated three times and one of the three consistent results was shown. For bleaching with the halogen lamp (Figure S12b), one set of measurements (n=6 for each sample) was performed.
Randomization	For microscopy of the cells (Figures 2d, S9, S10, S11, S13 and S14), the imaged fields were randomly selected.
Blinding	Blinding was not used.

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

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	hTERT RPE-1 cells (CRL-4000), HeLa (CCL-2) and HEK293T (CRL-3216) were all sourced from ATCC.
Authentication	None
Mycoplasma contamination	Cells used in this study were tested for mycoplasma with negative results
Commonly misidentified lines (See ICLAC register)	HEK293T