

Spatiotemporal functional assembly of split protein pairs through a light-activated SpyLigation

Received: 10 September 2021

Accepted: 26 January 2023

Published online: 17 April 2023

Emily R. Ruskowitz^{1,9}, Brizzia G. Munoz-Robles^{2,9}, Alder C. Strange³, Carson H. Butcher⁴, Sebastian Kurniawan¹, Jeremy R. Filteau¹ & Cole A. DeForest^{1,2,5,6,7,8} 

 Check for updates

Proteins provide essential functional regulation of many bioprocesses across all scales of life; however, new techniques to specifically modulate protein activity within living systems and in engineered biomaterials are needed to better interrogate fundamental cell signalling and guide advanced decisions of biological fate. Here we establish a generalizable strategy to rapidly and irreversibly activate protein function with full spatiotemporal control. Through the development of a genetically encoded and light-activated SpyLigation (LASL), bioactive proteins can be stably reassembled from non-functional split fragment pairs following brief exposure (typically minutes) to cytocompatible light. Employing readily accessible photolithographic processing techniques to specify when, where and how much photoligation occurs, we demonstrate precise protein activation of UnaG, NanoLuc and Cre recombinase using LASL in solution, biomaterials and living mammalian cells, as well as optical control over protein subcellular localization. Looking forward, we expect that these photoclick-based optogenetic approaches will find tremendous utility in probing and directing complex cellular fates in both time and three-dimensional space.

Biology comprises a series of well-orchestrated chemical reactions that are precisely controlled in time and three-dimensional (3D) space (that is, four dimensions). Proteins act as the key conductors of these reactions, providing essential and unmatched functional regulation of many bioprocesses across all scales of life. As proteins offer structural integrity, regulate gene expression and serve as the central language of cellular communication, there is little question as to why global research efforts continue to seek to improve existing techniques and develop new ones to regulate protein function within living systems.

While systematic edits to the genome can enable long-lasting over- and underexpression of proteins *in vitro* and *in vivo*, these efforts require long times ranging from many hours to weeks; critical need remains for systems that permit real-time modulation of protein function in a user-defined manner.

A growing and powerful trend towards the induction of bioactivity in living cells involves triggered protein reassembly from non-functional fragment pairs. In this strategy, proteins are genetically split into biologically inactive fragments whose negligible affinity

¹Department of Chemical Engineering, University of Washington, Seattle, WA, USA. ²Department of Bioengineering, University of Washington, Seattle, WA, USA. ³Department of Biochemistry, University of Washington, Seattle, WA, USA. ⁴Department of Biology, University of Washington, Seattle, WA, USA. ⁵Department of Chemistry, University of Washington, Seattle, WA, USA. ⁶Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA, USA. ⁷Molecular Engineering and Sciences Institute, University of Washington, Seattle, WA, USA. ⁸Institute for Protein Design, University of Washington, Seattle, WA, USA. ⁹These authors contributed equally: Emily R. Ruskowitz, Brizzia G. Munoz-Robles.  e-mail: profcole@uw.edu

prevents spontaneous reassembly but can be complexed into a functional species under specific conditions¹. Currently, controlled reassembly of split proteins is commonly achieved through genetic fusion of split fragments with inducible dimerizers, whereby exogenously triggered protein dimerization brings fragments into proximity to restore the function of the split parent species. Most frequently, these methods exploit small-molecule chemical ligands for inducible dimerization (for example, rapamycin-induced heterodimerization of FKBP/FRB² or coumermycin homodimerization of GyrB³). Although specification over when small-molecule inducers are added to culture affords temporal control over split protein activity, such chemical activation cannot be readily regulated in two-dimensional (2D) or 3D space⁴. To address this limitation, several optogenetic strategies have been developed that employ light-responsive proteins (for example, PhyB-PIF (ref. 5), Cry2/CIB1 (ref. 6) or magnets⁷) that bind under visible light irradiation⁸. While optogenetic approaches have already enabled a seemingly limitless collection of exciting new studies, they are not without limitations: (1) optogenetic proteins are often quite large (>500 amino acids per partner), limiting the expression efficiency and posing potential concern over reaction sterics; (2) previously developed systems are universally sensitive to visible light ($\lambda = 450\text{--}650\text{ nm}$), making them practically difficult to work with under standard laboratory lighting, and dramatically limiting their combined utility with common green-, orange- or red-type fluorophores; (3) reactions are non-covalent and quickly reverse under dark conditions (often with half-lives of seconds to minutes), necessitating continuous illumination for sustained protein activation that presents challenges with long-term culture or migrating systems; (4) the overall extent of the reaction is near-impossible to control, rendering intermediate activation states inaccessible; and (5) light-responsive proteins typically show minimal-to-no responsiveness to multiphoton activation, rendering full 3D spatial modulation largely out of reach.

Hypothesizing that many of the limitations to existing optogenetic approaches could be addressed through chemical advances, we sought to establish a genetically encoded protein–protein ligation reaction involving small partners that could be optically regulated in a rapid, irreversible, dose-dependent and highly specific manner. Towards this goal, we were inspired by the versatility of SpyLigation, in which genetically encoded SpyCatcher (SC; 113 amino acids; 12.1 kDa) and SpyTag (ST; 13 amino acids; 1.5 kDa) protein pairs undergo spontaneous covalent coupling with high yield in living cells (mammalian, bacterial and plant) that is maintained amid diverse conditions of pH, temperature and buffer⁹. Although recent efforts to shield ST with the light-responsive AsLOV2 protein have afforded some photocontrol over SpyLigation¹⁰, substantial dark reaction precludes the strategy's functional utilization in many contexts, including those within living cells and 3D materials. As SpyLigation involves isopeptide bond formation between a critical lysine on SC (Lys31) and an essential aspartic acid on ST, we postulated that complete optical control over SpyLigation could be obtained through molecular photocaging of either residue, chemically blocking the reactive side chains with a photoremovable moiety. Since genetic code expansion can be used to site specifically install non-canonical amino acids at user-defined locations within proteins during translation¹¹, and requisite transfer RNA (tRNA)/tRNA synthetase pairs have been previously evolved for efficient incorporation of photocaged lysine residues at the rarely utilized amber stop codon within bacterial and mammalian culture¹², we hypothesized that a photoactivatable SC (pSC) could be created through amber suppression in which the catalytic Lys31 residue was substituted with an *ortho*-nitrobenzyloxycarbonyl (*o*NB)-caged analogue (*N*^ε-(*o*-nitrobenzyloxycarbonyl)-L-lysine (Lys(*o*NB))). We expected that pSC's photocage would prevent isopeptide bond formation with ST but could be rapidly removed in response to cytocompatible near-ultraviolet light ($\lambda = 365\text{ nm}$)¹³ or near-infrared multiphoton stimulation to yield the functional SC in a dose-dependent manner and with four-dimensional

(4D) control. When used to photoligate split protein pairs, we anticipated that this light-activated SpyLigation (LASL) could be exploited to irreversibly assemble proteins and restore function in solution and biomaterials and intracellularly with spatiotemporal control, expanding current capabilities and opening new doors to probe and direct cellular fate (Fig. 1).

Results

Construction of a pSC for LASL

Success of the LASL strategy hinges on efficient incorporation of Lys(*o*NB) within pSC, necessitating a high-yielding synthesis of the photocaged lysine and effectively engineered orthogonal aminoacyl-tRNA synthetase/tRNA pairs for site-specific non-canonical amino acid installation in both bacterial and mammalian systems. Lys(*o*NB) was produced through an improved synthetic route, yielding sufficient gram quantities of the caged amino acid for several litres of protein expression from a single two-step synthesis (Supplementary Method 1). Upon in-solution light exposure and *o*NB photocage removal, native lysine was recovered following expected first-order photocleavage kinetics (Supplementary Fig. 1). For bacterial expression, we constructed a plasmid encoding two copies of a variant of *Methanosarcina mazei* pyrrollysyl-tRNA synthetase (MmPyIRS; Tyr306Met, Leu309Ala, Cys348Ala and Tyr349Phe amino acid substitutions)¹² previously evolved for Lys(*o*NB) incorporation and a single copy of its cognate Pyl-tRNA_{CUA} (Supplementary Method 2). To install Lys(*o*NB) site specifically in mammalian cells, we utilized a plasmid encoding a variant of *Methanosarcina bakeri* synthetase (MbPyIRS; Tyr271Ala and Tyr349Phe amino acid substitutions) compatible with derivatized lysine species¹⁴—but not previously with Lys(*o*NB)—and four copies of an engineered M15-tRNA_{CUA} recently reported to boost incorporation of non-canonical amino acids in pyrrollysine-based systems¹⁵ (Supplementary Method 2). Employing the relevant orthogonal aminoacyl-tRNA synthetase/tRNA plasmid for genetic code expansion in bacteria or mammalian cells, pSC was created through cotranslational incorporation of Lys(*o*NB) at SC's catalytic Lys31 residue via amber suppression (Supplementary Method 3).

Validation and characterization of LASL with purified proteins

To assess and quantify LASL's efficiency *in vitro*, we recombinantly expressed pSC, SC and a glutathione *S*-transferase–SpyTag fusion (GST-ST) in *Escherichia coli* and purified each species via affinity chromatography (Supplementary Methods 4 and 5). Highly pure samples with the expected molecular masses were obtained for all proteins, as indicated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) analysis and liquid chromatography mass spectrometry (Supplementary Fig. 2, Supplementary Method 6 and Supplementary Table 1). When exposed to near-ultraviolet light ($\lambda = 365\text{ nm}$; 10 mW cm⁻²; 30 min), pSC's *o*NB cage (delta mass = 179 Da) was photochemically removed to yield protein that precisely matched the mass of the native SC construct (13.2 kDa) (Fig. 2a). To assess whether this photoproduct was functional and could participate in SpyLigation, we reacted differentially light-exposed pSC ($\lambda = 365\text{ nm}$; 10 mW cm⁻²; 0–90 min) with excess GST-ST at physiological temperature (37 °C) for non-kinetically limited times (18 h) before SDS–PAGE analysis (Fig. 2b,c and Supplementary Method 7). The extent of photoligation was determined by quantifying the intensities of the disappearing bands from pSC and GST-ST, along with the appearing band corresponding to the SpyLigated product (Fig. 2d). These analyses revealed a first-order photocaging constant of $0.05 \pm 0.01\text{ min}^{-1}$ and a half-life of $15 \pm 2\text{ min}$, respectively, corresponding to $0.07 \pm 0.02\text{ cm}^2\text{ J}^{-1}$ and $9 \pm 1\text{ cm}^{-2}$ when accounting for the utilized light intensity. Uncaging photokinetics remained unchanged when pSC was exposed to light in the presence of GST-ST, and GST-ST bioactivity was unperturbed upon incubation and subsequent photoligation with pSC (Supplementary Figs. 3 and 4). Highly satisfied with the observed near-complete ligation

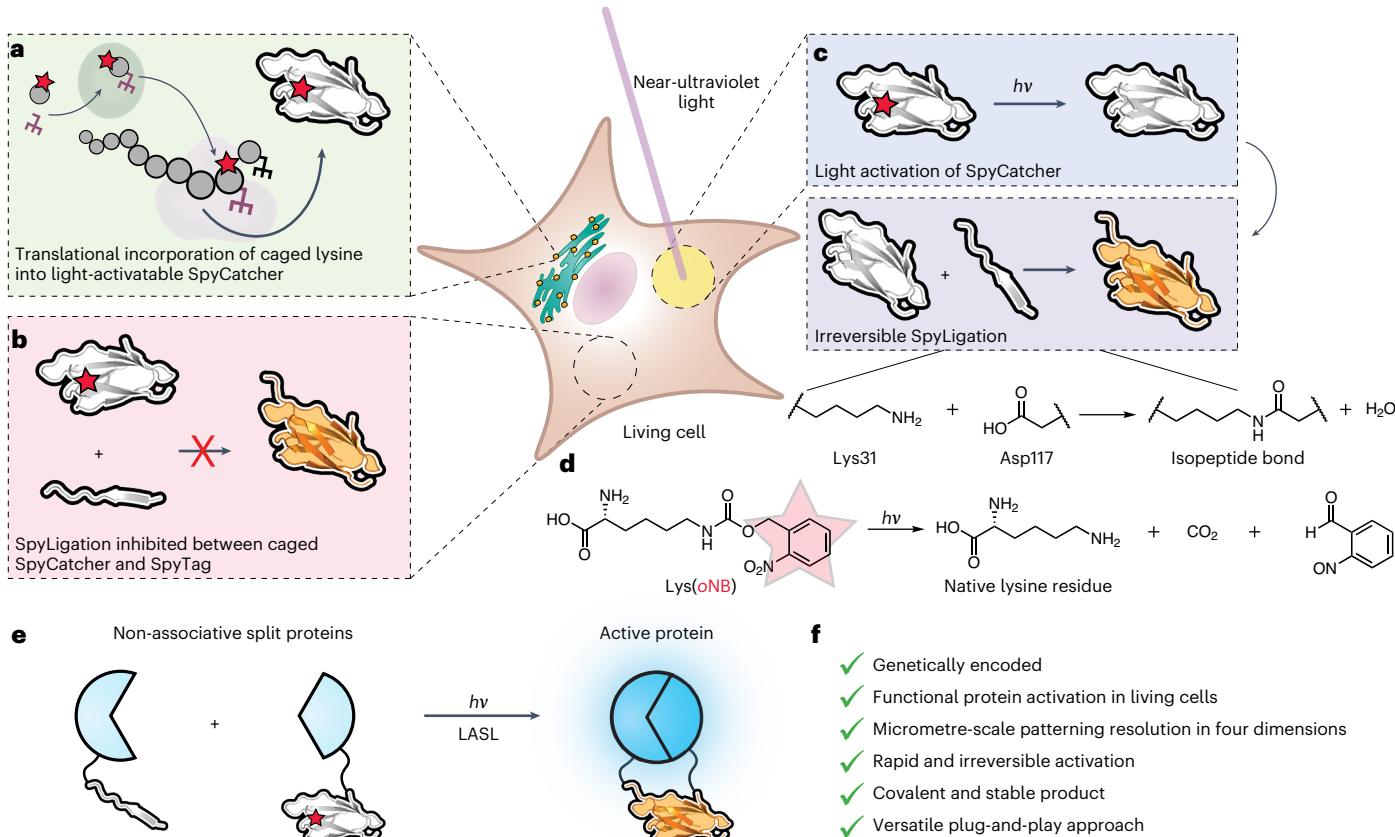


Fig. 1 | LASL affords complete spatiotemporal control over protein activation within living systems. **a**, A photocaged lysine is site specifically incorporated within the active site of SC during protein translation via an unnatural tRNA/tRNA synthetase pair, giving pSC. **b**, Owing to the bulky photocage masking its reactive amine, pSC remains inactive and unable to interact with or covalently bind ST. **c**, With user-directed light exposure, the critical lysine is liberated to generate newly uncaged SC, which is capable of spontaneous isopeptide bond formation

with ST. **d**, Photoactivation is imparted through an *o*NB moiety installed on the ϵ -amine of lysine, such that light exposure restores the native residue. **e**, When pSC and ST are genetically fused to otherwise non-associative split proteins, irreversible protein activation is photochemically regulated. **f**, LASL of the SC/ST pairs to irreversibly activate split proteins offers many distinct advantages over existing optogenetic strategies and can be used to interrogate a variety of biological functions in four dimensions.

in optically stimulated samples and an absence of dark reaction for unexposed pSC, we were encouraged to test LASL's efficiency in more complex environments. When performed in *E. coli* or human embryonic kidney 293T (HEK-293T) mammalian cell lysate, LASL yielded the expected ligated product band with seemingly no unintended reactions with other protein components (Fig. 2e, Supplementary Methods 8 and 9 and Supplementary Fig. 5). Collectively, these experiments establish that LASL is photocontrollable with high specificity and in a dosage-dependent manner.

Spatial control over LASL within hydrogel biomaterials

Photochemical reactions are unique in that they can be spatiotemporally initiated based on when and where light is directed onto reactants. This feature is now regularly exploited in many subfields, including by the biomaterials community to engineer cell culture platforms with user-defined and heterogeneous biochemistry¹⁶. To guide complex anisotropic cellular functions *in vitro*, our laboratory and others have utilized photochemistry to pattern full-length protein immobilization within polymeric hydrogels whose stiffness, water content and other essential features mimic those of native tissue^{17–23}. Although recent efforts have shown the importance of site-specific protein modification in maintaining their bioactivity upon tethering^{21,24,25}, genetically encoded photochemistries to control such immobilization have not yet been established. Towards filling this gap, we opted to employ LASL in the photopatterning of poly(ethylene glycol) (PEG)-based hydrogels

formed through bioorthogonal strain-promoted azide–alkyne cycloaddition (SPAAC)^{26–29}.

To immobilize the photoactivatable SpyCatcher uniformly within SPAAC gels, a pSC variant containing a carboxy (C)-terminal sortase recognition motif (that is, LPETG) was expressed, purified and chemoenzymatically modified³⁰ with an azido-polyglycine peptide probe (H-GGGGDDK(N₃)-NH₂) to yield the azide-monotagged pSC (pSC-N₃) (Supplementary Method 10). Photocaged SpyCatcher-decorated gels were formed through step-growth polymerization of PEG tetracyclononyne (PEG-tetraBCN; number-averaged molecular mass (M_n) = ~20 kDa), linear PEG-diazide (M_n = ~3.5 kDa) and pSC-N₃ (Fig. 3a and Supplementary Method 11). Upon mild near-ultraviolet irradiation, the *o*NB cage is cleaved, converting pSC into its active form and permitting localized conjugation with gel-swollen SpyTagged proteins via LASL. Diffusive removal of unbound proteins yields patterned gel substrates defined by user-selected light exposure locations and parameters. Gel modification was performed with a SpyTagged mRuby (mRuby-ST; Supplementary Method 4), whereby red fluorescence of the immobilized protein permitted visualization and quantification. Traditional photolithographic techniques were utilized to control the patterning of mask-defined shapes throughout gels with single micrometre-scale resolution; no background fouling or undesired conjugation was observed, highlighting minimal association of ST with pSC (Fig. 3b). Linear exposure gradients imposed across the gel surface using an opaque photomask moving at variable rates (0.13,

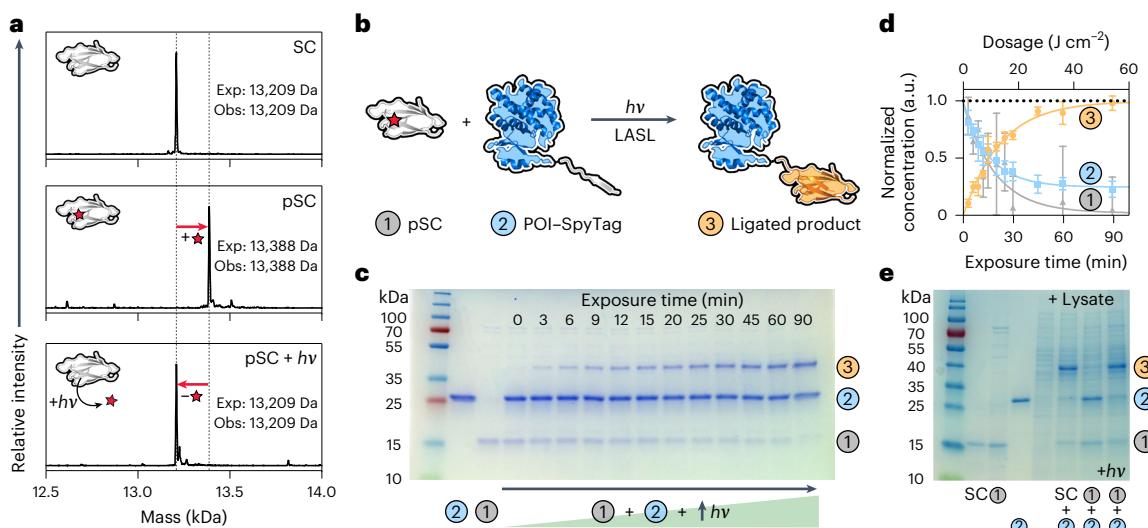


Fig. 2 | pSC provides user control of SpyLigation in solution. **a**, Intact protein mass spectrometry of SC, pSC and light-treated pSC (+hv; $\lambda = 365$ nm; 20 mW cm^{-2} ; 30 min). The dashed lines indicate the expected masses of SC and pSC containing a single Lys(oNB) (red stars). Exp, expected; Obs, observed. **b**, Covalent linkage of pSC (1; grey) and a protein of interest (POI)–SpyTag fusion (GST-ST; 2; blue) is inhibited by the presence of Lys(oNB) at a critical lysine residue. pSC photoactivation allows for formation of the covalently ligated product (3; orange). **c**, The extent of LASL between pSC and GST-ST ($10 \mu\text{M}$; 37°C) varies with light exposure ($\lambda = 365$ nm; 10 mW cm^{-2} ; 0–90 min),

as visualized by SDS–PAGE following 18 h of post-light incubation. **d**, pSC photoactivation and ligated product formation exhibit a light dose dependency, as determined by SDS–PAGE band intensity quantification for each species as labelled in **b**. Light dosage was calculated as the product of the light intensity and exposure time. The error bars correspond to s.d. about the mean for four experimental replicates. **e**, Reactivity of pSC and GST-ST ($10 \mu\text{M}$; 37°C ; 0.5 h) with and without light treatment (20 mW cm^{-2} ; $\lambda = 365$ nm; 0.75 h) in *E. coli* lysate analysed by SDS–PAGE. Similar results were achieved independently in three experimental replicates.

0.20 or 0.40 mm min^{-1}) yielded continuous exponential protein gradients of predicted shape (Fig. 3c,d). When accounting for gel position and mask translational speed, gradients collapsed onto a single dosage–response curve that revealed a pSC uncaging kinetic constant of $0.06 \pm 0.02 \text{ cm}^2 \text{ J}^{-1}$ and a half-life of $11 \pm 2 \text{ J cm}^{-2}$, both of which were statistically indistinguishable from values determined in solution (Fig. 3e). These experiments demonstrate a genetically encoded photochemistry to immobilize proteins site specifically within materials, as well as highlight LASL’s compatibility with other state-of-the-art protein modification chemistries and its ability to be performed with high spatiotemporal control.

Spatial control over LASL within living mammalian cells

Building on our success in spatially controlling biomacromolecular tethering within hydrogel biomaterials, we turned our efforts towards optically specifying protein subcellular location within living mammalian cells using LASL. Here we cloned a polycistronic construct consisting of a pSC-tagged mCherry³¹ (mCh), a self-cleaving P2A peptide sequence³² and an enhanced green fluorescent protein (eGFP)³³ fused with ST and a plasma membrane-localizing CAAX motif from K-Ras³⁴ (Supplementary Method 12 and Fig. 3f). HEK-293T cells were co-transfected with plasmids encoding the membrane-labelling components and the variant MbPyIRS/tRNA pair (Supplementary Method 13). When cultured in media supplemented with Lys(oNB), cells fluoresced red (mCh) throughout the cytosol and green (eGFP) at the plasma membranes, indicating successful read through of pSC’s in-frame amber stop codon and eGFP-ST membrane targeting with the CAAX motif (Fig. 3g). Following flood illumination ($\lambda = 365$ nm; 20 mW cm^{-2} ; 0–20 min), the uncaged pSC-mCh colocalized with eGFP-ST at the plasma membrane via LASL. As predicted, membrane labelling with and subcellular distribution of pSC-mCh scaled in a statistically significant manner with light dosage (Fig. 3h,i). Complementary experiments in which a membrane-bound pSC was phototagged with eGFP-ST yielded similar results (Supplementary Methods 12 and 13 and Supplementary Fig. 6). Together, these studies showcase LASL’s

ability to efficiently specify lasting changes to intracellular protein localization in a dose-dependent manner.

Split protein assembly and functional activation using LASL

Motivated by our unique ability to photoligate genetically encoded protein pairs with spatiotemporal control using LASL, we next sought to utilize the reaction to irreversibly restore protein function through covalent ligation of otherwise inactive split proteins. As an initial proof of concept, we selected UnaG—a green fluorescent protein derived from Japanese eel muscle whose activity can be optically assessed rapidly at single-cell and subcellular resolutions³⁵ (Fig. 4a). Since its fluorogenic chromophore, bilirubin, is non-covalently bound and readily available in sera and in vivo, UnaG exhibits virtually no long-term photobleaching in living systems³⁶, having an optical readout for protein function that would not be affected by LASL’s requisite light exposure led us to explore UnaG over alternative fluorescent proteins. Recognizing that split protein functional activation is dependent on factors that are often difficult to predict a priori (for example, geometry, sterics and intermediate folding), we exploited a previously validated split site for rapamycin-induced UnaG assembly (nUnaG = residues 1–84; cUnaG = residues 85–139)³⁷ and cloned bacterial expression vectors for all possible permutations wherein fragments were amino (N)- or C-terminally monofunctionalized with either ST or wild-type SC (Fig. 4b and Supplementary Method 14). 6xHis-tagged proteins were expressed in *E. coli* and purified by immobilized metal-ion affinity chromatography. Variants containing N-terminal ST moieties did not readily express in a soluble form and were thus abandoned; all other species were successfully obtained in good yield, validated as pure (SDS–PAGE; Supplementary Fig. 7) and determined to exhibit the expected molecular weight (liquid chromatography mass spectrometry; Supplementary Table 1). Although individual components lacked fluorescence, SpyLigated UnaG fragments brightly fluoresced after 30 min of reaction at levels approaching those from samples following extended incubation (24 h) (Fig. 4c and Supplementary Method 15), consistent with the rapid kinetics previously reported for the native SpyLigation. Conjugation of nUnaG-ST and SC-cUnaG gave

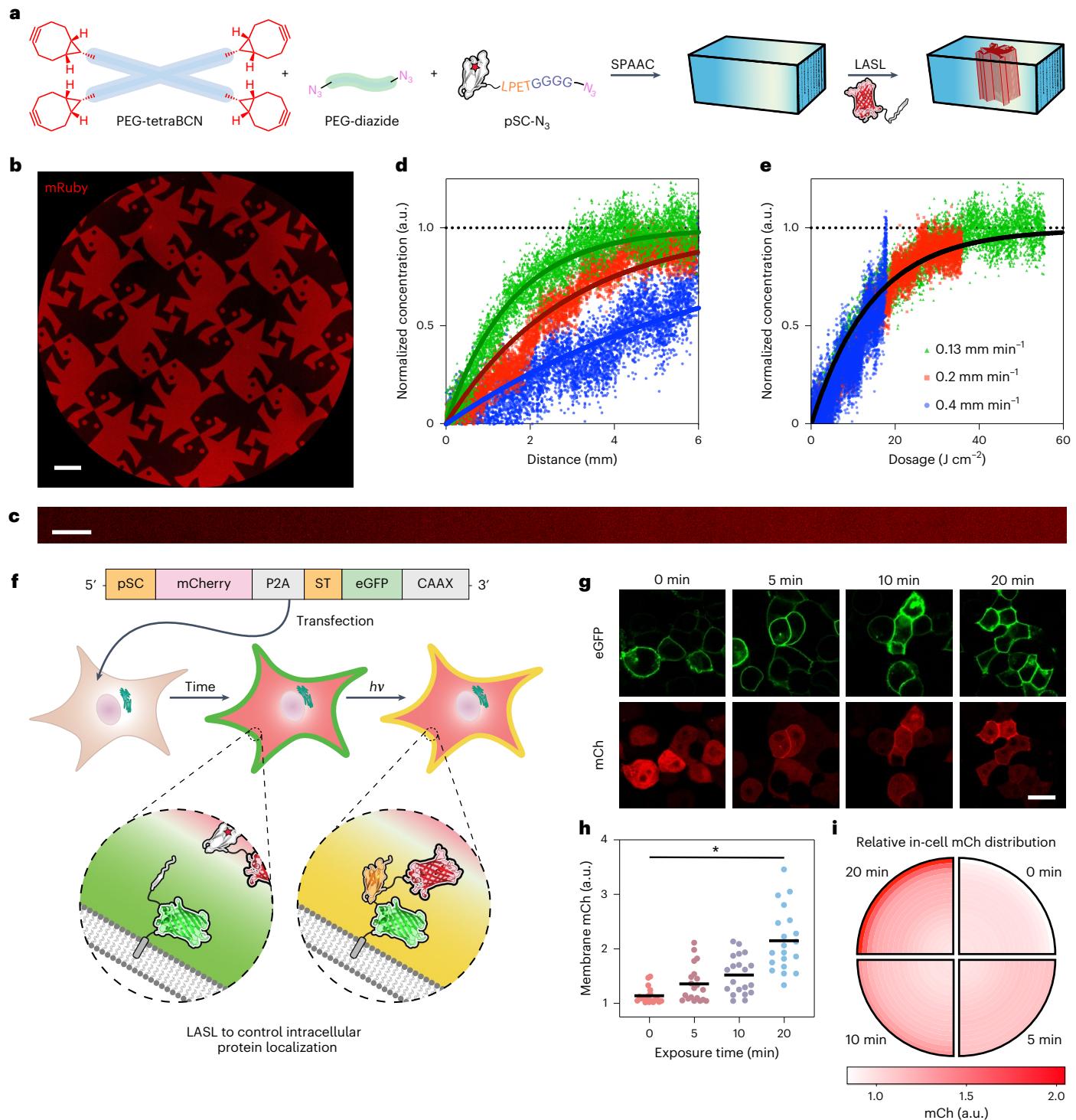


Fig. 3 | LASL enables site-specific patterned protein localization in 3D biomaterials and living cells. **a**, Step-growth polymerization of PEG-tetraBCN and PEG-diazide in the presence of pSC-N₃ form a SPAAC-based hydrogel uniformly decorated with the photocaged SpyCatcher. Photoactivation permits spatiotemporally defined immobilization of POI–SpyTag constructs via LASL. **b**, Mask-based photolithographic exposure ($\lambda = 365$ nm; 20 mW cm⁻²; 15 min) generated discrete patterns of immobilized mRuby-ST throughout the gel thickness. **c**, Representative fluorescent micrograph of immobilized mRuby gradient. **d**, Relative quantification of immobilized mRuby concentrations for varying sample coverage rates of 0.13 (green), 0.20 (red) and 0.40 mm min⁻¹ (blue). The solid lines depict the predicted concentrations based on

pSC photouncaging kinetics. **e**, Accounting for gel position and mask translational speed, mRuby immobilization gradients collapse onto a single dosage-responsive curve. **f–i**, Optogenetic specification of protein membrane tethering in mammalian cells. **f**, Schematic of the gene cassette used to prime cells for LASL-mediated plasma membrane labelling, where a CAAAX-anchored eGFP-ST was covalently ligated with cytosolic pSC-mCh upon light exposure. **g**, Representative fluorescence micrographs of transfected HEK-293T cells imaged 6 h following treatment with varied light exposures ($\lambda = 365$ nm; 20 mW cm⁻²; 0 – 20 min). **h**, Membrane labelling with mCh scaled in a statistically significant manner with light exposure. The relative fluorescence levels are visualized in a violin scatter plot. The asterisk denotes conditions with statistically significant differences in signal ($P < 0.0001$, two-tailed unpaired *t*-test). **i**, The intracellular mCh distribution transitioned through LASL from uniformly cytosolic to more membrane localized with increased light treatment duration. Scale bars, 250 μ m (**b** and **c**) and 20 μ m (**g**).

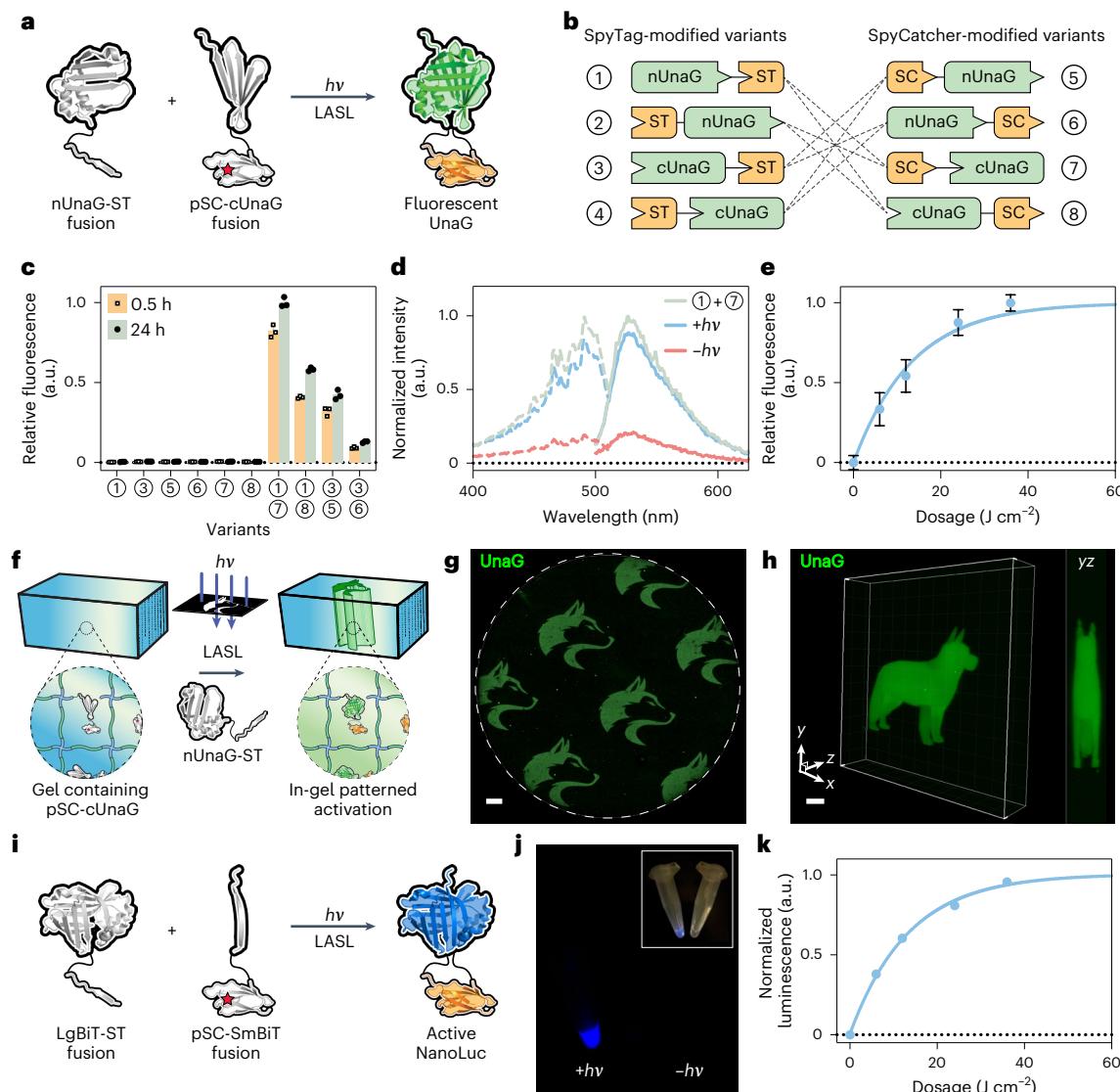


Fig. 4 | Assembly of UnaG and NanoLuc through LASL of split protein fragments in solution and biomaterials. **a**, UnaG is split into N- (nUnaG) and C-terminal (cUnaG) fragments genetically fused to ST and pSC. Fragments remain inactive until photoactivation of pSC and LASL-mediated functional assembly of split fragments to restore UnaG fluorescence. **b**, All possible fusion variants of ST/SC and UnaG fragments were cloned and recombinantly expressed to screen for maximum assembly fluorescence. **c** Individual and combined variant ($10 \mu\text{M}$) fluorescence after reaction for 0.5 (orange) and 24 h (grey). **d**, Excitation (dashed lines) and emission spectra (solid lines) of pSC-cUnaG kept in the dark ($-h\nu$; red) or exposed to light ($+h\nu$; blue; $\lambda = 365 \text{ nm}$; 20 mW cm^{-2} ; 20 min) and reacted with nUnaG-ST ((1)). Photoactivated pSC-cUnaG exhibited spectra similar to its wild-type counterpart (SC-cUnaG; (7)) when ligated to nUnaG-ST. **e**, UnaG reconstitution and the accompanying fluorescence from nUnaG-ST and

light-treated pSC-cUnaG exhibited dose dependency. **f**, UnaG can be spatiotemporally reassembled within hydrogel biomaterials functionalized with pSC-cUnaG and patterned with nUnaG-ST via LASL. **g**, Mask-based photolithographic exposure generated discrete patterns of active UnaG throughout the gel thickness. **h**, Multiphoton laser-scanning lithography affords patterned protein activation with full 3D control. **i–k**, Split NanoLuc is reassembled via LASL in a light dose-dependent manner. **i**, Schematic of reassembly. **j**, NanoLuc assembly and luminescence only observed upon light exposure. **k**, NanoLuc bioactivity is restored in a dose-dependent manner, mirroring the kinetic results for UnaG. The data represent means $\pm 1 \text{ s.d.}$ normalized to the experimental minimum/maximum ($n = 3$ experimental replicates). The error bars in **k** are substantially smaller than the symbols indicating mean luminescence. Scale bars, $500 \mu\text{m}$ (**g**) and $50 \mu\text{m}$ (**h**).

a fluorescent species that was between two- and fivefold brighter than all other fragment combinations—a result consistent with structural protein evaluation, as ST's N terminus is colocalized with the C terminus of SC³⁸. SpyLigated nUnaG-ST and SC-cUnaG exhibited ~15% maximal fluorescence compared with the wild-type UnaG species (Supplementary Method 16 and Supplementary Figs. 8 and 9); although reconstituted bioactivity depends just as much on the protein identity and split site as it does on the dimerizing chemistry, this finding is consistent with the best reported inducible-dimerizing split protein systems^{39–41}.

After identifying nUnaG-ST and SC-cUnaG as the brightest split protein combination, we cloned, expressed and purified

the photocaged SC variant (pSC-cUnaG) for LASL assembly (Supplementary Method 14). To assess whether LASL could be used to photochemically restore UnaG activation, stoichiometrically matched nUnaG-ST (0 – $10 \mu\text{M}$) and light-exposed pSC-cUnaG ($\lambda = 365 \text{ nm}$; 20 mW cm^{-2} ; 0 – 30 min) were combined (Supplementary Method 17). Sample fluorescence was measured immediately following light exposure and again after 24 h. Fluorescence excitation/emission spectra for the LASL product following the longest exposures (30 min) matched those of the spontaneous SpyLigation (that is, nUnaG-ST with SC-cUnaG) and wild-type UnaG³⁵ (Fig. 4d and Supplementary Method 18). LASL-restored UnaG fluorescence scaled linearly with

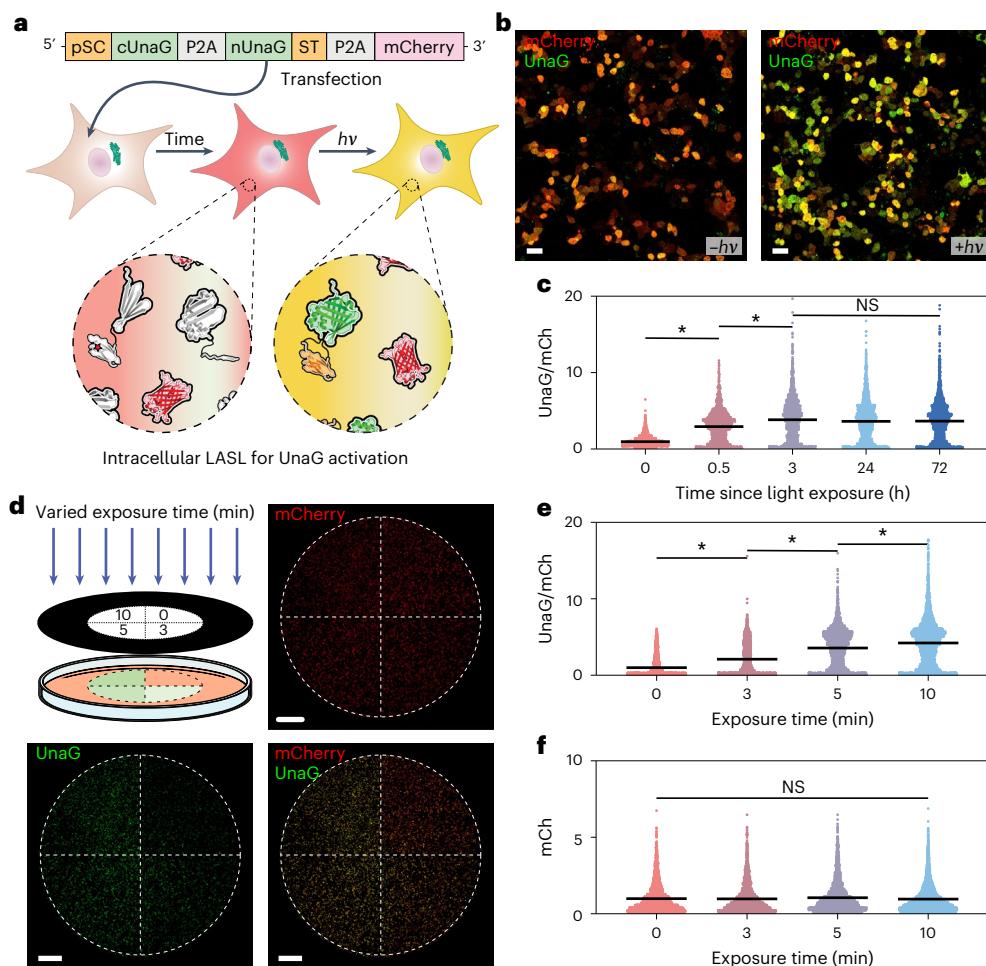


Fig. 5 | Photoactivation of split UnaG with spatiotemporal precision in living cells through intracellular LASL. **a**, Schematic of the pSC-UnaG gene cassette used to prime cells for LASL of non-associative UnaG fragments. **b**, Representative fluorescence images of transfected HEK-293T cells (red) with (+hv) and without (−hv) light illustrate UnaG photoactivation (green) by LASL. **c**, Time lapse of intracellular UnaG photoactivation after light treatment (+hv), normalized to initial UnaG/mCh ratios. **d–f**, Mask-based photolithography with spatially varied light exposure (0–10 min) yielded dose-dependent UnaG activation throughout mammalian culture. **d**, Fluorescence images of a culture dish, with dashed lines highlighting

the exposure pattern. **e**, Individual cell UnaG/mCh signal, quantified for each exposure subregion and normalized to the average unexposed UnaG/mCh ratio. **f**, Individual cell mCh signal, quantified for each exposure subregion. In **b–f**, the light treatment was 20 mW cm^{-2} and $\lambda = 365 \text{ nm}$. The asterisks denote conditions with statistically significant differences in signal ($P < 0.0001$, two-tailed unpaired *t*-tests). NS denotes conditions with no statistically significant difference in signal ($P = 0.09$ in **c** and $P = 0.11$ in **f**; one-way analysis of variance). Similar results were achieved independently in three experimental replicates. Scale bars, $20 \mu\text{m}$ (**b**) and 1 mm (**d**).

protein concentration and persisted well after initial light exposure (Supplementary Fig. 10). Critically, only small amounts of background fluorescence (~15% maximum) were observed in unexposed samples, which we attributed to non-covalent association of the UnaG fragments partially stabilized with bilirubin (Supplementary Methods 19 and 20 and Supplementary Fig. 11). UnaG assembly/activation followed LASL's predicted dosage dependency (Fig. 4e); here we observed a pSC uncaging kinetic constant of $0.07 \pm 0.02 \text{ cm}^2 \text{ J}^{-1}$ and a half-life of $10 \pm 2 \text{ J cm}^{-2}$ that was statistically indistinguishable from values determined in solution and biomaterials.

Building on our ability to photoassemble UnaG in solution, we extended efforts towards spatiotemporally controlling its activation in four dimensions. Although studies reported earlier (Fig. 3) and elsewhere have demonstrated photochemical immobilization of full-length proteins within hydrogels, patterned functional assembly of split proteins in biomaterials has not been reported. The generation of active protein only at desired locations within gels offers distinct advantages over strategies involving diffusive transport of active proteins before patterned tethering, particularly if these proteins are to

be used to guide embedded cell fate; the latter technique floods cells with active protein throughout the patterning process, whereas the former does not. Towards this goal, we appended pSC-cUnaG with a C-terminal sortase recognition motif; this protein (pSC-cUnaG-LPETG) was expressed, purified and sortagged with H-GGGGDDK(N₃)NH₂ to yield monotagged pSC-cUnaG-N₃ (Supplementary Method 21). SPAAC-based gels modified uniformly with pSC-cUnaG-N₃ were formed through reaction of PEG-tetraBCN and a linear PEG-diazide (Fig. 4f and Supplementary Method 22). When incubated with nUnaG-ST and exposed to photomasked light, UnaG activation was confined to 2D mask-defined shapes extending throughout the gel thickness (Fig. 4g). Multiphoton laser-scanning lithography, whereby programmed laser raster scanning within the gel specified photoactivation with full 3D control, afforded excellent 3D patterning at user-specified regions within the gel (Fig. 4h and Supplementary Fig. 12).

Having shown that LASL can be used to irreversibly assemble and activate UnaG, we sought to highlight the versatility of these methods through extension to another functional protein. We identified NanoLuc⁴² as a bioluminescent enzyme that has found great utility in

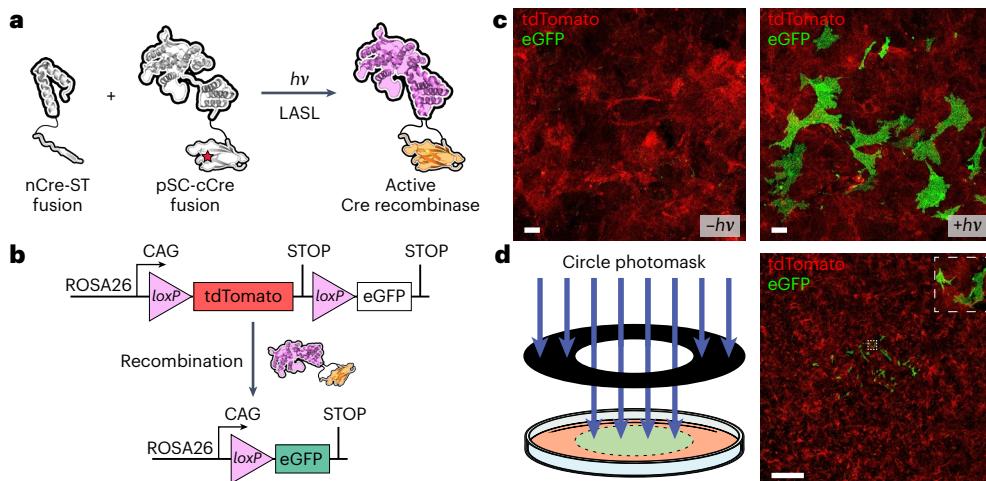


Fig. 6 | Spatially controlled photoactivation of primary cell genome editing via LASL. **a**, Cre recombinase split into inactive N- (nCre) and C-terminal (cCre) fragments and respectively genetically fused to ST and pSC can be functionally reassembled using LASL. **b**, Transgenic mouse dermal fibroblasts contain a dual-colour reporter for Cre activity; site-specific recombination of DNA between *loxP* sites results in tdTomato gene excision and expression of a downstream eGFP. **c**, Cells transfected with the pSC-Cre gene cassette and stimulated with

light ($\lambda = 365$ nm; 20 mW cm^{-2} ; 3 min) exhibit an irreversible red-to-green fluorescence switch associated with Cre-mediated genome editing. **d**, Left, using a circle photomask (2 mm diameter opening), recombination can be spatially regulated via photolithographically controlled LASL. Right, fluorescence image of patterned cells. Inset, magnification of the exposed region highlighted by a dashed box. Similar results were achieved independently in three experimental replicates. Scale bars, 50 μm (c) and 1 mm (d).

its split form for quantifying protein–protein interactions⁴³. Adopting the optimal protein arrangement for LASL, identified structurally and with UnaG (that is, nPOI-ST + pSC-cPOI, where nPOI and cPOI refer to the N- and C-terminal fragments of each protein of interest), we fused NanoLuc’s N-terminal fragment (LgBiT; 159 amino acids; 17.6 kDa) with ST and its C-terminal fragment (SmBiT; 11 amino acids; 1.6 kDa) to pSC. LgBiT-ST and pSC-SmBiT were expressed in *E. coli* and purified by immobilized metal-ion affinity chromatography (Fig. 4i, Supplementary Method 23 and Supplementary Fig. 13). When NanoLuc fragments were combined with equal stoichiometry (0–1 μM) in the presence of their substrate (furimazine), almost no luminescence (<5% maximum) was observed; this absence of dark activity is consistent with minimal reported association of LgBiT and SmBiT, as well as our findings that pSC and ST do not associate appreciably. Via LASL, protein bioactivity was restored in a light dose- and concentration-dependent manner following ultraviolet treatment ($\lambda = 365$ nm; 20 mW cm^{-2} ; 0–30 min) (Fig. 4j,k, Supplementary Method 24 and Supplementary Fig. 14) with uncaging kinetics (kinetic constant of $0.07 \pm 0.02 \text{ cm}^2 \text{ J}^{-1}$; half-life of $9 \pm 2 \text{ J cm}^{-2}$) matching those for mRuby immobilization, UnaG activation and in-solution Lys(oNB) photolysis. Taken together, these results establish our ability to precisely and irreversibly photoactivate functional proteins through LASL-mediated split fragment reconstitution with micrometre-scale 3D control in a plug-and-play manner.

LASL-mediated split protein activation in mammalian cells

Having demonstrated that LASL could be used to functionally assemble protein fragments with predictable dose dependence and well-defined kinetics *in vitro* using purified bacterial proteins, we pursued extension of the approach to irreversibly activate protein function spatiotemporally within living mammalian cells. Towards this goal, we cloned a polycistronic split UnaG LASL construct consisting of three components—pSC-cUnaG, nUnaG-ST and mCh—each separated by P2A (Supplementary Method 25); this design was selected to ensure similar expression levels of each UnaG fragment and to provide an internal standard (that is, red fluorescence of mCh) to account for transfectional variations on a cell-by-cell basis (Fig. 5a). HEK-293T cells were co-transfected with plasmids encoding for the variant MbPyLRS/tRNA pair and split UnaG components (Supplementary Method 26). When grown in media supplemented with Lys(oNB), virtually all

cells fluoresced red 24 h post-transfection, indicating high transfection efficiencies and amber codon read through (Fig. 5b). As predicted, cells subjected to flood illumination ($\lambda = 365$ nm; 20 mW cm^{-2} ; 20 min) gained green fluorescence, reflecting successful LASL-mediated functional assembly of UnaG intracellularly, while unexposed cells did not (Fig. 5b). UnaG/mCh fluorescence ratios from individual cells were determined through automated image analysis and quantified over time; statistically significant increases ($P < 0.0001$, unpaired *t*-test) for UnaG/mCh were observed for cells within 30 min of light exposure ($\lambda = 365$ nm; 20 mW cm^{-2} ; 7 min), with peak levels detected by 3 h, while UnaG/mCh did not increase over 72 h for unexposed cells (Fig. 5c and Supplementary Fig. 15). Although the individual fluorophore signal decreased slightly over time (attributed to expected protein degradative clearance and/or dilution accompanying cell division), maximal UnaG/mCh fluorescence ratios (around fivefold higher than unexposed samples) persisted in photostimulated cells for at least 72 h in a manner consistent with covalent UnaG assembly.

To further demonstrate LASL’s ability to spatiotemporally and optically regulate specific biomacromolecular function in living systems, we focused our efforts towards patterning covalent protein assembly and concomitant activation within mammalian cell culture. Co-transfected HEK-293T cells were lithographically subjected to collimated light ($\lambda = 365$ nm; 20 mW cm^{-2} ; 20 min) through an open-circle photomask (diameter = 7 mm); cells in the light-exposed region exhibited both green and red fluorescence, whereas unexposed cells appeared primarily red (Extended Data Fig. 1a). Quantification of individual cell UnaG/mCh fluorescence ratios as a function of radial distance from the mask’s centre revealed a clear activation pattern with micrometre-scale resolution (Extended Data Fig. 1b). Aggregated analysis for all cells (~35,000) demonstrated a roughly fivefold and statistically significant increase for UnaG/mCh in light-exposed regions relative to those in the unexposed outside ring ($P < 0.0001$, unpaired *t*-tests) with no significant mCh photobleaching (Extended Data Fig. 1c and Supplementary Fig. 16). Exposing different regions of culture to light for variable amounts of time (0, 3, 5 and 10 min), we observed the expected dose-dependent activation, as indicated by significantly increasing UnaG/mCh fluorescence ratios ($P < 0.0001$, unpaired *t*-tests) accompanying lengthened exposures (Fig. 5d–f).

Following the successful generation of UnaG fluorescence intracellularly via LASL, we expanded our efforts to a protein target whereby its activation would yield lasting cellular functional changes, specifically in the form of irreversible genome editing. Towards this aim, we selected Cre recombinase, a topoisomerase that can recognize and site specifically excise DNA between two *loxP*sites, which is commonly used for gene knock-in/out studies *in vivo*. Exploiting a previously validated split site for rapamycin-induced Cre assembly (nCre = residues 19–59; cCre = residues 60–343)⁴⁴, we created a polycistronic split-Cre LASL construct consisting of two components—nCre-ST and pSC-cCre—each separated by P2A and individually fused to nuclear localization sequences (Fig. 6a and Supplementary Method 27). This cassette was transfected into transgenic mouse dermal fibroblasts bearing a Cre-dependent dual-colour reporter in the safe harbour Rosa26 locus (Supplementary Method 28); cells constitutively express red fluorescent protein tdTomato³¹ but switch expression to eGFP upon Cre-mediated recombination (Fig. 6b)⁴⁵. As designed, isolated primary cells fluoresced red but not green under normal culture conditions; upon light exposure ($\lambda = 365$ nm; 20 mW cm⁻²; 3 min), a substantial percentage of cells stopped fluorescing red and turned green, reflecting success in optically restoring Cre activity via LASL and downstream genome editing (Fig. 6c). Directed illumination through a circle photomask (2 mm diameter opening) permitted spatial patterning of recombination and photoconditional gene regulation (Fig. 6d). Together, these experiments establish LASL’s unique ability to irreversibly activate intracellular protein function with spatiotemporal control and in a dose-dependent manner.

Discussion

In this manuscript, we have introduced LASL, a genetically encoded photoclick chemistry for protein–protein conjugation, and demonstrated its distinct utility in solution and biomaterials and within living mammalian cells. Infusing concepts from synthetic organic chemistry into the world of protein engineering, we demonstrate that reagent protection, deprotection and ligation schemes can be used for site-specific protein heterocoupling. Taking advantage of a translationally incorporated photocaged amino acid that undergoes rapid photolysis in response to cytocompatible light, we attain the unique ability to govern biomolecule conjugation with precise spatiotemporal control in a dose-dependent manner. Photoligation of biologically inactive fragments enables direct gain-of-function intracellularly and irreversibly, with 4D control.

Although our initial efforts have focused on photoregulation of SC/ST ligation, we anticipate that the introduced chemical strategy of caging reactive partners through genetic code expansion can be readily extended to trigger covalent assembly using other protein chemistries. In cases where even more rapid coupling is desired, evolved variants with dramatically enhanced transamination kinetics (for example, Spy-Catcher003 (ref. 46) and DogCatcher⁴⁷) could be similarly caged. In systems where a more minimal ligation scar is required, strategies based on three components (for example, SpyStapler⁴⁸, SpyLigase⁴⁹, Snoop-Ligase⁵⁰, transglutaminase factor XIII⁵¹ or sortase³⁰) could be readily envisioned. Complementarily, alternatively caged amino acids could yield induced protein assembly in a wavelength-dependent manner or with stimuli beyond light (for example, pH, enzymes, temperature or small molecules). Although its reliance on genetic code expansion may place some limits on its utilization *in vivo*, the introduced strategy outperforms those based solely on canonical amino acids (for example, BLISS¹⁰ and split inteins⁵²) in that it can be tightly controlled with high specificity in living systems and no leaky background reaction.

We have demonstrated that phototriggered covalent assembly of split protein pairs permits their rapid and sustained functional turn-on within complex biological settings. While we initially targeted photoactivation of UnaG, NanoLuc and Cre, the generalizable mechanism defining LASL-mediated reconstitution outlines a clear path towards

photoregulation of many other bioactive species. Like other protein activation schemes based on inducible split protein dimerization, LASL’s success hinges on the identification of fragments that exhibit minimal background association and conditionally associate into highly bioactive species. Nevertheless, countless proteins have been functionally split and conditionally activated through non-covalent interactions, providing an accessible working blueprint for irreversible activation of many different species via LASL. In cases where protein splits have not been identified, advanced computational tools for protein structure prediction (for example, AlphaFold⁵³ and RoseTTAFold⁵⁴) may inform effective split sites and fusion protein pair arrangements.

As established here, LASL represents a uniquely powerful and tunable tool to optically couple small protein pairs rapidly and irreversibly in a user-defined and dose-dependent manner. Since ligation is photochemically regulated with near-ultraviolet or pulsed near-infrared light, product formation can be spatiotemporally driven with micrometre- and minute-scale precision and using wavelengths compatible with cells and conventional fluorophores. As both reactive components are genetically encoded, LASL can be performed with high fidelity in living systems. When used to covalently link split protein pairs, LASL affords 4D control over their functional assembly and bioactivation through a versatile plug-and-play approach. Looking forward, we expect that these chemical strategies will find great utility in the quest to probe and direct biological systems, whereby precise specification over protein function can be used to interrogate or stimulate protein signalling pathways, as well as for applications in tissue engineering and regenerative medicine.

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/s41557-023-01152-x>.

References

1. Shekhawat, S. S. & Ghosh, I. Split-protein systems: beyond binary protein–protein interactions. *Curr. Opin. Chem. Biol.* **15**, 789–797 (2011).
2. Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. Controlling signal transduction with synthetic ligands. *Science* **262**, 1019–1024 (1993).
3. Farrar, M. A., Alberola-Illa, J. & Perlmutter, R. M. Activation of the Raf-1 kinase cascade by coumermycin-induced dimerization. *Nature* **383**, 178–181 (1996).
4. Fegan, A., White, B., Carlson, J. C. T. & Wagner, C. R. Chemically controlled protein assembly: techniques and applications. *Chem. Rev.* **110**, 3315–3336 (2010).
5. Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**, 997–1001 (2009).
6. Kennedy, M. J. et al. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* **7**, 973–975 (2010).
7. Kawano, F., Suzuki, H., Furuya, A. & Sato, M. Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. *Nat. Commun.* **6**, 6256 (2015).
8. Spiltoir, J. I. & Tucker, C. L. Photodimerization systems for regulating protein–protein interactions with light. *Curr. Opin. Struct. Biol.* **57**, 1–8 (2019).
9. Zakeri, B. et al. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl Acad. Sci. USA* **109**, E690–E697 (2012).
10. Hartzell, E. J., Terr, J. & Chen, W. Engineering a blue light inducible SpyTag system (BLISS). *J. Am. Chem. Soc.* **143**, 8572–8577 (2021).

11. Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).
12. Chen, P. R. et al. A facile system for encoding unnatural amino acids in mammalian cells. *Angew. Chem. Int. Ed.* **48**, 4052–4055 (2009).
13. Ruskowitz, E. R. & DeForest, C. A. Proteome-wide analysis of cellular response to ultraviolet light for biomaterial synthesis and modification. *ACS Biomater. Sci. Eng.* **5**, 2111–2116 (2019).
14. Yanagisawa, T. et al. Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode *N*^ε-(o-azidobenzylloxycarbonyl) lysine for site-specific protein modification. *Chem. Biol.* **15**, 1187–1197 (2008).
15. Serfling, R. et al. Designer tRNAs for efficient incorporation of non-canonical amino acids by the pyrrolysine system in mammalian cells. *Nucleic Acids Res.* **46**, 1–10 (2018).
16. Ruskowitz, E. R. & DeForest, C. A. Photoresponsive biomaterials for targeted drug delivery and 4D cell culture. *Nat. Rev. Mater.* **3**, 17087 (2018).
17. Wylie, R. G. et al. Spatially controlled simultaneous patterning of multiple growth factors in three-dimensional hydrogels. *Nat. Mater.* **10**, 799–806 (2011).
18. Mosiewicz, K. A. et al. In situ cell manipulation through enzymatic hydrogel photopatterning. *Nat. Mater.* **12**, 1071–1077 (2013).
19. DeForest, C. A. & Tirrell, D. A. A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. *Nat. Mater.* **14**, 523–531 (2015).
20. Grim, J. C. et al. A reversible and repeatable thiol–ene bioconjugation for dynamic patterning of signaling proteins in hydrogels. *ACS Cent. Sci.* **4**, 909–916 (2018).
21. Shadish, J. A., Benuska, G. M. & DeForest, C. A. Bioactive site-specifically modified proteins for 4D patterning of gel biomaterials. *Nat. Mater.* **18**, 1005–1014 (2019).
22. Broguiere, N. et al. Morphogenesis guided by 3D patterning of growth factors in biological matrices. *Adv. Mater.* **32**, 1908299 (2020).
23. Batalov, I., Stevens, K. R. & DeForest, C. A. Photopatterned biomolecule immobilization to guide three-dimensional cell fate in natural protein-based hydrogels. *Proc. Natl Acad. Sci. USA* **118**, e2014194118 (2021).
24. Hoyt, E. A., Cal, P. M. S. D., Oliveira, B. L. & Bernardes, G. J. L. Contemporary approaches to site-selective protein modification. *Nat. Rev. Chem.* **3**, 147–171 (2019).
25. Shadish, J. A. & DeForest, C. A. Site-selective protein modification: from functionalized proteins to functional biomaterials. *Matter* **2**, 50–77 (2020).
26. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A strain-promoted [3+2] azide–alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).
27. DeForest, C. A., Polizzotti, B. D. & Anseth, K. S. Sequential click reactions for synthesizing and patterning three-dimensional cell microenvironments. *Nat. Mater.* **8**, 659–664 (2009).
28. Arakawa, C. K., Badeau, B. A., Zheng, Y. & DeForest, C. A. Multicellular vascularized engineered tissues through user-programmable biomaterial photodegradation. *Adv. Mater.* **29**, 1703156 (2017).
29. Badeau, B. A., Comerford, M. P., Arakawa, C. K., Shadish, J. A. & DeForest, C. A. Engineered modular biomaterial logic gates for environmentally triggered therapeutic delivery. *Nat. Chem.* **10**, 251–258 (2018).
30. Guimaraes, C. P. et al. Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1787–1799 (2013).
31. Shaner, N. C. et al. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567–1572 (2004).
32. Szymczak, A. L. et al. Correction of multi-gene deficiency in vivo using a single ‘self-cleaving’ 2A peptide-based retroviral vector. *Nat. Biotechnol.* **22**, 589–594 (2004).
33. Sawano, A. & Miyawaki, A. Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. *Nucleic Acids Res.* **28**, E78 (2000).
34. Roberts, P. J. et al. Rho family GTPase modification and dependence on CAAX motif-signaled posttranslational modification. *J. Biol. Chem.* **283**, 25150–25163 (2008).
35. Kumagai, A. et al. A bilirubin-inducible fluorescent protein from eel muscle. *Cell* **153**, 1602–1611 (2013).
36. Kwon, J. et al. Bright ligand-activatable fluorescent protein for high-quality multicolor live-cell super-resolution microscopy. *Nat. Commun.* **11**, 273 (2020).
37. To, T. L., Zhang, Q. & Shu, X. Structure-guided design of a reversible fluorogenic reporter of protein–protein interactions. *Protein Sci.* **25**, 748–753 (2016).
38. Li, L., Fierer, J. O., Rapoport, T. A. & Howarth, M. Structural analysis and optimization of the covalent association between SpyCatcher and a peptide Tag. *J. Mol. Biol.* **426**, 309–317 (2014).
39. Gouzel, K., Josette, P., Agnes, U. & Daniel, L. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl Acad. Sci. USA* **95**, 5752–5756 (1998).
40. Paulmurugan, R., Umezawa, Y. & Gambhir, S. S. Noninvasive imaging of protein–protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. *Proc. Natl Acad. Sci. USA* **99**, 15608–15613 (2002).
41. Paulmurugan, R. & Gambhir, S. S. Monitoring protein–protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. *Anal. Chem.* **75**, 1584–1589 (2003).
42. Hall, M. P. et al. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **7**, 1848–1857 (2012).
43. Dixon, A. S. et al. NanoLuc complementation reporter optimized for accurate measurement of protein interactions in Cells. *ACS Chem. Biol.* **11**, 400–408 (2016).
44. Jullien, N., Sampieri, F., Enjalbert, A. & Herman, J. Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. *Nucleic Acids Res.* **31**, e131 (2003).
45. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593–605 (2007).
46. Keeble, A. H. et al. Approaching infinite affinity through engineering of peptide–protein interaction. *Proc. Natl Acad. Sci. USA* **116**, 26523–26533 (2019).
47. Keeble, A. H. et al. DogCatcher allows loop-friendly protein–protein ligation. *Cell Chem. Biol.* **29**, 339–350 (2021).
48. Wu, X.-L., Liu, Y., Liu, D., Sun, F. & Zhang, W.-B. An intrinsically disordered peptide–peptide stapler for highly efficient protein ligation both in vivo and in vitro. *J. Am. Chem. Soc.* **140**, 17474–17483 (2018).
49. Fierer, J. O., Veggiani, G. & Howarth, M. SpyLigase peptide–peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proc. Natl Acad. Sci. USA* **111**, E1176–E1181 (2014).
50. Buldun, C. M., Jean, J. X., Bedford, M. R. & Howarth, M. SnoopLigase catalyzes peptide–peptide locking and enables solid-phase conjugate isolation. *J. Am. Chem. Soc.* **140**, 3008–3018 (2018).

51. Lorand, L. & Graham, R. M. Transglutaminases: crosslinking enzymes with pleiotropic functions. *Nat. Rev. Mol. Cell Biol.* **4**, 140–156 (2003).
52. Wong, S., Mosabbir, A. A. & Truong, K. An engineered split intein for photoactivated protein *trans*-splicing. *PLoS ONE* **10**, e0135965 (2015).
53. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
54. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Methods

All pertinent experimental procedures and methods are available within this manuscript and its associated Supplementary Information. Data collection was performed with the assistance of NanoDrop2000/2000c 1.6.198, Bruker TopSpin 2.1, Gen5 v.3.09, Azure Biosystems 1.6.4.1229, DataMax v.2.2, Leica Application Suite X 4.5.0.25531, Analyst TF 1.7.1 and ScanImage Premium 2021.1.0. Data analysis was performed with the assistance of Microsoft Excel 2016, GraphPad Prism v.6, ImageJ v.1.53f51, Adobe Illustrator CS6 v.16.0.0, CellProfiler v.4.2.1, Imaris Viewer v.1.6.0, ChemDraw v.20.0 and PeakView v.2.2.

Reporting summary

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

Data availability

All pertinent experimental and characterization data are available within this manuscript and its associated Supplementary Information. Plasmids generated during the current study are available from the corresponding author upon reasonable request. Source data are provided with this paper.

Acknowledgements

We recognize T. Rapp and R. Francis for synthetic advice, J. Shadish for helpful discussion on molecular cloning, R. Gharios for forward-looking conversations, R. Bretherton for providing the SpyTag peptide and transgenic dermal fibroblasts, J. Davis for gifting the HEK-293T cells (originally acquired from the American Type Culture Collection; CRL-3216) and S. Edgar (in memoriam) for assistance with mass spectrometry. This work was supported by a CAREER Award (DMR 1652141 to C.A.D.) and grants (DMR 1807398 and CBET 1803054 to C.A.D.) from the National Science Foundation, as well as a Maximizing Investigators' Research Award (R35GM138036 to C.A.D.) from the National Institutes of Health. Student fellowship support was provided by the Institute for Stem Cell and Regenerative Medicine (to B.G.M.-R.) and Mary Gates

Endowment for Students (to A.C.S., C.H.B. and S.K.) at the University of Washington. Part of this work was conducted with instrumentation provided by the Joint Center for Deployment and Research in Earth Abundant Materials. The Thorlabs multiphoton microscope was acquired with and operated under support from the Washington Research Foundation and the University of Washington College of Engineering, Institute for Stem Cell and Regenerative Medicine and Departments of Chemical Engineering, Bioengineering, Chemistry and Biology.

Author contributions

E.R.R. and C.A.D. conceived of and designed the experiments. E.R.R., B.G.M.-R., A.C.S., C.H.B., S.K. and J.R.F. performed the experiments. E.R.R., B.G.M.-R., A.C.S., S.K., C.H.B. and C.A.D. analysed the data and prepared the figures. E.R.R., B.G.M.-R. and C.A.D. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

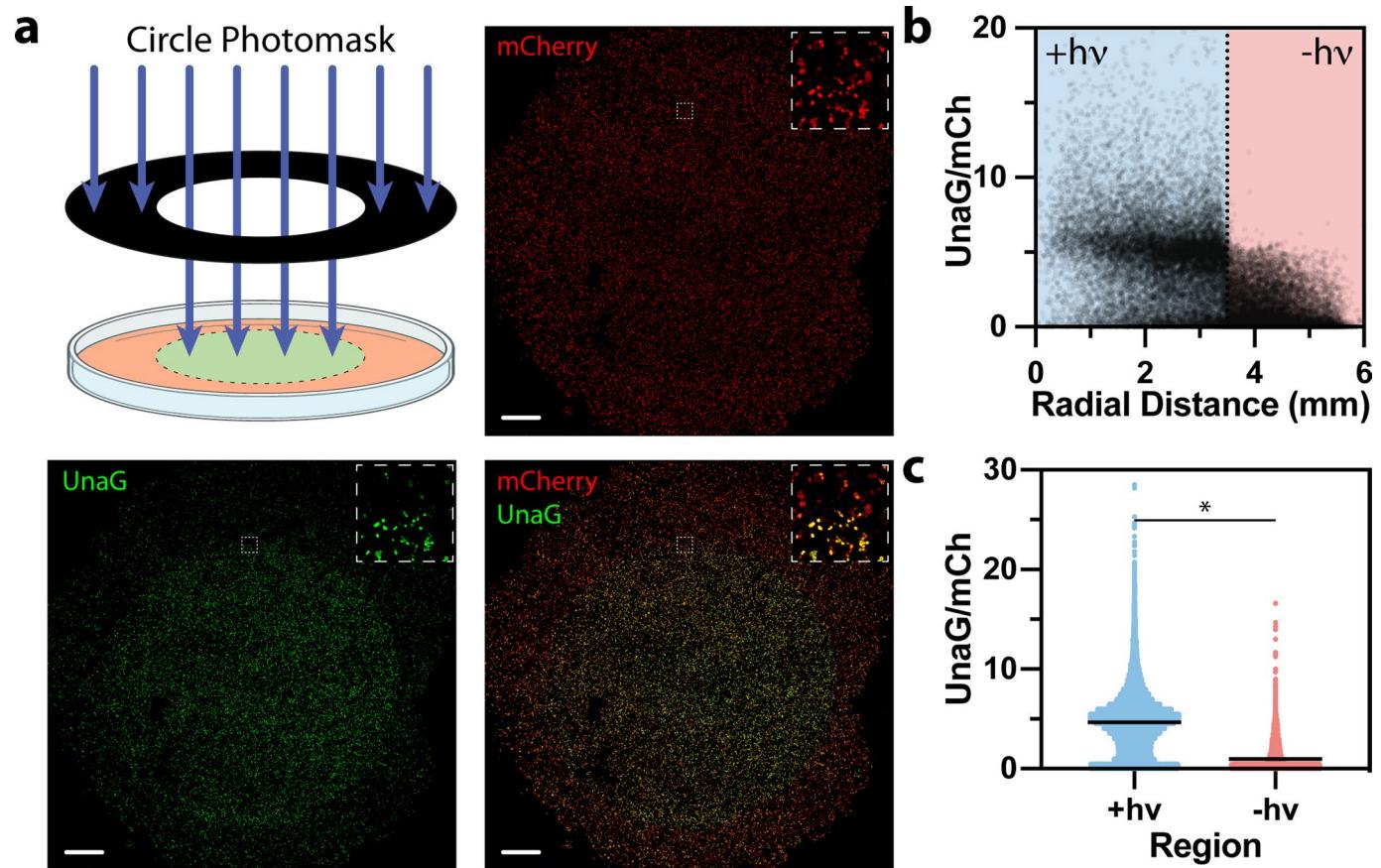
Extended data is available for this paper at <https://doi.org/10.1038/s41557-023-01152-x>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41557-023-01152-x>.

Correspondence and requests for materials should be addressed to Cole A. DeForest.

Peer review information *Nature Chemistry* thanks Cecile Echalier and the other, 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 | Photoactivation of split UnaG with spatiotemporal precision in living cells through intracellular LASL. **a-c**, Mask-based photolithography spatiotemporally directs UnaG reassembly within HEK-293T cell culture. **a**, Fluorescent images of culture dish with inlays of exposure boundary magnified. **b**, Individual cell UnaG/mCh signal quantified radially outwards from the photomask's center, normalized to the average UnaG/mCh

ratio in unexposed cells. Dashed line indicates exposure edge. **c**, Violin scatter plots of normalized UnaG/mCh ratios in light-(un)exposed regions. Light treatments, $\lambda = 365$ nm, 20 mW cm^{-2} , 20 min. Asterisks denote conditions with statistically significant differences in signal ($p < 0.0001$, two-tailed unpaired t-tests). Similar results were independently achieved in 3 experimental replicates. Scale bars, 1 mm (a).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see 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	Data collection was performed with the assistance of NanoDrop2000/2000c 1.6.198, Bruker Topspin 2.1, Gen5 Ver 3.09, Azure Biosystems 1.6.4.1229, DataMax V2.2, Leica Application Suite X 4.5.0.25531, Analyst TF 1.7.1, and ScanImage Premium 2021.1.0.
Data analysis	Data analysis was performed with the assistance of Microsoft Excel 2016, GraphPad Prism Version 6, ImageJ 1.53f51, Adobe Illustrator CS6 Version 16.0.0, CellProfiler 4.2.1, Imaris Viewer 1.6.0, ChemDraw 20.0, and Peak View 2.2.

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 Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All pertinent experimental and characterization data are available within this manuscript and its associated Supplementary Information. Source data for all main text Figures (2a, 2d, 3d, 3e, 3h, 3i, 4c, 4d, 4e, 4k, 5c, 5e, 5f), Extended Data Figures (1b, 1c), and Supplementary Figures (1, 3, 4, 6, 9, 10, 11, 14, 15, 16) are directly available. Plasmids generated during the current study are available from the corresponding author on reasonable request.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	All experiments were performed at least in triplicates. No statistical method was used to determine the sample size. Sufficient microscopy images were collected to ensure their representation of the sample.
Data exclusions	No data was excluded.
Replication	Multiple independent experiments were performed and exact numbers are stated in the figure captions.
Randomization	No randomizations were required for the experiments performed. A positive and negative control were used when necessary.
Blinding	Blinding was not relevant to the study as there were no relevant biases. There were no participants in any study so blinding was not necessary for these studies.

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
<input checked="" type="checkbox"/>	<input 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"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

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

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

HEK-293T cells were gifted by Dr. Jennifer Davis (University of Washington), who previously obtained the cells from ATCC (CRL-3216). Transgenic mouse dermal fibroblasts bearing a Cre-dependent dual-color reporter in the "safe harbor" Rosa26 locus were isolated and provided by Ross Bretherton (University of Washington).

Authentication

The cell lines were used without further authentication. Genetically modified cells lines clearly exhibited the expected labeling pattern.

Mycoplasma contamination

Cell line used in these studies were not tested for mycoplasma contamination.

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

HEK derivative cells were used for their ease of transfection.