Functional Assembly of Split Protein Pairs via a Chemically Activated SpyLigation

Yu-Ting Chang

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2024

Committee:

Cole A. DeForest

Andre Berndt

Program Authorized to Offer Degree:

Department of Bioengineering

University of Washington

Abstract

Functional Assembly of Split Protein Pairs via a Chemically Activated SpyLigation

Yu-Ting Chang

Chair of the Supervisory Committee:

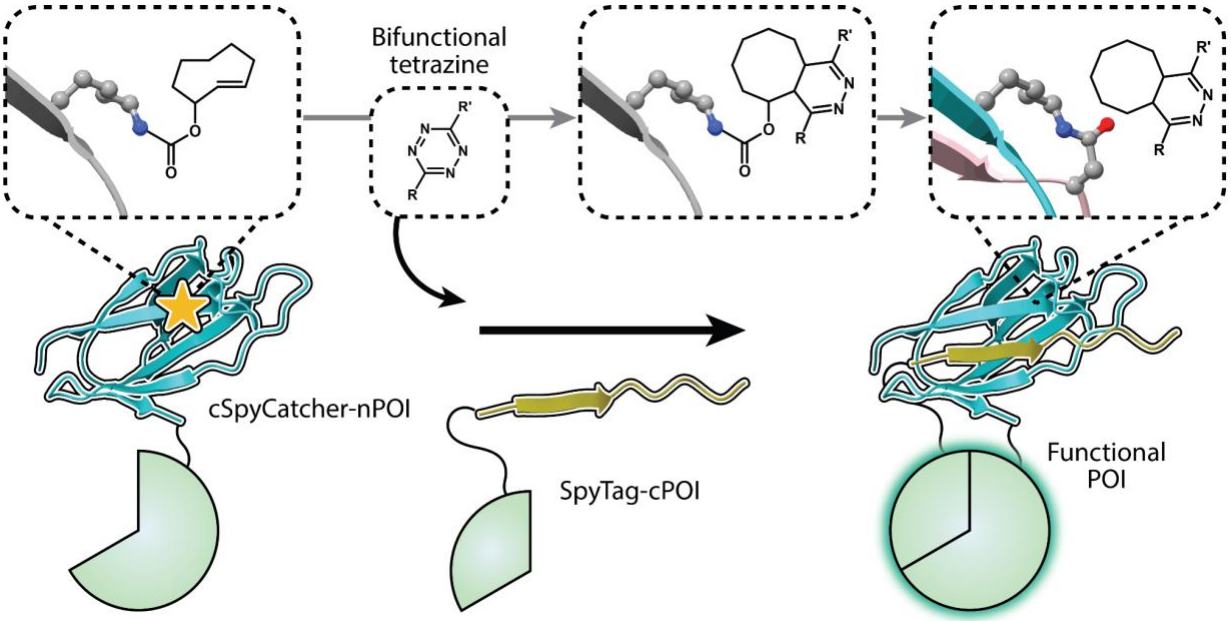
Cole A. DeForest

Department of Bioengineering

Given the critical roles proteins play in a broad variety of biological functions across all levels of life, it is in our best interest to develop innovative methods to control protein activity within biological systems. This is essential for expanding our understanding of cellular processes and facilitating novel therapeutic applications. Although genome editing technologies have advanced considerably, they still face limitations in terms of protein regulation and in providing real-time capabilities that allow us to probe biological processes across time scales. In this study, we present CASL (Chemically Activated SpyLigation) as a genetically encoded and chemically activated protein ligation method, which we utilize for triggered functional assembly of split

protein pairs. By incorporating a bioorthogonal trans-cyclooctene (TCO) cage at the isopeptide-forming lysine of SpyCatcher003, we achieve on-demand, permanent conjugation of non-functional split fragments in response to tetrazine treatment. The rapid TCO click-to-uncage process facilitates swift and covalent ligation, restoring functional activity both in solution and within living cells. We demonstrated that CASL's excellent specificity and bioorthogonality allow it to be easily adapted to cellular contexts for selective protein regulation without off-target effects or the requirement for high inducer concentrations. Looking ahead, we expect CASL to provide a powerful tool for real-time control of protein function, potentially serving as a complementary method to other inducible dimerization systems. This combined approach could lead to versatile applications for probing and directing complex cellular fates in 3D cell culture.

Graphical Abstract



Introduction

There is no doubt that proteins comprise a diverse group of biomolecules in terms of their functions, as enzymes, structural components, signaling molecules, and gene expression regulators. Their diverse roles make them crucial for cellular function, adaptation, and development. [1] [2] Temporal regulation of protein activity enables precise control over biological processes, ensuring appropriate reactions to environmental changes and guiding developmental phases. This control is essential for understanding disease mechanisms, and holds significant therapeutic potential, allowing for targeted treatments with minimal side effects.

To harness this potential, various strategies have been developed to temporally control protein activity. These strategies are fundamental in biological research and applications, providing the precision needed to study dynamic cellular processes, modulate cellular functions,

advance medicinal development, and create synthetic biological systems. Inducible expression systems, for example, are widely used to switch gene expression in response to specific inducers, enabling controlled protein synthesis. [3] Tetracycline-inducible systems (Tet-On/Tet-Off) regulate gene expression in eukaryotic cells, controlling gene activity for research and gene therapy through the administration of tetracycline or doxycycline. Complex circuits with temporal control are also possible by using ligand-binding domains that detect inducers and effector domains that regulate gene expression. [4] Photocaging is another widely used method for temporal protein function control. This involves chemically introducing light-sensitive groups into amino acids. [5] For instance, the *ortho*-nitrobenzyl (ONB) group can be attached to a specific amino acid, to control protein function by "caging" and "uncaging" proteins with UV light. [6] [7] Additionally, genetically encoded light-sensitive proteins including rhodopsins are utilized in optogenetics to manipulate neural activity and G-protein signaling inside cells in real-time. The light-sensitive cation channel ChR2 can trigger membrane depolarization and action potentials in mammalian cells and neurons when exposed to light, serving as an excitatory optogenetic tool. [8] CRISPR technology offers another approach for temporal control, such as with photoactivatable CRISPR-Cas9 for gene editing. In this method, a light-sensitive domain is attached to the CRISPR-Cas9 complex, allowing it to edit specific genes only when exposed to light, enabling precise spatiotemporal control. [9], [10] Similarly, Cre-LoxP recombination is a widely used site-specific method that enables precise genetic modifications, such as deletions, insertions, and inversions, at targeted locations. It employs Cre recombinase, an enzyme from the P1 bacteriophage, which facilitates recombination by recognizing and acting between LoxP sites. This system can be regulated by tissue-specific or inducible promoters, providing temporal and spatial control over gene editing. [11], [12] Transposases are crucial enzymes that facilitate the

movement of transposons, segments of DNA that can relocate within the genome, by catalyzing their excision and reinsertion. This process not only reshuffles genetic material, contributing to genetic variability and evolution but is also utilized in applications such as insertional mutagenesis and genetic delivery. [13] However, all these methods are naturally slow because they rely on cellular machinery to express modified or deleted genes, which can lack temporal control. Furthermore, because the changes they make are permanent, they lack the flexibility to be fine-tuned after the primary modification. [14]

Among these various approaches regulating protein activity, split protein reconstitution has emerged as a particularly powerful tool. Unlike other approaches, split protein systems provide a diverse platform that may be used in a variety of biological scenarios, ranging from researching protein-protein interactions to directing complicated cellular processes. [15] For instance, split luciferase reporters allow researchers to study protein-protein interactions and signaling pathways with high precision, detecting transient interactions by measuring luciferase activity upon fragment reassembly. [16], [17] Chemically induced dimerization (CID), a key strategy in split protein systems, offers distinct advantages over gene-editing techniques that rely on permanent genomic alterations. CID can directly activate protein signaling without modifying the genome or requiring new protein synthesis, enabling rapid, reversible control of protein activity. This flexibility allows for real-time manipulation, facilitating the study of interactions and pathways that might be challenging to investigate. The relative simplicity and independence from genetic alterations make CID highly desirable for therapeutic applications. Non-toxic bioorthogonal inducers can offer safer and more direct regulation of protein activity, therefore boosting the prospective use of CID in clinical settings. [18], [19], [20] CCID has driven significant innovations in protein engineering and synthetic biology, deepening our

understanding of cellular functions and advancing the development of novel therapeutic strategies. It is achieved by introducing two dimerization domains that bind to each other upon the addition of a small molecule inducer. Some common examples include the rapamycin-induced dimerization of FKBP and FRB [21], and the coumermycin-induced dimerization of proteins fused to the bacterial DNA gyrase subunit B. [22]

CID and split protein reassembly strategies do present challenges. Temporal control of split protein activity can be achieved by precise timing of the inducer addition; however, some inducers can affect other signaling pathways. For example, rapamycin-induced CID can affect the mTOR signaling pathway which leads to potential off-target effects. [23] Additionally, engineering CID domains into proteins without affecting their native folding or stability can be complicated. This process can lead to non-functional proteins or impede their interactions with other cellular processes. [21] Some CID systems need high inducer concentrations, posing risks of toxicity to organisms and potential disruption of cellular functions. For instance, excessive gibberellin application in the GID1-GAI system can induce cellular acidification, potentially leading to cytotoxic effects. [24] In addition, practical applications require scalability and efficiency of protein reassembly. Overcoming these challenges is crucial for developing split protein systems in both research and therapeutic settings. To address these limitations, we sought to develop an alternative CID strategy for programmable protein activation that features fast kinetics, bioorthogonality, and high specificity. Additionally, many CID interactions are non-covalent, making them unstable and reversible. The formation of stable, covalent bonds ensures more robust and long-lasting modifications. These systems aim to feature fast kinetics, true bioorthogonality, and ease of genetic engineering while being highly modular and versatile without unintentional background reaction.

To address these limitations, our group is exploring alternative strategies that offer improved control and precision in protein activation. Our hypothesis posits that recently developed click-to-release chemistries can significantly improve the control and precision of split protein reassembly. Here, we utilize genetic code expansion to create a caged proteins that can undergo small molecule-triggerable and specific protein-protein ligation into bioactive species. This system holds the potential for more efficient, dose-dependent, highly specific controlled, and irreversible protein reassembly. With the highly utilized versatile SpyLigation system, we are able to achieve our goals with SpyCatcher and SpyTag, which spontaneously form a covalent isopeptide bond with high efficiency and high yield within cells. These bonds are not affected by the reversibility issues often found with non-covalent CIDs. [25] Additionally, both SpyLigation pairs are small in size, which allows them to be easily fused to the N- or C-termini of POI without significantly altering their structure. To achieve even faster coupling, we employ the optimized SpyLigation003 version (SpyCatcher003, SC003; 113 amino acids; 15.6 kDa) and (SpyTag003, ST003; 16 amino acids; 1.9 kDa), featuring faster kinetics along with a more stable isopeptide bond formation, even at low concentrations. [26] SpyLigation 003 involves forming isopeptide bonds between Lys31 on SC003 and an aspartic acid side chain on ST003.

Recently reported light-activated SpyLigation (LASL) improves on traditional CID and optogenetic methods by using SpyTag/SpyCatcher tools with an *o*-nitrobenzyl photocage. [6], [27] Upon near-UV light exposure, it enables rapid, irreversible protein assembly with precise spatial and temporal control, offering effective, dose-dependent regulation in biomaterials and complex environments, including living cells. Despite the advantages offered by LASL, several challenges persist, particularly in photochemistry. One major issue is that light activation can be

slower than desired, often taking minutes to achieve full activation instead of producing instantaneous results. Another significant limitation is the poor penetration of light into biological tissues, which makes it difficult for light to reach deeper areas, hence limiting the activation of target proteins. The problem is exacerbated by employing low-wavelength light, which has a much lower capacity to penetrate skin in vivo. [28] [29] Although 365 nm light is generally considered less hazardous than other forms of UV radiation, improper calibration can still pose significant risks to tissues, potentially leading to in vivo damage and phototoxicity. Exposure to 365 nm UV radiation can induce both reversible and permanent damage to many cell types. [30] [31] [32] These limitations underscore the need for alternative strategies or enhanced technologies beyond optogenetics, particularly when targeting deep tissue stimulation or when immediate activation is essential.

In the pursuit of a fast, highly specific, bioorthogonal method to trigger split protein systems without complications, the concept of 'click chemistry' emerged, describing reactions that are rapid, reliable, and selectively applicable to synthesizing functional materials or biomolecule conjugation. [33] Among these, the "click-to-release" technique stands out, enabling the release of bioactive molecules from a carrier by cleaving a linker molecule. This bond cleavage elimination reaction, particularly effective in prodrug activation and instantaneous drug release in living systems, is powered by the ultra-fast inverse electron demand Diels-Alder (IEDDA) reaction. [34], [35], [36] A prime example is the tetrazine/trans-cyclooctene (TCO) ligation, this reaction forms two bonds and a six-membered ring through diene-dienophile cycloaddition with exceptionally high yields (second-order rate constants of 1 to 10^6 $M^{-1}s^{-1}$), even under dilute conditions, making it ideal for in vivo contexts. [37], [38], [39] Tetrazine and TCO groups are highly specific in biological media, showing no reactivity with amine or thiol

nucleophiles. IEDDA's quick kinetics at low concentrations underline its role as a leading tool for pre-targeting applications in bioorthogonal chemistry, providing accuracy in protein labeling and tracking within live systems. [5], [40], [41]

We hypothesized that SpyLigation could be chemically controlled by installing a TCO-caged lysine (TCOK) at the reactive site of SpyCatcher003 (cSC003) through genetic code expansion (GCE). This approach utilizes optimized tRNA/tRNA synthetase pairs for the efficient incorporation of TCOK at the amber stop codon, effectively blocking Lys31 on SC003. [42] [43] [44] This chemical cage keeps SpyCatcher inactive until the tetrazine "clicks and releases" the cage, allowing control over SpyLigation. We anticipated that the chemical cage on cSC003 would prevent isopeptide bond formation with ST003. However, treatment with tetrazine, known for its fast and selective bioorthogonal reactivity, could reverse this blockage. Upon uncaging, SC003 is activated in a dose-dependent manner with temporal control. We expect that this chemically-activated SpyLigation (CASL) could be used to irreversibly assemble split protein pairs and restore their function in a variety of contexts. This opens the door for future in vivo applications and broader research opportunities.

Result and Discussion

Construction of a cSC split protein for CASL

To implement CASL, the first key step is incorporating TCOK into SC003. This involves efficiently synthesizing caged lysine and effective GCE to install the non-canonical amino acid at specific sites in both bacterial and mammalian systems.

For bacterial expression, we used the pCDF-Pyl-AFx2 plasmid which contains two copies of *Methanosarcina mazei* Pyrrolysyl-tRNA synthetase (*Mm*PylRS) with Tyr306Ala and

Tyr384Phe mutations (AF) and three copies of the corresponding tRNA for amber codon suppression. These mutations were previously reported to facilitate the incorporation of bulky lysine derivatives including TCOK. [45] **(Supplementary Information Table S1)**

To achieve GCE for site-specific TCOK installation in mammalian cells, we utilized the pcDNA3.1(+)_U6 tRNAPyl_CMV NESPylRS(AF) plasmid containing the same *mM*PylRS with AF mutations fused to a nuclear export signal (NESPylRSAF). This plasmid includes four copies of an engineered M15 tRNA cassette and has previously been shown to be effective in incorporating lysine derivatives including TCOK.[46] **(Supplementary Information Table S1)** Toward the production of CASL proteins in either bacterial or mammalian cells, a plasmid containing the POI sequence is co-translated with the appropriate orthogonal aminoacyl-tRNA synthetase/tRNA plasmid, under TCOK-supplemented culture conditions for GCE. cSC003 is generated by incorporating TCOK at the catalytic Lys31 residue of SC through amber suppression. **(Supplementary Information Method S6)**

Chemically-Induced Cleavage and Native Lysine Recovery

Essential to creating the system for bioorthogonal decaging reaction in living cells was the development of non-canonical amino acids, ax-TCO-lysine as the compound that can be efficiently incorporated into the protein by GCE. TCOK was synthesized using an established method, yielding ample quantities of the caged amino acid to support several liters of protein expression (**Supplementary Information Method S1-4**). When treated with tetrazine, the TCO cage is removed via an inverse electron-demand Diels-Alder (IEDDA) reaction, restoring the native lysine activity. By directly mixing TCOK (0.3 mM in phosphate-buffered saline, PBS, pH = 7.4) with 5 mM tetrazine and monitoring the reaction using UV-Vis absorption spectrometry,

we were able to analyze the rapid reaction and confirm the molecule's functionality (**Fig. 1a**, **Supplementary Information Method S5**). During the reaction, absorbance at 525 nm, corresponding to the pink color of tetrazine, is significantly reduced, indicating the ligation of TCO and tetrazine. The IEDDA reactions between ax-TCO-K and tetrazine resulted in a significant absorbance shift at 525 nm. By fitting the recorded data to a second-order reaction kinetic equation, we performed a linear regression that yielded a rate constant (k) of 28.1 $M^{-1}s^{-1}$ and determined the reaction half-time to be approximately 7.11 seconds (**Fig. 1b**). These results are consistent with previous literature reports, confirming the rapid reaction kinetics for ligation.

Validation and Characterization of CASL Using Purified Proteins

Driven by the goal of assessing the effectiveness of CASL in vitro, constructs containing cSC003, SC003, and ST003 fused with the UnaG split protein were expressed in *Escherichia coli* and purified each species via affinity chromatography. UnaG, a fluorescent protein from Japanese eel muscle tissue, emits bright green fluorescence upon binding to bilirubin, a heme breakdown product. [47] UnaG's stability and photostability are attributed to its unique interaction with bilirubin, this binding not only activates the protein's fluorescence but also stabilizes it, making it less prone to photobleaching compared to other fluorescent proteins. This reaction enables rapid and noticeable green fluorescence at subcellular resolutions, serving as a proof of concept for the effectiveness of CASL with spatiotemporal control. [48] Functional activation can be influenced by various factors, so we used a previously validated method shown to work for rapamycin-induced assembly to split the UnaG protein. (nUnaG = residues 1-84; cUnaG = residues 85-139). [49] DDifferent combinations of ST003 and SC003 protein with UnaG fragment variants were previously expressed and tested, indicating that nUnaG-ST003 and

SC003-cUnaG protein pairs had better solubility and emitted the highest fluorescence for 24 hours compared to other combinations. [6] **(Fig. 2a)** Subsequently, the expressed proteins were purified via affinity chromatography. Protein samples were examined by analyzing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and mass spectrometry. (**Supplementary Information Figure S1**) To demonstrate the functionality of uncaging when fused with split proteins, the TCO cage on cSC003-cUnaG (TCO cage molecular weight: 152 Da) was chemically removed after tetrazine treatment. The resulting uncaged protein matched the mass of the native SC003-cUnaG species (molecular weight: 20158 Da), the mass shift indicating successful uncaging. (**Fig. 2b, Supplementary Information Table S2**) To assess the functionality of the caged product and its ability to perform SpyLigation, we subjected cSC003-cUnaG to reaction with excess nUnaG-ST003 at various tetrazine concentrations at physiological temperature (37°C) for 30 minutes before analyzing the samples using SDS-PAGE analysis. The amount of ligation was evaluated by measuring the intensities of the vanishing bands from cSC003-cUnaG and the emerging band corresponding to the SpyLigated complex. (**Fig. 2c, Supplementary Information Method S7**). To demonstrate that CASL can restore the split protein's functionality, we prepared a mixture of equal molar (1μM) CASL UnaG fragments alongside the substrate (bilirubin). Without tetrazine treatment, negligible fluorescence was observed, indicating a lack of covalent association between both UnaG fragments and the SpyLigation pair. Upon activation with tetrazine, CASL restores protein bioactivity. Green fluorescence becomes apparent and clearly visible even after 18 hours of treatment. **(Fig. 2d**) To show that CASL not only restores split protein functionality but also exhibits dose dependency with tetrazine, we treated the protein with varying concentrations of tetrazine. The results showed that the protein's fluorescence was restored in a dose-dependent manner, with higher

concentrations of tetrazine leading to greater fluorescence reconstitution, confirming the specificity and tunability of the reaction (**Fig. 2e, Supplementary Information Method S8**). These experiments conclusively demonstrate the controllability of CASL, exceptional specificity, and dose dependency. Together this confirms the ability to accurately and irreversibly activate functional proteins through CASL-mediated split fragment reconstitution.

Spatial control over CASL within living mammalian cells: mCherry membrane labeling

To aassess CASL's effectiveness in regulating protein pairs, uncaging, and successful ligation within mammalian cells. We cloned a membrane-labeling mammalian expression plasmid that contains a cSC003 fusion with mCherry, a P2A self-cleaving peptide sequence, and a ST003 tagged with an enhanced green fluorescent protein (eGFP) that localizes at the membrane through the CAAX motif derived from Ras sequences. [49] HEK-293T cells were co-translated with the membrane-labeling construct and the *Mm*PylRS/tRNA pair and cultured in TCOK-supplemented media to expand the genetic code.

Transfected cells exhibit green fluorescence at the plasma membrane due to the CAAX-EGFP-ST fusion and red fluorescence throughout the cytoplasm, indicating successful readthrough of the amber stop codon within cSC003. We additionally found that eGFP-ST003 with the CAAX motif is successfully localized on the cell membrane without fading even after an additional 24 hours of culture. (**Fig. 3a, Supplementary Information Method S9-11**) Upon tetrazine treatment, cSC003-mCherry is uncaged and subsequently accumulates at the plasma membrane along with eGFP-ST003 through CASL, cSC003-mCherry membrane labeling, and subcellular distribution scaled statistically significantly. (**Fig. 3b, c**) According to these studies,

CASL is an excellent method for determining long-term changes in intracellular protein localization.

CASL-mediated split protein functional activation in mammalian cells

We next explored whether we can permanently activate protein function in living mammalian cells in response to tetrazine treatment. To validate this hypothesis, we cloned a polycistronic construct that produces split proteins, consisting of cSC003-cUnaG, nUnaG-ST003, and mCherry, each separated by P2A. This design ensures uniform expression levels for all three protein components, with the mCherry component positioned at the 3' end of the transcriptional unit, the red fluorescence emitted by mCherry serves as an internal standard, providing a reference for the successful production of cSC003. ((**Fig. 4a, Supplementary Information Method S12-14**) After co-transfecting HEK-293T cells with the construct encoding split UnaG components and the variant MmPylRS/tRNA pair, the cells were cultured in TCOK-supplemented media. As anticipated, most cells exhibit red fluorescence after GCE which showcases the success of transfection and the read-through of the construct.

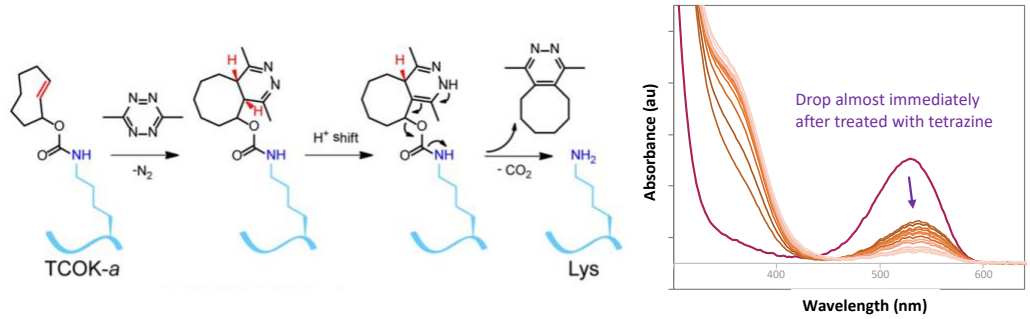
Transfected cells exhibited green fluorescence after overnight treatment with tetrazine, indicating successful functional assembly of UnaG via CASL, whereas unexposed cells retained red fluorescence only (**Fig. 4b**). The results reveal statistically significant increases in the UnaG/mCh fluorescence ratios, as determined through image analysis and quantification. Additionally, the green fluorescence persisted long-term for at least 24 hours post-tetrazine treatment, indicating the irreversible activation of UnaG protein function through SpyLigation (**Fig. 4c**).

Conclusions

In this study, we introduce CASL, an on-demand split protein assembly strategy that leverages genetically encoded, ultrafast click-to-release chemistry and covalent heterodimerization to drive split protein assembly in solution and in mammalian cells. By employing genetic code expansion to install a synthetic organic chemistry reactive TCO cage, we achieve site-specific, irreversible protein conjugation. This process involves the protection and deprotection of reactants and uses tetrazine, a bioorthogonal molecule, to precisely and irreversibly activate biologically inactive fragments within cells, allowing precise temporal control over activation. Although we have focused on chemical control with SpyCatcher003, we anticipate that this method could be easily adapted to other Tag/Catcher protein pairs (e.g., DogTag/Catcher, SnoopTag/Catcher) as each share an isopeptide-forming lysine; this would potentially allow for the orthogonal assembly of split pairs or better transamination kinetics for even faster coupling. [51] [52]. Additionally, this approach can be integrated with the light-activated system we previously reported, enabling dual complementary control. To extend this concept, developing different caged amino acids could also enable protein activation in response to other cues, such as pH, enzymes, or temperature. Another significant advantage of CASL is its usage of forming covalent bonds between SC003 and ST003, allowing for the design of inactive protein fragments with poor association. This ensures low background activity and reduces non-specific binding even under complex conditions. Further, once activated, the proteins activate permanent functionality, showcasing the robustness and reliability of the system. Based on the success of the Click Activated Protodrugs Against Cancer (CAPAC®) approach, which employed a tetrazine-modified biopolymer and TCO-linked doxorubicin prodrug for targeted

cancer therapy [53], CASL shows potential for in vivo applications as a precise, non-toxic, and effective therapeutic strategy.

a)



b)

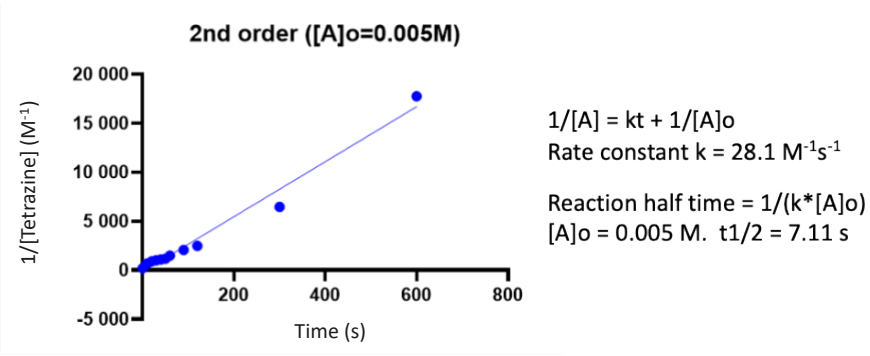


Fig1.

Click-to-release mechanism and ligation kinetics

(a) Left: Mechanism for the click-to-release reaction between TCOK and a model tetrazine. Right: Significant reduction in tetrazine absorbance at 525 nm is observed following ligation to TCOK (b) The time-dependent decrease in absorbance at 525 nm is fitted to a second-order reaction kinetics model, yielding a rate constant of 28 $M^{-1}s^{-1}$ and a half-life of approximately 7.11 seconds.

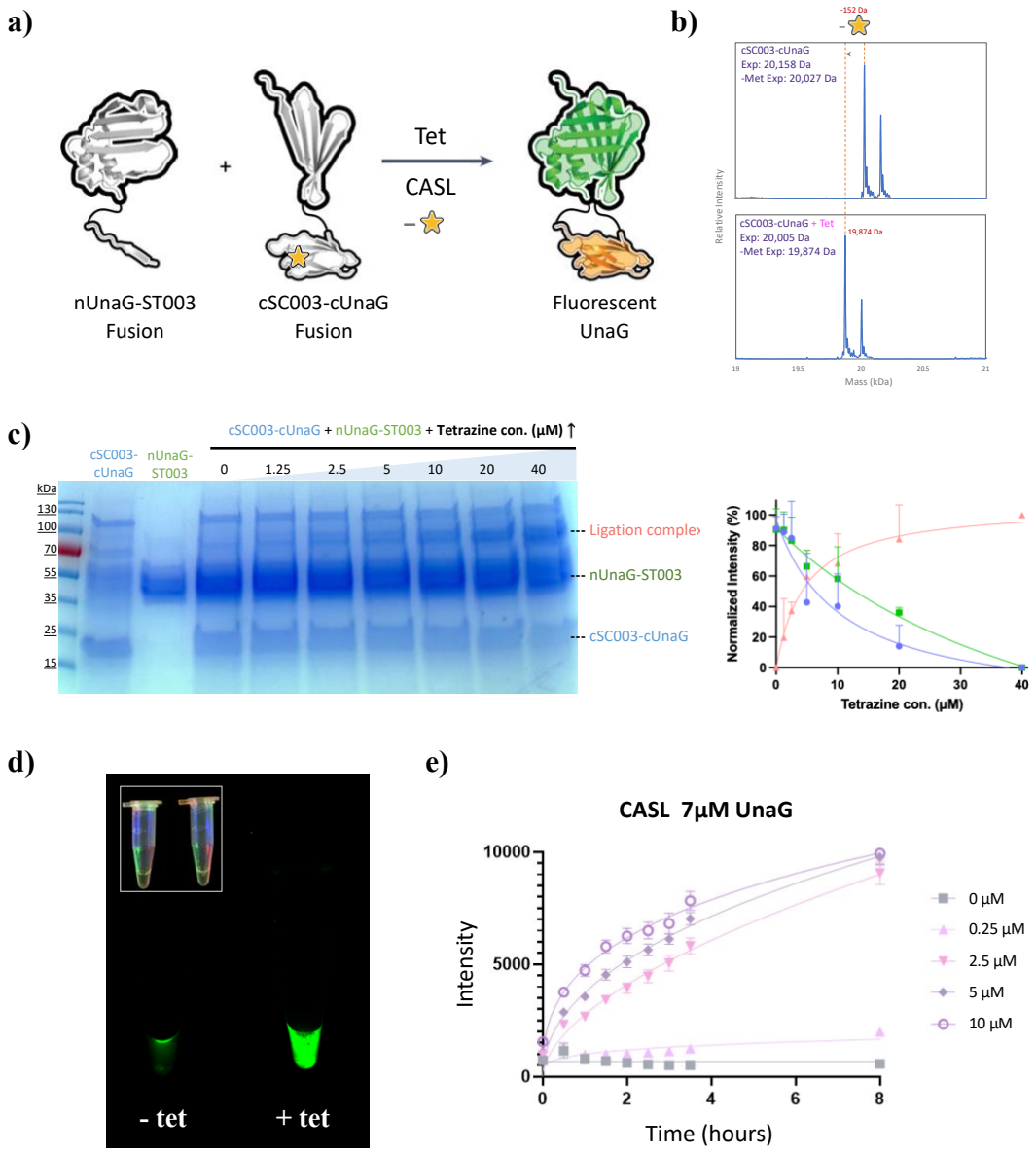


Fig 2.

Validation and characterization of CASL using split UnaG

(a) CASL UnaG Scheme (b) Intact protein mass spectrometry of purified cSC003-cUnaG before and after tetrazine treatment (final concentration: 40 μM). Dashed lines indicate the expected mass shift corresponding to TCOK uncaging (yellow stars). "Exp" denotes expected mass. (c) The ligation of CASL between cSC003-cUnaG and nUnaG-ST003 (35 mM; 37 °C) varies with

tetrazine treatment (final concentration: 0-40 μM), as visualized by SDS-PAGE after 30 minutes of incubation. cSC activation and ligated product formation show a tetrazine dose-dependency, as determined by SDS-PAGE band intensity quantification for each labeled species. Error bars represent the standard deviation (s.d.) from the mean of three experimental replicates. (d) UnaG assembly and fluorescence were only observed upon tetrazine exposure. (e) The fluorescence recovery of cSC003-UnaG demonstrates dose-dependency and functional activation in the presence of tetrazine in solution.

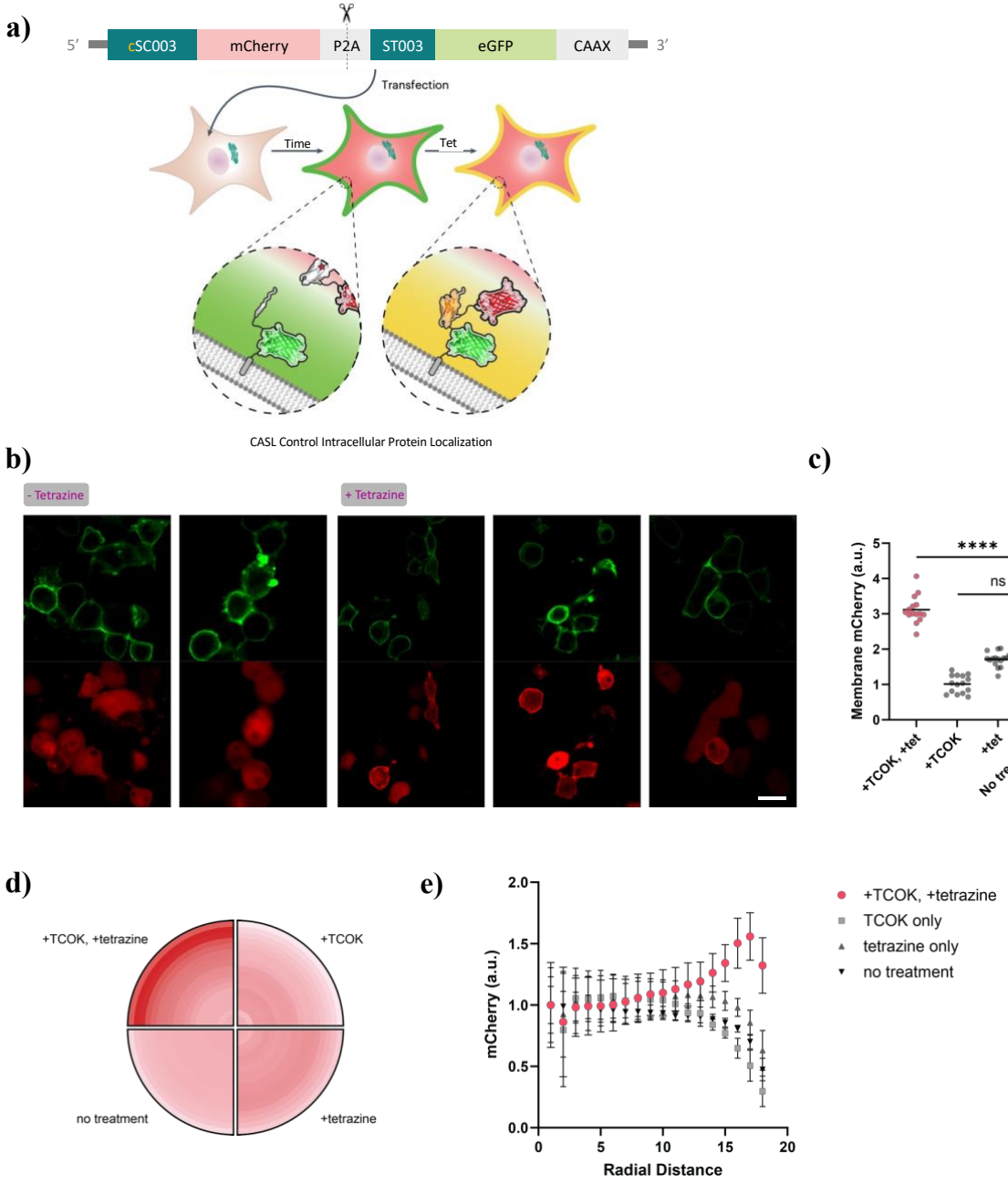


Fig 3.

CASL-mediated protein localization in HEK 293T cells

(a) Schematic of the gene cassette used for CASL-mediated plasma membrane labeling. In this construct, eGFP-ST is anchored to the membrane, where it forms a covalent bond with

cytoplasmic cSC-mCherry after tetrazine-induced uncaging. (b) Representative fluorescence micrographs of transfected HEK-293T cells post-overnight treatment with 100 μM tetrazine. (Scale bar: 20 μm) (c) Membrane labeling with mCherry increased significantly with tetrazine treatment, as shown by the relative fluorescence levels in the violin scatter plot. Conditions with statistically significant differences in signal are marked with an asterisk ($P < 0.0001$, two-tailed unpaired t-test). (d) The intracellular distribution of mCh shifted with CASL from being uniformly cytosolic to more membrane-localized following tetrazine treatment.

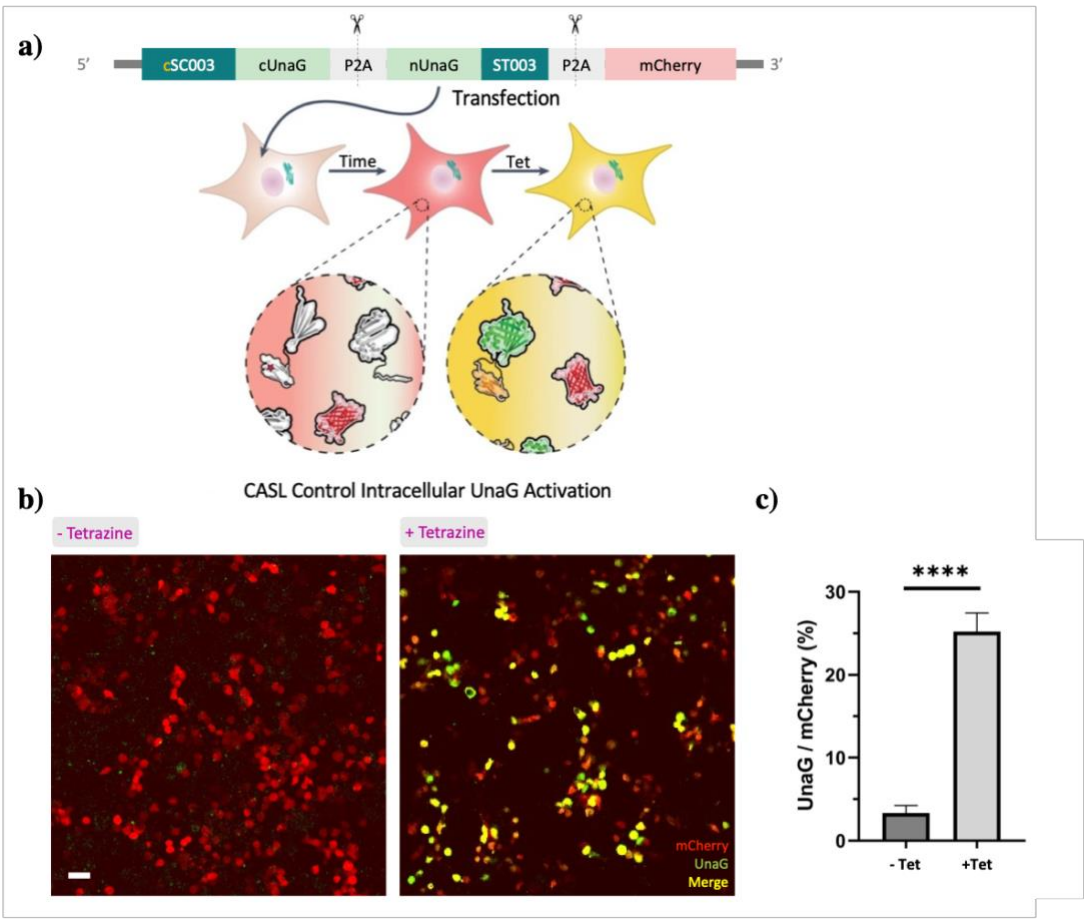


Fig 4.

Intracellular chemical activation control of split UnaG in living cells

(a) Schematic representation of the CASL UnaG plasmid constructs in mammalian cells, integrating non-associative UnaG fragments and mCherry as an internal control. (b) Representative fluorescence images of transfected HEK-293T cells (red), following overnight tetrazine treatment, displaying UnaG protein activation via CASL (green). (Scale bar: 20 µm) (c) Intracellular UnaG activation following tetrazine treatment, normalized to initial UnaG/mCh ratios.

Reference

[1] L. A. Banaszynski and T. J. Wandless, “Conditional Control of Protein Function,” *Chem. Biol.*, vol. 13, no. 1, pp. 11–21, Jan. 2006, doi: 10.1016/j.chembiol.2005.10.010.

[2] G. M. Cooper, “Regulation of Protein Function,” in *The Cell: A Molecular Approach. 2nd edition*, Sinauer Associates, 2000. Accessed: Aug. 29, 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK9923/

[3] M. Siddiqui, C. Tous, and W. W. Wong, “Small molecule-inducible gene regulatory systems in mammalian cells: progress and design principles,” *Curr. Opin. Biotechnol.*, vol. 78, p. 102823, Dec. 2022, doi: 10.1016/j.copbio.2022.102823.

[4] A. T. Das, L. Tenenbaum, and B. Berkhout, “Tet-On Systems For Doxycycline-inducible Gene Expression,” *Curr. Gene Ther.*, vol. 16, no. 3, pp. 156–167, Jun. 2016, doi: 10.2174/1566523216666160524144041.

[5] S. Mushtaq, S.-J. Yun, and J. Jeon, “Recent Advances in Bioorthogonal Click Chemistry for Efficient Synthesis of Radiotracers and Radiopharmaceuticals,” *Molecules*, vol. 24, no. 19, p. 3567, Oct. 2019, doi: 10.3390/molecules24193567.

[6] E. R. Ruskowitz *et al.*, “Spatiotemporal functional assembly of split protein pairs through a light-activated SpyLigation,” *Nat. Chem.*, 2023, doi: 10.1038/s41557-023-01152-x.

[7] J. W. Cheung *et al.*, “Genetic Encoding of a Photocaged Histidine for Light-Control of Protein Activity,” *ChemBioChem*, vol. 24, no. 7, p. e202200721, 2023, doi: 10.1002/cbic.202200721.

[8] E. S. Boyden, F. Zhang, E. Bamberg, G. Nagel, and K. Deisseroth, “Millisecond-timescale, genetically targeted optical control of neural activity,” *Nat. Neurosci.*, vol. 8, no. 9, pp.

1263–1268, Sep. 2005, doi: 10.1038/nn1525.

[9] J. Hemphill, E. K. Borchardt, K. Brown, A. Asokan, and A. Deiters, "Optical Control of CRISPR/Cas9 Gene Editing," *J. Am. Chem. Soc.*, vol. 137, no. 17, pp. 5642–5645, May 2015, doi: 10.1021/ja512664v.

[10] L. R. Polstein and C. A. Gersbach, "A light-inducible CRISPR/Cas9 system for control of endogenous gene activation," *Nat. Chem. Biol.*, vol. 11, no. 3, pp. 198–200, Mar. 2015, doi: 10.1038/nchembio.1753.

[11] B. Sauer and N. Henderson, "Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 85, no. 14, pp. 5166–5170, Jul. 1988, doi: 10.1073/pnas.85.14.5166.

[12] S. Turan *et al.*, "Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges," *J. Mol. Biol.*, vol. 407, no. 2, pp. 193–221, Mar. 2011, doi: 10.1016/j.jmb.2011.01.004.

[13] M. Muñoz-López and J. L. García-Pérez, "DNA Transposons: Nature and Applications in Genomics," *Curr. Genomics*, vol. 11, no. 2, pp. 115–128, Apr. 2010, doi: 10.2174/138920210790886871.

[14] H. Li, Y. Yang, W. Hong, M. Huang, M. Wu, and X. Zhao, "Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects," *Signal Transduct. Target. Ther.*, vol. 5, no. 1, pp. 1–23, Jan. 2020, doi: 10.1038/s41392-019-0089-y.

[15] S. S. Shekhawat and I. Ghosh, "Split-Protein Systems: Beyond Binary Protein-Protein Interactions," *Curr. Opin. Chem. Biol.*, vol. 15, no. 6, pp. 789–797, Dec. 2011, doi: 10.1016/j.cbpa.2011.10.014.

[16] J. Bae, J. Kim, J. Choi, H. Lee, and M. Koh, “Split Proteins and Reassembly Modules for Biological Applications,” *ChemBioChem*, vol. 25, no. 10, p. e202400123, May 2024, doi: 10.1002/cbic.202400123.

[17] S. Shen, “Development of Split-Protein Systems: From Binary to Ternary System,” *Adv. Biosci. Biotechnol.*, vol. 12, no. 3, Art. no. 3, Mar. 2021, doi: 10.4236/abb.2021.123006.

[18] S. Jana *et al.*, “Ultra-Fast Bioorthogonal Spin-Labeling and Distance Measurements in Mammalian Cells Using Small, Genetically Encoded Tetrazine Amino Acids,” *bioRxiv*, p. 2023.01.26.525763, Jan. 2023, doi: 10.1101/2023.01.26.525763.

[19] R. DeRose, T. Miyamoto, and T. Inoue, “Manipulating signaling at will: chemically-inducible dimerization (CID) techniques resolve problems in cell biology,” *Pflüg. Arch. - Eur. J. Physiol.*, vol. 465, no. 3, pp. 409–417, Mar. 2013, doi: 10.1007/s00424-012-1208-6.

[20] S. S. Shekhawat and I. Ghosh, “Split-Protein Systems: Beyond Binary Protein-Protein Interactions,” *Curr. Opin. Chem. Biol.*, vol. 15, no. 6, pp. 789–797, Dec. 2011, doi: 10.1016/j.cbpa.2011.10.014.

[21] D. M. Spencer, T. J. Wandless, S. L. Schreiber, and G. R. Crabtree, “Controlling signal transduction with synthetic ligands,” *Science*, vol. 262, no. 5136, pp. 1019–1024, Nov. 1993, doi: 10.1126/science.7694365.

[22] M. A. Farrar, J. Alberola-lla, and R. M. Perlmutter, “Activation of the Raf-1 kinase cascade by coumermycin-induced dimerization,” *Nature*, vol. 383, no. 6596, pp. 178–181, Sep. 1996, doi: 10.1038/383178a0.

[23] L. A. Banaszynski, C. W. Liu, and T. J. Wandless, “Characterization of the FKBP·Rapamycin·FRB Ternary Complex,” *J. Am. Chem. Soc.*, vol. 127, no. 13, pp. 4715–4721, Apr. 2005, doi: 10.1021/ja043277y.

[24] T. Miyamoto *et al.*, "Rapid and Orthogonal Logic Gating with a Gibberellin-induced Dimerization System," *Nat. Chem. Biol.*, vol. 8, no. 5, pp. 465–470, Mar. 2012, doi: 10.1038/nchembio.922.

[25] B. Zakeri *et al.*, "Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 12, pp. E690-697, Mar. 2012, doi: 10.1073/pnas.1115485109.

[26] A. H. Keeble *et al.*, "Approaching infinite affinity through engineering of peptide–protein interaction," *Proc. Natl. Acad. Sci.*, vol. 116, no. 52, p. 26523 LP – 26533, Dec. 2019, doi: 10.1073/pnas.1909653116.

[27] B. G. Munoz-Robles and C. A. DeForest, "Irreversible light-activated SpyLigation mediates split-protein assembly in 4D," *Nat. Protoc.*, vol. 19, no. 4, pp. 1015–1052, Apr. 2024, doi: 10.1038/s41596-023-00938-0.

[28] C. Ash, M. Dubec, K. Donne, and T. Bashford, "Effect of wavelength and beam width on penetration in light-tissue interaction using computational methods," *Lasers Med. Sci.*, vol. 32, no. 8, pp. 1909–1918, Nov. 2017, doi: 10.1007/s10103-017-2317-4.

[29] P. Avci *et al.*, "Low-level laser (light) therapy (LLLT) in skin: stimulating, healing, restoring," *Semin. Cutan. Med. Surg.*, vol. 32, no. 1, pp. 41–52, Mar. 2013.

[30] M. Klak *et al.*, "Irradiation with 365 nm and 405 nm wavelength shows differences in DNA damage of swine pancreatic islets," *PLOS ONE*, vol. 15, no. 6, p. e0235052, Jun. 2020, doi: 10.1371/journal.pone.0235052.

[31] S. J. Balk and the Council on Environmental Health and Section on Dermatology, "Ultraviolet Radiation: A Hazard to Children and Adolescents," *Pediatrics*, vol. 127, no. 3, pp. e791–e817, Mar. 2011, doi: 10.1542/peds.2010-3502.

[32] R. M. MacKie, “Long-term health risk to the skin of ultraviolet radiation,” *Prog. Biophys. Mol. Biol.*, vol. 92, no. 1, pp. 92–96, Sep. 2006, doi: 10.1016/j.pbiomolbio.2006.02.008.

[33] D. Svatunek, M. Wilkovitsch, L. Hartmann, K. N. Houk, and H. Mikula, “Uncovering the Key Role of Distortion in Bioorthogonal Tetrazine Tools That Defy the Reactivity/Stability Trade-Off,” *J. Am. Chem. Soc.*, vol. 144, no. 18, pp. 8171–8177, May 2022, doi: 10.1021/jacs.2c01056.

[34] E. G. Tomarchio *et al.*, “Tetrazine–*trans*-cyclooctene ligation: Unveiling the chemistry and applications within the human body,” *Bioorganic Chem.*, vol. 150, p. 107573, Sep. 2024, doi: 10.1016/j.bioorg.2024.107573.

[35] A. Rondon *et al.*, “Antibody PEGylation in bioorthogonal pretargeting with trans-cyclooctene/tetrazine cycloaddition: in vitro and in vivo evaluation in colorectal cancer models,” *Sci. Rep.*, vol. 7, no. 1, p. 14918, Nov. 2017, doi: 10.1038/s41598-017-15051-y.

[36] J. Kaur, M. Saxena, and N. Rishi, “An Overview of Recent Advances in Biomedical Applications of Click Chemistry,” *Bioconjug. Chem.*, vol. 32, no. 8, pp. 1455–1471, Aug. 2021, doi: 10.1021/acs.bioconjchem.1c00247.

[37] R. M. Versteegen, R. Rossin, W. ten Hoeve, H. M. Janssen, and M. S. Robillard, “Click to Release: Instantaneous Doxorubicin Elimination upon Tetrazine Ligation,” *Angew. Chem. Int. Ed.*, vol. 52, no. 52, pp. 14112–14116, 2013, doi: 10.1002/anie.201305969.

[38] M. Handula, K.-T. Chen, and Y. Seimbille, “IEDDA: An Attractive Bioorthogonal Reaction for Biomedical Applications,” *Molecules*, vol. 26, no. 15, p. 4640, Jul. 2021, doi: 10.3390/molecules26154640.

[39] M. L. Blackman, M. Royzen, and J. M. Fox, “Tetrazine ligation: fast bioconjugation

based on inverse-electron-demand Diels-Alder reactivity," *J. Am. Chem. Soc.*, vol. 130, no. 41, pp. 13518–13519, Oct. 2008, doi: 10.1021/ja8053805.

[40] X. Fan *et al.*, "Optimized Tetrazine Derivatives for Rapid Bioorthogonal Decaging in Living Cells," *Angew. Chem. Int. Ed.*, vol. 55, no. 45, pp. 14046–14050, 2016, doi: 10.1002/anie.201608009.

[41] N. K. Devaraj and R. Weissleder, "Biomedical Applications of Tetrazine Cycloadditions," *Acc. Chem. Res.*, vol. 44, no. 9, pp. 816–827, Sep. 2011, doi: 10.1021/ar200037t.

[42] J. Li, S. Jia, and P. R. Chen, "Diels-Alder reaction–triggered bioorthogonal protein decaging in living cells," *Nat. Chem. Biol.*, vol. 10, no. 12, pp. 1003–1005, 2014, doi: 10.1038/nchembio.1656.

[43] K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters, and J. W. Chin, "Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction," *Nat. Chem.*, vol. 4, no. 4, pp. 298–304, Apr. 2012, doi: 10.1038/nchem.1250.

[44] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, and S. Yokoyama, "Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases," *Biochem. Biophys. Res. Commun.*, vol. 371, no. 4, pp. 818–822, Jul. 2008, doi: 10.1016/j.bbrc.2008.04.164.

[45] T. Yanagisawa, M. Kuratani, E. Seki, N. Hino, K. Sakamoto, and S. Yokoyama, "Structural Basis for Genetic-Code Expansion with Bulky Lysine Derivatives by an Engineered Pyrrolysyl-tRNA Synthetase," *Cell Chem. Biol.*, vol. 26, no. 7, pp. 936-949.e13, 2019, doi: 10.1016/j.chembiol.2019.03.008.

[46] R. Serfling *et al.*, "Designer tRNAs for efficient incorporation of non-canonical amino

acids by the pyrrolysine system in mammalian cells," *Nucleic Acids Res.*, vol. 46, no. 1, pp. 1–10, Jan. 2018, doi: 10.1093/nar/gkx1156.

[47] A. Kumagai *et al.*, "A bilirubin-inducible fluorescent protein from eel muscle," *Cell*, vol. 153, no. 7, pp. 1602–1611, Jun. 2013, doi: 10.1016/j.cell.2013.05.038.

[48] J. Kwon *et al.*, "Bright ligand-activatable fluorescent protein for high-quality multicolor live-cell super-resolution microscopy," *Nat. Commun.*, vol. 11, no. 1, p. 273, Jan. 2020, doi: 10.1038/s41467-019-14067-4.

[49] T.-L. To, Q. Zhang, and X. Shu, "Structure-guided design of a reversible fluorogenic reporter of protein-protein interactions," *Protein Sci.*, vol. 25, no. 3, pp. 748–753, Mar. 2016, doi: 10.1002/pro.2866.

[50] P. J. Roberts *et al.*, "Rho Family GTPase modification and dependence on CAAX motif-signaled posttranslational modification," *J. Biol. Chem.*, vol. 283, no. 37, pp. 25150–25163, Sep. 2008, doi: 10.1074/jbc.M800882200.

[51] A. H. Keeble *et al.*, "DogCatcher allows loop-friendly protein-protein ligation," *Cell Chem. Biol.*, vol. 29, no. 2, pp. 339-350.e10, Feb. 2022, doi: 10.1016/j.chembiol.2021.07.005.

[52] D. Hatlem, T. Trunk, D. Linke, and J. C. Leo, "Catching a SPY: Using the SpyCatcher-SpyTag and Related Systems for Labeling and Localizing Bacterial Proteins," *Int. J. Mol. Sci.*, vol. 20, no. 9, p. 2129, Apr. 2019, doi: 10.3390/ijms20092129.

[53] J. M. M. Oneto and M. Royzen, "TCO conjugates and methods for delivery of therapeutic agents," US10806807B2, Oct. 20, 2020 Accessed: Aug. 29, 2024. [Online]. Available: https://patents.google.com/patent/US10806807B2/en?q=(university+of+California)&invento

r=Jose+Oneto&oq=Jose+Oneto+university+of+California

[54] M. Royzen, G. P. A. Yap, and J. M. Fox, "A Photochemical Synthesis of Functionalized trans-Cyclooctenes Driven by Metal Complexation," *J. Am. Chem. Soc.*, vol. 130, no. 12, pp. 3760–3761, Mar. 2008, doi: 10.1021/ja8001919.

[55] J. Li, S. Jia, and P. R. Chen, "Diels-Alder reaction–triggered bioorthogonal protein decaging in living cells," *Nat. Chem. Biol.*, vol. 10, no. 12, Art. no. 12, Dec. 2014, doi: 10.1038/nchembio.1656.

[56] A. Arsić, C. Hagemann, N. Stajković, T. Schubert, and I. Nikić-Spiegel, "Minimal genetically encoded tags for fluorescent protein labeling in living neurons," *Nat. Commun.*, vol. 13, p. 314, Jan. 2022, doi: 10.1038/s41467-022-27956-y.

[57] T. Yanagisawa, M. Kuratani, E. Seki, N. Hino, K. Sakamoto, and S. Yokoyama, "Structural Basis for Genetic-Code Expansion with Bulky Lysine Derivatives by an Engineered Pyrrolysyl-tRNA Synthetase," *Cell Chem. Biol.*, vol. 26, no. 7, pp. 936-949.e13, Jul. 2019, doi: 10.1016/j.chembiol.2019.03.008.

[58] C. McQuin *et al.*, "CellProfiler 3.0: Next-generation image processing for biology," *PLOS Biol.*, vol. 16, no. 7, p. e2005970, Jul. 2018, doi: 10.1371/journal.pbio.2005970.

Supplementary Information for

Functional Assembly of Split Protein Pairs via a Chemically Activated SpyLigation

Table of Contents

General Synthetic Information

Chemical reagents and solvents were procured from either Sigma-Aldrich or Fisher Scientific and utilized as received unless specified otherwise. Deionized water (dH2O) was produced using a U.S. Filter Corporation Reverse Osmosis System with a Desal membrane. Synthetic chemical reactions were conducted under a nitrogen atmosphere in glassware dried in an oven and stirred with a Teflon-coated magnetic stir bar unless otherwise stated. Solvent evaporation was carried out using a Büchi Rotovapor R-3 equipped with a V-700 vacuum pump and V-855 vacuum controller, along with a Welch 1400 DuoSeal Belt-Drive high vacuum pump. 1H nuclear magnetic resonance (NMR) data were collected at 298 K using Bruker instruments, with chemical shifts reported relative to tetramethylsilane (TMS, $\delta = 0$). Whole-protein mass spectrometry was performed on an AB Sciex 5600 QTOF instrument. Polymerase chain reaction (PCR) was conducted using an Applied Biosystems SimpliAmp. *E. coli* cultures were cultivated in a Thermo Scientific MaxQ 4000 shaker incubator. Cell lysis was achieved using a Fisher Scientific Model 505 Soni Dismembrator equipped with a 1.27 cm diameter probe. Nucleic acid ($\lambda = 260$ nm) and protein ($\lambda = 280$ nm) concentrations and UV-Vis absorption spectrometry were determined using a NanoDrop (Thermo Scientific). Protein fluorescence was measured using a BioTek Synergy H1M plate reader with Thermo Scientific Nunc black polypropylene 384-well plates. Protein gel electrophoresis was conducted using a Mini-PROTEAN tetra gel box equipped with a PowerPac basic power supply (BioRad). Fluorescence imaging was performed using a Leica Stellaris 5 confocal microscope equipped with a white light laser, live imaging chamber, and 10x CS2 APO dry objective. Fluorescent and true-color gel imaging was carried out using an Azure 600 AZI600 scanner. Mammalian cell culture was conducted in a NuAire LabGard ES NU-437 Class II Type A2 Biosafety Cabinet. Cells were maintained in a Sanyo

inCu saFe® MCO-17AC incubator at 37 °C and 5% CO2. Purchased plasmids were sourced from Addgene.

Preparation of Silver Nitrate Silica

Silver nitrate silica was prepared according to an established protocol. [54] Briefly, silver nitrate (10 g, 59 mmol) was dissolved in 100 mL of deionized water and added to a slurry of silica (90 g dry weight) in water. The slurry was mixed thoroughly, then the water was removed by rotary evaporation. The dry silica was washed twice with tetrahydrofuran which was also removed by rotary evaporation to yield free-flowing silica loaded with 10% w/w silver nitrate (0.59 mmol/g).

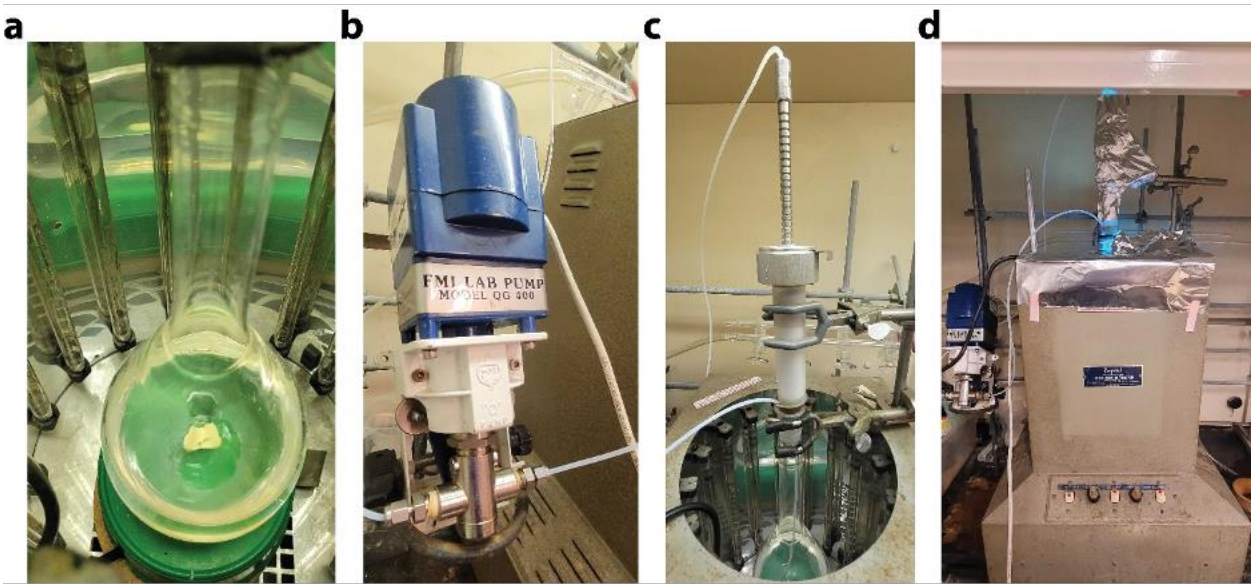


Figure 1. UV reactor design for cyclooctene isomerization. (a) A quartz Kjeldahl flask containing the reactants and solvent is stirred inside of the UV chamber. (b) FMI lab pump connected to Teflon tubing. (c) The silica column with the adapter is clamped above the reaction vessel. Compression against a rubber conical gasket is a sufficient seal. (d) The reactor is covered in foil for safety, and the column is covered to prevent unwanted reactions. A vent is left for airflow.

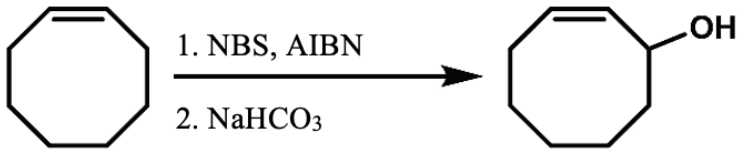
UV Reactor Configuration

For UV isomerization of cyclooctene, a quartz flask with stir bar was placed at the bottom of a Rayonet 254 nm UV reactor containing mercury arc lamps. A Teledyne RediSep® cartridge was packed with ~1 cm of dry silica followed by 40 g of silver nitrate silica (23.6 mmol $AgNO_3$), then sealed with an adapter salvaged from a CombiFlash Companion. Teflon tubing was installed from the reaction vessel through an FMI lab pump and into the top of the column adapter. The column was placed above the reaction vessel and connected by compressing against a conical rubber gasket. The pump flow was set to approximately 20 mL/min, and the silica was equilibrated with solvent before adding reactant. During the reaction, the lights and cooling fan remained on. The column and reactor top were shielded using aluminum foil. Some heat buildup occurred, but this did not seem to hinder the reaction as long as solvent evaporation was prevented.

Synthesis of TCOK (5)

1, **2**, and **5** were synthesized as reported previously. [55]

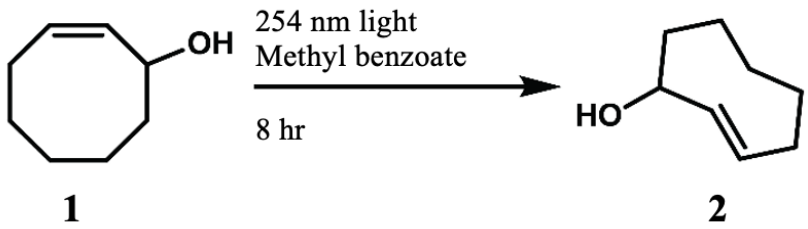
Synthesis of (Z)-Cyclooct-2-en-1-ol (1)



1

cis--Cyclooctene (30 mL, 0.23 mol), *N*-bromosuccinimide (30 g, 0.17 mol), and 2,2'-azobis(2-methylpropionitrile) (50 mg) were added to a flame-dried round bottom flask and dissolved in 120 mL of carbon tetrachloride. The flask was purged with N_2 gas, then the contents were refluxed for 2 hours. The reaction was cooled on ice, then the precipitate was removed by vacuum filtration. Solvent was removed *in vacuo* to give the brominated intermediate which was used without further purification. Brominated product was dissolved in 240 mL of 2:1 v/v acetone:water, then sodium bicarbonate (30 g, 0.36 mol) was added. The mixture was refluxed for 1 hour, then filtered. The filtrate was extracted with diethyl ether, then solvent was removed *in* vacuo to yield two liquid phases. The liquid products were cooled at -20 °C overnight to crystallize the undesired product, and **1** (19 g, 66%) was decanted as a clear, light-yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 1.32-1.70 (m, 7 H), 1.72-1.81 (s, 1 H), 1.85-1.97 (m, 1 H), 2.02-2.23 (m, 2 H), 4.59-4.70 (m, 1 H), 5.48-5.66 (m, 2 H).

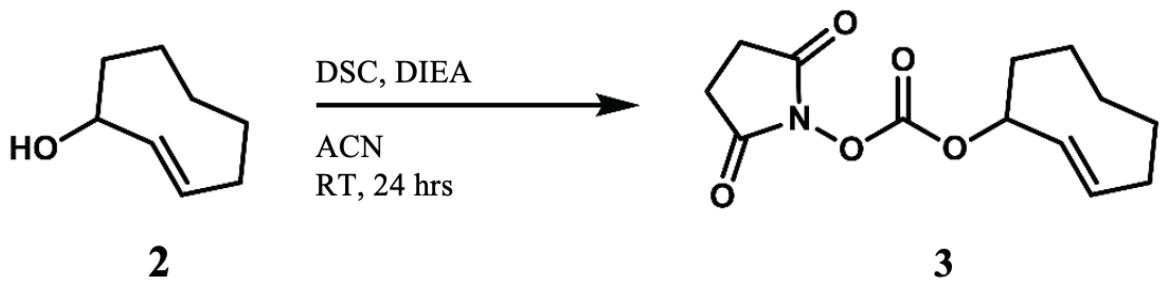
*Synthesis of ax-(E)-Cyclooct-2-en-1-ol (**2**)*



1 **2**

1 (3.0 g, 23.8 mmol) and methyl benzoate (4.0 g, 29.4 mmol) were dissolved in 450 mL of 2:1 petroleum ether:diethyl ether and added to a quartz Kjeldahl flask. The reaction was irradiated by 254 nm light while stirring and pumped continuously through a foil-wrapped column containing 40 g of silver nitrate silica for 16 hours. The column was removed and flushed with 250 mL of dichloromethane, then dried under air. The silica was transferred to a beaker containing 400 mL of 1:1 dichloromethane:ammonium hydroxide and stirred vigorously for 30 minutes. The organic layer was reserved and aqueous layer containing solids extracted twice with dichloromethane. Solvent was removed from the combined organics using a rotary evaporator, and the product isomers were separated by silica gel chromatography using a 16-37% gradient of ethyl acetate in hexanes to yield **2** (0.77 g, 26%) in the first fraction as a clear, colorless oil. R_f = 0.78, 1:1 EtOAc:hexanes. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 0.72-0.82 (m, 1 H), 1.06-1.17 (m, 1 H), 1.43-1.55 (m, 1 H), 1.55-1.64 (m, 2 H), 1.64-1.73 (m, 1 H), 1.80-1.90 (m, 1 H), 1.92-2.09 (m, 3 H), 2.43-2.53 (m, 1 H), 4.59-4.64 (br. s, 1 H), 5.53-5.62 (dd, 1 H, J = 2.2, 16.5 Hz), 5.91-6.00 (ddd, 1 H, J = 16.1, 4.0, 11.2 Hz). Equatorial isomer (1.30 g, 43%) was also collected in the second fraction, R_f = 0.64, 1:1 EtOAc:hexanes.

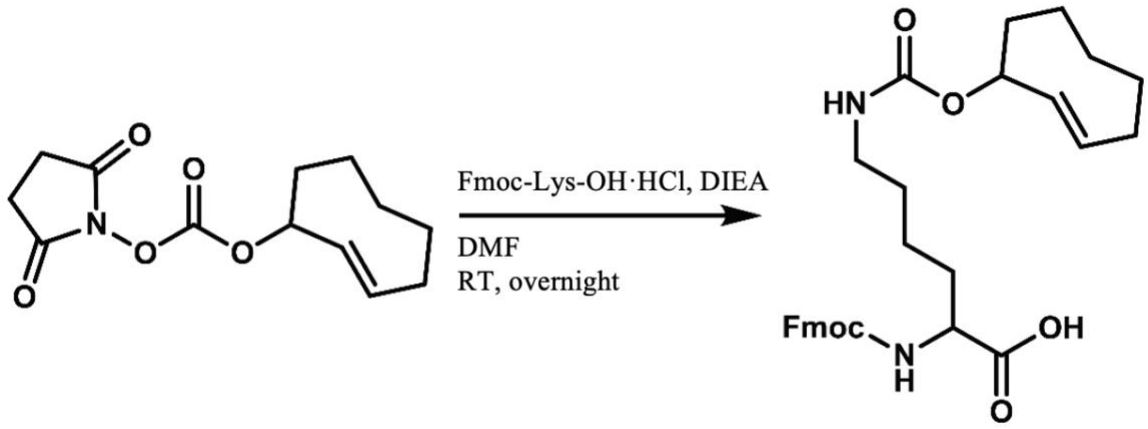
*Synthesis of ax-(2E)-2-Cycloocten-1-yl 2,5-dioxo-1-pyrrolidinyl carbonate (**3**)*



2 **3**

2 (0.75 g, 6.0 mmol),*N,N'*-disuccinimidyl carbonate (3.07 g, 12.0 mmol), and DIEA (2.09 mL, 12.0 mmol) were dissolved in 20 mL of dry acetonitrile in a flame-dried round bottom flask. The reaction continued overnight at room temperature under nitrogen, and solvent was removed by rotary evaporation. The resulting residue was dissolved in dichloromethane and washed twice with water and once with brine. Organics were dried over magnesium sulfate and filtered, then solvent was removed by rotary evaporation. The crude product was purified by silica chromatography using a 44-65% gradient of ethyl acetate in hexanes to yield **3** (0.83 g, 52%) as a white powder. R_f = 0.45 , 1:1 EtOAc:hexanes. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 0.73-0.93 (m, 1 H), 1.05-1.23 (m, 1 H), 1.43-2.12 (m, 6 H), 2.14-2.29 (dd, 1 H, J = 3.9, 15.3 Hz), 2.43-2.59 (m, 1 H), 2.78-2.87 (s, 4 H), 5.35-5.43 (br. s, 1 H), 5.43-5.55 (dd, 1 H, J = 2.0, 16.5 Hz), 5.88-6.05 (ddd, 1 H, J = 16.2, 3.9, 11.2 Hz). MS (ESI-MS): calculated for $C_{13}H_{17}NO_5$ ($[M+H]^+$): 268.119; found: 268.120.

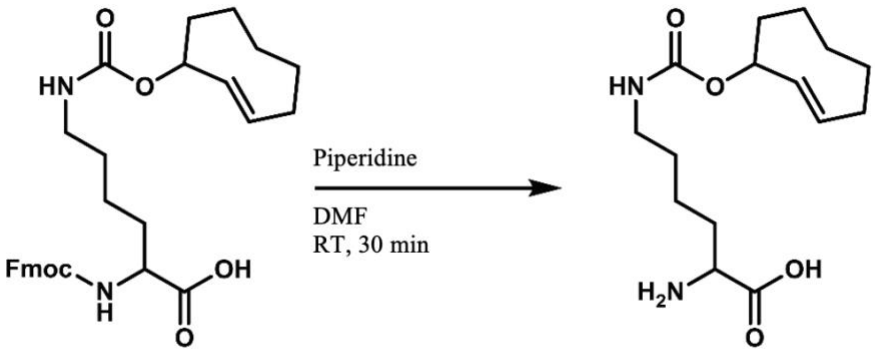
*Synthesis of ax-Nα-[(9H-Fluoren-9-ylmethoxy)carbonyl]-Nε-[({E}-cyclooct-2-enyl)oxycarbonyl]-L-lysine (**4**)*



3 **4**

3 (0.83 g, 3.1 mmol), Fmoc-Lys-OH·HCl (1.50 g, 3.7 mmol), and DIEA (0.80 g, 6.2 mmol) were dissolved in 15 mL of dry DMF in a flame-dried round bottom flask. The reaction continued overnight at room temperature. Solvent was removed by rotary evaporation and the resulting residue was purified by silica chromatography using a gradient of 0-5% methanol in dichloromethane to yield **4** (0.66 g, 41%) as a colorless oil. R_f = 0.49, 95:5 dichloromethane:methanol. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 0.70-0.85 (m, 1 H), 0.97-1.11 (m, 1 H), 1.35-2.14 (m, 14 H), 2.37-2.51 (m, 1 H), 3.10-3.22 (m, 2 H), 4.16-4.25 (t, 1 H, J = 6.7 Hz), 4.31-4.43 (m, 2 H), 5.27-5.34 (m, 1 H), 5.45-5.56 (m, 1 H), 5.63-5.70 (d, 1 H, J = 6.9 Hz), 5.74-5.84 (m, 1 H), 7.25-7.33 (t, 2 H, J = 7.4 Hz), 7.34-7.41 (t, 2 H, J = 7.4 Hz), 7.56-7.66 (m, 2 H), 7.71-7.77 (d, 2 H, J = 7.5 Hz), 7.98-8.04 (s, 2 H). MS (ESI-MS): calculated for $C_{30}H_{38}N_2O_6$ ($[M+H]^+$): 521.265; found: 521.266.

*Synthesis of ax-Nε-[({E}-Cyclooct-2-enyl)oxycarbonyl]-L-lysine (**TCOK**)*



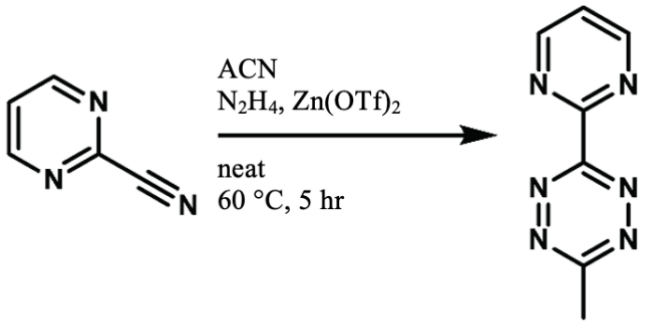
4 **TCOK**

-**4** (0.66 g, 1.3 mmol) was dissolved in 20 mL of 25% piperidine in dichloromethane and stirred at room temperature for 1 hour. Solvent was removed by rotary evaporation and the resulting residue dissolved in methanol, then precipitated in cold diethyl ether and filtered to yield **TCOK** (0.30 g, 77%) as a white powder. ^{1}H NMR (500 MHz, CD_3OD): δ = 0.80-0.94 (m, 1 H), 1.07-1.23 (m, 1 H), 1.38-2.10 (m, 12 H), 2.39-2.51 (m, 1 H), 3.12 (m, 2 H), 3.49-3.58 (m, 1 H), 5.23 (s, 1 H), 5.50-5.62 (dd, 1 H, J = 16.3, 1.9 Hz), 5.76-5.91 (ddd, 1 H, J = 15.8, 11.5, 3.9 Hz). MS (ESI-MS): calculated for $C_{15}H_{27}N_2O_4$ ($[M+H]^+$): 299.197; found: 299.197.

Synthesis Of Tetrazine Derivatives

Tetrazines were synthesized as previously reported. [40]

*Synthesis of 3-Methyl-6-(2-pyrimidinyl)-1,2,4,5-tetrazine (**6**)*

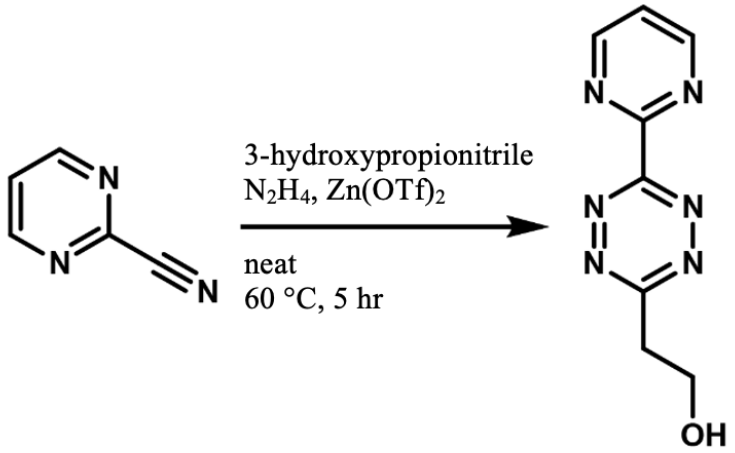


6

2-Cyanopyrimidine (220 mg, 2.1 mmol), anhydrous acetonitrile (522 µL, 10.0 mmol), zinc trifluoromethanesulfonate (183 mg, 0.5 mmol), and hydrazine monohydrate (500 mg, 10.0 mmol) were added to a nitrogen purged scintillation vial and heated to 60 °C. The reaction was allowed to continue for 24 hours then cooled to room temperature and diluted with 20 mL of 1 M sodium nitrite solution. The pH was slowly adjusted to 3.0 using 6 M HCl, and the mixture was stirred for an additional 30 minutes to allow NO_x gases to disperse. The aqueous solution was extracted with dichloromethane until all magenta color was removed from the aqueous phase. The combined organics were washed with brine and dried over anhydrous magnesium sulfate, then the solvent was removed *in vacuo*. The crude residue was purified by silica gel chromatography using a 0-8% gradient of methanol in dichloromethane to yield 6 (50 mg, 14%) as a magenta powder. R_f = 0.35, 95:5 dichloromethane:methanol. ^{1}H NMR (500 MHz, $CDCl_3$): δ

= 3.20 (s, 3 H), 7.56-7.63 (t, 1 H, J = 4.9 Hz), 9.09-9.16 (d, 2 H, J = 4.9 Hz). MS (ESI-MS): calculated for $C_7H_7N_6$ ($[M+H]^+$): 175.073; found: 175.072.

Synthesis of 6-(2-Pyrimidinyl)-1,2,4,5-tetrazine-3-ethanol (7)



7

2-Cyanopyrimidine (219 mg, 2.1 mmol), 3-hydroxypropionitrile (709 mg, 10.0 mmol), zinc trifluoromethanesulfonate (183 mg, 0.5 mmol), and hydrazine monohydrate (500 mg, 10.0 mmol) were added to a nitrogen purged scintillation vial and heated to 60 °C. The reaction was allowed to continue for 5 hours then cooled to room temperature and diluted with 20 mL of 1 M sodium nitrite solution. The pH was slowly adjusted to 3.0 using 6 M HCl, and the mixture was stirred for an additional 30 minutes to allow NO_x gases to disperse. The aqueous solution was extracted with dichloromethane until all magenta color was removed. The combined organics were washed with brine and dried over anhydrous sodium sulfate, then the solvent was removed *in vacuo*. The crude residue was purified using silica gel chromatography in 95:5 dichloromethane:methanol to yield **7** (47 mg, 11%) as a magenta powder. R_f = 0.25, 95:5 dichloromethane:methanol. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 3.74-3.78 (t, 2 H, J = 5.8 Hz), 4.34-

4.38 (t, 2 H, J = 5.8 Hz), 7.59-7.62 (t, 1 H, J = 5.0 Hz), 9.12-9.15 (d, 2 H, J = 4.7 Hz). MS (ESI-MS): calculated for $C_8H_9N_6O$ ($[M+H]^+$): 205.084; found: 205.084.

Construction of aaRS/tRNA Plasmids for TCOK Incorporation

Table 1

Name	Purpose	Source
pcDNA3.1(+)_U6 tRNAPyl_CMV NESPylRS(AF)	tRNA synthetase for TCOK incorporation in mammalian proteins	https://www.addgene.org/182652/, [4]
pCDF-Pyl-AFx2	tRNA synthetase for TCOK incorporation in bacterial proteins	[5]
pEVOL-NBK-1	tRNA synthetase for incorporation of bulky lysine derivatives in bacterial proteins	[6]

Chemical Uncaging of TCOK in Solution

TCO-Lys (0.3mM in phosphate-buffered saline, PBS, pH = 7.4) was subjected to 5 mM tetrazine for varying durations (0 – 10 min). The uncaging process was monitored using UV-Vis absorption spectrometry. The inverse electron-demand DielsAlder reactions between TCO-a and Tetrazine lead to a shift in absorbance, as indicated by the arrow. Through measurements at 525 nm, a second-order degradation constant of 28.1 $M^{-1}s^{-1}$ was determined.

Protein Expression and Purification

General protein expression and purification

Transform the desired plasmid into BL21(DE3) E. coli (Promega). Culture of Miller's lysogeny broth (LB) (0.25-0.5 L), supplemented with carbenicillin (100 μg mL^{-1}) or kanamycin (50 μg mL^{-1}), was inoculated with overnight cell culture (10 mL) and incubated at 37°C with agitation (250 rev min^{-1}). Once reaching an optical density at λ = 600 nm of 0.6, isopropyl β-D-1-thiogalactopyranoside (final concentration of 0.5 mM) was added, and expression was continued overnight at a reduced temperature (18 °C). Cells were harvested via centrifugation (4,000 g, 10 min, 4 °C). The cell pellets were suspended in lysis buffer (40 mL, 20 mM Tris, 50 mM NaCl, 10 mM imidazole) and sonicated on ice (6 cycles, each comprising 3 minutes at 30% amplitude, 33% duty cycle, followed by 3 minutes of rest). After centrifugation (5,000 x g, 20 minutes) to clarify the cell lysate, it was then applied to Ni-NTA resin (5 mL) and gently agitated at 4°C for at least 1 hour. After the removal of the flow-through, the resin was washed (6 mL, 20 mM Tris, 50 mM NaCl, 20 mM imidazole, pH 7.5) until minimal protein was observed in the flow-through (λ_{abs} = 280 nm). The protein was subsequently eluted with elution buffer (20 mM Tris, 50 mM NaCl, 250 mM imidazole, pH 7.5). The protein solution was dialyzed at 4°C for 24 hours against Tris buffer (20 mM Tris, 50 mM NaCl) using ThermoFisher SnakeSkin Dialysis Tubing with a molecular weight cut-off (MWCO 10 or 30 kDa) to remove any residual imidazole.

cSC003 fusion protein expression and purification

Expression and purification of cSC003 followed the previously outlined protocol with the following modifications. B95 Escherichia coli (from Promega) were co-transformed with equimolar amounts of the expression vector and the tRNA synthetase plasmid for incorporation of ncAA (pCDF-Pyl-AFx2). Cells were cultured overnight at 37°C with agitation at 250 rev min-

[1] in Terrific Broth (TB) supplemented with carbenicillin (100 μg mL^{-1}) and kanamycin (50 μg mL^{-1}). Large-scale cultures (125 mL, TB) were inoculated at a 1:25 ratio and grown until reaching an optical density of 1 (λ = 600 nm). TCOK dissolved in DMSO was added to a final concentration of 3 mM. Subsequently, cultures were induced by the addition of arabinose (final concentration of 0.2%) and IPTG (final concentration of 0.5 mM). Cultures were agitated overnight at 30°C, then pelleted by centrifugation (4,000 g, 10 min, 4 °C). Protein purification and dialysis were conducted as previously described. The expression of cSC003 yielded approximately 5 – 10 mg of pure protein per liter of culture. cSC003 was stored at -20°C in Tris buffer with 8% glycerol until needed

GST-ST003 purification

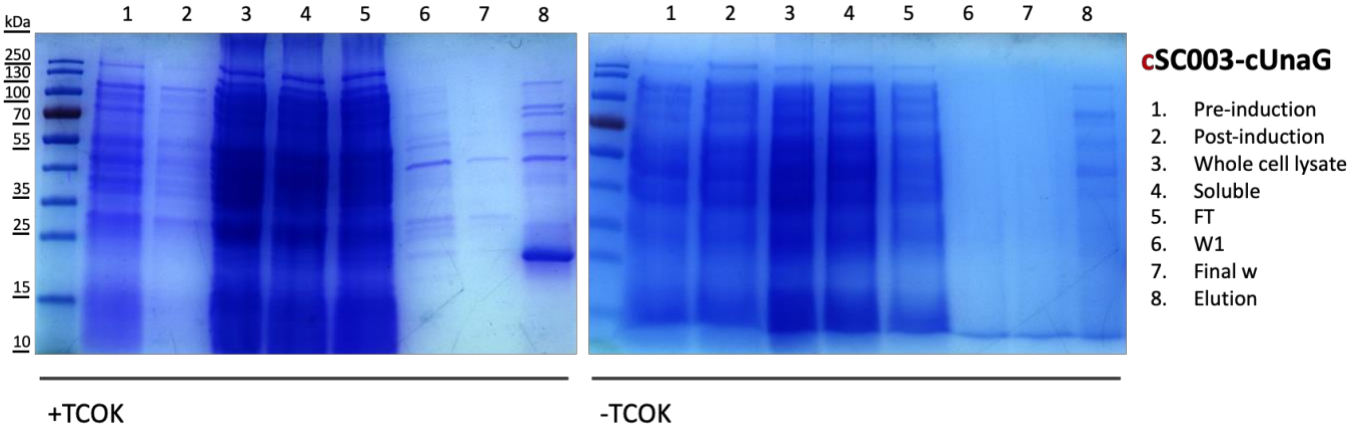
GST-ST003 was expressed in 500 mL LB supplemented with carbenicillin at 100 μg mL^{-1} as mentioned before with the following modifications. Pelleted cells were resuspended in GST Equilibration Buffer (40 mL; 50 mM Tris, 150 mM NaCl, pH 8.0). Cells were lysed via sonication and were cleared by centrifuge. Prewash the Pierce Glutathione Agarose resin (1 mL slurry, Thermo Scientific) with GST Equilibration Buffer (5 mL, 2x) before the cell lysate soluble fraction was added. The mixture of protein and resin was gently agitated at 4°C for one hour to ensure proper binding. Subsequently, the mixture was centrifuged (500 x g, 2 min), and the supernatant was discarded. The resin was then washed in fractions with 5 mL (GST Equilibration Buffer) followed by centrifugation (500 x g, 2 min) and discarded until the minimal protein was observed. GST- ST003 was obtained by washing the resin with elution buffer (2 mL, 50 mM Tris, 150 mM NaCl, 10 mM reduced glutathione, pH 8.0) and pellet the resin with centrifuging (500 x g, 2 min). The purified protein was decanted and applied to a Zeba Spin Desalting Column (7k MWCO, 10 mL, Thermo Scientific) equilibrated with GST

Equilibration Buffer to remove reduced glutathione. GST-ST expression yielded approximately 100 mg L^{-1} of culture and was stored at -20°C in GST Equilibration Buffer with 8% glycerol until needed.

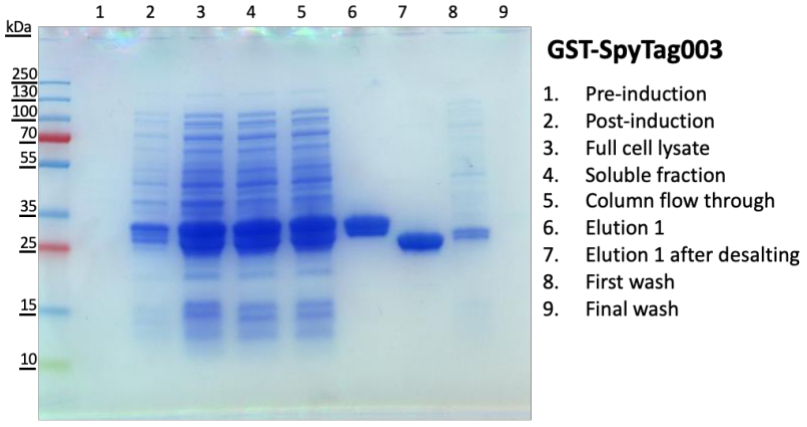
Assessing Purified Protein by SDS-Page Analysis

cSC003-cUnaG

Samples were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis with Coomassie Brilliant Blue visualization to examine the effectiveness of protein expression. Molecular weight markers are in 20kDa. The lack of the expected protein band in cSC expression without TCOK served as evidence of the essential role of TCOK in effective amber suppression. This proves the orthogonality of mutant tRNA/tRNA synthetase pair for proper readout and incorporation.



GST-ST003



Here, the GST-ST003 fusion shows a post-induction band of about ~29kDa. Fusion proteins were captured with high purity after being purified using the Pierce Glutathione Agarose resin.

Mass Spectrometric Analysis of Fusion Protein

Liquid chromatography-mass spectrometry (LC-MS) is used for protein identification. Each protein species (1 μg in buffer containing 20 mM Tris, 50 mM NaCl, pH 7.4) was injected into an LC-MS (AB Sciex 5600 QTOF) using an inline polymeric reversed-phase column (PL 1912-1503, Agilent). Protein solutions were separated over an 8-minute linear gradient from 5-95% acetonitrile in water with 0.1% formic acid at 0.5 mL min^{-1}. Mass spectrum scans were performed once every second in positive mode. Using Analyst (AB Sciex), the chromatogram was integrated, and the total molecular weight was determined. Table 2 lists the observed and predicted masses of every fusion protein.

Table 2

Protein Identity	Expected mass (Da)	Observation mass (Da)
GST-SpyTag003	29500	30416
nUnaG-SpyTag003	53907	41698
SpyCatcher003-cUnaG	20006	20006
cSpyCatcher003-cUnaG	20158	20158

Assessing CASL Does Dependency of Tetrazine with Premixed cSC And ST

Purified protein nUanG-ST003 (35 mM) and cSC003-cUnaG were mixed before adding varied amounts of tetrazine (final concentration: 0-40 μM) and reacted 30 minutes at 37 °C before analysis. Samples were diluted 3:1 with 4X Laemmli sample buffer and boiled for 10 min before SDS-PAGE analysis (8-16% TGX precast mini-gel, BioRad). The nUanG-ST003 band is identified with 1 in gray, cSC003-cUnaG with 2 in blue, and the covalent ligated product with 3 in orange. Experimental replicates (n = 3) were quantified by determining the band intensity of the two reactants (cSC003-cUnaG and nUanG-ST003) and the CASL product at each time point in FIJI.

Dose-Dependent Reconstitution of Split-UnaG via CASL

To assess fluorescence reconstitution relative to the concentration of tetrazine in CASL, nUnaG-ST003, and cSC003-cUnaG were mixed with varied amounts of tetrazine (final concentration: 0-10 μM). Protein pairs were combined at equimolar quantities (7 μM) in Tris Buffer containing bilirubin (10 μM). Fluorescence was detected every 15 min for 8 hours after the reaction started using a fluorescent plate reader ($\lambda_{excitation}$ = 495 nm, $\lambda_{emission}$ = 528 nm, at 37°C).

Polycistronic CASL Construct Designed to Visualize the Subcellular Localization of Proteins

The polycistronic cSC003-mCherry-P2A-ST003-eGFP-CAAX were made and passed along from the lab's former project. [6] The plasmid was then transformed into chemically competent Top10 E. coli. Overnight cultures (250 mL Miller's LB with 100 μg mL^{-1} carbenicillin) were harvested with centrifuging (4000 x g for 10 min.) Plasmid DNA was obtained using the ZymoPURE™ Plasmid Maxiprep kit (Zymo Research). Plasmid DNA was eluted using cell culture grade dH2O (Corning) and utilized in mammalian cell transfections.

Cell Culture and Transfection Condition and Visualize Membrane by CASL

General methods

HEK-293T cells were cultured in DMEM supplemented with 10% FBS and 1% P/S at 37°C with 5% CO2. Before transfection, cells were seeded at a density of 50,000 cells per square centimeter onto 35 mm Glass bottom dish with 14 mm micro-well coated with 0.1% gelatin, and incubated for 24 hours. Transfection was completed with complete DMEM containing 2.5 mM TCOK. Co-transfection was performed using Lipofectamine P3000 according to the manufacturer's protocol, with equal amounts (0.5 μg) of either pNeu-hMbPylRS-4xU6M15 or pcDNA3.1(+)_4x(U6 tRNA M15)_CMV NESPylRS(AF) synthetase plasmid and the polycistronic cSC003-mCherry-P2A-ST003-eGFP-CAAX expression vector.

Tetrazine treatment

After 24 hours of transfection, the medium was replaced with DMEM containing 100 μM Tetrazine and incubated overnight before imaging. Prior to imaging, Cells were transferred to HBSS and imaged with Stellaris 5 confocal microscope (Leica) with 20x objective or 40x objective Immersed in oil ($\lambda_{EGFP,excitation}$ = 489 nm; $\lambda_{mCherry,excitation}$ = 587 nm).

CellProfiler analysis

The IdentifyPrimaryObjects module initially detected cells using mCh or EGFP fluorescence. The MeasureObjectIntensityDistribution module was used to measure in-cell fluorescence, with 18 internal bins per cell. [7]

CASL-Mediated Intracellular Protein Localization with Membrane-Bound cSC

CASL was employed to precisely direct protein localization to the membrane of mammalian cells for membrane labeling. Co-transfected HEK-293T cells were treated with tetrazine overnight 24 hours post-transfection and photographed with 20x or 40x objective on Stellaris 5 confocal microscope (Leica). With the gene cassette utilized, the membrane-anchored ST003-eGFP-CAAX is conveniently linked to the cytosolic cSC003-mCh after tetrazine treatment. Intracellular cSC003-mCh distribution shifts from uniform cytosol to more membrane localization via CASL. Membrane labeling with mCh increases statistically significantly with tetrazine exposure.

Polycistronic Split-UnaG CASL Construct Designed for Intracellular Activation

For effective expression in mammalian cells, the cSC003-cUnaG-P2A-nUnaG-ST003-P2A-mCherry was cloned into the pTwist CMV vector with the puromycin resistance gene (Twist Bioscience).[6] The plasmid was then transformed into chemically competent Top10 E. coli. Overnight cultures (250ml Miller's LB with 100 μg mL^{-1} carbenicillin) were harvested with centrifuging (4000 x g for 10 min.) Plasmid DNA was obtained using the ZymoPURE™ Plasmid Maxiprep kit (Zymo Research). Plasmid DNA was eluted using cell culture grade dH2O (Corning) and utilized in mammalian cell transfections.

Cell Culture and Transfection Condition for Intracellular UnaG Activation

General methods

HEK-293T cells were cultured in DMEM supplemented with 10% FBS and 1% P/S at 37°C with 5% CO2. Before transfection, cells were seeded at a density of 50,000 cells per square centimeter onto 35 mm Glass bottom dish with 14 mm micro-well coated with 0.1% gelatin, and incubated for 24 hours. Transfection was completed with complete DMEM containing 2.5 mM TCOK. Co-transfection was performed using Lipofectamine P3000 according to the manufacturer's protocol, with equal amounts (0.5 μg) of either pNeu-hMbPylRS-4xU6M15 or pcDNA3.1(+)_4x(U6 tRNA M15)_CMV NESPylRS(AF) synthetase plasmid and the polycistronic cSC003-cUnaG-P2A-nUnaG-ST003-P2A-mCherry expression vector.

Tetrazine treatment

After 24 hours of transfection, the medium was replaced with DMEM containing 100 μM Tetrazine and incubated overnight before imaging. Prior to imaging, Cells were transferred to HBSS and imaged on Stellaris 5 confocal microscope (Leica) with 20x objective or 40x objective Immersed in oil ($\lambda_{EGFP,excitation}$ = 489 nm; $\lambda_{mCherry,excitation}$ = 587 nm).

CellProfiler analysis

CellProfiler10 was used to quantify the fluorescence and location of individual cells.
The IdentifyPrimaryObjects module initially detected cells using mCh or EGFP fluorescence.
The MeasureObjectIntensityDistribution module was used to measure in-cell fluorescence. [7]

Analyses of the UnaG/mCh Signal in Chemically (Un)exposed Cells

UnaG and mCh fluorescence levels were measured within each cell and quantified using automatic image analysis. Cells exposed to tetrazine exhibited statistically significant increases in UnaG/mCh fluorescence compared to cells that were not subjected to tetrazine treatment. Maximum UnaG/mCh fluorescence ratios (8-fold greater than untreated samples) remained for 48h, indicating covalent UnaG assembly. Due to the expected protein clearance and/or cell division, the mCh signal declined slightly with time. Asterisks denote conditions with statistically significant differences in signal ($p < 0.0001$, two-tailed unpaired t-tests).

DNA Sequences for All CASL Constructs

The open reading frames from 5' to 3' for the desired proteins in this study have been verified by Sanger sequencing. The following sequences are colored to represent the genes and characteristics in each construct: SC003 in cyan, cSC003 in yellow, the Lys29TAG mutation in red, SpyTag003 in orange, UnaG fragments in green, mCherry in dark red and the CAAX sequence in blue. The primers used to clone the plasmids are listed in Table 3.

GST-ST003

ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT
CTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGT
GATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTAT
TATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCT
GACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCT
TGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGA
CTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTT
CGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGA
CTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGAT
GCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGAT
AAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCAC
GTTTGGTGGTGGCGACCATCCTCCAAAATCGGATCTGGTTCCGCGTGGATCCGGTAG
CGGTGAAAGCGGTAGCCGCGGCGTGCCGCATATTGTGATGGTGGATGCGTATAAAC
GCTATAAA

nUnaG-ST003-MBP-6xHis

ATGGTTGAAAAGTTTGTAGGAACATGGAAAATCGCAGACAGCCATAATTTTGGTGA

GTATCTAAAAGCGATTGGTGCGCCAAAGGAGCTGTCGGACGGTGGTGATGCTACGA

CCCCGACGCTGTACATCTCCCAGAAGGACGGCGACAAGATGACCGTTAAGATTGAA

AACGGCCCACCGACCTTTCTGGATACCCAGGTGAAGTTTAAACTGGGCGAAGAGTTT

GATGAGTTTCCGAGCGACCGCCGTAAAGGCGGCTCTGGTGGCGGTAGTGGTGGTTCC

CGCGGCGTGCCGCACATCGTGATGGTGGACGCCTACAAAAGATATAAAAAACTTAT

GAAAATCGAAGAAGGTAAGTTAGTTATTTGGATCAACGGCGACAAGGGCTACAACG

GTCTGGCTGAAGTTGGCAAGAAGTTCGAAAAAGATACCGGCATTAAAGTGACCGTT

GAACATCCGGACAAGCTCGAAGAGAAATTCCCGCAGGTTGCAGCGACCGGTGATGG

ACCGGATATCATCTTTTGGGCACACGACCGTTTCGGCGGCTACGCGCAATCCGGCCT

GTTGGCGGAGATCACGCCGGACAAAGCCTTCCAGGACAAATTGTACCCGTTCACCT

GGGATGCGGTGCGTTATAACGGAAAACTGATTGCCTACCCGATTGCAGTTGAGGCCT

TATCTCTGATTTACAACAAGGACCTGCTGCCAAATCCGCCTAAAACCTGGGAGGAGA

TCCCGGCGCTGGACAAGGAATTGAAGGCGAAAGGTAAGAGCGCGCTGATGTTCAAT

CTGCAAGAGCCGTATTTCACGTGGCCCCTGATTGCGGCTGATGGCGGTTATGCATTC

AAATATGAAAATGGCAAGTACGATATCAAAGACGTGGGTGTCGATAATGCCGGTGC

GAAGGCTGGCCTGACCTTTTTGGTTGATCTGATCAAAAACAAGCACATGAATGCTGA

CACCGATTACAGCATTGCGGAGGCTGCGTTCAACAAGGGTGAGACCGCTATGACCA

TCAACGGTCCGTGGGCTTGGTCGAATATTGATACTTCAAAAGTGAACTACGGCGTGA

CCGTGCTGCCGACATTCAAAGGTCAGCCGAGCAAACCGTTTGTTGGCGTTCTGAGCG

CCGGTATTAACGCAGCGTCCCCGAACAAAGAGCTCGCAAAGGAGTTTCTGGAGAAC

TATCTGCTGACTGACGAAGGTCTGGAAGCGGTTAACAAGGATAAACCGCTTGGTGC

AGTCGCGCTGAAGTCTTATGAGGAAGAATTGGCCAAGGACCCGCGTATCGCAGCGA

CCATGGAAAATGCACAAAAAGGTGAAATCATGCCGAACATTCCGCAGATGAGCGCG

TTCTGGTACGCAGTACGCACCGCGGTCATCAATGCGGCTAGCGGCCGTCAAACGGT

GGATGAGGCGTTGAAGGACGCTCAAACCAACAGCAGCCTCGAGCACCACCACCACC

ACCAC

SC003-cUnaG-6xHis

ATGGTAACTACATTAAGTGGACTATCAGGGGAGCAGGGTCCGAGCGGTGACATGAC

CACAGAGGAAGACTCTGCAACGCATATCAAGTTTTCCAAACGTGATGAGGACGGCC

GTGAATTGGCTGGTGCGACGATGGAACTGCGTGATTCCAGCGGTAAAACCATCTCG

ACCTGGATTAGCGATGGTCACGTGAAGGACTTCTATCTCTACCCGGGTAAGTACACC

TTTGTTGAAACTGCGGCGCCAGATGGTTACGAAGTTGCGACTCCGATTGAATTCACC

GTCAATGAAGATGGTCAAGTTACCGTCGATGGGGAGGCGACGGAGGGTGACGCACA

CACCGGTGGCAGCGGCGGCGGCAGTGGCGGCGTGAAAAGCGTTGTGAACCTGGTTG

GTGAGAAATTGGTGTATGTGCAGAAATGGGATGGCAAAGAAACCACCTATGTTCGC

GAGATCAAGGACGGCAAGCTGGTCGTGACCCTGACCATGGGTGACGTGGTAGCCGT

TAGAAGCTACCGCCGTGCTACCGAGCTCGAGCACCACCACCACCACCACTGA

cSC003-cUnaG-6xHis

ATGGTAACTACATTAAGTGGACTATCAGGGGAGCAGGGTCCGAGCGGTGACATGAC

CACAGAGGAAGACTCTGCAACGCATATCAAGTTTTCCTAGCGTGATGAGGACGGCC

GTGAATTGGCTGGTGCGACGATGGAACTGCGTGATTCCAGCGGTAAAACCATCTCG

ACCTGGATTAGCGATGGTCACGTGAAGGACTTCTATCTCTACCCGGGTAAGTACACC

TTTGTTGAAACTGCGGCGCCAGATGGTTACGAAGTTGCGACTCCGATTGAATTCACC

GTCAATGAAGATGGTCAAGTTACCGTCGATGGGGAGGCGACGGAGGGTGACGCACA

CACCGGTGGCAGCGGCGGCGGCAGTGGCGGCGTGAAAAGCGTTGTGAACCTGGTTG

GTGAGAAATTGGTGTATGTGCAGAAATGGGATGGCAAAGAAACCACCTATGTTCGC

GAGATCAAGGACGGCAAGCTGGTCGTGACCCTGACCATGGGTGACGTGGTAGCCGT

TAGAAGCTACCGCCGTGCTACCGAGCTCGAGCACCACCACCACCACCACTGA

cSC003-mCh-P2A-ST003-EGFP-CAAX

ATGGTAACTACATTAAGTGGACTATCAGGGGAGCAGGGTCCGAGCGGTGACATGAC

CACAGAGGAAGACTCTGCAACGCATATCAAGTTTTCCTAGCGTGATGAGGACGGCC

GTGAATTGGCTGGTGCGACGATGGAACTGCGTGATTCCAGCGGTAAAACCATCTCG

ACCTGGATTAGCGATGGTCACGTGAAGGACTTCTATCTCTACCCGGGTAAGTACACC

TTTGTTGAAACTGCGGCGCCAGATGGTTACGAAGTTGCGACTCCGATTGAATTCACC

GTCAATGAAGATGGTCAAGTTACCGTCGATGGGGAGGCGACGGAGGGTGACGCACA

CACCGGGGGAAGTGGAGGGAGTGGGGGGAGCGGGGTGAGCAAGGGCGAGGAGGAT

AACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGT

GAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGC

ACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGA

CATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGA

CATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAT

GAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTCCAGGACG

GCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCG

TAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAG

GACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCC

ACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCC

GGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACAC

CATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACG

AGCTGTACAAGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGAC

GTGGAGGAGAACCCTGGACCTCGCGGCGTGCCGCACATCGTGATGGTGGACGCCTA

CAAAAGATATAAAGGGGGAAGTGGAGGGAGTGGGGGGAGCGGGGTGAGCAAGGGC

GAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA

CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGC

TGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCG

TGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGC

AGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT

TCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGAC

ACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT

CCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCG

ACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGAC

GGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCC

CGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACC

CCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATC

ACTCTCGGCATGGACGAGCTGTACAAGGGTGGATCTGGAGGTGGATCTGGAGGTAT

GAGCAAAGATGGCAAAAAGAAGAAAAAGAAGTCAAAGACAAAGTGTG TAATTATG

cSC003-cUnaG-P2A-nUnaG-ST003-P2A-mCh

ATGGTAACTACATTAAGTGGACTATCAGGGGAGCAGGGTCCGAGCGGTGACATGAC

CACAGAGGAAGACTCTGCAACGCATATCAAGTTTTCCTAGCGTGATGAGGACGGCC

GTGAATTGGCTGGTGCGACGATGGAACTGCGTGATTCCAGCGGTAAAACCATCTCG

ACCTGGATTAGCGATGGTCACGTGAAGGACTTCTATCTCTACCCGGGTAAGTACACC

TTTGTTGAAACTGCGGCGCCAGATGGTTACGAAGTTGCGACTCCGATTGAATTCACC

GTCAATGAAGATGGTCAAGTTACCGTCGATGGGGAGGCGACGGAGGGTGACGCACA

CACCGGTGGCTCTGGAGGTGGATCTGGCGGTGTTAAATCCGTAGTCAACTTGGTCGG

AGAAAAACTCGTATATGTCCAAAAATGGGATGGCAAAGAGACAACCTACGTGCGTG

AAATTAAAGACGGGAAACTGGTTGTGACGCTGACAATGGGCGACGTGGTGGCAGTG

CGCTCGTACCGGCGTGCAACCGAAGGAAGCGGTGCCACCAATTTCTCCCTCCTGAAG

CAGGCTGGGGACGTGGAGGAAAACCCTGGCCCCATGGTTGAAAAATTTGTTGGTAC

GTGGAAAATAGCGGATTCTCATAATTTTGGCGAATATCTGAAGGCTATCGGGGCGCC

TAAAGAGCTGAGTGATGGAGGCGATGCCACGACTCCAACTCTGTATATTTCACAGA

AGGACGGTGATAAAATGACCGTTAAGATCGAGAATGGCCCGCCCACCTTCCTGGAT

ACACAGGTAAAGTTTAAACTTGGTGAAGAATTTGATGAGTTCCCGAGCGACAGACG

CAAAGGAGGCTCTGGTGGAGGCTCTGGAGGTGGCTCTGGACGCGGCGTGCCGCACA

TCGTGATGGTGGACGCCTACAAAAGATATAAAGGAAGCGGAGCTACTAACTTCAGC

CTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTGTGAGCAAGGGCGA

GGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGG

GCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTAC

GAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGC

CTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCC

CGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCG

CGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTCC

AGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGAC

GGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTA

CCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGAC

GGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCA

GCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGG

ACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGC

ATGGACGAGCTGTACAAGTAA

Oligonucleotides for Molecular Cloning

Table 3

Primer Name	Sequence 5' to 3'
SpyCatcher003-F	ctaaaggggacgcccatacgggcagcggtgaaagc
SpyCatcher003-R	cgtatgggcgtcccctt
SpyCatcher003-SDM-F	ttcagctagcgcgatgagga
SpyCatcher003-SDM-R	gctgaatttgatatgggtcgc
SpyTag003-F	gacacatgcagctcccg
SpyTag003-R	ggatccacgcggaacc

References

[1] M. Royzen, G. P. A. Yap, and J. M. Fox, "A Photochemical Synthesis of Functionalized trans-Cyclooctenes Driven by Metal Complexation," *J. Am. Chem. Soc.*, vol. 130, no. 12, pp. 3760–3761, Mar. 2008, doi: 10.1021/ja8001919.

[2] J. Li, S. Jia, and P. R. Chen, "Diels-Alder reaction–triggered bioorthogonal protein decaging in living cells," *Nat. Chem. Biol.*, vol. 10, no. 12, Art. no. 12, Dec. 2014, doi: 10.1038/nchembio.1656.

[3] X. Fan *et al.*, "Optimized Tetrazine Derivatives for Rapid Bioorthogonal Decaging in Living Cells," *Angew. Chem. Int. Ed.*, vol. 55, no. 45, pp. 14046–14050, 2016, doi: 10.1002/anie.201608009.

[4] A. Arsić, C. Hagemann, N. Stajković, T. Schubert, and I. Nikić-Spiegel, "Minimal genetically encoded tags for fluorescent protein labeling in living neurons," *Nat. Commun.*, vol. 13, p. 314, Jan. 2022, doi: 10.1038/s41467-022-27956-y.

[5] T. Yanagisawa, M. Kuratani, E. Seki, N. Hino, K. Sakamoto, and S. Yokoyama, "Structural Basis for Genetic-Code Expansion with Bulky Lysine Derivatives by an Engineered Pyrrolysyl-tRNA Synthetase," *Cell Chem. Biol.*, vol. 26, no. 7, pp. 936-949.e13, Jul. 2019, doi: 10.1016/j.chembiol.2019.03.008.

[6] E. R. Ruskowitz *et al.*, "Spatiotemporal functional assembly of split protein pairs through a light-activated SpyLigation," *Nat. Chem.*, 2023, doi: 10.1038/s41557-023-01152-x.

[7] C. McQuin *et al.*, "CellProfiler 3.0: Next-generation image processing for biology," *PLOS Biol.*, vol. 16, no. 7, p. e2005970, Jul. 2018, doi: 10.1371/journal.pbio.2005970.