Genetic Incorporation of Two Mutually Orthogonal Bioorthogonal Amino Acids That Enable Efficient Protein Dual-Labeling in Cells

Riley M. Bednar[†], Subhashis Jana[†], Sahiti Kuppa[‡], Rachel Franklin[†], Joseph Beckman[†], Edwin Antony[‡], Richard B. Cooley[†] Ryan A. Mehl[†*]

[†]Department of Biochemistry and Biophysics, Oregon State University, 2011 Agricultural & Life Sciences Building, Corvallis, Oregon 97331-7305, United States

[‡]Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, Edward A. Doisy Research Center, 1100 South Grand Blvd., St. Louis, MO 63104, United States

[*]To whom correspondence should be addressed (E-mail: ryan.mehl@oregonstate.edu)

Abstract

The ability to site-specifically modify proteins at multiple sites *in vivo* will enable the study of protein function in its native environment with unprecedented levels of detail. Here, we present a versatile two-step strategy to meet this goal involving site-specific encoding of two distinct noncanonical amino acids bearing bioorthogonal handles into proteins *in vivo* followed by mutually orthogonal labeling. This general approach, that we call dual encoding and labeling (DEAL), allowed us to efficiently encoded tetrazine- and azide-bearing amino acids into a protein and demonstrate for the first time that the bioorthogonal labeling reactions with strained alkene and alkyne labels can function simultaneously and intracellularly with high yields when site-specifically encoded in a single protein. Using our DEAL system, we were able to perform topologically-defined protein-protein crosslinking, intramolecular stapling, and site-specific installation of fluorophores all inside living *Escherichia coli* cells, as well as study the DNA-binding properties of yeast Replication Protein A *in vitro*. By enabling the efficient dual modification of proteins *in vivo*, this DEAL approach provides a tool for the characterization and engineering of proteins *in vivo*.

Introduction

As the fields of molecular and cellular biology develop, the demand for tools that enable defined, independent, and residue-level manipulation of proteins at multiple sites in their native context continues to grow. Recent advances in bioorthogonal chemistry have the potential to meet this demand and thereby provide the ability to interrogate protein function in the complex chemical environment of living cells through dual labeling strategies[1]. For example, site-specific dual labeling could represent a general approach to enable the installation of FRET pairs to monitor protein function, allow the site-specific labeling of two distinct proteins for two-color super-resolution microscopy, or enable the construction of supramolecular protein complex to study and manipulate regulatory processes, all *in vivo* abilities that remain currently out of reach for most technologies, aside from highly specialized approaches. Many contemporary technologies such as self-labeling enzyme domains, chemoenzymatic labeling, and affinity tags have been used to encode bioorthogonal handles into proteins *in vivo*[2]. However, their bulky size, restrictive placement at the target protein's termini, and potential to disrupt protein function greatly impinge their utility[3]. In contrast, genetically encoded noncanonical amino acids (ncAAs) bearing bioorthogonal handles can be site-specifically installed at any location in a protein *in vivo* via genetic code expansion (GCE) with minimal perturbations to protein structure and function[4]. Moreover, this GCE approach has been adapted to allow simultaneous incorporation of multiple distinct ncAAs, including bioorthogonally reactive ncAAs, into a single protein *in vivo*[5–9]. Nevertheless, while the GCE approach enables site-specific attachment with small handles, the higher demands it places on efficient and orthogonal labeling reactions that can function inside cells have not been met.

An important requisite to dual modifying protein *in vivo* is that the two labeling reactions must be mutually orthogonal to both one another and to other functional groups present in the cell (labeling orthogonality)[10]. The reactions should also be free of catalysts, high yielding, and rapid under physiological conditions, thereby enabling effective labeling at low protein concentrations and on biologically relevant time frames[11]. Few reactions meet these criteria better than the strain-promoted azide-alkyne coupling (SPAAC) and inverse electron demand Diels-Alder (IEDDA) reactions. The SPAAC reaction is a [3+2] cycloaddition reaction originally described by Bertozzi and colleagues that commonly occurs between azides and dibenzoannulated cyclooctynes (DBCO)[12]. Despite modest reaction kinetics (~0.1 - 1 M^{-1} s^{-1}), its biocompatibility has made it one of the most extensively used bioorthogonal reactions for *in vivo* labeling[10,13]. The IEDDA reaction

is a [4+2] cycloaddition that occurs between an electron-deficient diene such as a 1,2,4,5-tetrazine and an electron-rich dienophile such as a strained alkene[1]. Bioorthogonal IEDDA reactions, such as those first described by Fox and colleagues, can reach rates upwards of 10^6 M^{-1} s^{-1}, allowing complete reaction within minutes at sub-micromolar concentrations[1,11,14–17]. Amino acid derivatives containing azide, cyclopropene, alkyne, *trans*-cyclooctene (TCO), and tetrazine functionalities have all been genetically encoded into proteins[18], albeit not yet in a manner conducive to dual encoding and subsequent intracellular labeling *in vivo*.

While density functional theory predicts that SPAAC and IEDDA reactions should be mutually orthogonal and the *in vivo* compatibility of these reactions has been shown, site-specific encoding of the handles into the same protein and dual labeling *in vivo* has not yet been demonstrated[19–21]. A recent attempt succeeded at dual encoding two different handles into a protein using GCE, however, the functional groups required the use of copper catalyst, thereby relegating labeling to the cell surface due to copper toxicity[22]. These examples highlight the challenge of developing efficient and orthogonal dual encoding and labeling approaches. An effective solution to this challenge should be encoding azide and tetrazine moieties, since they exhibit high biostability, participate in biocompatible, mutually orthogonal reactions, and are not anticipated to cross-react during the long incubation periods required for GCE[11,19,23]. Provided tetrazine- and azide-bearing amino acids can be encoded efficiently into the same protein, and the mutual orthogonality of the SPAAC and IEDDA reactions persists *in vivo*, this dual protein labeling approach would open avenues to manipulate and study molecular processes with greater scrutiny than is currently accessible.

To test this hypothesis, we developed the first genetic code expansion system that enables the simultaneous and site-specific incorporation of the two bioorthogonally reactive ncAAs, *para*-azidophenylalanine (pAzF) and a tetrazine-containing ncAA (Tet3.0), into proteins *in vivo*. Our optimized GCE system enables robust production of dual-ncAA containing proteins and allowed us to characterize the dual labeling efficiency and orthogonality on various proteins *in vivo* and *in vitro*. We demonstrate the utility and versatility of our dual encoding and labeling (DEAL) system by showcasing a diverse array of *in vivo* abilities, such as site-specific installation of FRET pairs, topologically-defined protein-protein crosslinking, and site-specific intramolecular protein stapling (Fig. 1), as well as applying our system to the study of the challenging *Saccharomyces*

cerevisiae Replication Protein A (*Sc*RPA) complex. These examples illustrate the precision and power of our optimized DEAL system to meet the demands of rapidly advancing biological fields.

Results and Discussion

Organization of Genetic Code Expansion Components

Generating a dual-ncAA suppression system capable of supporting the simultaneous encoding of two distinct ncAAs requires two mutually orthogonal suppression systems (subsystems), each comprised of four components: 1) the ncAA to be encoded, 2) the aminoacyl tRNA synthetase (aaRS) engineered to load the ncAA onto 3) the cognate tRNA and 4) codon that will be suppressed by the ncAA-tRNA during translation of the protein of interest. As such, we elected to use a popular *Methanocaldococcus jannaschii* tyrosyl aminoacyl tRNA synthetase/tRNA pair (*Mj*TyrRS/tRNA) that was originally generated for the incorporation of *para*-cyanophenylalanine, but that also shows polyspecificity towards pAzF[24], to encode this ncAA at amber stop codons (herein referred to as pAzFRS/$tRNA_{CUA}$). To drive incorporation at ochre stop codons, we turned to an engineered *Methanosarcina barkeri* Pyrrolysyl aminoacyl tRNA synthetase (*Mb*PylRS/$tRNA_{UUA}$ that efficiently recognizes a meta-substituted 1,2,4,5-tetrazine-containing phenylalanine derivative (Tet3.0RS/$tRNA_{UUA}$) previously selected by our laboratory[16]. The choice of anticodons was based on two factors. First, the tRNA anticodon loop is a known identity factor for *Mj*TyrRS and, as such has historically been utilized as a UAG suppression system with known success[25], while the tRNA anticodon is not an identity factor for *Mb*PylRS and can therefore tolerate codon reassignment of its cognate tRNA[26]. Second, The UGA codon was not considered because it is known to experience near-cognate suppression from endogenous $tRNA^{Trp}_{CCA}$ [27].

We performed optimization experiments to identify the most efficient combination of cell strains and vectors for hosting the two suppression subsystems and reporter gene under single stop codon suppression conditions (Fig. S2). For all experiments, the Tet3.0 and pAzF concentrations were set at 0.5 and 1.0 mM, respectively. These experiments revealed that the pAzFRS/$tRNA_{CUA}$ and Tet3.0RS/$tRNA_{UUA}$ are most efficient when in the pEVOL[28] and pUltraI[29] plasmids, respectively, and that these both function most optimally when the reporter gene is expressed from a pET28 vector in BL21(DE3) cells (for a detailed discussion of these results, see the supplemental section “Evaluation of UAG and UAA Single Site Suppression”). As such, our finalized dual

suppression system is a modular, three-vector configuration, with each suppression system placed on a separate plasmid, and the gene of interest contained on a third plasmid (Figure S1).

Dual incorporation of pAzF and Tet3.0

Earlier reports on dual ncAA incorporation have identified points in translation where orthogonality may become compromised, such as the substrate specificity of the aaRS, the interactions between aaRS and tRNA, and the specificity of tRNAs during decoding[29,30] (Fig. S1). The orthogonality of the individual aaRS/tRNA pairs in our systems, *Mj*TyrRS/$tRNA_{CUA}$ and *Mb*PylRS/$tRNA_{UUA}$, has previously been established[29]. Therefore, we characterized the orthogonality of our dual-encoding subsystems at the ncAA-aaRS and tRNA-codon levels. We observed that, in the absence of ncAA for the pAzFRS/$tRNA_{CUA}$ subsystem, UAG codons can be suppressed by near-cognate suppression by endogenous decoding systems, and that the UAA suppressor tRNA for the Tet3.0RS/$tRNA_{UUA}$ subsystem can decode UAG stop codons through wobble-pairing (Fig. S3)—both of which have previously been observed[29,31]. Importantly, these breaches in encoding orthogonality are situational, and do not manifest when both subsystems are present and functioning properly (see section "Evaluation of GCE Orthogonality" in the supplemental section for a full discussion).

Having established that our optimized pAzFRS/$tRNA_{CUA}$ and Tet3.0RS/$tRNA_{UUA}$ subsystems are contextually and mutually orthogonal, we sought to combine them for the dual encoding of pAzF and Tet3.0 into proteins. To do so, we paired a Small Ubiquitin-like Modifier-superfolder Green Fluorescent Protein (SUMO-sfGFP) fluorescent reporter possessing UAG and UAA codons at sites 35 and 102, respectively, with both subsystems and observed a 10% yield as compared to SUMO-$sfGFP^{WT}$ when both ncAAs were added (Fig. 2A). These positions for Tet3.0 minimize tetrazine-dependent sfGFP quenching due to their distance from the sfGFP fluorophore (see the supplemental section "*Evaluation of UAG and UAA Single Site Suppression*"). We also reversed the order of the nonsense codons by introducing UAA and UAG codons at positions 35 and 102 and observed a 22% yield (Fig. 2A), indicating that the order of the stop codons can be reversed and that this orientation may improve protein yields. Unsurprisingly, we also observed above-background fluorescence when pAzF and/or the pAzFRS/$tRNA_{CUA}$ subsystem were absent (Fig. 2A), which we attribute to the situational breaches in encoding orthogonality previously mentioned (Fig. S3). To confirm the fidelity of ncAA incorporation, sfGFP possessing both pAzF

and Tet3.0 at two commonly used positions 134[32] and 150[11] (sfGFP^{Dual}) was expressed and purified yielding 26 mg protein per L of expression media, approximately 12% of sfGFP^{WT} (Table S5). Electrospray ionization (ESI) mass spectrometry analysis on sfGFP^{Dual}, sfGFP^{pAzF}, sfGFP$^{Tet3.0}$, and sfGFP^{WT} expressed and purified under analogous conditions showed masses consistent with the expected ncAA-substitutions (Fig. 2B). These results verify that our dual-encoding system can be used to achieve efficient, simultaneous, and homogenous site-specific incorporation of pAzF and Tet3.0 into the same protein.

Evaluation of dual labeling *in vitro*

Dual encoding of pAzF and Tet3.0 into a protein should enable efficient one-pot dual labeling. To test this, we produced a series of SUMO-sfGFP fusion proteins with either pAzF or Tet3.0 individually in each domain at either site 35 or 253 (SUMO^{pAzF}-sfGFP or SUMO-sfGFP$^{Tet3.0}$) or together (SUMO^{pAzF}-sfGFP$^{Tet3.0}$) in yields ranging from ~100 to ~11 mg/L (Table S5). Upon exposure to a DBCO-TAMRA fluorophore, only proteins possessing pAzF exhibited labeling, as determined by SDS-PAGE in-gel fluorescence (Fig. 3B). Likewise, exposure to a strained *trans*-cyclooctene-terminated 5 kDa polyethylene glycol polymer (sTCO-PEG$_{5000}$) led to a mobility shift only for SUMO-sfGFP$^{Tet3.0}$ (Fig. 3B). When SUMO^{pAzF}-sfGFP$^{Tet3.0}$ was exposed to both DBCO-TAMRA and sTCO-PEG$_{5000}$ we observed the formation of a fluorescent, mobility-shifted product (Fig. 3B). To verify labeling orthogonality, we added ULP1 protease to cleave the labeled SUMO domain, and as expected, observed a fluorescently-labeled SUMO fragment and a mobility-shifted sfGFP fragment (Fig. S4).

To observe the labeling extent of SPAAC and IEDDA we utilized a modified blocking assay developed by Murrey and colleagues[33] with a series of sfGFP constructs that positions pAzF or Tet3.0 at site 150 (sfGFP^{pAzF} and sfGFP$^{Tet3.0}$) or together at sites 134 and 150, respectively (sfGFP^{Dual}). This assay consists of an initial non-visualizable labeling step with DBCO-NH_2 and/or sTCO-OH, followed by a quenching step with excessive DBCO-TAMRA and/or sTCO-JF669 dyes. Quantifying the loss in fluorescence during step 2 as a result of the reaction in step 1, we determined that SPAAC and IEDDA proceeded efficiently for these constructs *in vitro* (Fig. 4A-B), and that the labeling in step 2 is complete by the 2 h time point (Fig. S6). It should be noted that this assay only allows us to observe relative reaction extent, since compromised handles (such as azide reduction; see the supplemental section "*Mass Spectrometry Analysis*") are unreactive in

both labeling steps of the assay. In addition, we observed that 24-hour SPAAC labeling between pAzF and DBCO-NH_2 followed by 15-minute IEDDA labeling between Tet3.0 and sTCO-OH maximized *in vitro* labeling while minimizing cross-reactivity (Fig. S5A-B), consistent with observations made by Karver et al., who noted that azides exhibit slight reactivity towards TCO[19].

Using these labeling conditions, we reacted the aforementioned sfGFP constructs with DBCO-NH_2 and sTCO-OH *in vitro* and analyzed the products by mass spectrometry. Peaks corresponding to dual-labeled sfGFP were observed as the dominant product (Fig. S7), consistent with the observed orthogonality (Figs. 3 and S5). Minor peaks are consistent with pAzF reduction to *para*-aminophenylalanine, a common modification produced during protein expression[34] or salt adducts (see supplemental section "Mass Spectrometry Analysis"). Together these results indicate that we have identified optimal labeling conditions that enable the SPAAC and IEDDA reactions to proceed efficiently and orthogonally on proteins containing pAzF and Tet3.0.

Evaluation of dual labeling *in vivo*

Despite numerous examples of dual encoding of ncAAs into proteins *in vivo*[5–7,22], to our knowledge there are no reports of site-specific dual labeling of proteins inside a cell. To demonstrate *in vivo* DEAL with our system, we utilized our modified blocking assay to quantify labeling of our previously used sfGFP constructs in cells. Similar to *in vitro* labeling, *in vivo* labeling proceeded efficiently for SPAAC and IEDDA reactions, (Fig. 4C-D), confirming that the cellular environment does not pose a major impediment to dual-labeling with these ncAAs. Interestingly, we also noted that the SPAAC reaction seemed to proceed more efficiently *in vivo* (Fig. 4C-D). This may be explained by partitioning of the proteins and labeling reagents, as Tian and colleagues have observed rate enhancements for the SPAAC reaction in *in vivo* settings[35], among other effects. To evaluate labeling orthogonality of the SPAAC and IEDDA reactions *in vivo* we used our *in vitro* cross-reactivity analysis method by exposing cells containing $sfGFP^{pAzF}$ and $sfGFP^{Tet3.0}$ to sTCO-OH and DBCO-NH_2, respectively (Fig. S5C-D). Again, we detected minimal cross-reactivity, indicating that under these conditions the SPAAC and IEDDA reactions retain their orthogonality. Additionally, we performed mass spectrometry analysis on $sfGFP^{Dual}$ that was dual labeled *in vivo* and we observed similar peaks as with *in vitro* dual labeling. We did note a relative increase in the abundance of peaks corresponding to reduced pAzF and unreacted

Tet3.0, both of which likely result from reduced ncAA reduction formed during lengthy *in vivo* labeling periods. (Fig. S7F; see supplemental section "Mass Spectrometry Analysis").

Dual Encoding and Labeling *in vivo*

Having confirmed our new DEAL system functions efficiently *in vivo*, we next sought to explore the utility of our system by showcasing three *in vivo* capabilities: (1) intermolecular protein-protein crosslinking, (2) intramolecular protein stapling, (3) and dual-fluorophore labeling (Fig. 1). To aid in our analysis we developed a second fluorescent reporter protein, the blue fluorescent protein mTagBFP2, which has excellent spectral overlap with sfGFP[36]. We screened 6 sites for their ability to tolerate UAG and UAA suppression and selected site 105, with a yield of 79 mg/L when pAzF is encoded (Table S5, Fig. S8; see supplementary discussion "*mTagBFP2 Screening and Characterization*"). All constructs including mTagBFP2 encode pAzF at this relative position, while all constructs including sfGFP encode Tet3.0 at relative position 150.

Intermolecular Protein-Protein Crosslinking in vivo

Our DEAL system also offers a plethora of potential *in vivo* applications if dual encoding is applied to two distinct proteins, such as programmable supramolecular protein-complex formation, as was achieved using a combination of ncAA and biarsenical tags by Rutkowska et al.[37]. Here we demonstrate that a simplified dual-ncAA approach can be used to achieve a similar outcome with greater site-specificity through covalent intermolecular hetero-protein protein crosslinking between sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF} in *E. coli.* To do so, we generated a pETduet vector that most balances the expression of these constituent proteins under dual encoding conditions, and confirmed orthogonal labeling of the ncAAs therein (Fig. S8) (see the supplemental section "*sfGFP/mTagBFP2 and sfGFP-mTagBFP2 Expression and Characterization*" for a full discussion).

We first verified that these two ncAA-containing proteins could be successfully crosslinked *in vitro*, by incubating purified sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF} with a heterobifunctional sTCO-DBCO crosslinker followed by SDS-PAGE and analytical size-exclusion chromatography (Fig. 5A-C). A distinct higher molecular weight product was produced with a retention volume of 14 mL (compared to ~16 mL for the unlinked proteins) that absorbed at both 399 and 485 nm, the absorption maxima of mTagBFP2 and sfGFP, respectively (Fig. 5B-C). We isolated this product

and determined that it possesses spectral properties consistent with sfGFP and mTagBFP2, and displays detectable FRET between the two, supporting the formation of a crosslink between the two proteins (Fig. S10).

To track *in vivo* crosslinking of sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF} we initially attempted to observe ensemble FRET increase resulting from crosslinking in living *E. coli* cells; however, the signal was too diffuse to effectively capture this effect. Instead, we opted to directly observed the formation of the crosslinked product over time using SDS-PAGE in-gel fluorescence for sfGFP, which, in the absence of boiling retains fluorescence[38] (Fig. 5D-E). Immediately following the addition of sTCO-DBCO we observed a small downward mobility shift from the sfGFP$^{Tet3.0}$ as it reacted quickly with sTCO-DBCO followed by a slower coupling reaction with mTagBFP2^{pAzF} to form the crosslinked sfGFP-mTagBFP2 complex (Fig. 5D, see supplemental discussion "*sfGFP/mTagBFP2 and sfGFP-mTagBFP2 Expression and Characterization*"). The reaction appeared to reach 50% completion at about ~70 minutes after the addition of sTCO-DBCO to cells containing sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF} and approached completion after 6 hours (Fig. 5D-E). While the concentrations of the sTCO-DBCO crosslinker that was required was estimated empirically, it is possible that variables, such as reactant solubility and partitioning, or unoptimized sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} ratio may limit the extent of crosslinking[39]. Importantly, no detectable crosslinking was observed in the absence of sTCO-DBCO, or when the crosslinker was added to cells containing expressed proteins lacking ncAAs (Fig. 5D-E). A consistent decrease in overall fluorescence over time was observed, which is attributed to protein stability issues like proteolysis or photodegradation. These results confirm that our DEAL system can be used to achieve site-specific crosslinking between two different proteins, thereby enabling precise control over supramolecular topography in the complex milieu of live cells.

Intramolecular Protein Crosslinking in vivo

Our *in vivo* DEAL system should also be capable of intramolecular protein crosslinking often referred to as "protein stapling". While peptide stapling has been used extensively to alter the properties of peptides[40], no general protein stapling methodology has been demonstrated in live cells.

Prior to pursuing protein stapling in cells, we verified *in vitro* stapling was feasible using an sTCO-PEG$_4$-DBCO crosslinker on purified, dual-encoded protein. To do so, we evaluated the

extent of stapling for two sfGFP-mTagBFP2 fusion proteins (sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} and mTagBFP2^{pAzF}-sfGFP$^{Tet3.0}$; Fig. S11) when exposed to an sTCO-PEG$_4$-DBCO crosslinker (Figs. 6A-B). We found that in order to observe distinctive stapling via in-gel fluorescence the crosslinked product had to be cleaved by TEV protease at a cut site in the linker to resolve the linked and unlinked products (Figs. 6A-B, S11; for a more detailed description of this construct, see the supplemental section "*sfGFP/mTagBFP2 and sfGFP-mTagBFP2 Expression and Characterization*" and Table S4). Both constructs showed similar degrees of stapling, achieving ~50% completion after 120 minutes (Fig. 6A-B). Due to its higher levels of expression (Fig. S11B), we elected to proceed with the sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} orientation to explore *in vivo* stapling.

To demonstrate *in vivo* protein stapling with our DEAL system, *E. coli* cells containing sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} were exposed to sTCO-PEG$_4$-DBCO and the degree of stapling was determined at various time points. Stapling progress, as monitored by SDS-PAGE in-gel fluorescence, showed that *in vivo* stapling reached 50% of its maximal extent after about ~20 minutes and was nearly complete after 6 hours (Figs. 6C-D). The observed fastest migrating band is identified as sfGFP$^{Tet3.0}$ that was modified with sTCO-PEG$_4$-DBCO, as can be observed *in vitro* (Figure 6B, see the supplemental discussion section "*sfGFP/mTagBFP2 and sfGFP-mTagBFP2 Expression and Characterization"*), similarly to the phenomenon that was observed in *in vivo* crosslinking (Figure 5D). These results highlight the ability of our DEAL system to enable intramolecular "stapling" of proteins in live *E. coli* cells.

Dual Fluorescent Labeling in vivo

The ability to site-specifically install two different fluorescent dyes into a protein in their native context is highly desired to monitor conformational changes[5], track localization[41], and observe protein-protein interactions[42], among other applications. To demonstrate *in vivo* DEAL with fluorescent dyes, we simultaneously applied DBCO-TAMRA and sTCO-JF669 dyes (1 μM) to cells containing sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF}. After 60 and 5 minutes the SPAAC and IEDDA reactions were quenched and labeling was observed by SDS-PAGE in-gel fluorescence. As anticipated, we observed distinct and specific labeling with minimal cross-reactivity (Fig. 7). These results demonstrate that our DEAL system can be used to site-specifically and simultaneously dual-label proteins *in vivo* within biologically-relevant concentrations and times.

Dual labeling of *Saccharomyces cerevisiae* Replication Protein A DNA-Binding Domains *in vitro*

Next, we sought to determine if our DEAL system could address a more challenging and biologically relevant problem: monitoring the interaction between *Saccharomyces cerevisiae* Replication Protein A (RPA) and single-stranded DNA (ssDNA) via FRET. RPA is a heterotrimeric ssDNA-binding protein complex that is essential for almost all aspects of DNA metabolism[43]. RPA functions to sequester transiently open ssDNA in the cell and serves as a hub for the recruitment of over three dozen enzymes[44]. Structurally, RPA is composed of five ssDNA binding domains (DBDs) and two protein-protein interaction domains (PIDs) which are tethered by several flexible linkers[45] (Fig. 8A). Owing to these linkers, RPA can adopt multiple configurations on ssDNA and experimental tools to directly monitor transitions in these configurations are currently limited. Moreover, the RPA complex is replete with Cys residues and cannot be reconstituted *in vitro*, making installation of fluorophores by DEAL a necessary approach. We used DEAL to site-specifically introduce organic dyes onto the two terminal DBDs such that the change in configuration can be monitored by FRET upon binding to ssDNA (Fig. 8B).

To generate fluorescently dual labeled RPA, we encoded pAzF at site 211 on DBD-A (in RPA70) and Tet3.0 at site 101 on DBD-D (in RPA32)[46,47]. Following purification of full-length RPA-DBD-A^{pAzF}-DBD-$D^{Tet3.0}$ the complex was labeled with DBCO-Cy3 and TCO-Cy5. In-gel fluorescence was used to confirm DEAL of DBD-$A^{pAzF\text{-}Cy3}$ and DBD-$D^{Tet3.0\text{-}Cy5}$, respectively, with no detectable cross-reactivity (Fig. 8C). We next monitored FRET changes in dual-labeled RPA as a function of ssDNA concentration (Fig. 8D). In its unbound state RPA is expected to adopt an ensemble of configurations that give rise to a high FRET signal; however, upon ssDNA binding the complex is expected to stabilize into a splayed, low FRET configuration where the DBDs are linearly arranged on the ssDNA template[47] (Fig. 8B). In agreement with this prediction, we observed high of FRET between Cy3 and Cy5 at low ssDNA concentrations and a transition to low FRET as more DNA was added (Fig. 8D). These results highlight the robustness of our DEAL system to address challenging encoding and labeling problems on meaningful protein systems.

Conclusions

We have reported here that, when combined with GCE, the SPAAC and IEDDA bioorthogonal labeling reactions can be used to simultaneously and site-specifically label a protein in live cells. To do so, we developed and optimized a DEAL system that encodes pAzF and Tet3.0 into proteins and relies on their orthogonal reactivity to achieve labeling. Our DEAL system enables homogenous production of full-length proteins containing pAzF and Tet3.0 at ~10% of their natural counterparts, and affords mutually orthogonal reactions *in vitro* and *in vivo* with typical labeling yields upwards of 50% and 90% respectively. We showcased the utility and versatility of this DEAL system for *in vivo* applications through three vignettes. First, through the simultaneous incorporation of pAzF and Tet3.0 into two distinct proteins within the same cell, we demonstrated the first example of *in vivo* site-specific bioorthogonal protein-protein crosslinking via the addition of a heterobifunctional crosslinker. Second, we achieved *in vivo* protein stapling by performing intramolecular crosslinking between pAzF and Tet3.0 handles located within the same protein. Third, we demonstrated efficient dual labeling in *E. coli* using two fluorescent dyes with minimal cross-reactivities. Moreover, to tackle a traditionally intractable problem, DEAL was applied to monitor the relative configurational changes of *S. cerevisiae* RPA by FRET as it binds ssDNA. These applications conclusively demonstrate that the SPAAC and IEDDA reactions are efficient and mutually orthogonal on the same protein both *in vitro* and within living *E. coli* cells. While fluorescent proteins were used here to allow facile characterization of dual labeling, our approach should be generalizable to virtually any protein system.

The GCE systems were selected in this work to maximize DEAL on proteins from *E. coli* and to demonstrate the superior efficiency of combining these bioorthogonal handles on proteins. Due to the success of these applications, we hypothesize that transplanting the ncAAs and reactions developed here into eukaryotic cells using currently available systems[16,48] will have a similar level of mutual orthogonality and provided efficient DEAL for a broad range of applications. For example, while we illustrated dual *in vivo* labeling with two fluorophores, one could also install virtually any combination of probes or moieties, such as protein ligands and inhibitors to modulate protein function, photosensitive probes for spatiotemporal control, or secondary probes with shifted spectral properties, all of which could enable simultaneous monitoring and control of protein functions in real-time in their native context[49–51]. As demonstrated, the reactive handles interactions via FRET without needing bulky fluorescent proteins that may ablate interactions[52]. Likewise, DEAL could be used to construct supramolecular structures or to switch protein function

in living cells through installation of crosslinks within a protein structure with residue-specific precision, similar to how we were able to achieve bioorthogonal protein-protein crosslinking and protein stapling[53,54]. While this work represents the first examples of DEAL that are site-specific and function effectively *in vivo*, we expect DEAL system utility to increase with advances in GCE technology and bioorthogonal chemistry.

Materials and Methods

A detailed materials and methods section can be found in the supplementary materials.

Supplementary Materials

The supporting information includes a detailed description of the materials and methods; supplementary discussion as it pertains to single site suppression, evaluation of orthogonality, mass spectrometry analysis, mTagBFP2 screening and characterization, and characterization of sfGFP/mTagBFP2 and sfGFP-mTagBFP2 constructs; as well as all supplementary schemes, figures, tables, and references.

Author Contributions

R.M.B., R.A.M., and R.B.C. designed of the project. R.M.B. performed the experiments. The experiments specific to RPA were designed by E.A. and S.K and performed by S.K.. Organic reagents generated for the experiments were synthesized by S.J.. R.F. and J. B. analyzed the MS data. R.M.B., R.A.M., and R.B.C. wrote the manuscript with input from E.A., S.K., S.J., R.F., and J.B..

Funding

This work was supported in part by grants from the National Science Foundation (NSF-1518265 and NSF-2054824) and the National Institutes of Health awarded to R.A.M. (5R01GM131168-02), and E.A. (R01GM130746 and R01GM133967).

Acknowledgements

We would like to kindly thank L. Lavis of Janelia Research Campus, HHMI, for providing the JF669-CO_2H used in this study. We would also like to thank C. Johnson for use of his probe sonicator and fluorescence spectrometer.

Conflicts of interest

There are no conflicts to declare.

Figures

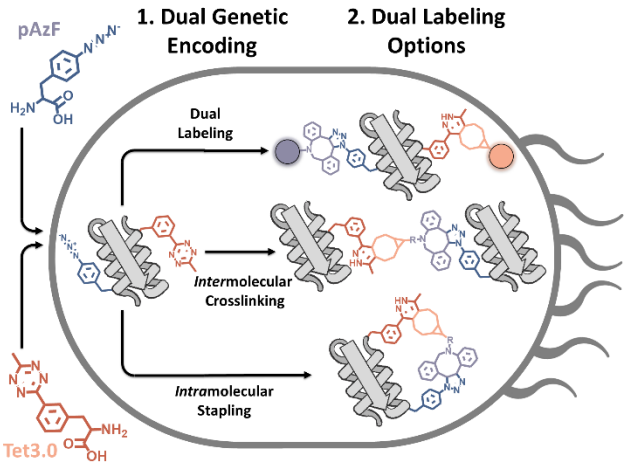

Figure 1. Schematic of *in vivo* dual encoding and labeling. Noncanonical amino acids (ncAAs) *para*-azidophenylalanine (pAzF; blue) and a meta-substituted tetrazine-containing phenylalanine derivative (Tet3.0; orange) are encoded into proteins using genetic code expansion via mutually orthogonal, dual nonsense suppression systems. Controlled bioorthogonal labeling at both sites within proteins enables site-specific dual labeling, topologically-defined intermolecular crosslinking and intramolecular stapling *in vivo*.

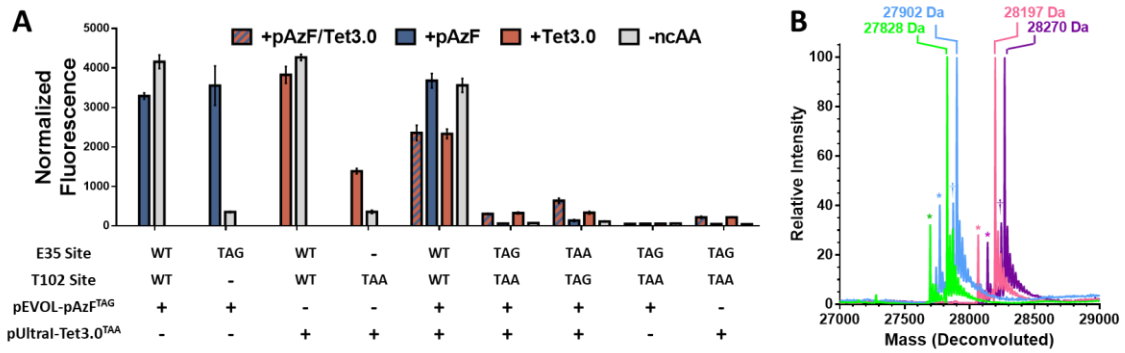

Figure 2. Validation of dual suppression. (A) Normalized SUMO-sfGFP fluorescence for single and dual UAG-UAA suppression at 24 hours in the presence of both pAzF and Tet3.0 at 1.0 and 0.5 mM (hashed blue and orange), pAzF alone (blue), Tet3.0 alone (orange), and in the absence of ncAAs (gray). (B) Overlaid ESI mass spectra of sfGFP containing either pAzF (blue) or Tet3.0 (pink) at site 150, or together at sites 134 (pAzF) and 150 (Tet3.0) (purple), or neither (WT; green). Observed masses are indicated above each major peak. Peaks corresponding to the loss of N-terminal methionine are indicated by "*", while peaks corresponding to pAzF reduction are indicated by "†". Mass measurement error is ± 1 Da, see Table S6 for expected masses.

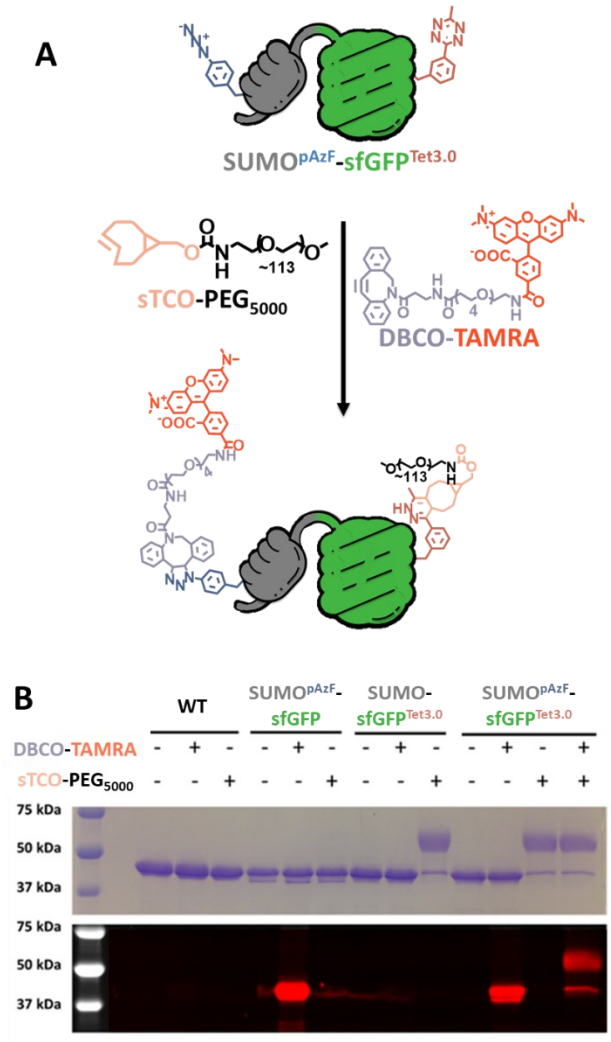

Figure 3. *In vitro* dual labeling of dual encoded SUMO-sfGFP. (A) Reaction scheme of SUMO^{pAzF}-sfGFP$^{Tet3.0}$ dual labeling. (B) SDS-PAGE of SUMO^{pAzF}-sfGFP$^{Tet3.0}$ (10 μM) exposed to DBCO-TAMRA and/or sTCO-PEG$_{5000}$ (100 μM) imaged by Coomassie staining (top) and in-gel fluorescence (bottom). In these experiments, pAzF is incorporated at position 35 (in the SUMO domain), and Tet3.0 at position 253 (in the sfGFP domain). The contents of each lane are indicated above.

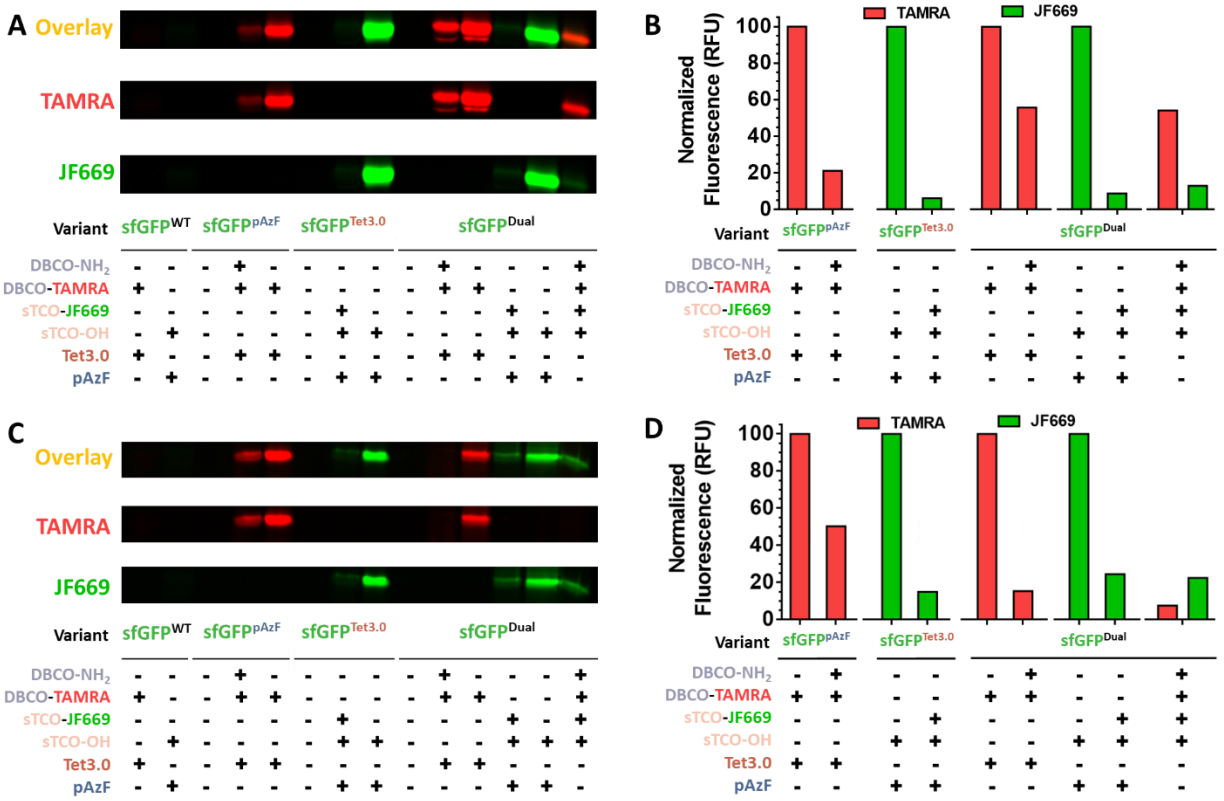

Figure 4. Quantification of dual labeling reactions *in vitro* and *in vivo* using two step blocking labeling (see methods). (A) SDS-PAGE of *in vitro* reactions on sfGFP (10 µM) variants labeled with sTCO-OH and/or DBCO-NH_2 (67 µM each for 15 min, and 24 h, respectively) followed by sTCO-JF669 and/or DBCO-TAMRA (667 µM, 2 h) and off-target reactions quenched with excess pAzF/Tet3.0 prior to in-gel fluorescence imaging; top panel is the overlay of TAMRA and JF669 signals followed by each individual channel, and Coomassie staining. Encoding positions for $sfGFP^{Tet3.0}$ and $sfGFP^{pAzF}$ are at site 150, while encoding of $sfGFP^{Dual}$ includes Tet3.0 at site 134, and pAzF at site 150. (B) Densitometry quantification of the fluorescent gels in panel A (red bars are quantifications of the TAMRA channel, and green bars are quantifications of the JF669 channel). See the materials and methods section for details about how these quantifications were performed and presented. (C) *In vivo* labeling analysis analogously presented to panel A. (D) Densitometry quantification of fluorescent gel channels in panel C, processed and presented analogously to panel B.

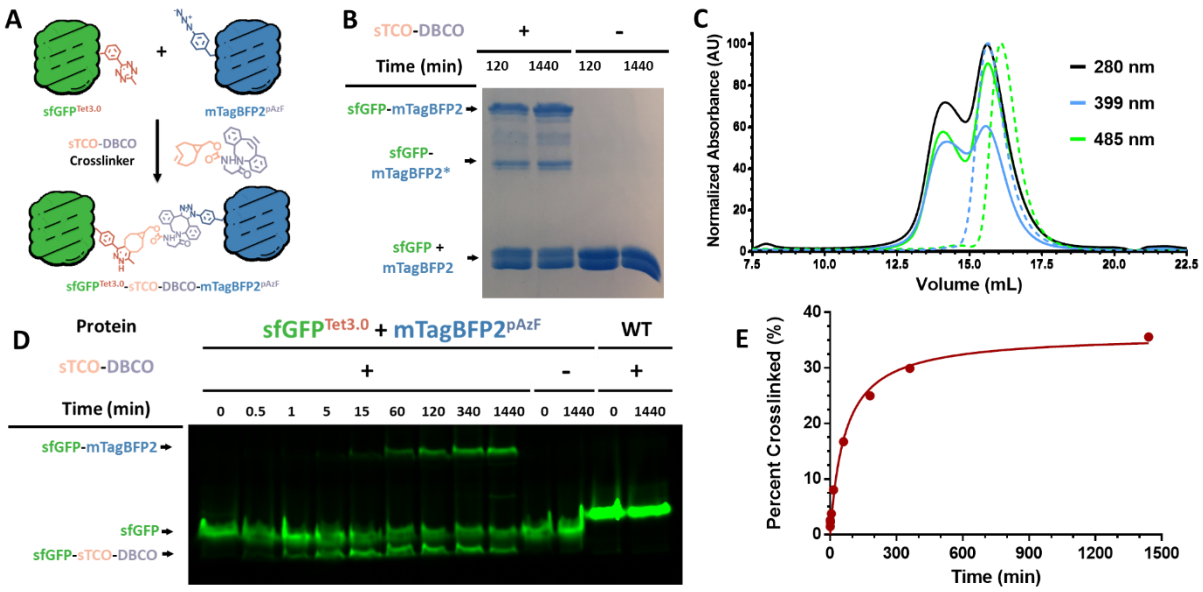

Figure 5. *In vitro* and *in vivo* intermolecular protein-protein crosslinking through DEAL. (A) Reaction scheme of crosslinking between sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF} via an sTCO-DBCO linker. (B) SDS-PAGE analysis of *in vitro* reaction between purified sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF} (10 μM each) at 2 and 24 hours after sTCO-DBCO addition (10 μM). The lower mobility crosslinked product, sfGFP-mTagBFP2*, matches the mass loss for photolyzed mTagBFP2* known to accumulate over time[36]. (C) Size exclusion chromatogram of crosslinking reaction mixture at 24 hours post sTCO-DBCO addition and monitored by absorbance at 280 nm (black), 399 nm (mTagBFP2 λ_{max}; blue), 485 nm (sfGFP λ_{max}; green). Dashed lines are elution profiles of purified sfGFP^{WT} (dashed green) and mTagBFP2^{WT} (dashed blue). (D) SDS-PAGE in-gel fluorescence image of *in vivo* crosslinking reaction time course following the addition of sTCO-DBCO (concentration empirically determined, see materials and methods) to *E. coli* cells containing expressed either sfGFP$^{Tet3.0}$ and mTagBFP2^{pAzF}, or sfGFP^{WT} and mTagBFP2^{WT}. (E) Densitometry quantification of panel D. The percent crosslinked was estimated as the percentage of fluorescence that the upper, crosslinked product band constitutes of the total fluorescence in each lane at each time point. The contents of each lane are indicated above each gel.

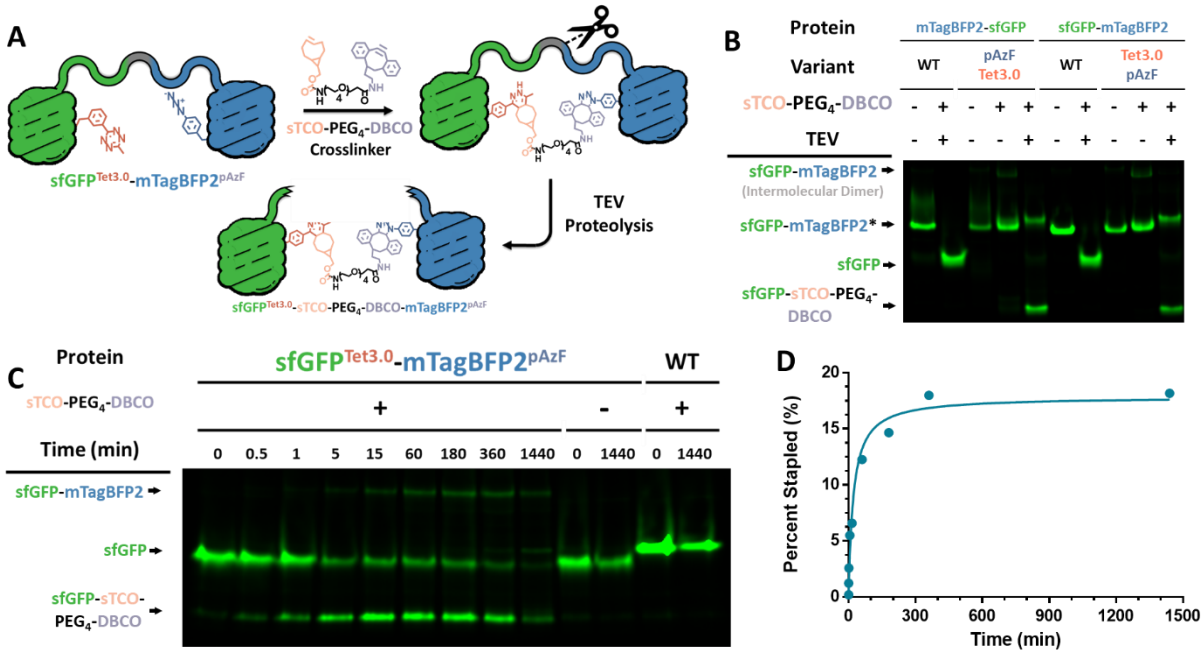

Figure 6. *In vitro* and *in vivo* intramolecular protein stapling via DEAL. (A) Reaction scheme of intramolecular stapling of sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} (10 μM) with sTCO-PEG$_4$-DBCO (10 μM) and subsequent TEV cleavage. (B) Analysis of *in vitro* protein stapling of sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} and mTagBFP2^{pAzF}-sfGFP$^{Tet3.0}$ with sTCO-PEG$_4$-DBCO, using SDS-PAGE fluorescent imaging. *refers to both sfGFP-mTagBFP2 (starting material) and sfGFP (stapled product), since they have similar mobility. (C) SDS-PAGE in-gel fluorescence image of *in vivo* reaction time course following the addition of sTCO-PEG$_4$-DBCO (20 μM) to *E. coli* cells containing expressed sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} or sfGFP^{WT}-mTagBFP2^{WT}. (D) Densitometry quantification of panel C. The percent stapled was estimated as the percentage of fluorescence that the upper, stapled product band constitutes of the total fluorescence in each lane at each time point. The contents of each lane are indicated above each gel.

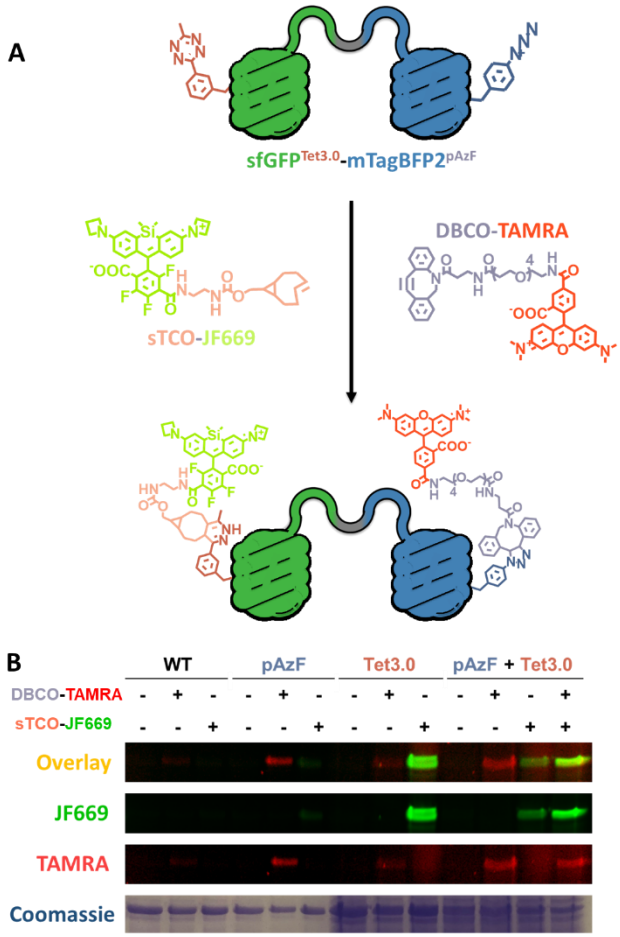

Figure 7. *In vivo* dual fluorophore labeling of dual encoded protein in *E. coli*. (A) Reaction scheme of dual fluorophore labeling of sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} with sTCO-JF669 and DBCO-TAMRA. (B) SDS-PAGE of resulting *E. coli* cell lysate after *in vivo* reactions in cells containing expressed sfGFP$^{Tet3.0}$-mTagBFP2^{pAzF} labeled with sTCO-JF669 and DBCO-TAMRA (1 μM for 5 and 60 min, respectively); top panel is the overlay of TAMRA and JF669 channels followed by each individual channel, and Coomassie staining. Cells exposed to fluorescent labels were quenched with excess pAzF and/or Tet3.0 prior to sample preparation.

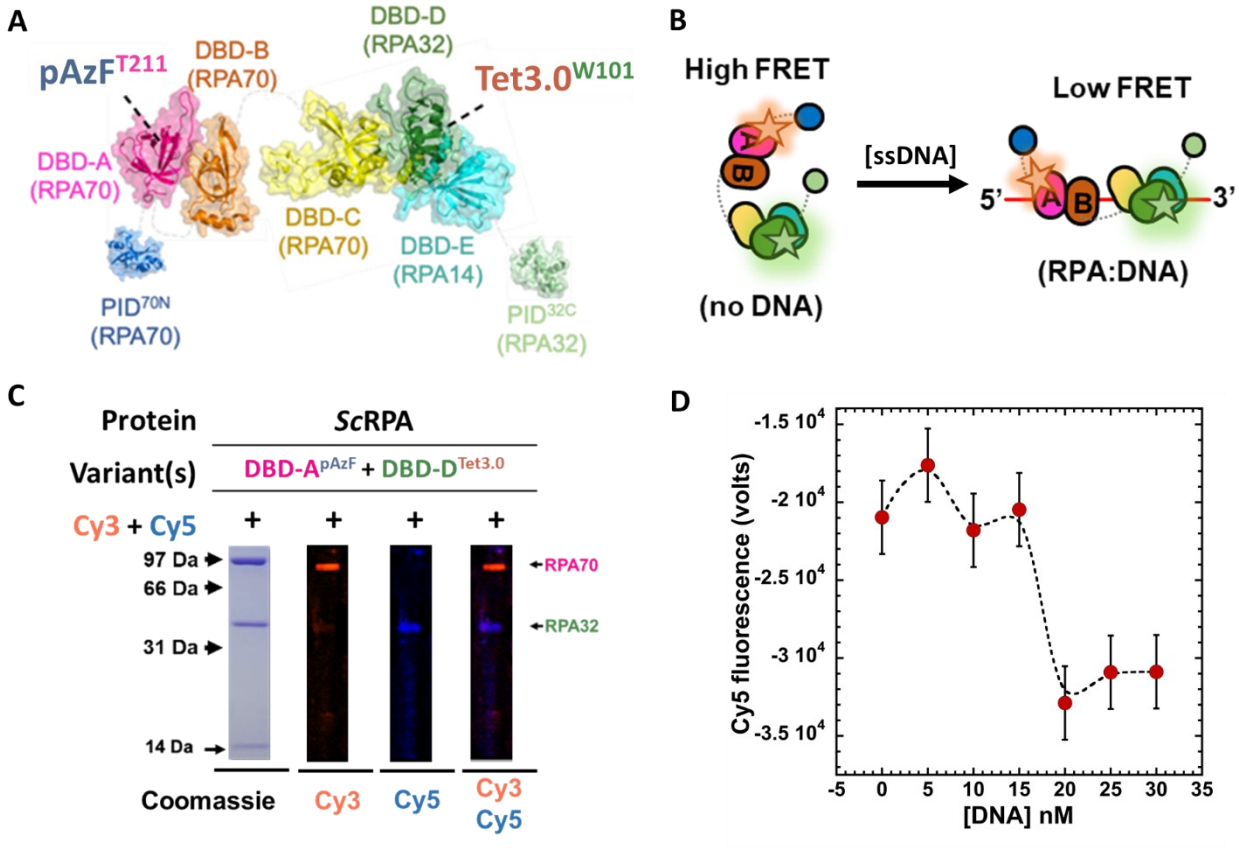

Figure 8. Characterization of *Sc*RPA labeling and FRET analysis of DNA binding. (A) Structural model of *S. cerevisiae* RPA complex (PDB: 6I52) highlighting the location of encoded pAzF and Tet3.0 in DBD-A and DBD-D, respectively. (B) Scheme depicting RPA binding ssDNA and the predicted FRET changes associated with these conformational states. (C) SDS-PAGE in-gel fluorescence analysis of *Sc*RPA after labeling by DBCO-Cy3 and TCO-Cy5. Each dye signal is imaged individually and as an overlay. (D) Resulting Cy5 fluorescence intensity from dual Cy3/Cy5 labeled *Sc*RPA^{Dual} as a function of DNA concentration.

References

(1) Oliveira, B. L.; Guo, Z.; Bernardes, G. J. L. Inverse Electron Demand Diels–Alder Reactions in Chemical Biology. *Chem. Soc. Rev.* **2017**, *46* (16), 4895–4950. https://doi.org/10.1039/C7CS00184C.

(2) Chen, X.; Wu, Y.-W. Selective Chemical Labeling of Proteins. *Organic & Biomolecular Chemistry* **2016**, *14* (24), 5417–5439. https://doi.org/10.1039/C6OB00126B.

(3) Zhang, G.; Zheng, S.; Liu, H.; Chen, P. R. Illuminating Biological Processes through Site-Specific Protein Labeling. *Chem Soc Rev* **2015**, *44* (11), 3405–3417. https://doi.org/10.1039/c4cs00393d.

(4) Young, D. D.; Schultz, P. G. Playing with the Molecules of Life. *ACS Chem Biol* **2018**, *13* (4), 854–870. https://doi.org/10.1021/acschembio.7b00974.

(5) Wang, K.; Sachdeva, A.; Cox, D. J.; Wilf, N. M.; Lang, K.; Wallace, S.; Mehl, R. A.; Chin, J. W. Optimized Orthogonal Translation of Unnatural Amino Acids Enables Spontaneous Protein Double-Labelling and FRET. *Nature Chemistry* **2014**, *6* (5), 393–403. https://doi.org/10.1038/nchem.1919.

(6) Wu, B.; Wang, Z.; Huang, Y.; Liu, W. R. Catalyst-Free and Site-Specific One-Pot Dual Labeling of a Protein Directed by Two Genetically Incorporated Noncanonical Amino Acids. *Chembiochem* **2012**, *13* (10), 1405–1408. https://doi.org/10.1002/cbic.201200281.

(7) Kim, J.; Seo, M.-H.; Lee, S.; Cho, K.; Yang, A.; Woo, K.; Kim, H.-S.; Park, H.-S. Simple and Efficient Strategy for Site-Specific Dual Labeling of Proteins for Single-Molecule Fluorescence Resonance Energy Transfer Analysis. *Anal. Chem.* **2013**, *85* (3), 1468–1474. https://doi.org/10.1021/ac303089v.

(8) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. Encoding Multiple Unnatural Amino Acids via Evolution of a Quadruplet-Decoding Ribosome. *Nature* **2010**, *464* (7287), 441–444. https://doi.org/10.1038/nature08817.

(9) Xiao, H.; Chatterjee, A.; Choi, S.; Bajjuri, K. M.; Sinha, S. C.; Schultz, P. G. Genetic Incorporation of Multiple Unnatural Amino Acids into Proteins in Mammalian Cells. *Angewandte Chemie International Edition* **2013**, *52* (52), 14080–14083. https://doi.org/10.1002/anie.201308137.

(10) Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. Finding the Right (Bioorthogonal) Chemistry. *ACS Chem. Biol.* **2014**, *9* (3), 592–605. https://doi.org/10.1021/cb400828a.

(11) Blizzard, R. J.; Backus, D. R.; Brown, W.; Bazewicz, C. G.; Li, Y.; Mehl, R. A. Ideal Bioorthogonal Reactions Using A Site-Specifically Encoded Tetrazine Amino Acid. *J Am Chem Soc* **2015**, *137* (32), 10044–10047. https://doi.org/10.1021/jacs.5b03275.

(12) Dommerholt, J.; Rutjes, F. P. J. T.; van Delft, F. L. Strain-Promoted 1,3-Dipolar Cycloaddition of Cycloalkynes and Organic Azides. *Top Curr Chem (Cham)* **2016**, *374* (2), 16. https://doi.org/10.1007/s41061-016-0016-4.

(13) Dommerholt, J.; Rutjes, F. P. J. T.; van Delft, F. L. Strain-Promoted 1,3-Dipolar Cycloaddition of Cycloalkynes and Organic Azides. *Top Curr Chem (Cham)* **2016**, *374* (2), 16. https://doi.org/10.1007/s41061-016-0016-4.

(14) Darko, A.; Wallace, S.; Dmitrenko, O.; Machovina, M. M.; Mehl, R. A.; Chin, J. W.; Fox, J. M. Conformationally Strained Trans-Cyclooctene with Improved Stability and Excellent Reactivity in Tetrazine Ligation. *Chem Sci* **2014**, *5* (10), 3770–3776. https://doi.org/10.1039/C4SC01348D.

(15) Blackman, M. L.; Royzen, M.; Fox, J. M. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels−Alder Reactivity. *J. Am. Chem. Soc.* **2008**, *130* (41), 13518–13519. https://doi.org/10.1021/ja8053805.

(16) Jang, H. S.; Jana, S.; Blizzard, R. J.; Meeuwsen, J. C.; Mehl, R. A. Access to Faster Eukaryotic Cell Labeling with Encoded Tetrazine Amino Acids. *Journal of the American Chemical Society* **2020**, 7245–7249. https://doi.org/10.1021/jacs.9b11520.

(17) Bednar, R. M.; Golbek, T. W.; Kean, K. M.; Brown, W. J.; Jana, S.; Baio, J. E.; Karplus, P. A.; Mehl, R. A. Immobilization of Proteins with Controlled Load and Orientation. *ACS Appl. Mater. Interfaces* **2019**, *11* (40), 36391–36398. https://doi.org/10.1021/acsami.9b12746.

(18) Dumas, A.; Lercher, L.; D. Spicer, C.; G. Davis, B. Designing Logical Codon Reassignment – Expanding the Chemistry in Biology. *Chemical Science* **2015**, *6* (1), 50–69. https://doi.org/10.1039/C4SC01534G.

(19) Karver, M. R.; Weissleder, R.; Hilderbrand, S. A. Bioorthogonal Reaction Pairs Enable Simultaneous, Selective, Multi-Target Imaging. *Angew Chem Int Ed Engl* **2012**, *51* (4), 920–922. https://doi.org/10.1002/anie.201104389.

(20) Macias-Contreras, M.; He, H.; Little, K. N.; Lee, J. P.; Campbell, R. P.; Royzen, M.; Zhu, L. SNAP/CLIP-Tags and Strain-Promoted Azide–Alkyne Cycloaddition (SPAAC)/Inverse Electron Demand Diels–Alder (IEDDA) for Intracellular Orthogonal/Bioorthogonal Labeling. *Bioconjugate Chemistry* **2020**, 1370–1381. https://doi.org/10.1021/acs.bioconjchem.0c00107.

(21) Plass, T.; Milles, S.; Koehler, C.; Szymański, J.; Mueller, R.; Wießler, M.; Schultz, C.; Lemke, E. A. Amino Acids for Diels–Alder Reactions in Living Cells. *Angewandte Chemie International Edition* **2012**, *51* (17), 4166–4170. https://doi.org/10.1002/anie.201108231.

(22) Meineke, B.; Heimgärtner, J.; Eirich, J.; Landreh, M.; Elsässer, S. J. Site-Specific Incorporation of Two NcAAs for Two-Color Bioorthogonal Labeling and Crosslinking of Proteins on Live Mammalian Cells. *Cell Reports* **2020**, *31* (12), 107811. https://doi.org/10.1016/j.celrep.2020.107811.

(23) Fang, Y.; Judkins, J. C.; Boyd, S. J.; am Ende, C. W.; Rohlfing, K.; Huang, Z.; Xie, Y.; Johnson, D. S.; Fox, J. M. Studies on the Stability and Stabilization of Trans-Cyclooctenes through Radical Inhibition and Silver (I) Metal Complexation. *Tetrahedron* **2019**, *75* (32), 4307–4317. https://doi.org/10.1016/j.tet.2019.05.038.

(24) Miyake-Stoner, S. J.; Miller, A. M.; Hammill, J. T.; Peeler, J. C.; Hess, K. R.; Mehl, R. A.; Brewer, S. H. Probing Protein Folding Using Site-Specifically Encoded Unnatural Amino

Acids as FRET Donors with Tryptophan. *Biochemistry* **2009**, *48* (25), 5953–5962. https://doi.org/10.1021/bi900426d.

(25) Kobayashi, T.; Nureki, O.; Ishitani, R.; Yaremchuk, A.; Tukalo, M.; Cusack, S.; Sakamoto, K.; Yokoyama, S. Structural Basis for Orthogonal TRNA Specificities of Tyrosyl-TRNA Synthetases for Genetic Code Expansion. *Nat Struct Biol* **2003**, *10* (6), 425–432. https://doi.org/10.1038/nsb934.

(26) Tharp, J. M.; Ehnbom, A.; Liu, W. R. TRNAPyl: Structure, Function, and Applications. *RNA Biol* **2017**, *15* (4–5), 441–452. https://doi.org/10.1080/15476286.2017.1356561.

(27) Engelberg-Kulka, H. UGA Suppression by Normal TRNA Trp in Escherichia Coli: Codon Context Effects. *Nucleic Acids Res* **1981**, *9* (4), 983–991.

(28) Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. An Enhanced System for Unnatural Amino Acid Mutagenesis in E. Coli. *J Mol Biol* **2010**, *395* (2), 361–374. https://doi.org/10.1016/j.jmb.2009.10.030.

(29) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia Coli. *Biochemistry* **2013**, *52* (10), 1828–1837. https://doi.org/10.1021/bi4000244.

(30) Dunkelmann, D. L.; Willis, J. C. W.; Beattie, A. T.; Chin, J. W. Engineered Triply Orthogonal Pyrrolysyl–TRNA Synthetase/TRNA Pairs Enable the Genetic Encoding of Three Distinct Non-Canonical Amino Acids. *Nature Chemistry* **2020**, *12* (6), 535–544. https://doi.org/10.1038/s41557-020-0472-x.

(31) O'Donoghue, P.; Prat, L.; Heinemann, I. U.; Ling, J.; Odoi, K.; Liu, W. R.; Söll, D. Near-Cognate Suppression of Amber, Opal and Quadruplet Codons Competes with Aminoacyl-TRNAPyl for Genetic Code Expansion. *FEBS Lett* **2012**, *586* (21), 3931–3937. https://doi.org/10.1016/j.febslet.2012.09.033.

(32) Bazewicz, C. G.; Liskov, M. T.; Hines, K. J.; Brewer, S. H. Sensitive, Site-Specific, and Stable Vibrational Probe of Local Protein Environments: 4-Azidomethyl-L-Phenylalanine. *J Phys Chem B* **2013**, *117* (30), 8987–8993. https://doi.org/10.1021/jp4052598.

(33) Murrey, H. E.; Judkins, J. C.; Am Ende, C. W.; Ballard, T. E.; Fang, Y.; Riccardi, K.; Di, L.; Guilmette, E. R.; Schwartz, J. W.; Fox, J. M.; Johnson, D. S. Systematic Evaluation of Bioorthogonal Reactions in Live Cells with Clickable HaloTag Ligands: Implications for Intracellular Imaging. *J Am Chem Soc* **2015**, *137* (35), 11461–11475. https://doi.org/10.1021/jacs.5b06847.

(34) Palzer, S.; Bantel, Y.; Kazenwadel, F.; Berg, M.; Rupp, S.; Sohn, K. An Expanded Genetic Code in Candida Albicans To Study Protein-Protein Interactions In Vivo. *Eukaryotic Cell* **2013**, *12* (6), 816–827. https://doi.org/10.1128/EC.00075-13.

(35) Tian, H.; Sakmar, T. P.; Huber, T. Micelle-Enhanced Bioorthogonal Labeling of Genetically-Encoded Azido Groups on the Lipid-Embedded Surface of a GPCR. *Chembiochem* **2015**, *16* (9), 1314–1322. https://doi.org/10.1002/cbic.201500030.

(36) Subach, O. M.; Cranfill, P. J.; Davidson, M. W.; Verkhusha, V. V. An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. *PLoS One* **2011**, *6* (12), e28674. https://doi.org/10.1371/journal.pone.0028674.

(37) Rutkowska, A.; Plass, T.; Hoffmann, J.-E.; Yushchenko, D. A.; Feng, S.; Schultz, C. T-CrAsH: A Heterologous Chemical Crosslinker. *ChemBioChem* **2014**, *15* (12), 1765–1768. https://doi.org/10.1002/cbic.201402189.

(38) Cooley, R. B.; O'Donnell, J. P.; Sondermann, H. Coincidence Detection and Bi-Directional Transmembrane Signaling Control a Bacterial Second Messenger Receptor. *eLife* **2016**, *5*, e21848. https://doi.org/10.7554/eLife.21848.

(39) Jacobs, J. F. M.; van der Molen, R. G.; Bossuyt, X.; Damoiseaux, J. Antigen Excess in Modern Immunoassays: To Anticipate on the Unexpected. *Autoimmunity Reviews* **2015**, *14* (2), 160–167. https://doi.org/10.1016/j.autrev.2014.10.018.

(40) Moore, E. J.; Zorine, D.; Hansen, W. A.; Khare, S. D.; Fasan, R. Enzyme Stabilization via Computationally Guided Protein Stapling. *PNAS* **2017**, *114* (47), 12472–12477. https://doi.org/10.1073/pnas.1708907114.

(41) Pellett, P. A.; Sun, X.; Gould, T. J.; Rothman, J. E.; Xu, M.-Q.; Corrêa, I. R.; Bewersdorf, J. Two-Color STED Microscopy in Living Cells. *Biomed Opt Express* **2011**, *2* (8), 2364–2371. https://doi.org/10.1364/BOE.2.002364.

(42) Margineanu, A.; Chan, J. J.; Kelly, D. J.; Warren, S. C.; Flatters, D.; Kumar, S.; Katan, M.; Dunsby, C. W.; French, P. M. W. Screening for Protein-Protein Interactions Using Förster Resonance Energy Transfer (FRET) and Fluorescence Lifetime Imaging Microscopy (FLIM). *Scientific Reports* **2016**, *6* (1), 28186. https://doi.org/10.1038/srep28186.

(43) Caldwell, C. C.; Spies, M. Dynamic Elements of Replication Protein A at the Crossroads of DNA Replication, Recombination, and Repair. *Critical Reviews in Biochemistry and Molecular Biology* **2020**, *55* (5), 482–507. https://doi.org/10.1080/10409238.2020.1813070.

(44) Maréchal, A.; Zou, L. RPA-Coated Single-Stranded DNA as a Platform for Post-Translational Modifications in the DNA Damage Response. *Cell Research* **2015**, *25* (1), 9–23. https://doi.org/10.1038/cr.2014.147.

(45) Fan, J.; Pavletich, N. P. Structure and Conformational Change of a Replication Protein A Heterotrimer Bound to SsDNA. *Genes Dev.* **2012**, *26* (20), 2337–2347. https://doi.org/10.1101/gad.194787.112.

(46) Pokhrel, N.; Caldwell, C. C.; Corless, E. I.; Tillison, E. A.; Tibbs, J.; Jocic, N.; Tabei, S. M. A.; Wold, M. S.; Spies, M.; Antony, E. Dynamics and Selective Remodeling of the DNA-Binding Domains of RPA. *Nature Structural & Molecular Biology* **2019**, *26* (2), 129–136. https://doi.org/10.1038/s41594-018-0181-y.

(47) Yates, L. A.; Aramayo, R. J.; Pokhrel, N.; Caldwell, C. C.; Kaplan, J. A.; Perera, R. L.; Spies, M.; Antony, E.; Zhang, X. A Structural and Dynamic Model for the Assembly of Replication Protein A on Single-Stranded DNA. *Nature Communications* **2018**, *9* (1), 5447. https://doi.org/10.1038/s41467-018-07883-7.

(48) Nödling, A. R.; Spear, L. A.; Williams, T. L.; Luk, L. Y. P.; Tsai, Y.-H. Using Genetically Incorporated Unnatural Amino Acids to Control Protein Functions in Mammalian Cells. *Essays in Biochemistry* **2019**, *63* (2), 237–266. https://doi.org/10.1042/EBC20180042.

(49) Widder, P.; Berner, F.; Summerer, D.; Drescher, M. Double Nitroxide Labeling by Copper-Catalyzed Azide–Alkyne Cycloadditions with Noncanonical Amino Acids for Electron Paramagnetic Resonance Spectroscopy. *ACS Chem. Biol.* **2019**, *14* (5), 839–844. https://doi.org/10.1021/acschembio.8b01111.

(50) Xue, L.; Prifti, E.; Johnsson, K. A General Strategy for the Semisynthesis of Ratiometric Fluorescent Sensor Proteins with Increased Dynamic Range. *J. Am. Chem. Soc.* **2016**, *138* (16), 5258–5261. https://doi.org/10.1021/jacs.6b03034.

(51) Tsai, Y.-H.; Essig, S.; James, J. R.; Lang, K.; Chin, J. W. Selective, Rapid and Optically Switchable Regulation of Protein Function in Live Mammalian Cells. *Nat Chem* **2015**, *7* (7), 554–561. https://doi.org/10.1038/nchem.2253.

(52) Cui, Y.; Zhang, X.; Yu, M.; Zhu, Y.; Xing, J.; Lin, J. Techniques for Detecting Protein-Protein Interactions in Living Cells: Principles, Limitations, and Recent Progress. *Sci. China Life Sci.* **2019**, *62* (5), 619–632. https://doi.org/10.1007/s11427-018-9500-7.

(53) Gwyther, R. E. A.; Jones, D. D.; Worthy, H. L. Better Together: Building Protein Oligomers Naturally and by Design. *Biochem Soc Trans* **2019**, *47* (6), 1773–1780. https://doi.org/10.1042/BST20190283.

(54) Stephanopoulos, N. Hybrid Nanostructures from the Self-Assembly of Proteins and DNA. *Chem* **2020**, *6* (2), 364–405. https://doi.org/10.1016/j.chempr.2020.01.012.

For Table of Contents Only