Published in final edited form as:
ACS Chem Biol. 2013 August 16; 8(8): 1664–1670. doi:10.1021/cb400267m.

A Genetically Encoded Acrylamide Functionality

Yan-Jiun C. Lee, **Bo Wu**, **Jeffrey Edward Raymond**, **Yu Zeng**, **Xinqiang Fang**, **Karen L. Wooley**, and **Wenshe R. Liu***
Department of Chemistry, Texas A&M University, College Station, TX 77843

Abstract

N^{ε}-acryloyl-L-lysine, a noncanonical amino acid with an electron deficient olefin, is genetically encoded in *Escherichia coli* using a pyrrolysyl-tRNA synthetase mutant in coordination with $\mathrm{tRNA}^{\mathrm{Pyl}}_{\mathrm{CUA}}$. The acrylamide moiety is stable in cells, whereas it is active enough to perform a diverse set of unique reactions for protein modifications *in vitro*. These reactions include 1,4-addition, radical polymerization, and 1,3-dipolar cycloaddition. We demonstrate that a protein incorporated with N^{ε}-acryloyl-L-lysine is efficiently modified with thiol-containing nucleophiles at slightly alkali conditions and the acrylamide moiety also allows rapid radical copolymerization of the same protein into a polyacrylamide hydrogel at physiological pH. At physiological conditions, the acrylamide functionality undergoes a fast 1,3-dipolar cycloaddition reaction with diaryl nitrile imine to show turn-on fluorescence. We have used this observation to demonstrate site-specific fluorescent labeling of proteins incorporated with N^{ε}-acryloyl-L-lysine both *in vitro* and in living cells. This critical development allows easy access to an array of modified proteins for applications where high specificity and reaction efficiency are needed.

Selective labeling has its great utility as a tool to interrogate physiological functions of proteins and develop biotechnological applications. To label a protein of interest site-specifically, the genetic noncanonical amino acid (NAA) incorporation approach has drawn great attention due to its potentially broad applications.[1, 2] Many NAAs have been genetically incorporated into proteins in bacteria, yeast, mammalian cells and even animals.[3–8] Of these NAAs, a few bioorthogonal functionalities are available, including ketone, alkyne, and azide that undergo biocompatible reactions such as hydrazone/oxime formation and azide-alkyne cycloaddition reactions for site-selective protein modifications.[9–18] Although ketone, alkyne, and azide are involved in diverse organic reaction types beyond the aforementioned bioorthogonal reactions, most of these reactions either cannot be realized in physiological conditions or lack selectivity in the biological system. A genetically encoded and simple chemical moiety that undergoes multiple types of biocompatible and selective reactions could have broad implications for selective labeling strategies *in vitro* and *in vivo*. One functional group that shows many facets of organic reactivity is the carbon-carbon double bond. Given the versatility in organic transformation, olefins have found extensive application in synthetic organic chemistry. Various biocompatible olefin reactions have been demonstrated elsewhere. These include 1,4-addition, olefin metathesis, thiol-ene click, Diels-Alder, and 1,3-dipolar cycloaddition reactions.[19, 20] The olefin functionality has been genetically incorporated into proteins. Using an evolved amber-suppressing aminoacyl-tRNA synthetase-tRNA pair that was

***Corresponding author** wliu@chem.tamu.edu.

Authors declare no competing interest.

Supporting Information

NAA synthesis, plasmid constructions, clone identification, protein expression, kinetic analysis, protein labeling, confocal imaging, and additional ESI-MS spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

derived from *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase-$tRNA^{Tyr}$ pair, Zhang *et al*. were able to incorporate *O*-allyl-L-tyrosine into proteins in *E. coli*.[21] The same NAA and two other aliphatic olefin-containing tyrosine analogs were also genetically encoded in *E. coli* using a mutant pyrrolysyl-tRNA synthetase (PylRS) – $tRNA^{Pyl}_{CUA}$ pair.[22] A photoclick reaction that exploits the nitrile imine-alkene cycloaddition reaction was applied to fluorescently label proteins with *O*-allyl-L-tyrosine incorporated in *E. coli*, albeit the labeling efficiency was low.[23] Five other aliphatic olefin-containing NAAs have also been encoded in *Saccharomyces cerevisiae* and used to undergo olefin metathesis on proteins.[24] In general, these aliphatic olefin-containing NAAs have low reactivity toward nitrile imine and are not capable of performing 1,4-addition under physiological conditions. Recently, several groups have used wild-type and evolved *PylRS* - $tRNA^{Pyl}_{CUA}$ pairs to site-specifically install NAAs with a strain-promoted alkene into proteins.[25–30] Reacting rapidly with tetrazine or a tetrazole photo-cleavage product, these NAAs contain a norbornene, transcyclooctene, or cyclopropene moiety that is a relatively large functional group. These NAAs may not be optimal options when minimal structural perturbation of proteins is necessary. Acrylamide is a small electron deficient olefin that undergoes several potentially biocompatible reactions, e.g., those shown in Scheme 1. Here, we wish to report its genetic incorporation and its mediated diverse reactions for protein modifications. We envision that the acrylamide moiety would serve as a user-defined multifunctional chemical handle for different purposes of applications.

Due to the electrophilic nature of acrylamide, one may suspect it undergoes 1,4-addition with thiol nucleophiles such as intracellular glutathione. The kinetics of 1,4-addition of β-mercaptoethanol (β-ME) to acrylamide was investigated. β-ME instead of glutathione was used in our study due to the simple NMR spectrum of its 1,4-addition product with acrylamide. At pH 7.4, the second order rate constant determined was 0.0040±0.0003 $M^{-1}s^{-1}$ (Supplementary Figure 1) that is twice a reported rate constant (0.002 $M^{-1}s^{-1}$) of the reaction between glutathione and acrylamide at the same pH.[31] Providing the intracellular reduced glutathione concentration at 5 mM,[32] acrylamide will have at least a half-life of 10 h that is ample for protein expression. A parallel kinetic analysis at pH 8.0 inferred a second order rate constant of 0.013±0.001 M^{-1} s^{-1} (Supplementary Figure 2), indicating a linear dependence of the logarithm of the rate constant on pH. Therefore, an acrylamide-containing protein could be produced in cells and then modified at slightly alkali conditions *in vitro*. Encouraged by the initial kinetic studies, we synthesized N^{ε}-acryloyl-L-lysine (AcrK) and N^{ε}-crotonyl-L-lysine (CrtK) whose structures are shown in Figure 1A and then searched mutant PylRS – $tRNA^{Pyl}_{CUA}$ pairs for their genetic encoding in *E. coli*. Two more stable analogous NAAs, N^{ε}-propionyl-L-lysine (PrK) and N^{ε}-butyryl-L-lysine (BuK), were also synthesized and used during the initial search of mutant PylRS – $tRNA^{Pyl}_{CUA}$ pairs. Due to the structural similarity between N^{ε}-acetyl-L-lysine (AcK), PrK and AcrK, MmAcKRS1 that was evolved from *Methanosarcina mazei* PylRS for the AcK incorporation also recognizes PrK and AcrK.[33] To improve its efficiency, we constructed a single-site MmAcKRS1 library that randomized Y384. Mutations at Y384 have been shown in mutant PylRS variants for other lysine derivatives.[34, 35] Clones from this library were screened against PrK using a protocol described previously.[36] Most of the positive clones converged to a same MmAcKRS1 variant that has the Y384W mutation and is coined as PrKRS. The following testing with AcrK indicated that the PrKRS – $tRNA^{Pyl}_{CUA}$ pair also mediates a high-level incorporation of AcrK at an amber codon (Supplementary Figure 3). It was demonstrated previously that wild-type PylRS recognizes CrtK albeit at a low level.[37, 38] To improve the binding, a single site PylRS library with randomization at Y384 was constructed. Clones of this library were screened against BuK. Most positive clones

converged to a single PylRS variant that has the Y384W mutation and is defined as BuKRS. BuKRS also recognizes CrtK (Supplementary Figure 3).

For further characterization of the genetic incorporation of AcrK and CrtK, PrKRS and BuKRS genes were cloned separately into an optimized pEVOL vector under control of an *araBAD* promoter to afford plasmids pEVOL-PrKRS and pEVOL-BuKRS that also contain a $\mathrm{tRNA}^{\mathrm{Pyl}}_{\mathrm{CUA}}$ gene under control of the *proK* promoter. Transforming BL21(DE3) cells that contained a reporter plasmid pET-sfGFP2TAG with pEVOL-PrKRS or pEVOL-BuKRS afforded cell strains for further protein expression. pET-sfGFP2TAG has a His-tagged superfolder green fluorescent protein (sfGFP) gene that contains an amber mutation at its S2 position and is under control of a IPTG-inducible T7 promoter.[22] Growing cells and inducing with the addition of 0.2% arabinose, 1 mM IPTG, and 2 mM AcrK or CrtK resulted in the overexpression of sfGFP. When AcrK or CrtK is absent, only a very small amount of sfGFP could be observed (Figure 1B). Purified proteins were then subjected to electrospray ionization mass spectrometry (ESI-MS) analysis that showed the expected molecular weights (For AcrK, calculated: 27,764 Da; detected: 27,762 Da. For CrtK, calculated: 27,778 Da; detected: 27,778 Da). Although proteins were expressed for 8 h at 37 °C, there was no apparent 1,4-addition product that could be detected by ESI-MS, indicating intracellular stability for both AcrK and CrtK (Figure 1C).

To demonstrate 1,4-addition for protein labeling, we chose to work with sfGFP that had AcrK incorporated at S2 (sfGFP-S2AcrK), due to the known steric effect of the methyl group to hinder the addition of a thiol nucleophile to CrtK. sfGFP-S2AcrK (39 μM) was incubated with β-ME (40 mM) at 37 °C overnight. pH 8.8 was used to elevate the reaction rate. The reaction proceeded smoothly with little original sfGFP-S2AcrK left after an overnight incubation (Supplementary Figure 4). Besides the expected β-ME addition adduct, the ESI-MS spectrum also displayed two additional peaks that are 76 and 151 Da greater than the expected product, indicating that one and two additional β-ME molecules were covalently linked to sfGFP-S2AcrK. Since the β-ME adducts of proteins through the disulfide bond linkage were observed previously[33] and sfGFP-S2AcrK has two cysteine residues, these two additional peaks clearly resulted from the disulfide covalent linkage of β-ME to sfGFP-S2AcrK. These extra β-ME additions could be removed by further treatment with 5 mM dithiothreitol (DTT) for 1 h (Supplementary Figure 4). A similar reaction was carried out to covalent link sfGFP-S2AcrK to methoxypolyethylene glycol thiol 5 kDa (mPEGSH5k). sfGFP-S2AcrK (50 μM) was incubated with mPEGSH5k (40 mM) at pH 8.8, 37 °C for 8 h and then treated with DTT (5 mM) for an additional hour. The following SDS-PAGE analysis showed two protein bands, with one around 5 kDa larger than the other. The relative intensities of the two protein bands indicated roughly 50% conversion (Figure 2). The same reaction with wild-type (WT) sfGFP did not lead to a modified protein with a significant signal that could be detected by SDS-PAGE.

Radical polymerization has been applied to undergo selective protein modifications in physiological conditions.[39, 40] One typical chemical moiety for radical polymerization is acrylamide. To demonstrate that the acrylamide moiety in sfGFP-S2AcrK can serve as a monomer for radical polymerization, a copolymer hydrogel was prepared by mixing sfGFP-S2AcrK (110 μM), acrylamide (15%), and bis-acrylamide (0.5%) and initiating polymerization with the addition of tetramethylethylenediamine (0.3%) and ammonium persulfate (3%). The reaction solution was loaded to a well of a native PAGE gel and allowed to solidify. By this process, sfGFP-S2AcrK was covalently immobilized in the hydrogel and did not migrate during the following electrophoresis process (Figure 3). However, a similar test with WT sfGFP showed that the majority of the protein did not form covalent linkages with hydrogel. Moreover, as indicated by the fluorescence maintenance of

the immobilized sfGFP-S2AcrK, this radical copolymerization reaction is compatible with preservation of the native protein state of sfGFP.

It was demonstrated previously that olefin reacts with diaryl nitrile imine to generate a fluorescent pyrazoline at an appreciable rate in mild aqueous conditions.[41] Using a hydrazonoyl chloride **1** as a precursor that undergoes rapid dissociation in water to form a diaryl nitrile imine (**2**), Kaya *et al.* demonstrated a genetically encoded norbornene moiety could be used to fluorescently label proteins.[29] However, a systematic kinetic investigation of the olefin reaction with diaryl nitrile imine is absent. **1** was synthesized according to the literature procedures. Reactions of **1** (5 μM) with acrylamide were examined by detecting the fluorescent emission from the product. The time-dependent fluorescent emission increase at different acrylamide concentrations is presented in Figure 4A. They all show single exponential curvatures. The determined apparent reaction rate constants have a linear dependence on the acrylamide concentrations with a significant y-axis intercept (Figure 4B). The reactions between **1** and other olefins shown in Figure 4C all display equivalent kinetic features (Supplementary Figures 5–9). Figure 4A also shows very different fluorescent intensities for the final products at different acrylamide concentrations. All these observations indicate that **2** underwent two parallel reactions in an aqueous buffer, one with olefin and the other with water (Figure 4C).[41] Based on the mechanism in Figure 4C, the pyrazoline formation follows the kinetic Equation 1, where k_1 is the second order rate constant of the reaction between **2** and an olefin, k_2 the hydrolysis rate constant of **2**, [O] the olefin concentration, and [NI] the concentration of **2**. Using this equation to analyze all the data resulted in k_1 and k_2 for various olefins shown in the table in Figure 4C. All determined k_2 values are similar, supporting the parallel reaction mechanism. The results also indicated that acrylamide reacted with diaryl nitrile imine much faster than with all the other tested olefins, specifically 3 times faster than norbornene. This observation is consistent with a previous report.[42] The hydrolysis of **2** was also analyzed by examining the UV absorbance decay of **2** at 315 nm in a buffer that contained acetonitrile and PBS with a ratio of 1:1 (Supplementary Figure 10) and gave a first order rate constant of 0.00034 ± 0.00001 s^{-1}, similar to the k_2 values shown in Figure 4C.

To demonstrate the nitrile imine-olefin cycloaddition for the fluorescent protein labeling, we expressed sfGFP variants that were incorporated with AcrK, CrtK, and three other olefin-containing tyrosine analogs, illustrated in Figure 5A as **3**–**5**, at the sfGFP S2 position. The genetic incorporation of **3** and **4** was demonstrated previously using a rationally designed PylRS mutant.[22] This same mutant also recognizes **5**. sfGFP variants (14 μM) were incubated with **1** (70 μM) in a Tris-HCl buffer (pH 7.4) at RT for 1 h, fully denatured, and then analyzed by SDS-PAGE. The cyclization product of sfGFP-S2AcrK and sfGFP with CrtK incorporated at S2 (sfGFP-S2CrtK) displayed detectable fluorescent emission when irradiated by UV light (Figure 5A). The signal from sfGFP-S2CrtK is relatively weak and was estimated as 17% of that from sfGFP-S2AcrK. This was consistent with the simulation results for reactions of **1** with acrylamide and crotonamide (Supplementary Figure 11). However, sfGFP variants incorporated with **3**–**5** did not show significant fluorescent signals after their reactions with **1**. Prolonging the reaction time did not appreciably improve the labeling efficiency. Giving the low cyclization reactivity of **3**–**5** with **1**, **1** preferentially underwent hydrolysis, leading to low labeling efficiency (Supplementary Figure 11). To verify the cyclization product of sfGFP-S2AcrK, it was subjected to ESI-MS analysis. The major detected peak (28,085 Da) agrees well with the calculated labeling product (28,083 Da) (Figure 5B). The original protein could be barely detected.

To explore the possibility of labeling the acrylamide moiety with **1** in living cells, plasmid pETDuet-OmpXTAG was constructed for expressing the *E. coli* outer membrane protein OmpX incorporated with AcrK (OmpX-AcrK). The OmpX gene in pETDuet-OmpXTAG

has an AAAAXAA (A denotes alanine; X denotes an amber mutation) insertion between its two extracellular OmpX residues S53 and S54. Transforming BL21(DE3) cells with pEVOL-PrKRS and pETDuet-OmpXTAG and growing the transformed cells in LB medium supplemented with 1 mM IPTG and 5 mM AcrK afforded OmpX-AcrK overexpression (Figure 6A). Incubating the lysate of cells expressing OmpX-AcrK with **1** (0.5 mM) at RT for 1 h led to a detectable fluorescent band in a SDS-PAGE gel that corresponded well with the Coomassie blue stained OmpX-AcrK in the same gel. While many proteins in the lysate were present at higher expression levels, in-gel fluorescence imaging showed that only the expressed OmpX-AcrK could be specifically labeled, giving nearly no detectable background (Figure 6A). With the success of the cell lysate labeling, experiments were conducted to fluorescently label living cells with **1** by incubating the PBS-washed *E. coli* that overexpressed OmpX-AcrK with **1** (0.5 mM) at RT for 1 h. The cells were then washed with an isotonic saline solution and subjected to confocal microscopy imaging. *E. coli* cells expressing OmpX-AcrK showed detectable fluorescence (Figure 6B), indicating the cycloaddition is biocompatible and highly selective. The low variation in florescence lifetime profiles for all cells indicates a uniformity of dye states consistent with single-site covalent binding of the dye to the cell membrane and not of multiple dye environments typical of electrostatically bound or internalized dyes (Supplementary Figure 12). We also noticed that the labeled OmpX-AcrK is not uniformly distributed on the *E. coli* cell membrane. This preferential accumulation of OmpX at the two poles of *E. coli* needs to be further investigated. The same reaction to label *E. coli* cells without OmpX-AcrK did not yield any fluorescent cells that could be detected by confocal fluorescent microscopy (Supplementary Figure 13).

In summary, we report a method for the genetic installation of a small electron deficient olefin, acrylamide, into recombinant proteins in *E. coli*. Being a multifaceted chemical group for functionalizing proteins, the acrylamide moiety is relatively stable in physiological conditions and undergoes a number of biocompatible and unique reactions. We demonstrate that acrylamide-containing AcrK can be genetically encoded using an evolved $\mathrm{PrKRS} - \mathrm{tRNA}^{\mathrm{Pyl}}_{\mathrm{CUA}}$ pair and the genetically incorporated AcrK mediated efficient 1,4-addition, radical copolymerization, and 1,3-diplor cycloaddition reactions on proteins. These unique reactions provide a direct route to applications that include synthesizing PEGylated therapeutic proteins for disease intervention and generating protein-immobilizing hydrogels for ELISA, microarray/microfluidic detections. One reaction is proven viable for site-selective protein labeling in living cells for protein functional investigations. The 1,4-addition reaction to PEGylate an acrylamide in a protein produces a selective and simple thioether linkage that is likely well tolerated in the biological system, potentially providing an advantage over the existing PEGylation techniques. The nitrile imine-acrylamide cycloaddition reaction presented in this study displays fast kinetics and high selectivity and is performed without catalyst or light activation. This reaction can be potentially used as an excellent alternative for *in vivo* protein labeling to the most popular click reaction type, the azide-alkyne cycloaddition that requires a Cu(I) catalyst and therefore may disrupt certain cellular functions.[43, 44] The current study also identified mutant PylRS variants that allow the synthesis of proteins incorporated with PrK, BuK, and CrtK. Propionylation, butylation, and crotonylation are naturally existing posttranslational lysine modifications of proteins.[45, 46] In line with others,[37, 38] we have provided approaches that can be applied to investigate these novel posttranslational modifications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

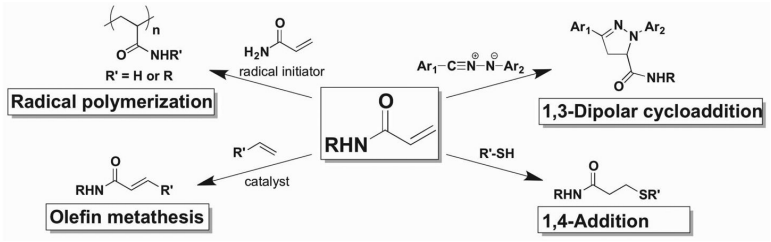
This work was supported in part by the National Institute of Health (grant 1R01CA161158 to W.R.L.), the National Science Foundation (grant CHEM-1148684), and the Welch Foundation (grant A-1715 to W.R.L.). We also thank Dr. Y. H. Rezenom from Laboratory for Biological Mass Spectrometry at Texas A&M University for characterizing proteins with electrospray ionization mass spectrometry.

References

1. Wang L, Schultz PG. Expanding the genetic code. Angew. Chem. Int. Ed. 2004; 44:34–66.
2. Liu WR, Wang YS, Wan W. Synthesis of proteins with defined posttranslational modifications using the genetic noncanonical amino acid incorporation approach. Mol. Biosyst. 2011; 7:38–47. [PubMed: 21088799]
3. Liu CC, Schultz PG. Adding new chemistries to the genetic code. Annu. Rev. Biochem. 2010; 79:413–444. [PubMed: 20307192]
4. Liu W, Brock A, Chen S, Chen S, Schultz PG. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. Nat. Methods. 2007; 4:239–244. [PubMed: 17322890]
5. Greiss S, Chin JW. Expanding the genetic code of an animal. J. Am. Chem. Soc. 2011; 133:14196–14199. [PubMed: 21819153]
6. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 2001; 292:498–500. [PubMed: 11313494]
7. Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG. An expanded eukaryotic genetic code. Science. 2003; 301:964–967. [PubMed: 12920298]
8. Parrish AR, She X, Xiang Z, Coin I, Shen Z, Briggs SP, Dillin A, Wang L. Expanding the genetic code of Caenorhabditis elegans using bacterial aminoacyl-tRNA synthetase/tRNA pairs. ACS Chem. Biol. 2012; 7:1292–1302. [PubMed: 22554080]
9. Wang L, Zhang Z, Brock A, Schultz PG. Addition of the keto functional group to the genetic code of Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 2003; 100:56–61. [PubMed: 12518054]
10. Zeng H, Xie J, Schultz PG. Genetic introduction of a diketone-containing amino acid into proteins. Bioorg. Med. Chem. Lett. 2006; 16:5356–5359. [PubMed: 16934461]
11. Zhang Z, Smith BA, Wang L, Brock A, Cho C, Schultz PG. A new strategy for the site-specific modification of proteins in vivo. Biochemistry. 2003; 42:6735–6746. [PubMed: 12779328]
12. Huang Y, Wan W, Russell WK, Pai PJ, Wang Z, Russell DH, Liu W. Genetic incorporation of an aliphatic keto-containing amino acid into proteins for their site-specific modifications. Bioorg. Med. Chem. Lett. 2010; 20:878–880. [PubMed: 20074948]
13. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. J. Am. Chem. Soc. 2002; 124:9026–9027. [PubMed: 12148987]
14. Deiters A, Schultz PG. In vivo incorporation of an alkyne into proteins in Escherichia coli. Bioorg. Med. Chem. Lett. 2005; 15:1521–1524. [PubMed: 15713420]
15. Deiters A, Cropp TA, Mukherji M, Chin JW, Anderson JC, Schultz PG. Adding amino acids with novel reactivity to the genetic code of Saccharomyces cerevisiae. J. Am. Chem. Soc. 2003; 125:11782–11783. [PubMed: 14505376]
16. Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin JW. Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. J. Am. Chem. Soc. 2009; 131:8720–8721. [PubMed: 19514718]
17. Wang YS, Fang X, Chen HY, Wu B, Wang ZU, Hilty C, Liu WR. Genetic Incorporation of Twelve meta-Substituted Phenylalanine Derivatives Using a Single Pyrrolysyl-tRNA Synthetase Mutant. ACS Chem. Biol. 2013; 8:405–415. [PubMed: 23138887]
18. Hao Z, Song Y, Lin S, Yang M, Liang Y, Wang J, Chen PR. A readily synthesized cyclic pyrrolysine analogue for site-specific protein “click” labeling. Chem. Commun. 2011; 47:4502–4504.

19. Sletten EM, Bertozzi CR. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew. Chem. Int. Ed. 2009; 48:6974–6998.
20. Li YM, Yang MY, Huang YC, Song XD, Liu L, Chen PR. Genetically encoded alkenyl-pyrrolysine analogues for thiol-ene reaction mediated site-specific protein labeling. Chem. Sci. 2012; 3:2766–2770.
21. Zhang Z, Wang L, Brock A, Schultz PG. The selective incorporation of alkenes into proteins in Escherichia coli. Angew. Chem. Int. Ed. 2002; 41:2840–2842.
22. Wang YS, Fang X, Wallace AL, Wu B, Liu WR. A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. J. Am. Chem. Soc. 2012; 134:2950–2953. [PubMed: 22289053]
23. Wang Y, Song W, Hu WJ, Lin Q. Fast alkene functionalization in vivo by Photoclick chemistry: HOMO lifting of nitrile imine dipoles. Angew. Chem. Int. Ed. 2009; 48:5330–5333.
24. Ai HW, Shen W, Brustad E, Schultz PG. Genetically encoded alkenes in yeast. Angew. Chem. Int. Ed. 2010; 49:935–937.
25. Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. Genetic Encoding of bicyclononynes and transcyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. J. Am. Chem. Soc. 2012; 134:10317–10320. [PubMed: 22694658]
26. Lang K, Davis L, Torres-Kolbus J, Chou C, Deiters A, Chin JW. Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. Nat. Chem. 2012; 4:298–304. [PubMed: 22437715]
27. Borrmann A, Milles S, Plass T, Dommerholt J, Verkade JM, Wiessler M, Schultz C, van Hest JC, van Delft FL, Lemke EA. Genetic encoding of a bicyclo[6.1.0]nonyne-charged amino acid enables fast cellular protein imaging by metal-free ligation. Chembiochem. 2012; 13:2094–2099. [PubMed: 22945333]
28. Plass T, Milles S, Koehler C, Szymanski J, Mueller R, Wiessler M, Schultz C, Lemke EA. Amino acids for Diels-Alder reactions in living cells. Angew. Chem. Int. Ed. 2012; 51:4166–4170.
29. Kaya E, Vrabel M, Deiml C, Prill S, Fluxa VS, Carell T. A genetically encoded norbornene amino acid for the mild and selective modification of proteins in a copper-free click reaction. Angew. Chem. Int. Ed. 2012; 51:4466–4469.
30. Yu Z, Pan Y, Wang Z, Wang J, Lin Q. Genetically encoded cyclopropene directs rapid, photoclick-chemistry-mediated protein labeling in mammalian cells. Angew. Chem. Int. Ed. 2012; 51:10600–10604.
31. Giege R, Sissler M, Florentz C. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res. 1998; 26:5017–5035. [PubMed: 9801296]
32. Chapman JS, Diehl MA, Lyman RC. Biocide Susceptibility and Intracellular Glutathione in Escherichia-Coli. J. Ind. Microbiol. 1993; 12:403–407.
33. Umehara T, Kim J, Lee S, Guo LT, Soll D, Park HS. N-Acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess N-acetyl lysine specificity in vitro and in vivo. FEBS Lett. 2012; 586:729–733. [PubMed: 22289181]
34. Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N-epsilon-(o-Azidobenzyloxycarbonyl) lysine for Site-Specific Protein Modification. Chem. Biol. 2008; 15:1187–1197. [PubMed: 19022179]
35. Nguyen DP, Elliott T, Holt M, Muir TW, Chin JW. Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. J. Am. Chem. Soc. 2011; 133:11418–11421. [PubMed: 21736333]
36. Wang YS, Wu B, Wang Z, Huang Y, Wan W, Russell WK, Pai PJ, Moe YN, Russell DH, Liu WR. A genetically encoded photocaged Nepsilon-methyl-L-lysine. Mol. Biosyst. 2010; 6:1557–1560. [PubMed: 20711534]
37. Gattner MJ, Vrabel M, Carell T. Synthesis of epsilon-N-propionyl-, epsilon-N-butyryl-, and epsilon-N-crotonyl-lysine containing histone H3 using the pyrrolysine system. Chem. Commun. 2013; 49:379–381.

38. Kim CH, Kang M, Kim HJ, Chatterjee A, Schultz PG. Site-specific incorporation of epsilon-N-crotonyllysine into histones. Angew. Chem. Int. Ed. 2012; 51:7246–7249.
39. Brueggemeier SB, Wu D, Kron SJ, Palecek SP. Protein-acrylamide copolymer hydrogels for array-based detection of tyrosine kinase activity from cell lysates. Biomacromolecules. 2005; 6:2765–2775. [PubMed: 16153117]
40. Brueggemeier SB, Kron SJ, Palecek SP. Use of protein-acrylamide copolymer hydrogels for measuring protein concentration and activity. Anal. Biochem. 2004; 329:180–189. [PubMed: 15158476]
41. Song W, Wang Y, Qu J, Madden MM, Lin Q. A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. Angew. Chem. Int. Ed. 2008; 47:2832–2835.
42. Bast K, Christl M, Huisgen R, Mack W. 1,3-Dipolar Cycloadditions .73. Relative Dipolarophile Activities in Cycloadditions of Benzonitrile Oxide. Chem. Ber. 1973; 106:3312–3344.
43. Kolb HC, Finn MG, Sharpless KB. Click chemistry: Diverse chemical function from a few good reactions. Angew. Chem. Int. Ed. 2001; 40:2004–2021.
44. Link AJ, Tirrell DA. Cell surface labeling of Escherichia coli via copper(I)-catalyzed [3+2] cycloaddition. J. Am. Chem. Soc. 2003; 125:11164–11165. [PubMed: 16220915]
45. Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng J, Gu W, Zhao Y. Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol. Cell. Proteomics. 2007; 6:812–819. [PubMed: 17267393]
46. Xie Z, Dai J, Dai L, Tan M, Cheng Z, Wu Y, Boeke JD, Zhao Y. Lysine succinylation and lysine malonylation in histones. Mol. Cell. Proteomics. 2012; 11:100–107. [PubMed: 22389435]



Scheme 1.

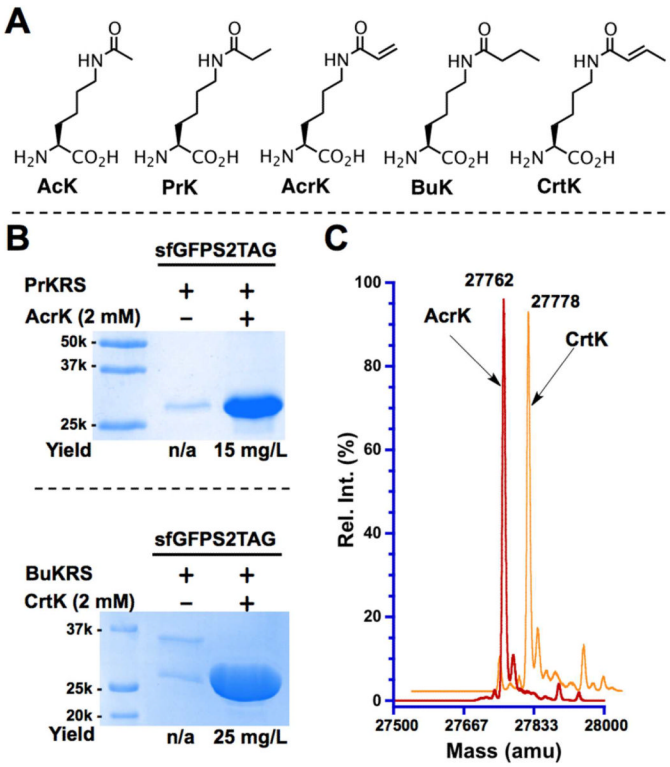


Figure 1.

(**A**) Structures of AcK, PrK, AcrK, BuK, and CrtK. (**B**) The site-selective incorporation of AcrK and CrtK at the S2 position of sfGFP. (**C**) ESI-MS of the purified proteins.

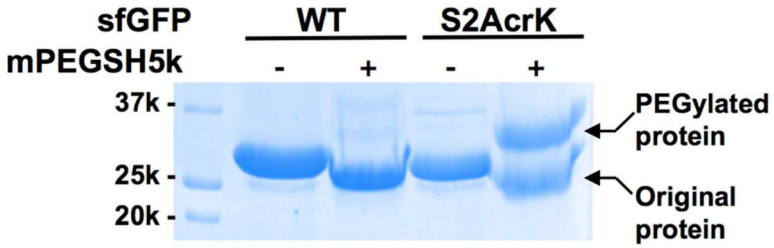


Figure 2.
SDS-PAGE analysis of WT sfGFP and sfGFP-S2AcrK before and after their reactions with mPEGSH5k.

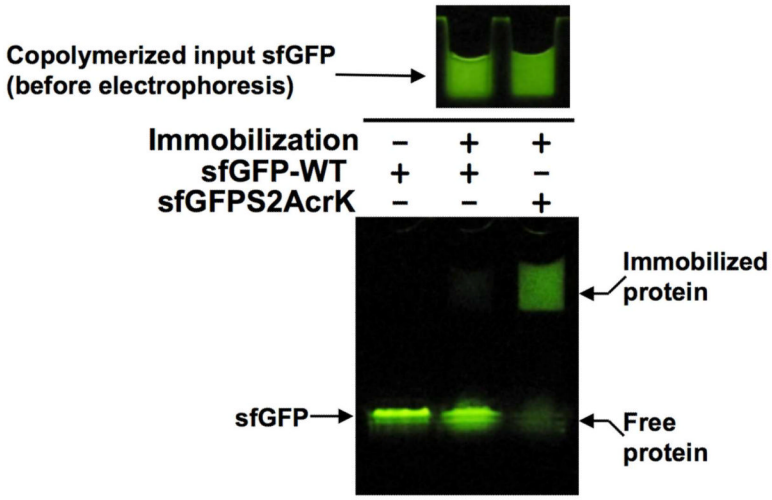
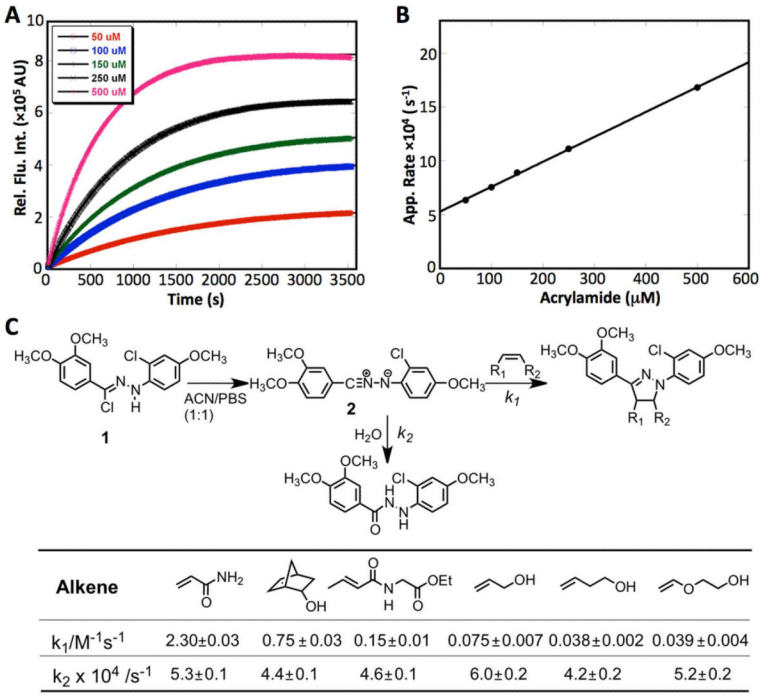


Figure 3.
A native PAGE analysis of coimmobilized sfGFP-S2AcrK in an acrylamide hydrogel. WT sfGFP was loaded to the first lane to serve as the standard.



Alkene						
$k_1/M^{-1}s^{-1}$	2.30±0.03	0.75 ± 0.03	0.15±0.01	0.075±0.007	0.038±0.002	0.039 ±0.004
$k_2 \times 10^4/s^{-1}$	5.3±0.1	4.4±0.1	4.6±0.1	6.0±0.2	4.2±0.2	5.2±0.2

Figure 4.

(**A**) The pyrazoline formation in reactions between 5 μM **1** and varying concentrations of acrylamide. Reactions were monitored by fluorescence emission at 480 nm, with an excitation light at 318 nm. (**B**) The linear dependence of apparent rates of the reaction of **1** with acrylamide on the acrylamide concentrations. (**C**) Mechanism of the reaction between **1** and olefin in aqueous conditions and the finally determined rate constants for different olefins.

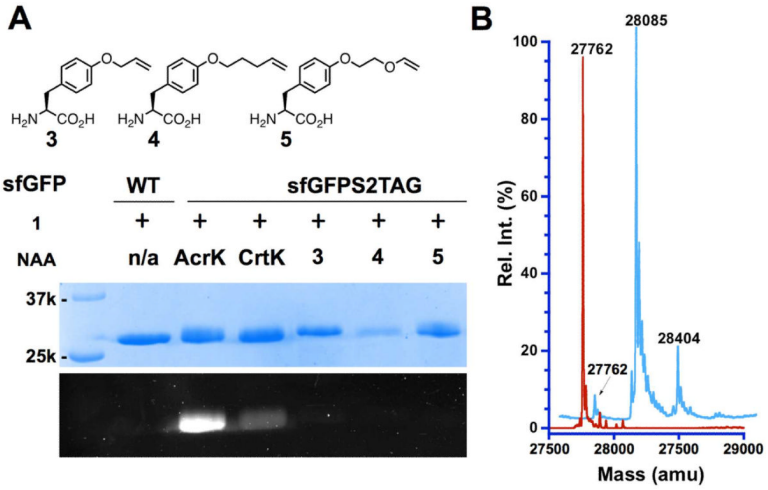


Figure 5.

(**A**) Fluorescent labeling of sfGFP variants with **1**. The image at the top panel was the Coomassie blue stained gel; the bottom panel was the fluorescent image of the same gel before Coomassie blue staining. The image was detected by BioRad ChemiDoc XRS+ system under UV irradiation. (**B**) The ESI-MS spectra of the original (in red) and labeled (in blue) sfGFP-S2AcrK proteins. Calculated molecular weight of the labeled protein is 28,083 Da.

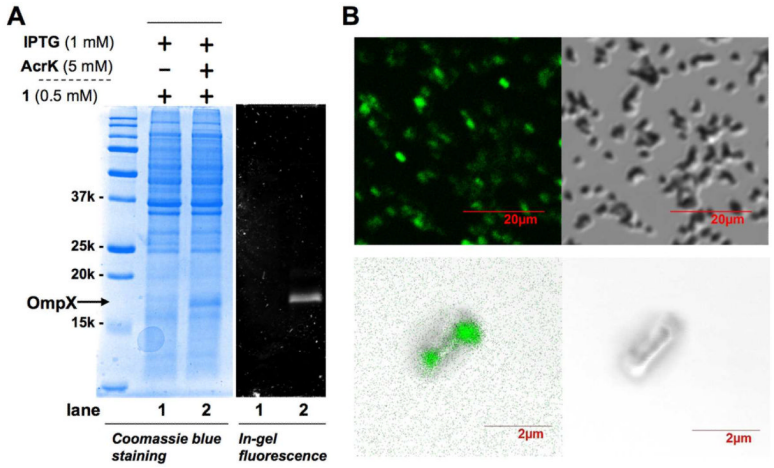


Figure 6.

(**A**) Labeling OmpX-AcrK in a cell lysate with **1**. The first lane shows a control without OmpX-AcrK expressed. (**B**) Confocal fluorescent imaging of cells overexpressing OmpX-AcrK and labeled with **1**. The top left image shows the fluorescent imaging of multiple cells; the top right is the corresponding DIC image; the bottom right is the DIC image of a single cell; the bottom left is the overlay of the DIC and the confocal fluorescent image of a single *E. coli*.

$$p = \frac{k_1 \times [O] \times [N1]}{k_1 \times [O] + k_2} \times \left(1 - e^{-\left(k_1 \times [O] + k_2\right) \times t}\right)$$

Equation 1.