Molecules and Cells

Minireview

Site-Specific Labeling of Proteins Using Unnatural Amino Acids

Kyung Jin Lee[1,2,3], Deokhee Kang[1,3], and Hee-Sung Park[1,*]

[1]Department of Chemistry, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea, [2]Present address: CHA Advanced Research Institute, Seongnam 13488, Korea, [3]These authors contributed equally to this work.
*Correspondence: hspark@kaist.ac.kr
https://doi.org/10.14348/molcells.2019.0078
www.molcells.org

Labeling of a protein with a specific dye or tag at defined positions is a critical step in tracing the subtle behavior of the protein and assessing its cellular function. Over the last decade, many strategies have been developed to achieve selective labeling of proteins in living cells. In particular, the site-specific unnatural amino acid (UAA) incorporation technique has gained increasing attention since it enables attachment of various organic probes to a specific position of a protein in a more precise way. In this review, we describe how the UAA incorporation technique has expanded our ability to achieve site-specific labeling and visualization of target proteins for functional analyses in live cells.

Keywords: bioorthogonal, click chemistry, genetic code expansion, site-specific labeling, unnatural amino acid

INTRODUCTION

Visualization of a protein of interest in living cells is becoming increasingly important, not only to precisely address cellular functions of the individual target protein, but also to understand the complicated biological system as a whole (Crivat and Taraska, 2012; Kim and Heo, 2018; Liu et al., 2015; Stephens and Allan, 2003). In particular, recent rapid advances in live-cell imaging techniques, including nanoscopy and super-resolution microscopy (SRM), have augmented our ability to address the cellular functions of biomolecules at the nanoscale level (10-30 nm) (Sengupta et al., 2012; Toomre and Bewersdorf, 2010). Numerous dyes and tag-labeling approaches have been developed for efficiently visualizing cellular target proteins (Crivat and Taraska, 2012). Fusion of fluorescent protein tags has been used routinely owing to its ease and simplicity (Shaner et al., 2007). In addition, diverse self-labeling enzymes, such as Halo-tag (Los et al., 2008), SNAP-tag (Keppler et al., 2003) and CLIP-tag (Gautier et al., 2008), have also been explored for labeling target proteins using chemical fluorescent probes, which usually exhibit better photophysical properties than fluorescent proteins (van de Linde et al., 2012). However, such protein-based tagging methods suffer from intrinsic limitations. Fused proteins or enzymes can only be added to N- or C-termini of the target protein (Crivat and Taraska, 2012; Tsien, 1998). More critically, they can perturb the cellular function of the target protein, such as localization, activity or protein-protein interaction, owing to their relatively large size (> 20 kDa) and tendency to form multimeric structures (Prescher and Bertozzi, 2005; Vreja et al., 2015).

Attempts to reduce the size of the tagging unit have given rise to the development of various peptide-based approaches. For example, phage display of peptide libraries has led to isolation of the 12-residue peptide tags, S6 and A1, which enable orthogonal protein labeling catalyzed by the phosphopantetheinyl transferases, Sfp and AcpS, respectively (Zhou et al., 2007). A tetracysteine-based labeling scheme uses an even smaller peptide, a 6-amino acid motif (CCXXCC), that

Received 15 April, 2019; accepted 2 May, 2019; published online 16 May, 2019

eISSN: 0219-1032

selectively binds to fluorescein arsenical hairpin (FlAsH) (Adams et al., 2002). Such peptide-tag labeling methods, despite certain favorable characteristics, still suffer from critical problems, including high background and cellular toxicity of the organic fluorophores. Strictly speaking, peptide-based methods are not site-specific, since they do not allow residue-specific protein labeling (Crivat and Taraska, 2012; Fernandez and Freed, 2017; Lotze et al., 2016). On the other hand, the genetic incorporation of unnatural amino acid (UAA) approach using an engineered orthogonal aminoacyl-tRNA synthetase (AARS) and tRNA pair can ideally lead to site-specific protein labeling with minimal perturbations, since this method allows installation of special chemical groups into the desired residue of the target protein (Chatterjee et al., 2013; Lang et al., 2015; Prescher and Bertozzi, 2005). Details of UAA-enabled approaches and diverse labeling schemes have been summarized by some excellent reviews (Lang and Chin, 2014b; Liu et al., 2007; Oliveira et al., 2017; Sletten and Bertozzi, 2009; Young and Schultz, 2018). In this review, we briefly discuss recent progress in UAA-based site-specific protein labeling methods and their promising applications. UAAs described in this report are listed in Figure 1 and Table 1.

EXPANDING THE GENETIC CODE BEYOND NATURE'S REPERTOIRE

In addition to the standard 20 amino acids, biological systems have evolved to genetically encode two additional amino acids, selenocysteine (Sec) and pyrrolysine (Pyl), and incorporate them into proteins (Ambrogelly et al., 2007). These two amino acids are thus sometimes referred to as the 21st and 22nd amino acids. In the case of Sec, serine is first attached to selenocysteinyl tRNA by SerRS (Leinfelder et al., 1988), and then converted into Sec by the sequential action of phosphoseryl-tRNA kinase (PSTK) (Carlson et al., 2004) and Sep (O-phosphoserine) tRNA:Sec tRNA synthase (SepSecS) (Yuan et al., 2006). Sec is then co-translationally inserted into proteins at a TGA stop codon with the help of special elongation factor (SelB) and a stem-loop Sec insertion sequence element (SECIS) (Forchhammer et al., 1989). Pyl is directly attached to pyrrolysyl tRNA ($tRNA^{pyl}$) by pyrrolysyl-tRNA synthetase (PylRS) and incorporated into proteins at a TAG stop codon (Blight et al., 2004; Polycarpo et al., 2004). Inspired by this natural expansion of the genetic code, researchers have developed a site-specific UAA incorporation technique (also called genetic

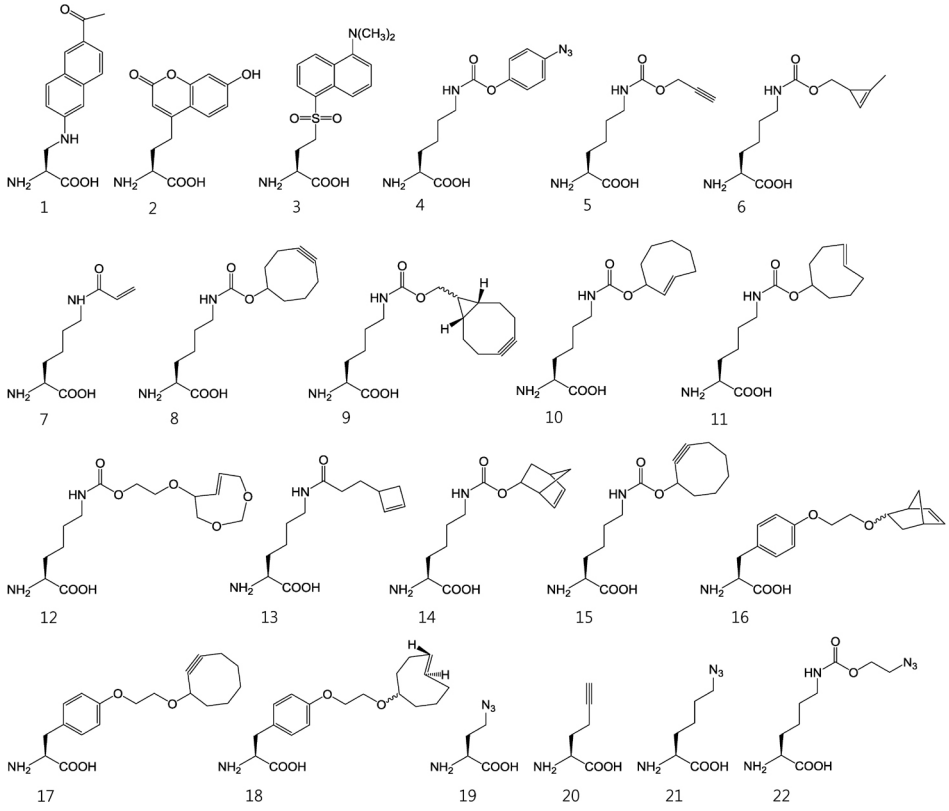

Fig. 1. Unnatural amino acids used for site-specific protein labeling. 1: 3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (Anap), 2: (S)-1-carboxy-3-(7-hydroxy-2-oxo-*2H*-chromen-4-yl)propan-1-aminium (CouAA), 3: 3-(5-(dimethylamino)naphthalene-1-sulfonamide) propanoic acid (Dansylalanine), 4: N^{ε}-*p*-azidobenzyloxycarbonyl lysine (PABK), 5: Propargyl-L-lysine (PrK), 6: N^{ε}-(1-methylcycloprop-2-enecarboxamido)lysine (CpK), 7: N^{ε}-acryllysine (AcrK), 8: N^{ε}-(cyclooct-2-yn-1-yloxy)carbonyl)L-lysine (CoK), 9: bicyclo[6.1.0]non-4-yn-9-ylmethanol lysine (BCNK), 10: trans-cyclooct-2-ene lysine (2'-TCOK), 11: trans-cyclooct-4-ene lysine (4'-TCOK), 12: dioxo-TCO lysine (DOTCOK), 13: 3-(2-cyclobutene-1-yl)propanoic acid (CbK), 14: N^{ε}-5-norbornene-2-yloxycarbonyl-L-lysine (NBOK), 15: cyclooctyne lysine (SCOK), 16: 5-norbornen-2-ol tyrosine (NOR), 17: cyclooct-2-ynol tyrosine (COY), 18: (E)-2-(cyclooct-4-en-1-yloxyl)ethanol tyrosine (DS1/2), 19: azidohomoalanine (AHA), 20: homopropargylglycine (HPG), 21: azidonorleucine (ANL), 22: N^{ε}-2-azideoethyloxycarbonyl-L-lysine (NEAK).

Table 1. Unnatural amino acids and bioorthogonal conjugation schemes

UAA	Orthogonal AARS	Mutations	Reaction	References
Fluorescent unnatural amino acids				
1, Anap	EcLeuRS	L38F, M40G, L41P, Y499V, Y500L, Y527A, H537E, L538S, F541C, A560V		Chatterjee et al., 2013; Lee et al., 2009
2, CouAA	MjTyrRS	Y32E, L65H, A67G, H70G, F108Y, Q109H, D158G, L162G		Charbon et al., 2011a; 2011b; Wang et al., 2006
3, Dansylalanine	EcLeuRS	M40A, L41N, T252A, S497C, Y499I, Y527G, H537T		Summerer et al., 2006
Lysine derivatives				
4, PABK	MmPylRS		CuAAC	Chen and Wu, 2016
5, PrK	MmPylRS		CuAAC	Greiss and Chin, 2011; Milles et al., 2012; Swiderska et al., 2017; Vreja et al., 2015; Zhang et al., 2018
6, CpK	MbPylRS		Photoclick	Ramil and Lin, 2014; Yu et al., 2012
7, AcrK	MbPylRS	D76G, L266M, L270I, Y271F, L274A, C313F	Photoclick	Li et al., 2013
8, CoK	MmPylRS	Y306A, Y384F	SPAAC	Alamudi et al., 2016; Plass et al., 2011
9, BCNK	MbPylRS	Y271M, L274G, C313A	IEDDA	Borrmann et al., 2012; Lang et al., 2012b; Nikic et al., 2015; Peng and Hang, 2016
10, 2'-TCOK	MmPylRS, MbPylRS	Y306A, Y384F Y271A, L274M, C313A, respectively.	IEDDA	Lang et al., 2012b; Nikic et al., 2014; Peng and Hang, 2016
11, 4'-TCOK	MmPylRS, MbPylRS	Y306A, Y384F Y271A, L274M, C313A, respectively.	IEDDA	Lang et al., 2012b; Nikic et al., 2014; Peng and Hang, 2016
12, DOTCOK	MmPylRS	Y306A, Y384F	IEDDA	Kozma et al., 2016
13, CbK	MbPylRS	L274M, C313A	IEDDA	Liu et al., 2017
14, NBOK	MbPylRS		IEDDA	Lang et al., 2012a
15, SCOK	MmPylRS	Y306A, Y384F	IEDDA	Nikic et al., 2015; Plass et al., 2012
22, NEAK	MmPylRS		SPAAC	Wu et al., 2016
Tyrosine derivatives				
16, NOR	MmPylRS	Y306A, N346A, C348A, Y384F	IEDDA	Kurra et al., 2014
17, COY	MmPylRS	Y306A, N346A, C348A, Y384F	IEDDA	Kurra et al., 2014
18, DS1/2	MmPylRS	Y306A, N346A, C348A, Y384F	IEDDA	Kurra et al., 2014
Methionine analogs				
19, AHA	MetRS			Calve et al., 2016; Dieterich et al., 2006; Stone et al., 2017; Tom Dieck et al., 2012
20, HPG	MetRS			Calve et al., 2016; Stone et al., 2017; Su Hui Teo et al., 2016; Tom Dieck et al., 2012
21, ANL	MetRS	L13G		Erdmann et al., 2015; Stone et al., 2017

code expansion) for reprogramming the universal genetic code by redesigning and evolving pairs of AARS and their cognate tRNAs (Mukai et al., 2017). The discovery of the PylRS/$tRNA^{pyl}$ pair in methanogenic archaea, in particular, has helped rapidly advance this field (Wan et al., 2014). PylRS/$tRNA^{pyl}$ is a naturally occurring orthogonal pair that functions efficiently in diverse living organisms from bacteria (Blight et al., 2004) and yeast (Hancock et al., 2010) to mice (Han et al., 2017). PylRS has a flexible amino acid binding pocket that is easily modified to accommodate a wide range of UAAs with diverse chemical structures. In addition, PylRS does not recognize the anticodon of $tRNA^{pyl}$, thus enabling suppression of three different stop codons, TAG, TGA and TAA, without sacrificing PylRS/$tRNA^{pyl}$ interactions (Suzuki et al., 2017; Wan et al., 2014).

DIRECT INCORPORATION OF FLUORESCENT UAAs

The major strength of the UAA incorporation technique is its ability to allow direct addition of a designed UAA into a specific position in a protein of interest. Researchers have explored this site-specific protein-labeling advantage by undertaking efforts to directly incorporate UAAs bearing various fluorescent chemical groups. The fluorescent UAA, 3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (Anap) (Chatterjee et al., 2013; Lee et al., 2009), was designed based on 6-propionyl-2-(N,N-dimethyl)aminonaphthalene (Prodan), a fluorophore that has been widely used in cell imaging (Weber and Farris, 1979). For genetic incorporation of Anap in mammalian cells, the Anap-specific AARS, AnapRS, was developed from *Escherichia coli* leucyl-tRNA synthetase (LeuRS) (Chatterjee et al., 2013). Multiple mutations (L38F, M40G, L41P, Y499V, Y500L, Y527A, H537E, L538S, F541C, and A560V) were introduced into the amino acid-binding

pocket of LeuRS through directed evolution to accommodate the large hydrophobic fluorescent UAA. Using the resulting Anap site-specific incorporation system, researchers demonstrated subcellular localization of target proteins to the nucleus or endoplasmic reticulum in mammalian cells (Chatterjee et al., 2013). Another fluorescent UAA is the coumarin derivative, (S)-1-carboxy-3-(7-hydroxy-2-oxo-2*H*-chromen-4-yl)propan-1-aminium (CouAA) (Charbon et al., 2011a; 2011b; Wang et al., 2006). A CouAA-specific synthetase was evolved from *Methanococcus jannaschii* tyrosyl-tRNA synthetase (TyrRS) through incorporation of eight mutations, Y32E, L65H, A67G, H70G, F108Y, Q109H, D158G, and L162G (Wang et al., 2006). This CouAA incorporation system was successfully used to visualize the molecular chaperone, GroEL (Charbon et al., 2011b), and the bacterial tubulin, FtsZ (Charbon et al., 2011a), without affecting their cellular functions. 2-Amino-3-(5-[dimethylamino]naphthalene-1-sulfonamide) propanoic acid is a fluorescent UAA derived from the dansyl group, and is thus also called dansylalanine (Summerer et al., 2006). A dansylalanine-specific synthetase containing seven mutations (M40A, L41N, T252A, S497C, Y499I, Y527G, and H537T), was engineered from *E. coli* LeuRS, exhibiting relatively low protein production yield (0.29 mg/L culture) in yeast (Summerer et al., 2006).

Ideally, direct genetic incorporation of fluorescent UAAs would minimize possible functional disruptions caused by fluorescence tagging. This method can also minimize the distance between the fluorophore and the protein of interest, which is essential for super-resolution imaging (Lang and Chin, 2014b). However, on a discouraging note, there are some concerns regarding the direct incorporation of fluorescent UAAs that impede their widespread use in cell biology (Lang and Chin, 2014b). Developing a specialized binding pocket in an orthogonal AARS for the unusually bulky chemical structures of fluorescent UAAs is quite difficult (Bryson et al., 2017). In addition, the incorporation efficiency of intrinsically fluorescent UAAs is considerably lower than that of UAAs designed for attachment of a fluorescent group, listed below (Summerer et al., 2006). Furthermore, most genetically incorporated fluorescent UAAs suffer from low quantum yield (Summerer et al., 2006). For example, CouAA has a quantum yield of 0.63 (Wang et al., 2006), which is comparable to that of enhanced green fluorescent protein (EGFP) (Tsien, 1998), whereas the quantum yield of Anap, at 0.48 (Chatterjee et al., 2013), is even lower. Such unfavorable characteristics of direct fluorescent UAA incorporation systems have limited their use in biological systems and live-cell visualization.

BIOORTHOGONAL CONJUGATION SCHEMES

An alternative to directly incorporating fluorescent UAAs that has been widely attempted is bio-conjugation–mediated fluorescent tagging. For this, UAAs that enable bioorthogonal reactions need to be incorporated into proteins and then conjugated with specially designed dyes to achieve site-specific fluorescence labeling of the protein of interest (Lang and Chin, 2014a; Liu and Schultz, 2010; McKay and Finn, 2014). Numerous bioorthogonal schemes for conjugating chemical probes or tags to target biomolecules have been developed. The most popular schemes that are applicable to selective fluorescence labeling under physiological conditions without interfering with biological systems include copper-catalyzed azide alkyne cycloaddition (CuAAC) (Kolb et al., 2001; Tornoe et al., 2002), photoclick cycloaddition (Wang et al., 2007), strain-promoted azide alkyne cycloaddition (SPAAC) (Agard et al., 2004), and inverse electron-demand Diels–Alder cycloadditions (IEDDA) (Boger, 1986; Thalhammer et al., 1990) (Fig. 2).

CuAAC is a cycloaddition reaction between azide and alkyne groups catalyzed by copper. Because of its rapid (~3 M^{-1} S^{-1}) and highly specific reaction, CuAAC has been applied in numerous contexts (McKay and Finn, 2014), from organic synthesis to biomolecule conjugation—even in diagnostic and therapeutic settings (He et al., 2016). However, despite the speed of the reaction and stability of the conjugation products, cytotoxicity caused by the Cu(I) metal catalyst has remained a lingering concern and has limited the widespread use of CuAAC in certain applications, including live-cell imaging (Brewer, 2010; Chen and Wu, 2016). The development of the photoclick cycloaddition scheme, which does not require copper, circumvents this toxicity issue and thus enables the expanded use of CuAAC in living organisms (Yu et al., 2012). One example of such a reaction is photoinducible dipolar cycloaddition, which induces conjugation between an alkene and a nitrile imine, activated from tetrazole (Herner and Lin, 2016; Song et al., 2008). Photoclick chemistry offers some additional advantages; the reaction is relatively fast (up to 50 M^{-1} S^{-1}) and does not require toxic metals or ligands. Notably, the photo-induced reaction can be controlled in a spatiotemporal manner (Li et al., 2013). The highly reactive tetrazole group is readily activated by a low-power UV lamp, LED or laser beam, all of which are less harmful to living cells than Cu(I) (Ramil and Lin, 2014).

SPAAC induces bioorthogonal conjugation in the absence of metal catalysts or irradiation by introducing ring strain into an alkyne group (Agard et al., 2004). Many early generation strained alkynes, such as cyclooctyne, exhibited a slow conjugation reaction (~2×10^{-3} M^{-1} S^{-1})—clearly an undesirable attribute. Subsequent studies have improved the stability of the alkyne and reaction rate by modifying the cycloalkyne chemical structure (Chen and Wu, 2016; Dommerholt et al., 2016). Because of these improvements and inherently low cytotoxicity, SPAAC has come to be broadly used in live mammalian cells (Baskin et al., 2007; Mbua et al., 2011), and even in animals (Baskin et al., 2007; Laughlin et al., 2008). IEDDA, which enables cycloaddition between a dienophile and a diene, such as 1,2,4,5-tetrazine (Boger, 1986; Thalhammer et al., 1990), is a more recently developed bioorthogonal conjugation strategy (Blackman et al., 2008). Because of its extremely fast (up to 10^6 M^{-1} S^{-1}) and selective reaction, IEDDA has become a prominent reaction scheme for bioorthogonal conjugation that enables chemical modifications and fluorescent dye labeling of proteins of interest in live mammalian cells (Lang and Chin, 2014a; Lang et al., 2012a; Nikic et al., 2015).

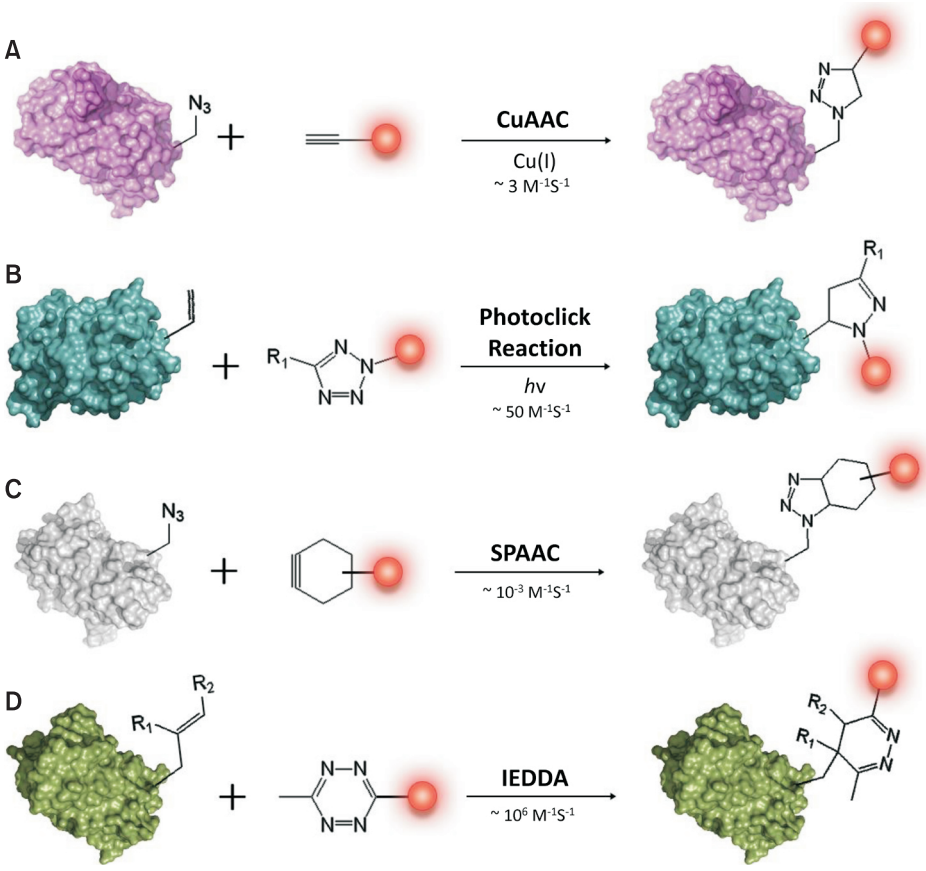

Fig. 2. Bioorthogonal reactions for labeling of a protein of interest. CuAAC, copper-catalyzed azide alkyne cycloaddition; Photoclick cycloaddition; SPAAC, strain-promoted azide alkyne cycloaddition; IEDDA, inverse electron-demand Diels-Alder cycloadditions.

SITE-SPECIFIC LABELING VIA BIOORTHOGONAL CONJUGATION USING UAAs

CuAAC

Since its first application in a biological system in 2002 (Kolb et al., 2001; Tornoe et al., 2002), the CuAAC approach has been applied to a variety of research areas, including drug development and diagnostics (He et al., 2016). However, a practical impediment to its application in living systems is the induction of cytotoxic reactive oxygen species (ROS) by Cu(I) oxidation (Brewer, 2010). A variety of water-soluble ligands have been designed to reduce the amount of Cu(I) metal catalyst needed to achieve efficient conjugation with minimal cytotoxicity. These include tris-(benzyltriazolylmethyl)amine (TBTA) (Chan et al., 2004), tris-(3-hydroxypropyltriazolylmethyl)-amine (THPTA) (Hong et al., 2009), bis(L-histidine) (Kennedy et al., 2011), and 2-[4-{(bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino)methyl}-1H-1,2,3-triazol-1-yl]acetic acid (BTTAA) (Besanceney-Webler et al., 2011).

To date, several UAA incorporation systems utilizing CuAAC have been successfully applied to label target proteins. For example, the epidermal growth factor receptor (EGFR), a cell surface receptor tyrosine kinase, was visualized in live mammalian cells by site-specifically incorporating the lysine-based UAA, N^{ε}-*p*-azidobenzyloxycarbonyl lysine (PABK), into the receptor using a PABK-specific synthetase derived from *Methanococcus barkeri* PylRS (MbPylRS). The expressed EGFR bearing PABK was labeled with alkyne-cy5 in the presence of low $CuSO_4$ (50 μM) with the help of the ligand BTTAA (300 μM) (Chen and Wu, 2016). N^{ε}-(propargyloxycarbonyl)-l-lysine, also known as propargyl lysine (PrK), another lysine-based UAA, can be incorporated into proteins using wild-type *Methanococcus mazei* PylRS (MmPylRS) (Milles et al., 2012). PrK has been widely used for bio-conjugation of fluorescent probes in diverse organisms from *E. coli* to mammalian cells (Swiderska et al., 2017; Vreja et al., 2015)—even in animals (Greiss and Chin, 2011)—because of the simplicity of the synthesis reaction and high incorporation efficiency. With this approach, imaging of target proteins using super resolution microscopy is achievable after PrK-mediated conjugation of probe via a CuAAC scheme and fixation of the cells (Vreja et al., 2015). Such a PrK-based labeling scheme has also been applied to hyperspectral stimulated Raman scattering (SRS) microscopy. This study demonstrated labeling of PrK-carrying cellular proteins, such as Sec61β, Htt74Q and the histone H3 variant H3.3, with a sensitive Raman tag by click chemistry for molecular hyperspectral SRS imaging (Zhang et al., 2018). However, despite the reliability of UAA incorporation and conjugation reaction, the CuAAC scheme frequently requires a high concentration of $CuSO_4$ (~millimolar levels), which continues to limit its wider application to live-cell imaging (Chen and Wu, 2016).

Photoclick chemistry

The photoactivated 1,3-dipolar cycloaddition reaction between diphenyltetrazole and methyl crotonate in benzene,

first reported in 1967 (Huisgen et al., 1967), has been further expanded to cover nitrile oxide, azomethine ylide, and azide cycloaddition in aqueous media (Molteni et al., 2000). 1,3-Dipolar cycloaddition between alkene dipolarophiles and nitrile imine dipoles has also been developed (Molteni et al., 2000). Moreover, an efficient photo-inducible bioorthogonal reaction based on photoclick chemistry that enables tetrazole-alkene 1,3-dipolar cycloaddition was recently reported in living cells (Yu et al., 2012).

The lysine-based UAAs, N^{ε}-(1-methylcycloprop-2-enecarboxamido)-lysine (CpK) and N^{ε}-5-norbornene-2-yloxycarbonyl-L-lysine (NorK), which enable such cycloadditions with photoinduced tetrazole, were designed and co-translationally inserted into proteins by wild-type MbPylRS and a mutant form of MbPylRS carrying five mutations (L266M, L270I, Y271L, L274A, and C313I), respectively (Ramil and Lin, 2014; Yu et al., 2012). The yield of cycloaddition reactions with CpK was higher than that with NorK. Using this system, GFP carrying CpK was expressed and labeled with photoreactive tetrazole by photoirradiation for 2 min with 365 nm UV light (Ramil and Lin, 2014; Yu et al., 2012). Another photoclickable lysine analog, N^{ε}-acryllysine (AcrK), which requires fewer synthetic steps than CpK, has also been designed. AcrK was co-translationally inserted into target proteins using an engineered MbPylRS bearing six mutations (D76G, L266M, L270I, Y271F, L274A, and C313F). Applying this system to the tubulin-like bacterial cytoskeleton protein, FtsZ, carrying AcrK led to efficient expression and labeling with diaryltetrazole using a 1-minute exposure to 365 nm UV irradiation (Li et al., 2013).

SPAAC

SPAAC is a simple and spontaneous conjugation between azide and cycloalkyne groups that generates a triazole product without the requirement for a catalyst (Agard et al., 2004). The early stage [3 + 2] azide-alkyne cycloaddition using strained cyclooctyne was considerably slower than CuAAC, but various subsequently developed strained alkyne chemical structures have increased the reaction rate (Chen and Wu, 2016). Thanks to such efforts, SPAAC has been effectively applied to visualize diverse biomolecules, including nucleotides (Singh and Heaney, 2011) and lipids (Neef and Schultz, 2009), in various living systems (Alamudi et al., 2016). Tetra-acetylated N-azidoacetyl-D-mannosamine (Ac-4ManNAz), an azide-functionalized sugar, can be metabolically incorporated into the cell surface membrane through the sialic acid biosynthetic pathway (Baskin et al., 2007; Kayser et al., 1992; Saxon and Bertozzi, 2000). The resulting azide-containing glycans can be visualized by labeling with cyclooctyne fluorophores such as difluorinated cyclooctyne (DIFO) (Baskin et al., 2007), dibenzocyclooctyne (DIBO) (Ning et al., 2008), azadibenzylcyclooctyne (DIBAC) (Debets et al., 2010) and bicyclo[6.1.0]nonyne (BCN) fluorescent dyes, allowing three-dimensional visualization of living melanoma cells (Dommerholt et al., 2010).

The SPAAC scheme has been further expanded to protein molecules in living organisms (Plass et al., 2011). N^{ε}-(cyclooct-2-yn-1-yloxy)carbonyl)L-lysine (CoK) was site-specifically incorporated into red fluorescent protein mCherry by MmPylRS carrying two mutations, Y306A and Y384F (MmPylRS-AF). mCherry carrying CoK was labeled with coumarin azide and subsequently studied by high-resolution single-molecule Förster resonance energy transfer (FRET) analysis in live *E. coli* cells, illustrating the potential of SPAAC for *in vivo* labeling and visualization of target proteins (Plass et al., 2011). The use of the CoK-based fluorescent tagging scheme was recently extended to labeling intracellular proteins in live mammalian cells (Alamudi et al., 2016). In this application, alpha tubulin carrying CoK was expressed using mutant MmPylRS-AF and labeled with an azide-bearing, background-free "tame" BODIPY-based probe (AzG-1). Despite such efforts, the application of SPAAC is still limited by its relatively slow conjugation reaction (Li and Zhang, 2016). To increase reaction rate, researchers have designed diverse cyclooctyne derivatives with enhanced water solubility for labeling of biomolecules in live cells (Li and Zhang, 2016).

IEDDA

IEDDA is a [4 + 2] cycloaddition reaction between 1,2,4,5-tetrazine and dienophiles that is highly specific and extremely fast under physiological conditions (Thalhammer et al., 1990). In addition, it does not require any catalyst or reagent for the conjugation reaction. Because of such favorable characteristics, IEDDA has become one of the most popular bioorthogonal schemes for biological applications. Thus far, a wide range of dienophiles has been designed for IEDDA reactions, and numerous UAAs bearing such dienophile chemical groups have been developed for IEDDA-based bioorthogonal conjugation (Oliveira et al., 2017; Plass et al., 2012; Prokhorov and Kozhevnikov, 2012).

Bicyclo[6.1.0]non-4-yn-9-ylmethanol (BCN) is considered an attractive dienophile owing to its excellent reactivity with tetrazine. The BCN-bearing UAA, bicyclo[6.1.0]non-4-yn-9-ylmethanol lysine (BCNK), was first designed in 2012 and subsequently genetically incorporated using a mutant form of MbPylRS containing three amino acid substitution (Y271M, L274G, and C313A), generated by directed evolution. This triply mutated MbPylRS exhibited good protein production yield (6-12 mg/L) (Lang et al., 2012b). The double mutant, MmPylRS-AF, was also found to incorporate isoforms of BCNK (endo-BCNK and exo-BCNK) with high yield (Borrmann et al., 2012; Nikic et al., 2015; Peng and Hang, 2016). Exo-BCNK showed higher incorporation efficiency and faster labeling than its isomer when conjugated with monosubstituted tetrazine (H-Tetrazine) fluorophores, which are generally superior to methyl-substituted tetrazine (Me-Tetrazine) moieties (Peng and Hang, 2016). The BCNK incorporation and tetrazine-based fluorophore conjugation scheme has been successfully applied to label diverse cellular proteins, including cell surface and cytoskeletal proteins.

Trans-cyclooctene (TCO) is a highly reactive dienophile that exhibits about a 10-times higher reaction rate compared with BCN owing to its strained chemical structure (Lang et al., 2012b). TCO-bearing lysines (TCOKs) were designed and found to efficiently react with tetrazines. MmPylRS-AF and MbPylRS carrying three mutations (Y271A, L274M, and C313A) were shown to genetically incorporate 2'-TCOK and 4'-TCOK (Nikic et al., 2014; Peng and Hang, 2016). In

addition to engineering PylRS, $tRNA^{pyl}$ was also evolved to increase the incorporation of TCOKs and other related bulky UAAs that previously exhibited low incorporation efficiency (Serfling et al., 2018). The engineered tRNA, $tRNA^{M15}$, carrying a canonical hinge between the D- and T-loop that is highly conserved in human tRNAs, was found to improve UAA incorporation several fold compared to the original $tRNA^{pyl}$. Combining MbPylRS-AF and the engineered tRNA enabled the extracellular target protein, CRF1R (corticotropin releasing factor type 1 receptor), carrying 2'-TCOK, to be efficiently produced and labeled with cy3-tetrazine dye for *in situ* visualization (Serfling et al., 2018).

The efficiency of genetic incorporation and labeling of TCOKs, like BCNKs, is highly dependent on the chemical structure (Peng and Hang, 2016; Uttamapinant et al., 2015). 2'-TCOK showed higher incorporation by MmPylRS-AF and better labeling with tetrazines than 4'-TCOK since it is less reactive towards cysteine residues, which are exposed in many proteins in mammalian cells and might otherwise cause interference (Nikic et al., 2014). Moreover, a side-by-side comparison of axial and equatorial isomers of 2'- and 4'-TCOKs showed better labeling efficiency of 2'-aTCOK (axial) and 4'-eTCOK (equatorial) with tetrazines compared with their diastereomers, probably owing to greater chemical stability or a more rapid reaction rate (Peng and Hang, 2016). This TCOK incorporation and tetrazine-based labeling system has been applied to a wide range of cellular targets, including EGFR (Lang et al., 2012b) and actin (Peng and Hang, 2016).

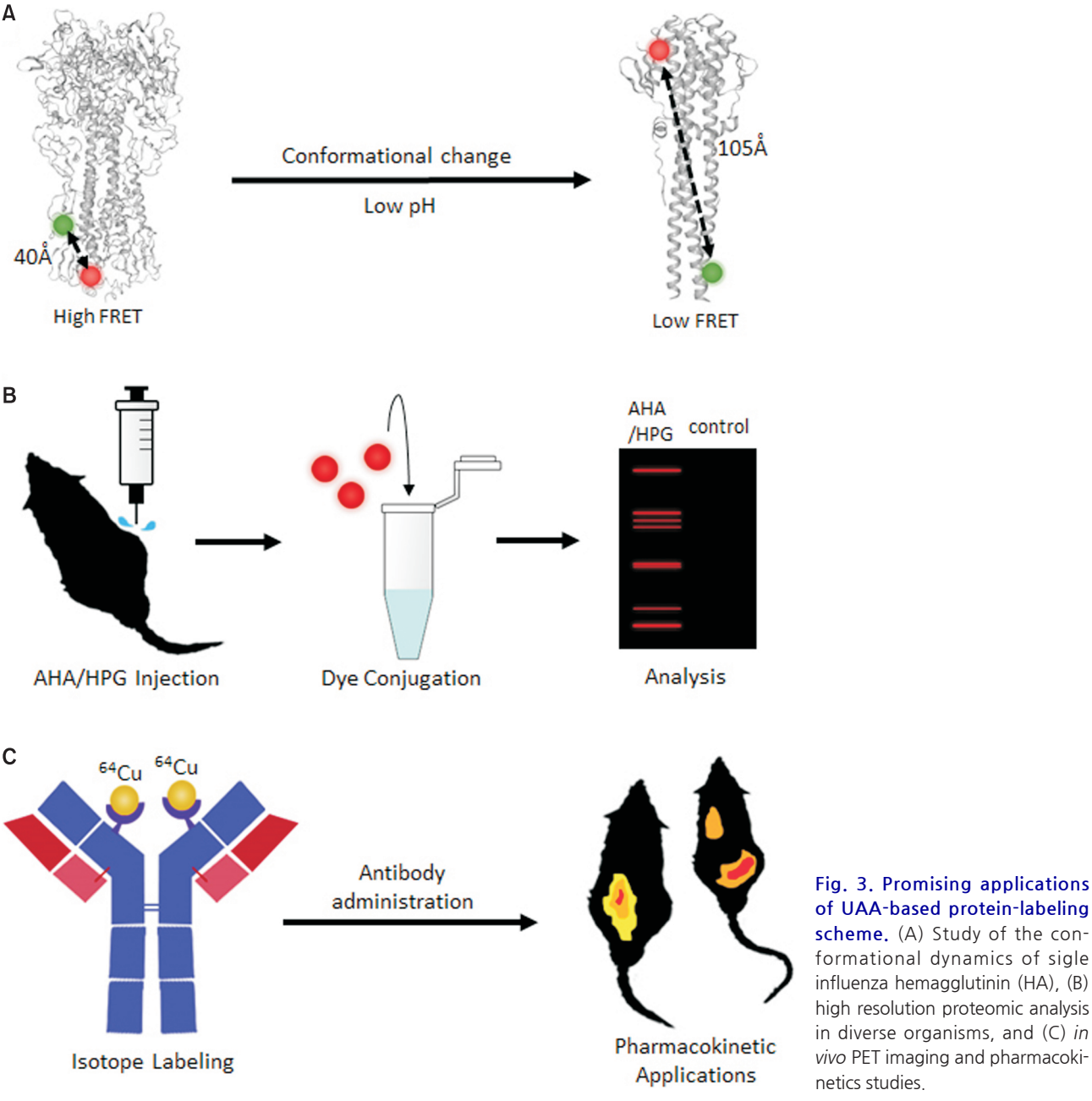

Fig. 3. Promising applications of UAA-based protein-labeling scheme. (A) Study of the conformational dynamics of sigle influenza hemagglutinin (HA), (B) high resolution proteomic analysis in diverse organisms, and (C) *in vivo* PET imaging and pharmacokinetics studies.

In particular, this scheme was found to be highly useful for visualizing small transmembrane proteins like interferon-inducible transmembrane protein 3 (IFITM3), which could not easily be tracked using the classic fluorescent protein fusion method (Peng and Hang, 2016).

Background fluorescence resulting from insufficient washing of unreacted probes, which is often caused by the hydrophobicity of UAAs, can be troublesome, especially in visualizing intracellular proteins (Kozma et al., 2016). To further improve labeling efficiency of TCOK toward cytosolic proteins, researchers synthesized the more hydrophilic TCOK derivative, dioxo-TCOK (DOTCOK). This UAA was genetically incorporated into proteins by MmPylRS-AF with good yield. Excess DOTCOK is easily removed using considerably shorter washing periods (several minutes) compared with BCNK and TCOK (6 h or longer), which is clearly advantageous for visualizing cytosolic proteins (Kozma et al., 2016).

Other useful UAAs that support IEDDA include cyclobutene-containing lysine (CbK) (Liu et al., 2017), norbornene-2-yloxycarbonyl-lysine (NBOK) (Lang et al., 2012a), and strained cyclooctyne-lysine (SCOK) (Nikic et al., 2015; Plass et al., 2012). CbK is small and chemically stable; it also has a higher conjugation rate with tetrazine than most cycloprene, small alkene or norbornene derivatives, but a lower rate than TCO. CbK was genetically incorporated into OmpX (outer membrane protein X) of *E. coli* using mutant MbPylRS (L274M and C313A) and subsequently labeled with fluorescent tetrazine (Liu et al., 2017). Wild-type MbPylRS efficiently incorporated NBOK with a good production yield of target protein (~4 mg/L of culture) in *E. coli* and mediated a highly selective cycloaddition reaction with tetrazine *in vitro* (Lang et al., 2012a). NBOK also allowed *in vivo* imaging of a membrane-localized target protein using the tetrazine dye, tetramethylrhodamine (TAMRA), although a much longer incubation time (2-16 h) was needed owing to the slow reaction rate (9 M^{-1} S^{-1}) (Lang et al., 2012a). Notably, SCOK enabled a faster conjugation reaction with tetrazine dye (400 M^{-1} S^{-1}) (Plass et al., 2012), and was incorporated by mutant MmPylRS-AF into the insulin receptor (IR) with high yield (10 mg/L of culture) and was efficiently labeled with H-tetrazine dye (Nikic et al., 2015).

In addition to lysine analogues, dienophile-carrying tyrosine derivatives have also been utilized. 5-Norbornen-2-ol (NOR), cyclooct-2-ynol (COY), and two isomers of (E)-2-(cyclooct-4-en-1-yloxyl)-ethanol tyrosines (DS1 and DS2) were designed and synthesized. These UAAs were genetically incorporated into target proteins in the outer membrane of *E. coli* using mutant MmPylRS (Y306A, N346A, C348A, and Y384F) and efficiently labeled with fluorescent tetrazine. Among these four dienophiles, cyclooctene-based DS1 and DS2 showed extremely fast conjugation reaction rates (0.6-2.9 × 10^5 M^{-1} S^{-1}) and high labeling efficiency (~97%). *In vivo* labeling of DS1 and DS2 could be completed within 100 s in the presence of 1 μM fluorescent tetrazine. Collectively, these results demonstrate that genetic UAA incorporation techniques are highly effective in achieving site-specific labeling of cellular target proteins for functional analyses (Kurra et al., 2014).

INSPIRING APPLICATIONS AND OUTLOOK

Beyond its straightforward use for protein labeling and visualization, the UAA-based protein-labeling scheme is beginning to find extended applications in diverse research areas. One of the major drawbacks of the conventional fluorescent protein fusion method is the difficulty in probing component proteins that form higher-order complexes, such as the virus capsid and virus-like particles, owing to their substantial size (Costantini and Snapp, 2015). Such difficult target proteins can be accessed and analyzed at the single-molecule level using UAA incorporation and a bioorthogonal probe tagging system (Das et al., 2018). Influenza hemagglutinin (HA) is a canonical type I viral envelope glycoprotein that is responsible for fusion of the viral envelope with the endosome membrane after entry into host cells. Acidification of the late endosome was thought to trigger a conformational change in HA, although this had not been directly observed. Recombinant influenza HA carrying two 2'-TCOKs (at positions 17 and 127) was produced using MmPylRS-AF and then labeled with Tet-Cy3 and (Me-)Tet-Cy5 via the IEDDA reaction. This dual-labeling scheme made it possible to study the conformational dynamics of single influenza HA molecules triggered by acidification using single-molecule FRET (smFRET) analysis and total internal reflection fluorescence (TIRF) microscopy (Das et al., 2018) (Fig. 3A). UAA-based bioorthogonal chemistry can also be applied to analyze differential gene and protein expression. Wild-type methionyl-tRNA synthetase (MetRS) and mutant $MetRS^{LtoG}$ are known to incorporate diverse clickable methionine analogs, including azidohomoalanine (AHA), homopropargylglycine (HPG), and azidonorleucine (ANL), which can be efficiently used for bioorthogonal noncanonical amino acid tagging (BONCAT) technique (Dieterich et al., 2006; Erdmann et al., 2015). Proteins carrying Met analogs containing azide or alkyne can be selectively labeled with fluorescence tags for visualization or with affinity tags for enrichment by CuAAC (Fig. 3B). This method has been used to visualize newly synthesized proteins in diverse organisms, including mouse (Calve et al., 2016) and Drosophila (Erdmann et al., 2015), in specific contexts, such as viral infection and stress conditions (Su Hui Teo et al., 2016). This approach also allows spatiotemporal analysis of proteomics after drug treatment or electrophysiological stimulation (Tom Dieck et al., 2012). The clear benefit of this technique compared with conventional isotope-based methods is that it confers a higher degree of temporal resolution without the need for isotopes (Stone et al., 2017).

The UAA incorporation technique can be further extended to *in vivo* positron-emission tomography (PET) imaging and pharmacokinetics using pre-targeting components such as antibodies (Wu et al., 2016). An azido-bearing UAA, N^{ε}-2-azideoethyloxycarbonyl-L-lysine (NEAK), was incorporated into the heavy chain of the anti-CD20 monoclonal antibody, rituximab, using MmPylRS. The heavy chain-carrying NEAK was then linked with bifunctional 4-dibenzocyclooctynol-1,4,7,10-tetraazacyclotetradecane-1,4,7,10-tetraacetic acid (DIBO-DOTA) via copper-free SPAAC. Lastly, the positron-emitting isotope ^{64}Cu was chelated to the attached DOTA, generating [^{64}Cu]-labeled rituximab. Injection of the

labeled antibody into mice allowed efficient visualization of CD20-positive tumors by PET imaging for *in vivo* pharmacokinetics studies (Fig. 3C). These inspiring reports illustrate critical applications of the UAA-based labeling scheme in a wide range of fields, from basic science to clinical research. We are still in the early stages of developing important applications of UAA-based techniques. We anticipate that such imaginative uses of UAA-based strategies will further embolden efforts to investigate complex cellular processes and facilitate more systematic assessments of drug efficacy and pharmacokinetics.

Disclosure of potential conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We apologize to all scientists in the field whose important work could not be cited in this report due to space limitations.

This work was supported by grants from the National Research Foundation of Korea (2014M3A6A4075060 and 2017R1A2B3011543) and from Samsung Science & Technology Foundation (SSTF-BA1702-09).

ORCID

Hee-Sung Park https://orcid.org/0000-0001-8645-0405

REFERENCES

Adams, S.R., Campbell, R.E., Gross, L.A., Martin, B.R., Walkup, G.K., Yao, Y., Llopis, J., and Tsien, R.Y. (2002). New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. J. Am. Chem. Soc. *124*, 6063-6076.

Agard, N.J., Prescher, J.A., and Bertozzi, C.R. (2004). A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. J. Am. Chem. Soc. *126*, 15046-15047.

Alamudi, S.H., Satapathy, R., Kim, J., Su, D., Ren, H., Das, R., Hu, L., Alvarado-Martínez, E., Lee, J.Y., Hoppmann, C., et al. (2016). Development of background-free tame fluorescent probes for intracellular live cell imaging. Nat. Commun. *7*, 11964.

Ambrogelly, A., Palioura, S., and Söll, D. (2007). Natural expansion of the genetic code. Nat. Chem. Biol. *3*, 29-35.

Baskin, J.M., Prescher, J.A., Laughlin, S.T., Agard, N.J., Chang, P.V., Miller, I.A., Lo, A., Codelli, J.A., and Bertozzi, C.R. (2007). Copper-free click chemistry for dynamic in vivo imaging. Proc. Natl. Acad. Sci. U. S. A. *104*, 16793-16797.

Besanceney-Webler, C., Jiang, H., Zheng, T., Feng, L., Soriano del Amo, D., Wang, W., Klivansky, L.M., Marlow, F.L., Liu, Y., and Wu, P. (2011). Increasing the efficacy of bioorthogonal click reactions for bioconjugation: a comparative study. Angew. Chem. Int. Ed. Engl. *50*, 8051-8056.

Blackman, M.L., Royzen, M., and Fox, J.M. (2008). Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. J. Am. Chem. Soc. *130*, 13518-13519.

Blight, S.K., Larue, R.C., Mahapatra, A., Longstaff, D.G., Chang, E., Zhao, G., Kang, P.T., Green-Church, K.B., Chan, M.K., and Krzycki, J.A. (2004). Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo. Nature *431*, 333-335.

Boger, D.L. (1986). Diels-Alder reactions of heterocyclic aza dienes. Scope and applications. Chem. Rev. *86*, 781-793.

Borrmann, A., Milles, S., Plass, T., Dommerholt, J., Verkade, J.M., Wiessler, M., Schultz, C., van Hest, J.C., van Delft, F.L., and Lemke, E.A. (2012). Genetic encoding of a bicyclo[6.1.0]nonyne-charged amino acid enables fast cellular protein imaging by metal-free ligation. Chembiochem *13*, 2094-2099.

Brewer, G.J. (2010). Risks of copper and iron toxicity during aging in humans. Chem. Res. Toxicol. *23*, 319-326.

Bryson, D.I., Fan, C., Guo, L.T., Miller, C., Söll, D., and Liu, D.R. (2017). Continuous directed evolution of aminoacyl-tRNA synthetases. Nat. Chem. Biol. *13*, 1253-1260.

Calve, S., Witten, A.J., Ocken, A.R., and Kinzer-Ursem, T.L. (2016). Incorporation of non-canonical amino acids into the developing murine proteome. Sci. Rep. *6*, 32377.

Carlson, B.A., Xu, X.M., Kryukov, G.V., Rao, M., Berry, M.J., Gladyshev, V.N., and Hatfield, D.L. (2004). Identification and characterization of phosphoseryl-tRNA[Ser]Sec kinase. Proc. Natl. Acad. Sci. U. S. A. *101*, 12848-12853.

Chan, T.R., Hilgraf, R., Sharpless, K.B., and Fokin, V.V. (2004). Polytriazoles as copper(I)-stabilizing ligands in catalysis. Org. Lett. *6*, 2853-2855.

Charbon, G., Brustad, E., Scott, K.A., Wang, J., Løbner-Olesen, A., Schultz, P.G., Jacobs-Wagner, C., and Chapman, E. (2011a). Subcellular protein localization by using a genetically encoded fluorescent amino acid. Chembiochem *12*, 1818-1821.

Charbon, G., Wang, J., Brustad, E., Schultz, P.G., Horwich, A.L., Jacobs-Wagner, C., and Chapman, E. (2011b). Localization of GroEL determined by in vivo incorporation of a fluorescent amino acid. Bioorg. Med. Chem. Lett. *21*, 6067-6070.

Chatterjee, A., Guo, J., Lee, H.S., and Schultz, P.G. (2013). A genetically encoded fluorescent probe in mammalian cells. J. Am. Chem. Soc. *135*, 12540-12543.

Chen, X., and Wu, Y.W. (2016). Selective chemical labeling of proteins. Org. Biomol. Chem. *14*, 5417-5439.

Costantini, L.M., and Snapp, E.L. (2015). Going viral with fluorescent proteins. J. Virol. *89*, 9706-9708.

Crivat, G., and Taraska, J.W. (2012). Imaging proteins inside cells with fluorescent tags. Trends Biotechnol. *30*, 8-16.

Das, D.K., Govindan, R., Nikić-Spiegel, I., Krammer, F., Lemke, E.A., and Munro, J.B. (2018). Direct visualization of the conformational dynamics of single influenza hemagglutinin trimers. Cell *174*, 926-937.e12.

Debets, M.F., van der Doelen, C.W., Rutjes, F.P., and van Delft, F.L. (2010). Azide: a unique dipole for metal-free bioorthogonal ligations. Chembiochem *11*, 1168-1184.

Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc. Natl. Acad. Sci. U. S. A. *103*, 9482-9487.

Dommerholt, J., Rutjes, F.P.J.T., and van Delft, F.L. (2016). Strain-promoted *1*,3-dipolar cycloaddition of cycloalkynes and organic azides. Top. Curr. Chem. (Cham.) *374*, 16.

Dommerholt, J., Schmidt, S., Temming, R., Hendriks, L.J., Rutjes, F.P., van Hest, J.C., Lefeber, D.J., Friedl, P., and van Delft, F.L. (2010). Readily accessible bicyclononynes for bioorthogonal labeling and three-dimensional imaging of living cells. Angew. Chem. Int. Ed. Engl. *49*, 9422-9425.

Erdmann, I., Marter, K., Kobler, O., Niehues, S., Abele, J., Müller, A., Bussmann, J., Storkebaum, E., Ziv, T., Thomas, U., et al. (2015). Cell-selective labelling of proteomes in Drosophila melanogaster. Nat. Commun. *6*, 7521.

Fernandez, M.V., and Freed, E.O. (2017). "Expand and Click": a new method for labeling HIV-1 envelope glycoproteins. Cell Chem. Biol. *24*, 548-550.

Forchhammer, K., Leinfelder, W., and Böck, A. (1989). Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. Nature *342*, 453-456.

Gautier, A., Juillerat, A., Heinis, C., Corrêa, I.R., Jr., Kindermann, M., Beaufils, F., and Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. Chem. Biol. *15*, 128-136.

Greiss, S., and Chin, J.W. (2011). Expanding the genetic code of an animal. J. Am. Chem. Soc. *133*, 14196-14199.

Han, S., Yang, A., Lee, S., Lee, H.W., Park, C.B., and Park H.S. (2017). Expanding the genetic code of Mus musculus. Nat. Commun. *8*, 14568.

Hancock, S.M., Uprety, R., Deiters, A., and Chin, J.W. (2010). Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. J. Am. Chem. Soc. *132*, 14819-14824.

He, X.P., Zeng, Y.L., Zang, Y., Li, J., Field, R.A., and Chen, G.R. (2016). Carbohydrate CuAAC click chemistry for therapy and diagnosis. Carbohydr. Res. *429*, 1-22.

Herner, A., and Lin, Q. (2016). Photo-triggered click chemistry for biological applications. Top. Curr. Chem. (Cham.) *374*, 1.

Hong, V., Presolski, S.I., Ma, C., and Finn, M.G. (2009). Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. Angew. Chem. Int. Ed. Engl. *48*, 9879-9883.

Huisgen, R., Szeimies, G., and Möbius, L. (1967). 1.3-Dipolare Cycloadditionen, XXXII. Kinetik der additionen organischer Azide an CC-Mehrfachbindungen. Chemische Berichte *100*, 2494-2507.

Kayser, H., Zeitler, R., Kannicht, C., Grunow, D., Nuck, R., and Reutter, W. (1992). Biosynthesis of a nonphysiological sialic acid in different rat organs, using N-propanoyl-D-hexosamines as precursors. J. Biol. Chem. *267*, 16934-16938.

Kennedy, D.C., McKay, C.S., Legault, M.C., Danielson, D.C., Blake, J.A., Pegoraro, A.F., Stolow, A., Mester, Z., and Pezacki, J.P. (2011). Cellular consequences of copper complexes used to catalyze bioorthogonal click reactions. J. Am. Chem. Soc. *133*, 17993-18001.

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat. Biotechnol. *21*, 86-89.

Kim, J., and Heo, W.D. (2018). Synergistic ensemble of optogenetic actuators and dynamic indicators in cell biology. Mol. Cells *41*, 809-817.

Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001). Click chemistry: diverse chemical function from a few good reactions. Angew. Chem. Int. Ed. Engl. *40*, 2004-2021.

Kozma, E., Nikić, I., Varga, B.R., Aramburu, I.V., Kang, J.H., Fackler, O.T., Lemke, E.A., and Kele, P. (2016). Hydrophilic trans-cyclooctenylated noncanonical amino acids for fast intracellular protein labeling. Chembiochem *17*, 1518-1524.

Kurra, Y., Odoi, K.A., Lee, Y.J., Yang, Y., Lu, T., Wheeler, S.E., Torres-Kolbus, J., Deiters, A., and Liu, W.R. (2014). Two rapid catalyst-free click reactions for in vivo protein labeling of genetically encoded strained alkene/alkyne functionalities. Bioconjug. Chem. *25*, 1730-1738.

Lang, K., and Chin, J.W. (2014a). Bioorthogonal reactions for labeling proteins. ACS Chem. Biol. *9*, 16-20.

Lang, K., and Chin, J.W. (2014b). Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. Chem. Rev. *114*, 4764-4806.

Lang, K., Davis, L., and Chin, J.W. (2015). Genetic encoding of unnatural amino acids for labeling proteins. Methods Mol. Biol. *1266*, 217-228.

Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J.W. (2012a). Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. Nat. Chem. *4*, 298-304.

Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D.J., Blackman, M.L., Fox, J.M., and Chin, J.W. (2012b). Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. J. Am. Chem. Soc. *134*, 10317-10320.

Laughlin, S.T., Baskin, J.M., Amacher, S.L., and Bertozzi, C.R. (2008). In vivo imaging of membrane-associated glycans in developing zebrafish. Science *320*, 664-667.

Lee, H.S., Guo, J., Lemke, E.A., Dimla, R.D., and Schultz, P.G. (2009). Genetic incorporation of a small, environmentally sensitive, fluorescent probe into proteins in Saccharomyces cerevisiae. J. Am. Chem. Soc. *131*, 12921-12923.

Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.A., and Böck, A. (1988). Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. Nature *331*, 723-725.

Li, F., Zhang, H., Sun, Y., Pan, Y., Zhou, J., and Wang, J. (2013). Expanding the genetic code for photoclick chemistry in E. coli, mammalian cells, and A. thaliana. Angew. Chem. Int. Ed. Engl. *52*, 9700-9704.

Li, L., and Zhang, Z. (2016). Development and applications of the copper-catalyzed azide-alkyne cycloaddition (CuAAC) as a bioorthogonal reaction. Molecules *21*, 1393.

Liu, C.C., and Schultz, P.G. (2010). Adding new chemistries to the genetic code. Annu. Rev. Biochem. *79*, 413-444.

Liu, K., Enns, B., Evans, B., Wang, N., Shang, X., Sittiwong, W., Dussault, P.H., and Guo, J. (2017). A genetically encoded cyclobutene probe for labelling of live cells. Chem. Commun. (Camb.) *53*, 10604-10607.

Liu, W., Brock, A., Chen, S., Chen, S., and Schultz, P.G. (2007). Genetic incorporation of unnatural amino acids into proteins in mammalian cells. Nat. Methods *4*, 239-244.

Liu, Z., Lavis, L.D., and Betzig, E. (2015). Imaging live-cell dynamics and structure at the single-molecule level. Mol. Cell *58*, 644-659.

Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., et al. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem. Biol. *3*, 373-382.

Lotze, J., Reinhardt, U., Seitz, O., and Beck-Sickinger, A.G. (2016). Peptide-tags for site-specific protein labelling in vitro and in vivo. Mol. Biosyst. *12*, 1731-1745.

Mbua, N.E., Guo, J., Wolfert, M.A., Steet, R., and Boons, G.J. (2011). Strain-promoted alkyne-azide cycloadditions (SPAAC) reveal new features of glycoconjugate biosynthesis. Chembiochem *12*, 1912-1921.

McKay, C.S., and Finn, M.G. (2014). Click chemistry in complex mixtures: bioorthogonal bioconjugation. Chem. Biol. *21*, 1075-1101.

Milles, S., Tyagi, S., Banterle, N., Koehler, C., VanDelinder, V., Plass, T., Neal, A.P., and Lemke, E.A. (2012). Click strategies for single-molecule protein fluorescence. J. Am. Chem. Soc. *134*, 5187-5195.

Molteni, G., Orlandi, M., and Broggini, G. (2000). Nitrilimine cycloadditions in aqueous media. J. Chem. Soc. *1*, 3742-3745.

Mukai, T., Lajoie, M.J., Englert, M., and Söll, D. (2017). Rewriting the genetic code. Annu. Rev. Microbiol. *71*, 557-577.

Neef, A.B., and Schultz, C. (2009). Selective fluorescence labeling of lipids in living cells. Angew. Chem. Int. Ed. Engl. *48*, 1498-1500.

Nikic, I., Kang, J.H., Girona, G.E., Aramburu, I.V., and Lemke, E.A. (2015). Labeling proteins on live mammalian cells using click chemistry. Nat. Protoc. *10*, 780-791.

Nikic, I., Plass, T., Schraidt, O., Szymański, J., Briggs, J.A., Schultz, C., and Lemke, E.A. (2014). Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. Angew. Chem. Int. Ed. Engl. *53*, 2245-2249.

Ning, X., Guo, J., Wolfert, M.A., and Boons, G.J. (2008). Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast huisgen cycloadditions. Angew. Chem. Int. Ed. Engl. *47*, 2253-2255.

Oliveira, B.L., Guo, Z., and Bernardes, G.J.L. (2017). Inverse electron demand Diels-Alder reactions in chemical biology. Chem. Soc. Rev. *46*, 4895-4950.

Peng, T., and Hang, H.C. (2016). Site-specific bioorthogonal labeling for fluorescence imaging of intracellular proteins in living cells. J. Am. Chem.

Soc. *138*, 14423-14433.

Plass, T., Milles, S., Koehler, C., Schultz, C., and Lemke, E.A. (2011). Genetically encoded copper-free click chemistry. Angew. Chem. Int. Ed. Engl. *50*, 3878-3881.

Plass, T., Milles, S., Koehler, C., Szymański, J., Mueller, R., Wiessler, M., Schultz, C., and Lemke, E.A. (2012). Amino acids for Diels-Alder reactions in living cells. Angew Chem. Int. Ed. Engl. *51*, 4166-4170.

Polycarpo, C., Ambrogelly, A., Bérubé, A., Winbush, S.M., McCloskey, J.A., Crain, P.F., Wood, J.L., and Söll, D. (2004). An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. Proc. Natl. Acad. Sci. U. S. A. *101*, 12450-12454.

Prescher, J.A., and Bertozzi, C.R. (2005). Chemistry in living systems. Nat. Chem. Biol. *1*, 13-21.

Prokhorov, A.M., and Kozhevnikov, D.N. (2012). Reactions of triazines and tetrazines with dienophiles. Chem. Heterocycl. Compd. *48*, 1153-1176.

Ramil, C.P., and Lin, Q. (2014). Photoclick chemistry: a fluorogenic light-triggered in vivo ligation reaction. Curr. Opin. Chem. Biol. *21*, 89-95.

Saxon, E., and Bertozzi, C.R. (2000). Cell surface engineering by a modified Staudinger reaction. Science *287*, 2007-2010.

Sengupta, P., Van Engelenburg, S., and Lippincott-Schwartz, J. (2012). Visualizing cell structure and function with point-localization superresolution imaging. Dev. Cell *23*, 1092-1102.

Serfling, R., Lorenz, C., Etzel, M., Schicht, G., Böttke, T., Mörl, M., and Coin, I. (2018). Designer tRNAs for efficient incorporation of non-canonical amino acids by the pyrrolysine system in mammalian cells. Nucleic Acids Res. *46*, 1-10.

Shaner, N.C., Patterson, G.H., and Davidson, M.W. (2007). Advances in fluorescent protein technology. J. Cell Sci. *120*, 4247-4260.

Singh, I., and Heaney, F. (2011). Solid phase strain promoted "click" modification of DNA via [3+2]-nitrile oxide-cyclooctyne cycloadditions. Chem. Commun. (Camb.) *47*, 2706-2708.

Sletten, E.M., and Bertozzi, C.R. (2009). Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew. Chem. Int. Ed. Engl. *48*, 6974-6998.

Song, W., Wang, Y., Qu, J., Madden, M.M., and Lin, Q. (2008). A photoinducible *1*,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. Angew. Chem. Int. Ed. Engl. *47*, 2832-2835.

Stephens, D.J., and Allan, V.J. (2003). Light microscopy techniques for live cell imaging. Science *300*, 82-86.

Stone, S.E., Glenn, W.S., Hamblin, G.D., and Tirrell, D.A. (2017). Cell-selective proteomics for biological discovery. Curr. Opin. Chem. Biol. *36*, 50-57.

Su Hui Teo, C., Serwa, R.A., and O'Hare, P. (2016). Spatial and temporal resolution of global protein synthesis during HSV infection using bioorthogonal precursors and click chemistry. PLoS Pathog. *12*, e1005927.

Summerer, D., Chen, S., Wu, N., Deiters, A., Chin, J.W., and Schultz, P.G. (2006). A genetically encoded fluorescent amino acid. Proc. Natl. Acad. Sci. U. S. A. *103*, 9785-9789.

Suzuki, T., Miller, C., Guo, L.T., Ho, J.M.L., Bryson, D.I., Wang, Y.S., Liu, D.R., and Söll, D. (2017). Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. Nat. Chem. Biol. *13*, 1261-1266.

Swiderska, K.W., Szlachcic, A., Czyrek, A., Zakrzewska, M., and Otlewski, J. (2017). Site-specific conjugation of fibroblast growth factor 2 (FGF2) based on incorporation of alkyne-reactive unnatural amino acid. Bioorg. Med. Chem. *25*, 3685-3693.

Thalhammer, F., Wallfahrer, U., and Saue, J. (1990). Reaktivität einfacher offenkettiger und cyclischer dienophile bei Diels-Alder-reaktionen mit inversem elektronenbedarf. Tetrahedron Lett. *31*, 6851-6854.

Tom Dieck, S., Müller, A., Nehring, A., Hinz, F.I., Bartnik, I., Schuman, E.M., and Dieterich, D.C. (2012). Metabolic labeling with noncanonical amino acids and visualization by chemoselective fluorescent tagging. Curr. Protoc. Cell Biol. *56*, 7.11.1-7.11.29.

Toomre, D., and Bewersdorf, J. (2010). A new wave of cellular imaging. Annu. Rev. Cell Dev. Biol. *26*, 285-314.

Tornoe, C.W., Christensen, C., and Meldal, M. (2002). Peptidotriazoles on solid phase: [*1*,*2*,3]-triazoles by regiospecific copper(i)-catalyzed *1*,3-dipolar cycloadditions of terminal alkynes to azides. J. Org. Chem. *67*, 3057-3064.

Tsien, R.Y. (1998). The green fluorescent protein. Annu. Rev. Biochem. *67*, 509-544.

Uttamapinant, C., Howe, J.D., Lang, K., Beránek, V., Davis, L., Mahesh, M., Barry, N.P., and Chin, J.W. (2015). Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. J. Am. Chem. Soc. *137*, 4602-4605.

van de Linde, S., Heilemann, M., and Sauer, M. (2012). Live-cell super-resolution imaging with synthetic fluorophores. Annu. Rev. Phys. Chem. *63*, 519-540.

Vreja, I.C., Nikić, I., Göttfert, F., Bates, M., Kröhnert, K., Outeiro, T.F., Hell, S.W., Lemke, E.A., and Rizzoli, S.O. (2015). Super-resolution microscopy of clickable amino acids reveals the effects of fluorescent protein tagging on protein assemblies. ACS Nano *9*, 11034-11041.

Wan, W., Tharp, J.M., and Liu, W.R. (2014). Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. Biochim. Biophys. Acta *1844*, 1059-1070.

Wang, J., Xie, J., and Schultz, P.G. (2006). A genetically encoded fluorescent amino acid. J. Am. Chem. Soc. *128*, 8738-8739.

Wang, Y., Vera, C.I., and Lin, Q. (2007). Convenient synthesis of highly functionalized pyrazolines via mild, photoactivated *1*,3-dipolar cycloaddition. Org. Lett. *9*, 4155-4158.

Weber, G., and Farris, F.J. (1979). Synthesis and spectral properties of a hydrophobic fluorescent probe: 6-propionyl-2-(dimethylamino) naphthalene. Biochemistry *18*, 3075-3078.

Wu, Y., Zhu, H., Zhang, B., Liu, F., Chen, J., Wang, Y., Wang, Y., Zhang, Z., Wu, L., Si, L., et al. (2016). Synthesis of site-specific radiolabeled antibodies for radioimmunotherapy via genetic code expansion. Bioconjug. Chem. *27*, 2460-2468.

Young, D.D., and Schultz, P.G. (2018). Playing with the molecules of life. ACS Chem. Biol. *13*, 854-870.

Yu, Z., Pan, Y., Wang, Z., Wang, J., and Lin, Q. (2012). Genetically encoded cyclopropene directs rapid, photoclick-chemistry-mediated protein labeling in mammalian cells. Angew. Chem. Int. Ed. Engl. *51*, 10600-10604.

Yuan, J., Palioura, S., Salazar, J.C., Su, D., O'Donoghue, P., Hohn, M.J., Cardoso, A.M., Whitman, W.B., and Söll, D. (2006). RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. Proc. Natl. Acad. Sci. U. S. A. *103*, 18923-18927.

Zhang, J., Yan, S., He, Z., Ding, C., Zhai, T., Chen, Y., Li, H., Yang, G., Zhou, X., and Wang, P. (2018). Small unnatural amino acid carried raman tag for molecular imaging of genetically targeted proteins. J. Phys. Chem. Lett. *9*, 4679-4685.

Zhou, Z., Cironi, P., Lin, A.J., Xu, Y., Hrvatin, S., Golan, D.E., Silver, P.A., Walsh, C.T., and Yin, J. (2007). Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. ACS Chem. Biol. *2*, 337-346.