CHAPTER SEVEN

Novel Fluorescence-Based Biosensors Incorporating Unnatural Amino Acids

Wei Niu[1], **Jiantao Guo**[1]
University of Nebraska–Lincoln, Lincoln, NE, United States
[1]Corresponding authors: e-mail address: wniu2@unl.edu; jguo4@unl.edu

Contents

Abstract

Fluorescent proteins of different colors are useful probes to study protein structure and function, and to investigate cellular events and conditions. Large efforts have focused on engineering new properties into fluorescent proteins via rational design and directed evolution. In addition to applications in imaging of protein expression level and subcellular localization, fluorescent proteins have been increasingly engineered to act as biosensors to track concentrations of small-molecule metabolites, enzyme activities, and protein conformational changes in living cells. Unlike small-molecule fluorescence biosensors, fluorescent proteins are genetically encodable, and thus can be expressed inside living cells. Attachment of organelle-specific signals to the proteins allows their localization to be specified. Recently, a new class of fluorescent protein biosensors has been developed to include unnatural amino acids as the sensing element.

Methods in Enzymology, Volume 589
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2017.01.012

The unique chemical and physical properties of the unnatural amino acids enable sensor designs that cannot be realized by using the standard genetic code with the 20 canonical amino acids. In this chapter, we detail the general procedure for the genetic incorporation of unnatural amino acids. We further present two protocols for the in vitro and in vivo detection of hydrogen peroxide (H_2O_2) using a fluorescent protein biosensor that contains an unnatural amino acid, *p*-boronophenylalanine.

1. INTRODUCTION

Fluorescent proteins are useful probes that have greatly facilitated the study of biological molecules and processes (Chudakov, Matz, Lukyanov, & Lukyanov, 2010; Frommer, Davidson, & Campbell, 2009; Newman, Fosbrink, & Zhang, 2011; Tsien, 1998; Wang, Nakata, & Hamachi, 2009; Zimmer, 2002). One important application of fluorescent proteins is to tag and visualize target proteins in living systems under physiological conditions using fluorescence microscopy (Chudakov et al., 2010; Frommer et al., 2009; Newman et al., 2011; Shaner et al., 2004; Tsien, 1998; Wang et al., 2009; Zimmer, 2002). In order to monitor multiple biomolecules simultaneously, many variants of fluorescent proteins with different emission spectral profiles have been isolated or engineered (Matz et al., 1999; Miyawaki, Shcherbakova, & Verkhusha, 2012; Shaner et al., 2004). In addition, fluorescent proteins with photocontrollable properties have been used in superresolution microscopy (Chi, 2009). Fluorescent proteins have also been employed as biosensors to actively track changes in concentrations of small-molecule metabolites, enzyme activities, protein conformations, and signaling pathways in living cells (Frommer et al., 2009; Newman et al., 2011). The beauty of fluorescent protein biosensors lies in their genetic encodability, which allows the easy introduction into biological subjects, including cells, tissues, or whole animals, and the facile targeting into different cell organelles.

The traditional design of fluorescent protein biosensors is restricted to the standard genetic code of the common 20 amino acids. The side-chain functional groups of these amino acids, which are key elements in biosensing, are limited to amine, amide, acid, alcohol, thiol, thioether, and alkyl/aryl groups. Consequently, the genetic incorporation of amino acids containing unnatural functional groups with novel chemical and/or physical properties can potentially enhance one's ability to engineer novel biosensors. This notion is demonstrated by several recent examples in which fluorescent

protein engineering is combined with powerful unnatural amino acid mutagenesis methodologies (Ayyadurai et al., 2011; Chen, Chen, Ren, & Ai, 2012; Chen, Ren, Wright, & Ai, 2013; Groff, Wang, Jockusch, Turro, & Schultz, 2010; Liu et al., 2013; Niu & Guo, 2013; Wang, Niu, Guo, & Schultz, 2012).

2. SITE-SPECIFIC INCORPORATION OF UNNATURAL AMINO ACIDS INTO PROTEINS IN LIVE CELLS

A number of methods have been reported for the introduction of unnatural amino acids into proteins in living cells (Budisa et al., 1999; Datta, Wang, Carrico, Mayo, & Tirrell, 2002; Döring et al., 2001; Dougherty, 2000; Hohsaka & Sisido, 2002; Liu & Schultz, 2010). One of the most widely used approaches was developed in the laboratory of Dr. Peter G. Schultz at the Scripps Research Institute (Liu & Schultz, 2010). This general approach has been employed to site-specifically incorporate unnatural amino acids into proteins in living bacterial, yeast, and mammalian cells. The incorporation system features the introduction of two new components into the endogenous protein biosynthetic machinery, including an orthogonal tRNA-codon pair and an aminoacyl-tRNA synthetase that together specifically charge the orthogonal tRNA with the desirable unnatural amino acid.

The orthogonal tRNA must not be recognized by the native aminoacyl-tRNA synthetases of the host cell. In the meantime, after the aminoacylation by its cognate tRNA synthetase, the aminoacylated orthogonal tRNA can function efficiently in translation. This orthogonal tRNA should also be able to decode a noncoding codon (also called "blank codon," which does not encode a natural amino acid). The most utilized noncoding codon in this system is the amber nonsense codon (UAG), which is the least utilized stop codon in both *Escherichia coli* and in *Saccharomyces cerevisiae*. Introduction of the amber suppression system, which consists of an engineered aminoacyl-tRNA synthetase and amber suppressor $tRNA_{CUA}$ pair that can decode the amber codon with an unnatural amino acid, does not significantly affect cell growth. In fact, several common lab *E. coli* strains (e.g., DH5α and XL-1 blue) contain natural amber suppressors. On the other hand, overexpression of amber suppressor $tRNA_{CUA}$ from a plasmid containing multiple copies of $tRNA_{CUA}$ genes could lead to notable slow cell growth and, in some cases, the loss of $tRNA_{CUA}$ genes from the plasmid. In order to improve amber suppression efficiency and to reduce

potential adverse effects of $tRNA_{CUA}$ expression, efforts have been made to further engineer $tRNA_{CUA}$ (Guo, Melancon, Lee, Groff, & Schultz, 2009). In recent years, quadruplet codons have also been used to encode unnatural amino acids (Anderson et al., 2004; Neumann, Wang, Davis, Garcia-Alai, & Chin, 2010; Niu, Schultz, & Guo, 2013; Rodriguez, Lester, & Dougherty, 2006; Taki, Matsushita, & Sisido, 2006; Wang et al., 2014). In order to employ quadruplet codons in protein biosynthesis, tRNAs need to be engineered to bear extended anticodon loops (8-nucleotide instead of normal 7-nucleotide loop) and to form apparent quadruplet anticodons (Neumann et al., 2010; Niu et al., 2013; Wang et al., 2014; Wang, Shang, Cerny, Niu, & Guo, 2016).

An engineered orthogonal aminoacyl-tRNA synthetase is the other essential component of the system. The orthogonal synthetase should only aminoacylate the orthogonal tRNA, its cognate partner, but not any of the endogenous tRNAs. In addition, the aminoacylation reaction should only use the desired unnatural amino acid as the substrate, but not any of the 20 common amino acids. Likewise, the unnatural amino acid should not be utilized by any of the endogenous aminoacyl-tRNA synthetases.

The above methodology has enabled high-efficiency incorporations of over 100 unnatural amino acids with a variety of side-chain structures and functions, including spectroscopic probes (Cellitti et al., 2008; Deiters, Geierstanger, & Schultz, 2005; Groff, Thielges, Cellitti, Schultz, & Romesberg, 2009; Jackson, Hammill, & Mehl, 2007; Schultz et al., 2006; Wang, Xie, & Schultz, 2006), metal chelators (Lee, Spraggon, Schultz, & Wang, 2009; Liu et al., 2013, 2012; Xie, Liu, & Schultz, 2007; Ye et al., 2008), posttranslational modification analogs (Guo, Wang, Lee, & Schultz, 2008; Liu & Schultz, 2006; Xie, Supekova, & Schultz, 2007), redox-active groups (Alfonta, Zhang, Uryu, Loo, & Schultz, 2003; Seyedsayamdost, Xie, Chan, Schultz, & Stubbe, 2007; Tippmann & Schultz, 2007), and probes with unique chemical reactivities (Brustad, Bushey, Brock, Chittuluru, & Schultz, 2008; Brustad, Lemke, Schultz, & Deniz, 2008; Chen et al., 2014; Chin et al., 2002; Lang, Davis, Torres-Kolbus et al., 2012; Lang, Davis, Wallace et al., 2012; Wang, Zhang, Brock, & Schultz, 2003; Xiang et al., 2014; Zhang, Wang, Brock, & Schultz, 2002). One example to demonstrate the utility of the incorporated unnatural amino acid is the genetic incorporation of sulfotyrosine into proteins in live cells. In nature, sulfotyrosine is formed through posttranslational modification of tyrosine residues in

proteins. The introduction of negatively charged sulfate group plays crucial roles in extracellular biomolecular interactions that dictate various cellular processes including cell adhesion, leukocyte trafficking, hormone activities, and immune responses (Huttner, 1988; Stone, Chuang, Hou, Shoham, & Zhu, 2009). The ability to site-specifically incorporate sulfotyrosine residues into proteins of interest greatly facilitates biochemical studies of protein tyrosine sulfation (Ju et al., 2013; Ju, Niu, & Guo, 2016; Liu, Brustad, Liu, & Schultz, 2007). More commonly, the incorporated unnatural amino acid serves as a chemical handle in protein-labeling experiments. An exciting recent development is to achieve the fluorogenic protein labeling through bioorthogonal reactions enabled by genetically incorporated unnatural amino acids, such as ones bearing alkene groups (Shang et al., 2017; Song, Wang, Qu, Madden, & Lin, 2008). Alkenes can specifically react with either tetrazole (Song et al., 2008) or tetrazine (Shang et al., 2017) to generate fluorescent products. The fluorogenic property of these reactions leads to low background fluorescence interference, which is more desirable in studies of protein folding and function under physiological conditions. In recent years, efforts have also been made to further improve the efficiency and fidelity of unnatural amino acid mutagenesis through optimizing properties of the aminoacyl-tRNA synthetase-$tRNA_{CUA}$ pair (Amiram et al., 2015; Wang, Ju, Niu, & Guo, 2015).

A general scheme for the genetic incorporation of unnatural amino acids into proteins in live *E. coli* cells is shown in Fig. 1. All genetic components are encoded on plasmid(s). Commonly, a two-plasmid system (Fig. 2) is employed. Strains that are able to naturally suppress amber codons should not be used, e.g., XL1-Blue or DH5α. Suitable and nonamber suppressor strains include BL21, BL21(DE3), GeneHogs, DH10B, and TOP10. After transforming the two plasmids into a proper *E. coli* host strain, cells are grown and selected using two compatible antibiotic markers. The protein expression is induced at the exponential phase of cell growth ($OD_{600} = 0.6–1.0$). The desirable unnatural amino acid can either be included in the initial culture media or be added to the media together with the induction reagent for protein expression. In general, a chemically and metabolically stable unnatural amino acid is preferably included in initial media in order to suppress the residual misincorporation of natural amino acids by engineered aminoacyl-tRNA synthetase. On the other hand, a relatively unstable unnatural amino acid should be added at the induction step.

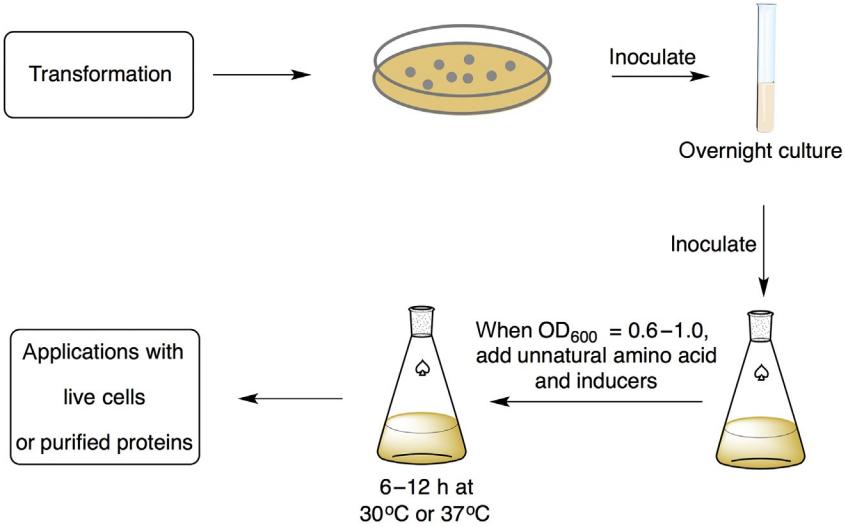

Fig. 1 A general scheme of genetic incorporation of unnatural amino acids into proteins in bacteria.

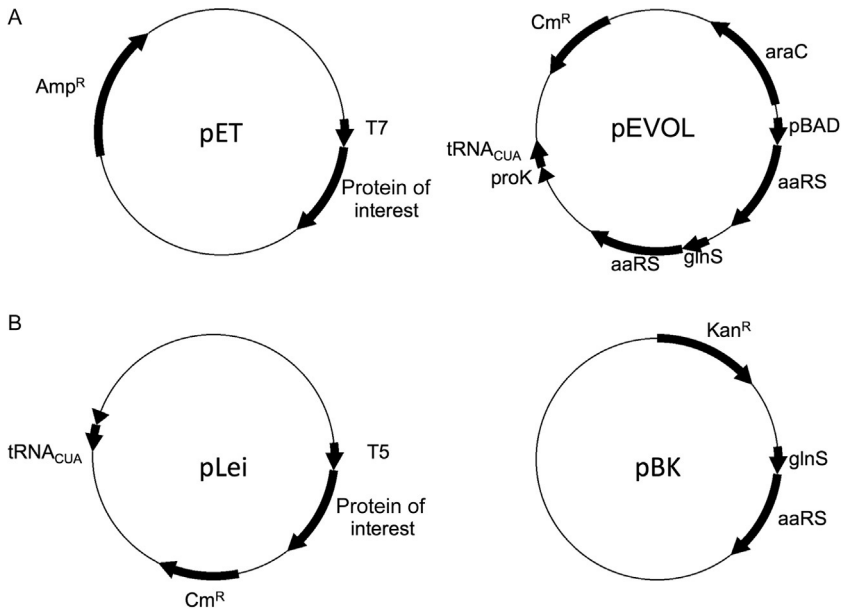

Fig. 2 General plasmid pairs for unnatural amino acid mutagenesis. (A) The pEVOL and pET system. (B) The pBK and pLei system.

2.1 Materials

- 37°C and/or 30°C incubator
- 37°C and/or 30°C shaker
- 42°C water bath
- UV–vis spectrophotometer
- Refrigerated centrifuge
- Autoclave for sterilization
- pEVOL plasmid containing genes encoding amber suppressor tRNA and aminoacyl-tRNA synthetase (available from Professor Peter G. Schultz, The Scripps Research Institute, La Jolla, CA)
- pBK plasmid containing the gene encoding aminoacyl-tRNA synthetase (available from Professor Peter G. Schultz, The Scripps Research Institute, La Jolla, CA)
- pET30 (Invitrogen) or similar expression plasmid containing the encoding gene of the target protein
- pLei plasmid containing the encoding gene of the target protein and the gene of amber suppressor tRNA (Vector is available from Professor Peter G. Schultz, The Scripps Research Institute, La Jolla, CA)
- High-fidelity DNA polymerase (New England Biolabs)
- Antarctic phosphatase (New England Biolabs)
- T4 DNA ligase (New England Biolabs)
- Restriction enzymes (New England Biolabs)
- Chemically competent DH5α (Thermo Fisher Scientific) or similar *E. coli* hosts for plasmid propagation and isolation
- Chemically competent BL21(DE3) (Thermo Fisher Scientific), GeneHogs (Thermo Fisher Scientific), DH10B (Thermo Fisher Scientific), or other similar *E. coli* strains that can be used for the expression of recombinant proteins
- Chloramphenicol (Cm), 34 mg/mL (1000 ×) stock solution in ethanol, stored at –20°C
- Ampicillin, sodium salt (Amp), 100 mg/mL (1000 ×) stock solution in sterile H_2O, stored at –20°C
- Kanamycin sulfate (Kan), 50 mg/mL (1000 ×) stock solution in sterile H_2O, stored at –20°C
- LB liquid medium
- LB Agar plates
- *p*-Azido phenylalanine hydrochloride (pAzPhe) (Bachem)
- *p*-Nitro phenylalanine hydrochloride (pNitroPhe) (Bachem)

- *p*-Boronophenylalanine hydrochloride (pBoPhe) (Sigma-Aldrich)
- Isopropyl β-D-1-thiogalactopyranoside (IPTG), 1 *M* (1000 ×) stock solution in H_2O, sterile filtered through a 0.2-μm membrane, stored at –20°C
- L-(+)-Arabinose, 20 mg/mL (100 ×) stock solution in H_2O, sterile filtered through a 0.2-μm membrane, stored at –20°C
- Ni Sepharose 6 Fast Flow resin (GE Healthcare) and buffers for protein purification
- Fluorometer or fluorescence plate reader
- Home-made or commercial SDS-PAGE gel (e.g., NuPAGE Novex 4%–12% SDS gels from Thermo Fisher Scientific)

2.2 Plasmids for Protein Expression

Two general plasmid systems are commonly used (Fig. 2).

2.2.1 The pEVOL and pET Combination

The pEVOL plasmid contains genes that encode amber suppressor tRNA and aminoacyl-tRNA synthetase (aaRS). There are two copies of aaRS gene on the pEVOL plasmid. One aaRS gene is under the control of the constitutive *glnS* promoter and the second one is under the control of the arabinose-inducible P_{BAD} promoter and the *rrnB* terminator. Simultaneous overexpression of both copies of aaRS gene enhances the production of the full-length target protein (Cellitti et al., 2008). The expression of amber suppressor tRNA is under the control of a *proK* promoter and terminator. The pEVOL plasmid also contains a chloramphenicol acetyltransferase marker (Cm^R), an *araC* repressor gene, and the p15A origin of replication (Fig. 2A).

The target protein is encoded by a pET plasmid. The expression of the target protein is under the control of an inducible *T*7 promoter and the *T*7 terminator. The pET30 plasmid contains a pBR322 origin of replication and a kanamycin resistance gene (Kan^R). The expressed target protein contains a C-terminal His_6 tag for detection and purification. Other similar vectors can be used to encode the target protein of interest. The target protein vector should have compatible replication origin and selection maker to the pEVOL plasmid so that these two plasmids can be simultaneously maintained in the *E. coli* host. Typical hosts for the pEVOL-pET system are BL21(DE3) or other similar strains that express the T7 RNA polymerase.

A C-terminal His_6 tag is typically employed for unnatural amino acid incorporation using the amber suppression approach. Such design simplifies the purification process, since only full-length target protein, which contains the unnatural amino acid of interest, has a His_6 tag. Inability to decode the

amber stop codon leads to translational termination, which produces truncated proteins without a His_6 Tag. On the other hand, if the His_6 tag is fused to the N-terminus of the protein of interest, truncated protein can be copurified with the full-length protein.

2.2.2 The pBK and pLei Combination

The pBK plasmid contains the aaRS gene, under the control of a constitutive *glnS* promoter and terminator. The pBK plasmid also encodes a kanamycin resistance protein (Kan^R) and a ColE1 origin of replication (Fig. 2B).

The target protein is encoded by the pLei plasmid. The expression of the target protein is under the control of an inducible *T5* promoter and a lambda t_0 transcriptional terminator. The pLei plasmid also encodes the amber suppressor tRNA that is under the control of a constitutive *lpp* promoter and *rrnC* terminator. This plasmid contains a p15A origin of replication and a chloramphenicol acetyltransferase marker (Cm^R).

The expressed target protein contains a C-terminal cleavable His_6 tag and S-tag for detection and purification. Other similar vectors can be used to encode the target protein of interest. Such vector should have a compatible replication origin and selection marker to the pBK plasmid so that these two plasmids can be simultaneously maintained in the *E. coli* host. Typical hosts for the pBK-pLei system include DH10B and GeneHogs.

2.3 Expression of Fluorescent Proteins Containing Unnatural Fluorophores

2.3.1 Mutagenesis and Verification

In order to generate an unnatural fluorophore, the fluorophore-forming tyrosine (e.g., Tyr66 in GFP) residue can be replaced with an unnatural amino acid of one's interest. To this end, the codon (TAC) of Tyr66 needs to be mutated into an amber nonsense codon (TAG). Two general methods can be employed for the site-directed mutagenesis. One is the QuikChange (Agilent Technologies) method, which uses a pair of complementary mutagenic primers to amplify the entire plasmid in a thermocycling reaction. A free online program can facilitate the primer design (http://www.genomics.agilent.com/primerDesignProgram.jsp). The second method uses overlapping PCR where a specific primer is designed to contain the desirable amber mutation. After the mutagenesis experiment, the presence of the amber TAG mutation in the gene of interest should be verified by DNA sequencing.

2.3.2 Cotransformation of Host E. coli Cells

Standard transformation protocols can be used in this step. If the pEVOL and pET combination is used, the transformants are plated onto LB-Agar plates containing Cm (34 μg/mL) and Amp (100 μg/mL). If the pBK and pLei combination is used, the transformants are plated onto LB-Agar plates containing Kan (50 μg/mL) and Cm (34 μg/mL).

2.3.3 Protein Expression

A single colony from the Agar plate is used to inoculate 5 mL LB liquid media containing appropriate antibiotics at 37°C. The seed culture (1 mL) is used to inoculate fresh LB media (50 mL) containing appropriate antibiotics at 37°C. For the pBK and pLei system, the protein expression is induced at OD_{600} of 0.6–1.0 by the additions of IPTG (0.2–0.5 m*M*) and unnatural amino acid of interest (1 m*M*). For the pEVOL and pET system, the protein expression is induced at OD_{600} of 0.6–1.0 by the additions of IPTG (0.2–0.5 m*M*), L-arabinose (0.2% (w/v)), and unnatural amino acid of interest (1 m*M*). Following an additional 4–12 h of cultivation, cells are collected by centrifugation at 5000 g and 4°C for 15 min. Harvested cells are resuspended in lysis buffer containing potassium phosphate (20 m*M*, pH 7.4), NaCl (150 m*M*), and imidazole (10 m*M*). Cells are subsequently disrupted by sonication. Cellular debris is removed by centrifugation (21,000 × *g*, 30 min, 4°C). The cell-free lysate is applied to Ni Sepharose 6 Fast Flow resin (GE Healthcare). Protein purification follows the manufacturer's instructions. Protein concentrations are determined by Bradford assay (Bio-Rad). Purified protein can be used directly or desalted prior to next step.

The expression of the target protein using the pLei system is leaky. Significant amounts of protein can be obtained without the addition of IPTG. The optimal IPTG concentration is usually between 0.2 and 0.5 m*M*. If tight control of protein expression is needed (e.g., the target protein is toxic to the host), a host strain that overexpresses the LacI protein can be used. Besides LB, minimal medium, such as GMML (1 × M9, 1 m*M* $MgSO_4$, 0.1 m*M* $CaCl_2$, 8.5 m*M* NaCl, 5 μ*M* $FeSO_4$, 1% glycerol), can be used. If the host is a nutrient auxotroph, appropriate supplements need to be included. In general, higher fidelity can be achieved in minimal medium, but lower protein yield is also observed in some cases due to low cell density. When high cell density is desirable, media that are rich in nutrients, such as 2YT medium, can also be used in the incorporation experiments.

The solubility of some unnatural amino acids is quite low in water. To solve this problem, a proper amount of unnatural amino acid of interest is first dissolved in a small amount of water by adjusting the pH value of the solution. Either base (e.g., 1 *M* NaOH in water) or acid (e.g., 1 *M* HCl in water) can be used depending on the pH stability of the unnatural amino acid of interest. Examples of basic pH-sensitive amino acids include 3-amino-L-tyrosine and 4-borono-L-phenylalanine. An examples of acidic pH-sensitive amino acid is N^{ϵ}-(*tert*-butyloxy-carbonyl)-L-lysine. The unnatural amino acid solution is subsequently added into the culture media, followed by the addition of an equimolar amount of base or acid to adjust the pH back to the original value.

To improve the yield of protein expression, an *E. coli* strain (C321.ΔA) can be used (Lajoie et al., 2013). This strain lacks amber nonsense codons and does not express release factor 1 (RF1; RF1 decodes UAG nonsense codon as the termination signal of protein translation). Such modifications abolish amber nonsense codon-mediated translation termination in this strain. The benefits of using RF1-minus strain in the production of unnatural amino acid-containing proteins have been reported (Chatterjee, Lajoie, Xiao, Church, & Schultz, 2014; Johnson et al., 2011; Wang et al., 2016).

To estimate the fidelity of unnatural amino acid incorporation into a fluorescent protein of interest, two parallel protein expression experiments are commonly conducted. In one experiment, the expression of the target fluorescent protein is carried out in the presence of the unnatural amino acid. In a second experiment, the expression is conducted in the absence of the unnatural amino acid. A good fidelity of incorporation is obtained if full-length fluorescent protein can only be detected when protein expression is conducted in the presence of unnatural amino acid.

2.3.4 Protein Purification and Characterization

Proteins can be purified by commonly used procedures, such as affinity chromatography using Ni-NTA resin. Generally, unnatural amino acids are stable under the affinity purification conditions. If needed, imidazole and salts in the solution of purified protein can be removed using a desalting column or by dialysis. The protein concentration can be quantified by Bradford assay (Bio-Rad). The purified protein can be analyzed by SDS-PAGE. To further confirm the site-specific incorporation of unnatural amino acid into the fluorescent protein of interest, mass spectrometry analysis is commonly conducted. The analysis can be based on the full-length protein or on protein fragments from trypsin digestion. The data of a typical analysis

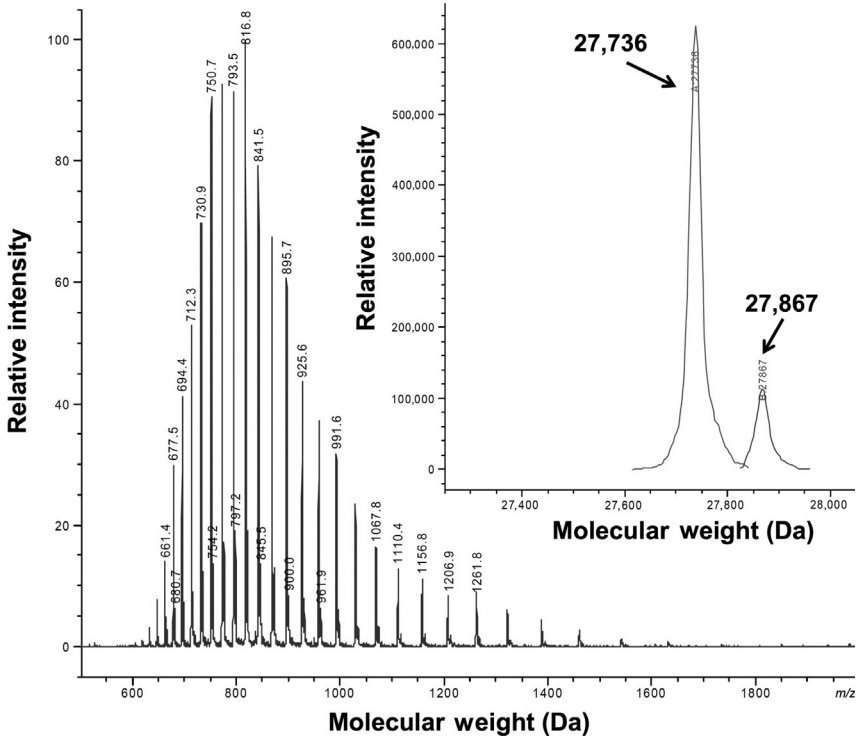

Fig. 3 Mass spectra of GFP mutant that contains *p*-boronophenylalanine (pBoPhe) at position 66. Expected mass: 27,738 Da (without N-terminal methionine) and 27,869 Da (with N-terminal methionine); observed mass: 27,736 Da (without N-terminal methionine) and 27,867 Da (with N-terminal methionine).

of full-length protein are shown in Fig. 3. In a typical analysis of trypsin-digested fragments, the band of the fluorescent protein of interest is cut from SDS-PAGE, then subjected to in-gel digestion with trypsin (in 50 m*M* ammonium bicarbonate, pH 8.0) overnight at 37°C. The resulting peptide fragments are extracted with 0.1% formic acid in 75% acetonitrile, and then analyzed by LC/MS/MS. Database searches are subsequently performed.

The fluorescence property of the unnatural amino acid-containing fluorescent proteins can be examined by a fluorometer or a fluorescence plate reader. The unnatural amino acid-containing fluorescent proteins may have similar or very different emission and excitation wavelength in comparison to their parent fluorescent proteins (Fig. 4).

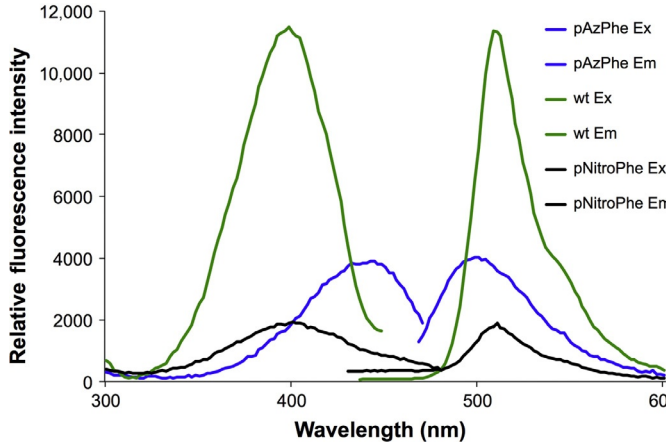

Fig. 4 Fluorescence excitation (Ex) and emission (Em) spectra of wild-type GFP and GFP mutants containing unnatural amino acids at position 66. Fluorescence spectra were measured in phosphate buffer (50 mm, pH 7.4). The emission spectra were recorded with maximum excitation wavelengths of each fluorescent proteins.

3. AN UNNATURAL FLUORESCENT PROTEIN BIOSENSOR FOR HYDROGEN PEROXIDE

3.1 General Design Principles

Since the report of a green fluorescent protein (GFP) biosensor for Ca^{2+} in 1997 (Miyawaki et al., 1997), more than 100 fluorescent protein biosensors have been developed to enable the monitoring of intracellular pH, redox potential, enzyme activities, and concentrations of small molecules (Frommer et al., 2009; Newman et al., 2011).

Several design approaches (Fig. 5) have been developed to engineer fluorescent protein biosensors. The first design (Fig. 5A) focuses on the protonation state of the chromophore. Under different pH, chromophores of fluorescent proteins can adopt altered protonation states, which makes fluorescent proteins naturally sensitive to pH changes (Bizzarri, Serresi, Luin, & Beltram, 2009). While four protonation states (protonated/deprotonated phenol group and alpha nitrogen from Tyr66) can exist in the GFP chromophore, the proton exchange reaction of the phenol group is the main factor in the design of pH sensors that work under physiological conditions. In the model chromophores, the pK_a values of the imidazolinone ring nitrogen (N^{66}; Fig. 5A) and the phenolic hydroxyl group are in the range of

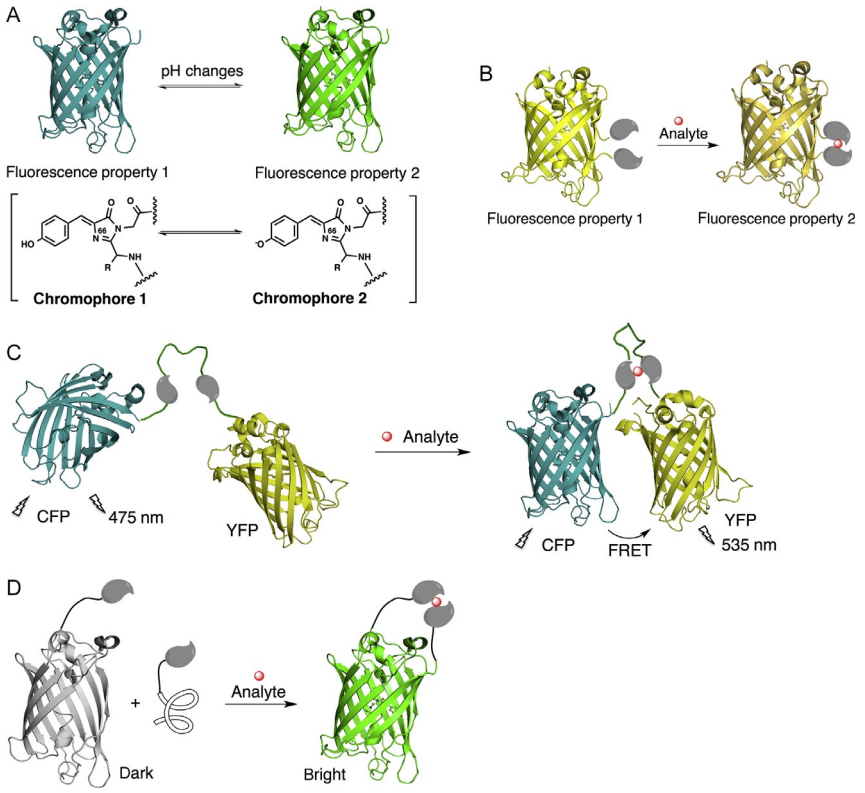

Fig. 5 General design principles of fluorescent protein biosensors. (A) Single fluorescent protein pH sensors that are based on the protonation state of the chromophore. (B) Sensors that are based on conformational change of circularly permuted fluorescent proteins (cpFP). (C) FRET-based fluorescent protein biosensors. (D) Biosensors that are based on bimolecular fluorescence complementation (BiFC).

1.8–2.4 and 8.1–8.5 (Bizzarri et al., 2009), respectively. The estimated pK_a value of the phenolic hydroxyl group of the chromophore in the folded protein is around 5.2–7.5 (Bizzarri et al., 2009). With a protonated phenol group of Tyr66, GFP absorbs at 398 nm and emits at 508 nm. On the other hand, GFP containing a deprotonated phenol group on Tyr66 absorbs at 475 nm and emits at 503 nm (Tsien, 1998).

The second general design is also a class of single fluorescent protein-based biosensors (Fig. 5B) (Frommer et al., 2009; Newman et al., 2011; Wang et al., 2009). It is known that the surrounding molecular environment of a fluorescent protein's chromophore defines its fluorescence signature. Therefore, chemical and/or physical changes to nonchromophore

residues can affect the optical properties of a fluorescent protein. For example, the reversible formation of a disulfide bond between two cysteine residues can influence the fluorescence signals of fluorescent proteins. This strategy has been employed to monitor redox potential changes in living cells (Hanson et al., 2004; Ostergaard, Tachibana, & Winther, 2004). In another example, a fluorescent protein, usually a circularly permuted fluorescent protein (cpFP), is fused to a protein scaffold(s) that act as a sensing domain. Upon recognition of the target, the sensing domain causes a conformational change of cpFP and leads to a shift in the fluorescence intensity or the fluorescence hue. This strategy was demonstrated by a genetically encoded hydrogen peroxide (H_2O_2) sensor, Hyper (Belousov et al., 2006), which is based on a cpYFP and the prokaryotic H_2O_2-sensing transcription regulator, OxyR.

In the third design strategy, two fluorescent proteins with overlapping excitation and emission spectra are, respectively, fused to the N- and the C-terminus of a protein scaffold of interest that acts as a sensing domain for the target recognition (Fig. 5C). The sensing domain undergoes a conformational change in response to the target, which leads to the change in the relative distance between the two fused fluorescent proteins. As a consequence, the efficiency of Förster resonance energy transfer (FRET) changed, which often results in the shift in emission (Frommer et al., 2009; Newman et al., 2011; Wang et al., 2009). A well-known example is the "cameleon" class of Ca^{2+} sensors (Mank & Griesbeck, 2008; Miyawaki et al., 1997), in which the fluorescent protein FRET pair is fused to a sensing domain containing calmodulin and a peptide that binds to calmodulin in the presence of Ca^{2+} ions.

The fourth design strategy is based on bimolecular fluorescence complementation (BiFC; Fig. 5D) (Cabantous et al., 2005; Wang et al., 2009), which is commonly known as "split" fluorescent protein. This approach entails the coexpression of two fusion proteins. Each of them contains one fragment of the split fluorescent protein and a part of sensing domains. Association of the sensing domains in the presence of target analyte brings the complementary fragments of the fluorescent protein reporter within proximity, which promotes the reconstitution of the reporter into its native or near native three-dimensional structure and the emission of fluorescence signal. A classic example of such a split fluorescent protein biosensor is the split-pericam (Nagai, Sawano, Park, & Miyawaki, 2001). The association of calmodulin and the M13 peptide in response to Ca^{2+} leads to the formation of an intact GFP.

3.2 General Design Principles for Unnatural Amino Acid-Containing Fluorescent Protein Biosensors

While design principles in the above section have led to the construction of many useful biosensors, significant challenges remain to improve the performance and to expand the diversity of fluorescent protein biosensors. Sensing elements of current fluorescent protein biosensors are either naturally existing proteins, such as transcription factors, or engineered sensing segments derived from natural amino acid (e.g., cysteine residues for disulfide bond formation), which not only limit the specificity and the affinity of fluorescent protein biosensors but also restrict the types of analytes that can be effectively detected. Recently, efforts have been made to develop a new group of fluorescent protein biosensors with expanded chemical/structural space through the incorporation of genetically encoded unnatural amino acids with unique chemical and/or physical properties.

The main focus of above engineering efforts is the chromophore of fluorescent proteins, such as GFP. Since the chemical properties and the structure of the chromophore are the major factors that define the photophysical properties of GFP, a maximum signal output may be achieved by directly modifying the composition of the chromophore. The chromophore of GFP is derived from a tripeptide sequence through the spontaneous post-translational cyclization of the three consecutive amino acid residues (Ser65-Tyr66-Gly67), followed by a dehydration reaction to form an imidazolin-5-one intermediate, and a final dehydrogenation reaction to yield the chromophore (Barondeau, Putnam, Kassmann, Tainer, & Getzoff, 2003; Chalfie, Tu, Euskirchen, Ward, & Prasher, 1994; Cubitt et al., 1995; Pouwels, Zhang, Chan, Dorrestein, & Wachter, 2008; Rao, Kemple, & Prendergast, 1980; Shu, Shaner, Yarbrough, Tsien, & Remington, 2006; Tsien, 1998; Verkhusha & Lukyanov, 2004; Zhang, Patel, Lappe, & Wachter, 2006). This process is self-catalyzed by the protein fold in the presence of molecular oxygen. As Tyr66 forms the major part of the chromophore, the replacement of Tyr66 with an unnatural analog generates an artificial chromophore that cannot be realized in nature. It has been shown previously that these unnatural chromophores displayed altered fluorescent properties from that of the wild-type GFP (Wang et al., 2012; Wang, Xie, et al., 2003).

A general approach to design fluorescent protein biosensors that contain unnatural amino acids is to replace Tyr66 by an unnatural tyrosine analog with a side chain that either selectively reacts with or has high affinity toward

the analyte of interest (Fig. 6). The chromophore of the fluorescent protein biosensor changes after the desirable reaction or binding event with the analyte, which leads to fluorescence property changes of the biosensor. These unnatural amino acid-containing fluorescent protein biosensors can generally lead to a large signal output. In order to be highly specific, these unnatural amino acids should only react or bind with the analyte of interest and are chemically orthogonal to other biomolecules in a living cell. Several unnatural amino acid-containing fluorescent protein biosensors have been

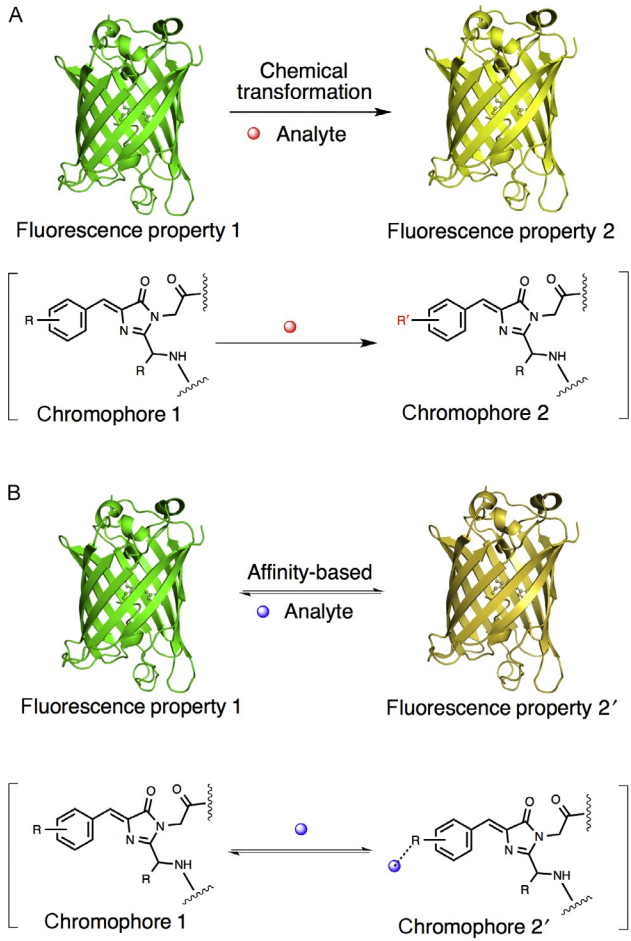

Fig. 6 Basic design principles of green fluorescent protein biosensors containing unnatural amino acid at position 66. (A) Chemical reaction-based biosensors. (B) Affinity/chelation-based biosensors.

developed by following above design principles. These include a transition metal (Cu^{2+} and Zn^{2+}) sensor (Ayyadurai et al., 2011; Liu et al., 2013), a hydrogen peroxide (H_2O_2) sensor (Wang et al., 2012), a peroxynitrite sensor (Chen et al., 2013), a hydrogen sulfide (H_2S) sensor (Chen et al., 2012), and a light sensor (Groff et al., 2010).

To construct this class of fluorescent protein biosensors, an unnatural amino acid of interest is genetically incorporated in response to an amber nonsense codon. Since the chromophore of a fluorescent protein is buried at the center of the β-barrel, it can only be accessed by analytes of small sizes. To enable the analysis of larger analytes and to improve the kinetics of detection, cpFP variants can be employed. Circular permutation is achieved by connecting the original N- and C-termini of fluorescent proteins through flexible peptide linkers and introducing new N- and C-termini at sites in close proximity to the chromophore (Baird, Zacharias, & Tsien, 1999; Nagai et al., 2001). While we discussed GFP specifically here, the above approach can be applied to other fluorescent proteins by replacing functionally equivalent Tyr residues.

3.3 Reaction-Based Fluorescent Protein Biosensor for Hydrogen Peroxide

Reactive oxygen species such as hydrogen peroxide (H_2O_2) acts as a second messenger that selectively modifies cysteine residues on target proteins and, therefore, participates in the control of many intracellular events in response to external stimuli (Paulsen & Carroll, 2010; Poole & Nelson, 2008; Rhee, 2006; Stone & Yang, 2006; Veal, Day, & Morgan, 2007). A prominent example is the reversible inactivation of protein tyrosine phosphatases in receptor tyrosine kinase signaling (Denu & Tanner, 1998; Lee, Kwon, Kim, & Rhee, 1998). Given the importance of H_2O_2 in biology, methods that can be used to selectively image the intracellular concentrations of H_2O_2 and do not interfere with cellular redox events in living systems are highly desirable. A few small-molecule fluorescent probes have been designed based on the chemoselective reaction of arylboronates with H_2O_2 (Lippert, Van de Bittner, & Chang, 2011). These probes are highly specific for H_2O_2 and are bioorthogonal to most cellular processes. In order to generate intracellular H_2O_2 biosensors that have good spatial resolution, a family of genetically encoded H_2O_2 sensor, Hyper (Belousov et al., 2006), Hyper-2 (Markvicheva et al., 2011), and Hyper-3 (Bilan et al., 2012) were constructed by fusing the regulatory domain of bacterial H_2O_2-sensing protein (OxyR) with a circularly permuted yellow fluorescent protein (cpYFP)

variant. The output signal is produced through the conformational change of cpYFP, which is induced by the intramolecular disulfide bond formation between two cysteine residues of OxyR in the presence of H_2O_2. A second class of genetically encoded H_2O_2 sensor, roGFP2-Orp1, is based on peroxidase-roGFP2 relay, in which the peroxidase Orp1-mediated oxidation of roGFP2 by H_2O_2 results to signal output (Gutscher et al., 2009). Both the Hyper and roGFP2-Orp1 families of sensors selectively respond to H_2O_2 with very good sensitivity, but lead to less than a fivefold signal enhancement. These hydrogen peroxide sensors are reversible due to the reversibility of disulfide bond formation. Therefore, they can be potentially affected by the redox state of cells, and thus cannot accurately image the accumulated H_2O_2 signal over time.

Here we present a single GFP-based H_2O_2 sensor, in which the Tyr66 is replaced with pBoPhe. The GFP mutant (GFP-Tyr66pBoPhe) is not fluorescent. Based on the crystal structure of GFP-Tyr66pBoPhe (Wang et al., 2012), the $-B(OH)_2$ group forms hydrogen bonds with the backbone nitrogen of His148 and the hydroxyl side chain of Ser205, which affects the planarity of the chromophore. In addition, the bulky $-B(OH)_2$ side chain pushes His148 away from the chromophore (Fig. 3A). The dislocation of His148 (Seifert et al., 2003) may also account for the lack of any observed fluorescence of GFP-Tyr66pBoPhe. In the presence of H_2O_2, arylboronate is oxidized into phenol to restore the native chromophore of GFP (Fig. 7A).

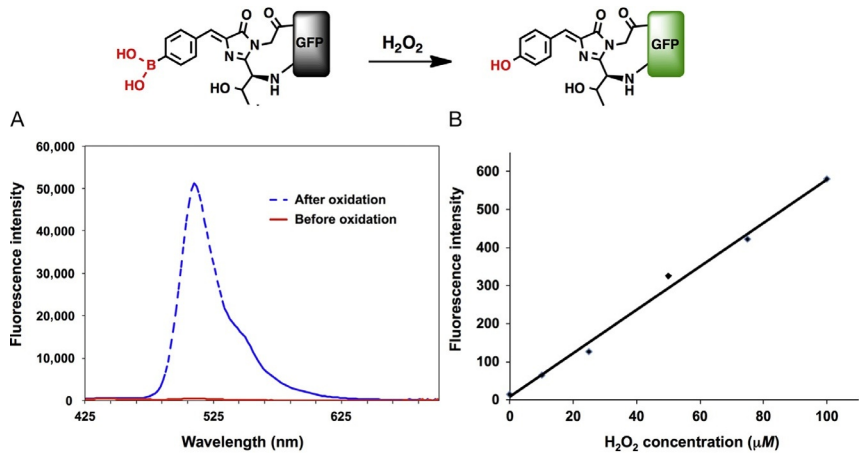

Fig. 7 A fluorescent protein-based H_2O_2 sensor. (A) Fluorescence of GFP-Tyr66pBoPhe before (*red trace*) and after (*blue trace*) H_2O_2 treatment. (B) Changes in fluorescence of GFP-Tyr66pBoPhe after the incubation with different concentrations of H_2O_2.

Therefore, GFP-Tyr66pBoPhe can be used as an H_2O_2 sensor. Since the oxidation of an arylboronate-based fluorescent protein biosensor by H_2O_2 is irreversible, the output signal of the sensor is insensitive to the rapidly recovered cellular redox state after a transient H_2O_2 burst. Therefore, the arylboronate-based fluorescent protein biosensor can be applied to monitoring the cumulative H_2O_2 signal in biological systems.

3.4 Materials

- Fluorescence plate reader
- Inverted confocal fluorescence microscope
- Refrigerated microcentrifuge
- Purified GFP-Tyr66BoF
- Purified wild-type GFP
- *E. coli* cells expressing GFP-Tyr66BoF
- *E. coli* cells expressing wild-type GFP
- 30% H_2O_2 (Sigma)
- PBS buffer
- Assay buffer (50 m*M* phosphate buffer, pH 7.4, containing 10% glycerol)
- 96-Well fluorescence plate (polystyrene plate, black, flat bottomed, and nontransparent; Costar 3915)
- Glass slides for imaging

3.5 Sample Preparation

3.5.1 With Purified Protein

After the purification by affinity chromatography, the isolated protein, GFP-Tyr66BoF (the Tyr66 of GFP is replaced by BoF), is dialyzed against the assay buffer (50 m*M* phosphate buffer, pH 7.4; 10% glycerol).

3.5.2 With Live Cells

Cells (1 mL) expressing GFP-Tyr66BoF are collected using benchtop centrifuge at $5000 \times g$ for 8 min. After removal of the supernatant, cells are resuspended in 1 mL PBS buffer and collected by centrifugation. This process is repeated for three times in order to remove residual culture media. After the final washing step, cells are resuspended in 1 mL PBS buffer and are ready for the next step of the experiment.

3.6 Biosensing

3.6.1 With Purified Protein

To a solution of GFP-Tyr66BoF (0.1–1 mg/mL) in assay buffer, a desirable amount of H_2O_2 (0.01–10 m*M*) is added and the resulting mixture is

incubated at room temperature for a desirable period of time (3–60 min). As one control, the GFP-Tyr66BoF solution is incubated under the same conditions without the addition of H_2O_2 (replaced with the same amount of water). Fluorescence intensities of the above samples are then measured using a fluorescence plate reader. As shown in Fig. 7A, strong fluorescence can be only observed with the protein sample that is treated with H_2O_2. The fluorescence properties of the H_2O_2-treated protein sample is identical to those of the wild-type GFP $\left(\lambda_{em}^{max} = 508\,nm\right)$. In general, significant fluorescence can be observed after 3 min of incubation with 1 m*M* H_2O_2. Longer time is needed when H_2O_2 concentration is low.

To determine if the fluorescence intensity has a linear correlation with the concentration of H_2O_2, GFP-Tyr66BoF samples are incubated with a range of different concentrations of H_2O_2 (0.01–10 m*M* final concentration). As shown in Fig. 7B, the fluorescence intensities of H_2O_2-treated protein samples correlate well with the concentrations (0.01–0.1 m*M*) of H_2O_2. The linear relationship is observed when H_2O_2 concentration is lower than 10 m*M*. Within a broader range of H_2O_2 concentration (0.01–100 m*M*), a biphasic curve of fluorescence vs H_2O_2 concentration is observed. This likely results from the oxidation of wild-type GFP chromophore at high concentrations of H_2O_2. Indeed, a decrease in fluorescence intensity can be observed when wild-type GFP is incubated with 100 m*M* H_2O_2.

3.6.2 With Live Cells

The fluorescence intensity of washed cells expressing GFP-Tyr66BoF in PBS buffer (100 μL with an OD_{600} of 0.5–1.0) is measured using a fluorescence plate reader. The observed fluorescence intensity is very low (RFU = ~50–100) and is comparable to that of cells that do not express GFP-Tyr66BoF. The same cell sample is incubated with 100 m*M* H_2O_2 at room temperature. The fluorescence intensity is monitored every minute in a 15-min period. As shown in Fig. 8A, the fluorescence intensity of the sample increases as the incubation time increases. In this experiment, a high extracellular concentration of H_2O_2 is needed to attain a sufficient intracellular H_2O_2 concentration. Degradation of H_2O_2 is observed (the appearance of gas bubbles), likely due to the presence of catalase activity in the *E. coli* host. *E. coli* cells are still viable after H_2O_2 treatment at 100 m*M* concentration. As a control, no fluorescence intensity change can be detected with the host *E. coli* cells that are incubated with 100 m*M* H_2O_2 under the same conditions (Fig. 8B).

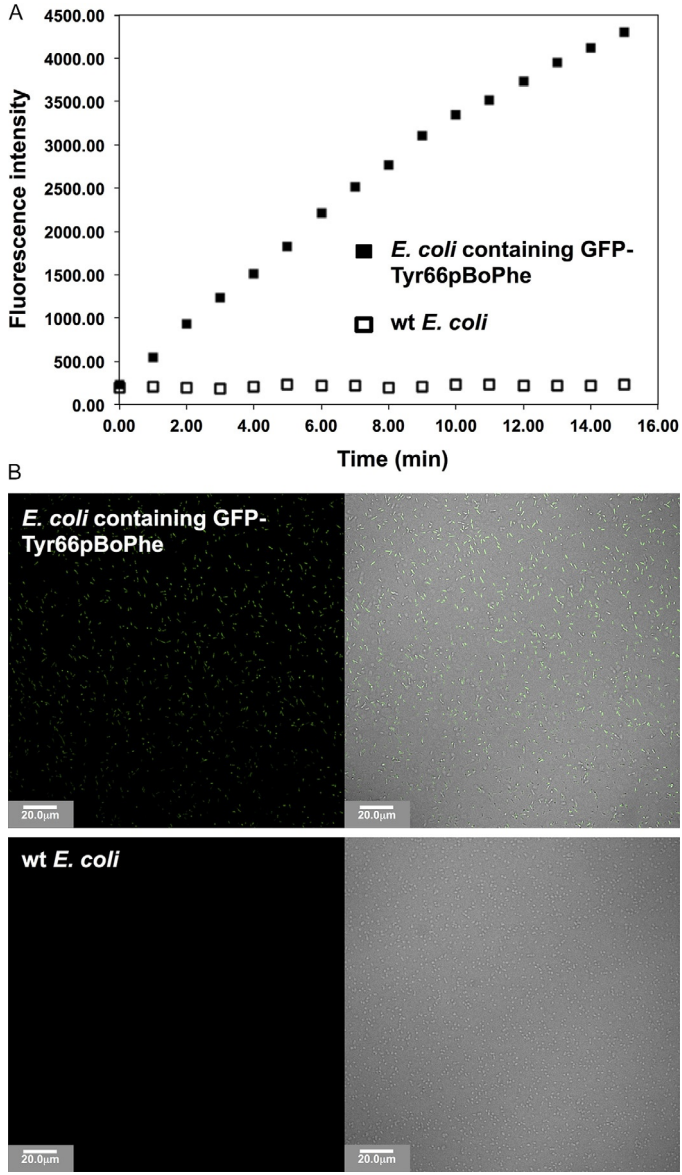

Fig. 8 Sensing H_2O_2 in live cells. (A) Oxidation of wild-type *E. coli* and *E. coli* expressing GFP-Tyr66pBoPhe in the presence of 100 m*M* H_2O_2. (B) Overlapped bright-field and fluorescence images of wild-type *E. coli* and *E. coli* expressing GFP-Tyr66pBoPhe in the presence of 100 m*M* H_2O_2.

In addition to fluorescence intensity measurement, cells can be visualized using a confocal fluorescence microscope. In this set of experiments, a 50 μL of washed cells (with an OD_{600} of one's choice) expressing GFP-Tyr66BoF are resuspended in PBS buffer and added to a glass slide, which is visualized under a confocal fluorescence microscope. As shown in Fig. 8B, no fluorescent cells can be detected. The same batch of cell sample is then incubated with 100 m*M* H_2O_2 at room temperature for 5 min and visualized again. Strong fluorescent signals can be detected. The fluorescence is nicely localized inside cells. Because GFP can be genetically encoded, this approach may allow spatial resolution of redox chemistry when fused to a target protein of interest.

4. CONCLUDING REMARKS

In summary, we report the construction of a H_2O_2 biosensor that combines the desirable properties of the arylboronate-based small-molecule fluorescent probes and fluorescent protein-based biosensors. This new class of H_2O_2 biosensor design aims to reach a maximum signal output by directly modifying the chemical composition of the chromophore. A chemical transformation of the unnatural amino acid (pBoPhe)-containing chromophore leads to dramatic change in fluorescence intensity. The properties of this H_2O_2 biosensor, such as the detection limit and the sensing kinetics, still need to be further optimized in order to be applied to real biological studies. Nevertheless, the construction of this H_2O_2 sensor and other unnatural amino acid-containing fluorescent protein biosensors represent an important step toward the generation of new sensors with high sensitivity and selectivity. Introduction of additional functional groups into fluorescent proteins allows one to explore new designs that are not possible by using the 20 common amino acids. The integration of the two powerful methodologies, fluorescent protein and unnatural amino acid mutagenesis, should lead to the design and construction of novel tools for biological studies in living systems. The design principles of the reaction-based fluorescent protein biosensor for H_2O_2 can be generalized to the construction of other fluorescent protein biosensors. We expect that many well-established aqueous phase chemical reactions can be used for the design of fluorescent protein biosensors and thus may expand the repertoire of analytes that can be detected.

ACKNOWLEDGMENTS

This work was supported by the New Faculty Startup Funds (to J.G. and W.N.) from the University of Nebraska–Lincoln and by an NSF Grant MCB1553041 (to J.G.) and an NIH Grant1R01AI111862 (to J.G. and W.N.). The authors thank Dr. Ronald Cerny in the Nebraska Center for Mass Spectrometry at University of Nebraska–Lincoln for help in protein mass spectrometry. The authors thank Dr. You Zhou in the Microscopy facility at the Center for Biotechnology at University of Nebraska–Lincoln for help in fluorescence microscope analysis.

REFERENCES

Alfonta, L., Zhang, Z., Uryu, S., Loo, J. A., & Schultz, P. G. (2003). Site-specific incorporation of a redox-active amino acid into proteins. *Journal of the American Chemical Society*, *125*, 14662–14663.

Amiram, M., Haimovich, A. D., Fan, C., Wang, Y.-S., Aerni, H.-R., Ntai, I., et al. (2015). Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nature Biotechnology*, *33*, 1272–1279.

Anderson, J. C., Wu, N., Santoro, S. W., Lakshman, V., King, D. S., & Schultz, P. G. (2004). An expanded genetic code with a functional quadruplet codon. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 7566–7571.

Ayyadurai, N., Prabhu, N. S., Deepankumar, K., Lee, S.-G., Jeong, H.-H., Lee, C.-S., et al. (2011). Development of a selective, sensitive, and reversible biosensor by the genetic incorporation of a metal-binding site into green fluorescent protein. *Angewandte Chemie (International ed. in English)*, *50*, 6534–6537.

Baird, G. S., Zacharias, D. A., & Tsien, R. Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 11241–11246.

Barondeau, D. P., Putnam, C. D., Kassmann, C. J., Tainer, J. A., & Getzoff, E. D. (2003). Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 12111–12116.

Belousov, V. V., Fradkov, A. F., Lukyanov, K. A., Staroverov, D. B., Shakhbazov, K. S., Terskikh, A. V., et al. (2006). Genetically encoded fluorescent indicator for intracellular hydrogen peroxide. *Nature Methods*, *3*, 281–286.

Bilan, D. S., Pase, L., Joosen, L., Gorokhovatsky, A. Y., Ermakova, Y. G., Gadella, T. W. J., et al. (2012). HyPer-3: A genetically encoded H2O2 probe with improved performance for ratiometric and fluorescence lifetime imaging. *ACS Chemical Biology*, *8*, 535–542.

Bizzarri, R., Serresi, M., Luin, S., & Beltram, F. (2009). Green fluorescent protein based pH indicators for in vivo use: A review. *Analytical and Bioanalytical Chemistry*, *393*, 1107–1122.

Brustad, E., Bushey, M. L., Brock, A., Chittuluru, J., & Schultz, P. G. (2008). A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins. *Bioorganic and Medicinal Chemistry Letters*, *18*, 6004–6006.

Brustad, E. M., Lemke, E. A., Schultz, P. G., & Deniz, A. A. (2008). A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *Journal of the American Chemical Society*, *130*, 17664–17665.

Budisa, N., Minks, C., Alefelder, S., Wenger, W., Dong, F., Moroder, L., et al. (1999). Toward the experimental codon reassignment in vivo: Protein building with an expanded amino acid repertoire. *FASEB Journal*, *13*, 41–51.

Cabantous, S., Pedelacq, J.-D., Mark, B. L., Naranjo, C., Terwilliger, T. C., & Waldo, G. S. (2005). Recent advances in GFP folding reporter and split-GFP solubility reporter

technologies. Application to improving the folding and solubility of recalcitrant proteins from mycobacterium tuberculosis. *Journal of Structural and Functional Genomics*, *6*, 113–119.

Cellitti, S. E., Jones, D. H., Lagpacan, L., Hao, X., Zhang, Q., Hu, H., et al. (2008). In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *Journal of the American Chemical Society*, *130*, 9268–9281.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science*, *263*, 802–805.

Chatterjee, A., Lajoie, M. J., Xiao, H., Church, G. M., & Schultz, P. G. (2014). A bacterial strain with a unique quadruplet codon specifying non-native amino acids. *Chembiochem: A European Journal of Chemical Biology*, *15*, 1782–1786.

Chen, S., Chen, Z.-J., Ren, W., & Ai, H.-W. (2012). Reaction-based genetically encoded fluorescent hydrogen sulfide sensors. *Journal of the American Chemical Society*, *134*, 9589–9592.

Chen, Z.-J., Ren, W., Wright, Q. E., & Ai, H.-W. (2013). Genetically encoded fluorescent probe for the selective detection of peroxynitrite. *Journal of the American Chemical Society*, *135*, 14940–14943.

Chen, X.-H., Xiang, Z., Hu, Y. S., Lacey, V. K., Cang, H., & Wang, L. (2014). Genetically encoding an electrophilic amino acid for protein stapling and covalent binding to native receptors. *ACS Chemical Biology*, *9*, 1956–1961.

Chi, K. R. (2009). Super-resolution microscopy: Breaking the limits. *Nature Methods*, *6*, 15–18.

Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., & Schultz, P. G. (2002). Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. *Journal of the American Chemical Society*, *124*, 9026–9027.

Chudakov, D. M., Matz, M. V., Lukyanov, S., & Lukyanov, K. A. (2010). Fluorescent proteins and their applications in imaging living cells and tissues. *Physiological Reviews*, *90*, 1103–1163.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., & Tsien, R. Y. (1995). Understanding, improving and using green fluorescent proteins. *Trends in Biochemical Sciences*, *20*, 448–455.

Datta, D., Wang, P., Carrico, I. S., Mayo, S. L., & Tirrell, D. A. (2002). A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins. *Journal of the American Chemical Society*, *124*, 5652–5653.

Deiters, A., Geierstanger, B. H., & Schultz, P. G. (2005). Site-specific in vivo labeling of proteins for NMR studies. *Chembiochem: A European Journal of Chemical Biology*, *6*, 55–58.

Denu, J. M., & Tanner, K. G. (1998). Specific and reversible inactivation of protein tyrosine phosphatases by hydrogen peroxide: Evidence for a sulfenic acid intermediate and implications for redox regulation. *Biochemistry*, *37*, 5633–5642.

Döring, V., Mootz, H. D., Nangle, L. A., Hendrickson, T. L., de Crécy-Lagard, V., Schimmel, P., et al. (2001). Enlarging the amino acid set of Escherichia coli by infiltration of the valine coding pathway. *Science*, *292*, 501–504.

Dougherty, D. A. (2000). Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology*, *4*, 645–652.

Frommer, W. B., Davidson, M. W., & Campbell, R. E. (2009). Genetically encoded biosensors based on engineered fluorescent proteins. *Chemical Society Reviews*, *38*, 2833–2841.

Groff, D., Thielges, M. C., Cellitti, S., Schultz, P. G., & Romesberg, F. E. (2009). Efforts toward the direct experimental characterization of enzyme microenvironments: Tyrosine100 in dihydrofolate reductase. *Angewandte Chemie (International ed. in English)*, *48*, 3478–3481. S3478/3471-S3478/3475.

Groff, D., Wang, F., Jockusch, S., Turro, N. J., & Schultz, P. G. (2010). A new strategy to photoactivate green fluorescent protein. *Angewandte Chemie (International ed. in English)*, *49*, 7677–7679.

Guo, J., Melancon, C. E., III, Lee, H. S., Groff, D., & Schultz, P. G. (2009). Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing non-natural amino acids. *Angewandte Chemie (International ed. in English)*, *48*, 9148–9151. S9148/9141-S9148/9123.

Guo, J., Wang, J., Lee, J. S., & Schultz, P. G. (2008). Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angewandte Chemie (International ed. in English)*, *47*, 6399–6401.

Gutscher, M., Sobotta, M. C., Wabnitz, G. H., Ballikaya, S., Meyer, A. J., Samstag, Y., et al. (2009). Proximity-based protein thiol oxidation by H_2O_2-scavenging peroxidases. *Journal of Biological Chemistry*, *284*, 31532–31540.

Hanson, G. T., Aggeler, R., Oglesbee, D., Cannon, M., Capaldi, R. A., Tsien, R. Y., et al. (2004). Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *Journal of Biological Chemistry*, *279*, 13044–13053.

Hohsaka, T., & Sisido, M. (2002). Incorporation of non-natural amino acids into proteins. *Current Opinion in Chemical Biology*, *6*, 809–815.

Huttner, W. B. (1988). Tyrosine sulfation and the secretory pathway. *Annual Review of Physiology*, *50*, 363–376.

Jackson, J. C., Hammill, J. T., & Mehl, R. A. (2007). Site-specific incorporation of a 19F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *Journal of the American Chemical Society*, *129*, 1160–1166.

Johnson, D. B. F., Xu, J., Shen, Z., Takimoto, J. K., Schultz, M. D., Schmitz, R. J., et al. (2011). RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nature Chemical Biology*, *7*, 779–786.

Ju, T., Niu, W., Cerny, R., Bollman, J., Roy, A., & Guo, J. (2013). Molecular recognition of sulfotyrosine and phosphotyrosine by the Src homology 2 domain. *Molecular BioSystems*, *9*, 1829–1832.

Ju, T., Niu, W., & Guo, J. (2016). Evolution of Src homology 2 (SH2) domain to recognize sulfotyrosine. *ACS Chemical Biology*, *11*, 2551–2557.

Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., et al. (2013). Genomically recoded organisms expand biological functions. *Science*, *342*, 357–360.

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

Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D. J., Blackman, M. L., et al. (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. *Journal of the American Chemical Society*, *134*, 10317–10320.

Lee, S.-R., Kwon, K.-S., Kim, S.-R., & Rhee, S. G. (1998). Reversible inactivation of protein-tyrosine phosphatase 1B in A431 cells stimulated with epidermal growth factor. *Journal of Biological Chemistry*, *273*, 15366–15372.

Lee, H. S., Spraggon, G., Schultz, P. G., & Wang, F. (2009). Genetic incorporation of a metal-ion chelating amino acid into proteins as a biophysical probe. *Journal of the American Chemical Society*, *131*, 2481–2483.

Lippert, A. R., Van de Bittner, G. C., & Chang, C. J. (2011). Boronate oxidation as a bioorthogonal reaction approach for studying the chemistry of hydrogen peroxide in living systems. *Accounts of Chemical Research*, *44*, 793–804.

Liu, C. C., Brustad, E., Liu, W., & Schultz, P. G. (2007). Crystal structure of a biosynthetic sulfo-hirudin complexed to thrombin. *Journal of the American Chemical Society*, *129*, 10648–10649.

Liu, X., Li, J., Hu, C., Zhou, Q., Zhang, W., Hu, M., et al. (2013). Significant expansion of the fluorescent protein chromophore through the genetic incorporation of a metal-chelating unnatural amino acid. *Angewandte Chemie (International ed. in English)*, *52*, 4805–4809.

Liu, C. C., & Schultz, P. G. (2006). Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nature Biotechnology*, *24*, 1436–1440.

Liu, C. C., & Schultz, P. G. (2010). Adding new chemistries to the genetic code. *Annual Review of Biochemistry*, *79*, 413–444.

Liu, X., Yu, Y., Hu, C., Zhang, W., Lu, Y., & Wang, J. (2012). Significant increase of oxidase activity through the genetic incorporation of a tyrosine-histidine cross-link in a myoglobin model of heme-copper oxidase. *Angewandte Chemie (International ed. in English)*, *51*, 4312–4316.

Mank, M., & Griesbeck, O. (2008). Genetically encoded calcium indicators. *Chemical Reviews*, *108*, 1550–1564.

Markvicheva, K. N., Bilan, D. S., Mishina, N. M., Gorokhovatsky, A. Y., Vinokurov, L. M., Lukyanov, S., et al. (2011). A genetically encoded sensor for H2O2 with expanded dynamic range. *Bioorganic and Medicinal Chemistry*, *19*, 1079–1084.

Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., et al. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology*, *17*, 969–973.

Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., et al. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature*, *388*, 882–887.

Miyawaki, A., Shcherbakova, D. M., & Verkhusha, V. V. (2012). Red fluorescent proteins: Chromophore formation and cellular applications. *Current Opinion in Structural Biology*, *22*, 679–688.

Nagai, T., Sawano, A., Park, E. S., & Miyawaki, A. (2001). Circularly permuted green fluorescent proteins engineered to sense Ca^{2+}. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 3197–3202.

Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., & Chin, J. W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature*, *464*, 441–444.

Newman, R. H., Fosbrink, M. D., & Zhang, J. (2011). Genetically encodable fluorescent biosensors for tracking signaling dynamics in living cells. *Chemical Reviews*, *111*, 3614–3666.

Niu, W., & Guo, J. (2013). Expanding the chemistry of fluorescent protein biosensors through genetic incorporation of unnatural amino acids. *Molecular BioSystems*, *9*, 2961–2970.

Niu, W., Schultz, P. G., & Guo, J. (2013). An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS Chemical Biology*, *8*, 1640–1645.

Ostergaard, H., Tachibana, C., & Winther, J. R. (2004). Monitoring disulfide bond formation in the eukaryotic cytosol. *Journal of Cell Biology*, *166*, 337–345.

Paulsen, C. E., & Carroll, K. S. (2010). Orchestrating redox signaling networks through regulatory cysteine switches. *ACS Chemical Biology*, *5*, 47–62.

Poole, L. B., & Nelson, K. J. (2008). Discovering mechanisms of signaling-mediated cysteine oxidation. *Current Opinion in Chemical Biology*, *12*, 18–24.

Pouwels, L. J., Zhang, L., Chan, N. H., Dorrestein, P. C., & Wachter, R. M. (2008). Kinetic isotope effect studies on the de novo rate of chromophore formation in fast- and slow-maturing GFP variants. *Biochemistry*, *47*, 10111–10122.

Rao, B. D., Kemple, M. D., & Prendergast, F. G. (1980). Proton nuclear magnetic resonance and fluorescence spectroscopic studies of segmental mobility in aequorin and a green fluorescent protein from aequorea forskalea. *Biophysical Journal*, *32*, 630–632.

Rhee, S. G. (2006). Cell signaling. H2O2, a necessary evil for cell signaling. *Science*, *312*, 1882–1883.

Rodriguez, E. A., Lester, H. A., & Dougherty, D. A. (2006). In vivo incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 8650–8655.

Schultz, K. C., Supekova, L., Ryu, Y., Xie, J., Perera, R., & Schultz, P. G. (2006). A genetically encoded infrared probe. *Journal of the American Chemical Society*, *128*, 13984–13985.

Seifert, M. H. J., Georgescu, J., Ksiazek, D., Smialowski, P., Rehm, T., Steipe, B., et al. (2003). Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution. *Biochemistry*, *42*, 2500–2512.

Seyedsayamdost, M. R., Xie, J., Chan, C. T. Y., Schultz, P. G., & Stubbe, J. (2007). Site-specific insertion of 3-aminotyrosine into subunit α2 of E. Coli ribonucleotide reductase: Direct evidence for involvement of Y730 and Y731 in radical propagation. *Journal of the American Chemical Society*, *129*, 15060–15071.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology*, *22*, 1567–1572.

Shang, X., Song, X., Faller, C., Lai, R., Li, H., Cerny, R., et al. (2017). Fluorogenic protein labeling using a genetically encoded unstrained alkene. *Chemical Science*, *8*, 1141–1145. http://dx.doi.org/10.1039/C1036SC03635J.

Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., & Remington, S. J. (2006). Novel chromophores and buried charges control color in mFruits. *Biochemistry*, *45*, 9639–9647.

Song, W., Wang, Y., Qu, J., Madden, M. M., & Lin, Q. (2008). A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. *Angewandte Chemie (International ed. in English)*, *47*, 2832–2835.

Stone, M. J., Chuang, S., Hou, X., Shoham, M., & Zhu, J. Z. (2009). Tyrosine sulfation: An increasingly recognized post-translational modification of secreted proteins. *New Biotechnology*, *25*, 299–317.

Stone, J. R., & Yang, S. (2006). Hydrogen peroxide: A signaling messenger. *Antioxidants & Redox Signaling*, *8*, 243–270.

Taki, M., Matsushita, J., & Sisido, M. (2006). Expanding the genetic code in a mammalian cell line by the introduction of four-base codon/anticodon pairs. *Chembiochem: A European Journal of Chemical Biology*, *7*, 425–428.

Tippmann, E. M., & Schultz, P. G. (2007). A genetically encoded metallocene containing amino acid. *Tetrahedron*, *63*, 6182–6184.

Tsien, R. Y. (1998). The green fluorescent protein. *Annual Review of Biochemistry*, *67*, 509–544.

Veal, E. A., Day, A. M., & Morgan, B. A. (2007). Hydrogen peroxide sensing and signaling. *Molecular Cell*, *26*, 1–14.

Verkhusha, V. V., & Lukyanov, K. A. (2004). The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nature Biotechnology*, *22*, 289–296.

Wang, N., Ju, T., Niu, W., & Guo, J. (2015). Fine-tuning interaction between aminoacyl-tRNA synthetase and tRNA for efficient synthesis of proteins containing unnatural amino acids. *ACS Synthetic Biology*, *4*, 207–212.

Wang, H., Nakata, E., & Hamachi, I. (2009). Recent progress in strategies for the creation of protein-based fluorescent biosensors. *Chembiochem: A European Journal of Chemical Biology*, *10*, 2560–2577.

Wang, F., Niu, W., Guo, J., & Schultz, P. G. (2012). Unnatural amino acid mutagenesis of fluorescent proteins. *Angewandte Chemie (International ed. in English)*, *51*, 10132–10135.
Wang, K., Sachdeva, A., Cox, D. J., Wilf, N. M., Lang, K., Wallace, S., et al. (2014). Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labeling and FRET. *Nature Chemistry*, *6*, 393–403.
Wang, N., Shang, X., Cerny, R., Niu, W., & Guo, J. (2016). Systematic evolution and study of UAGN decoding tRNAs in a genomically recoded bacteria. *Scientific Reports*, *6*, 21898.
Wang, L., Xie, J., Deniz, A. A., & Schultz, P. G. (2003). Unnatural amino acid mutagenesis of green fluorescent protein. *Journal of Organic Chemistry*, *68*, 174–176.
Wang, J., Xie, J., & Schultz, P. G. (2006). A genetically encoded fluorescent amino acid. *Journal of the American Chemical Society*, *128*, 8738–8739.
Wang, L., Zhang, Z., Brock, A., & Schultz, P. G. (2003). Addition of the keto functional group to the genetic code of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 56–61.
Xiang, Z., Lacey, V. K., Ren, H., Xu, J., Burban, D. J., Jennings, P. A., et al. (2014). Proximity-enabled protein crosslinking through genetically encoding haloalkane unnatural amino acids. *Angewandte Chemie (International ed. in English)*, *53*, 2190–2193.
Xie, J., Liu, W., & Schultz, P. G. (2007). A genetically encoded bidentate, metal-binding amino acid. *Angewandte Chemie (International ed. in English)*, *46*, 9239–9242.
Xie, J., Supekova, L., & Schultz, P. G. (2007). A genetically encoded metabolically stable analog of phosphotyrosine in Escherichia coli. *ACS Chemical Biology*, *2*, 474–478.
Ye, S., Koehrer, C., Huber, T., Kazmi, M., Sachdev, P., Yan, E. C. Y., et al. (2008). Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. *Journal of Biological Chemistry*, *283*, 1525–1533.
Zhang, L., Patel, H. N., Lappe, J. W., & Wachter, R. M. (2006). Reaction progress of chromophore biogenesis in green fluorescent protein. *Journal of the American Chemical Society*, *128*, 4766–4772.
Zhang, Z., Wang, L., Brock, A., & Schultz, P. G. (2002). The selective incorporation of alkenes into proteins in Escherichia coli. *Angewandte Chemie (International ed. in English)*, *41*, 2840–2842.
Zimmer, M. (2002). Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. *Chemical Reviews*, *102*, 759–781.