

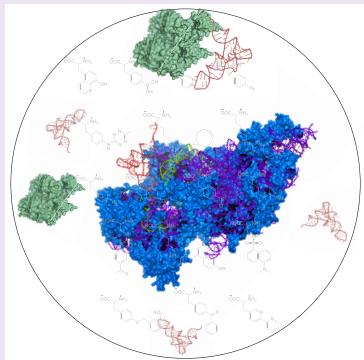
Playing with the Molecules of Life

Douglas D. Young^{*,†,‡} and Peter G. Schultz^{*,†,‡}

[†]Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037, United States

[‡]Department of Chemistry, College of William & Mary, P.O. Box 8795, Williamsburg, Virginia 23187, United States

ABSTRACT: Our understanding of the complex molecular processes of living organisms at the molecular level is growing exponentially. This knowledge, together with a powerful arsenal of tools for manipulating the structures of macromolecules, is allowing chemists to harness and reprogram the cellular machinery in ways previously unimagined. Here we review one example in which the genetic code itself has been expanded with new building blocks that allow us to probe and manipulate the structures and functions of proteins with unprecedented precision.



The genetic code is conserved across virtually every species—a total of 61 triplet codons provide the genetic information to encode the 20 canonical amino acids. It is remarkable that the limited set of functional groups contained in these building blocks enable proteins to carry out most of the complex processes of living organisms. The ability to add noncanonical amino acids (ncAAs) with novel structures and properties to the code might therefore allow the biosynthesis of proteins with new or enhanced functions and provides the opportunity to probe the structure and functions of these macromolecules with unparalleled chemical precision.

This review overviews the recombinant technology developed for the site-specific incorporation of ncAAs into proteins, as well as recent advances that expand its scope. We then highlight illustrative examples of the use of ncAAs to control, evolve, and understand protein function.

■ EXPANDING THE GENETIC CODE

Generating Bio-orthogonal Translational Machinery. Several key components are required to genetically encode noncanonical amino acids with high efficiency and fidelity in a host organism (Figure 1A).^{1–4} The first component is a reassigned codon that encodes the ncAA. Frequently, the amber nonsense codon (TAG) is selected due to its less frequent utilization as one of the three stop codons; however, other nonsense or four base frameshift codons have also been used for this purpose.^{5–10} Another requirement is an orthogonal aminoacyl-tRNA synthase (aaRS)/tRNA pair capable of recognizing the desired ncAA and incorporating it in response to the appropriate codon.^{11–16} The orthogonality of this pair must be such that the aaRS does not recognize endogenous tRNAs (84 in *Escherichia coli*) or amino acids and only aminoacylates its cognate tRNA; additionally, the tRNA must not act as a substrate for any endogenous aaRSs.^{3,17} The final key component is the ncAA itself, which can often be

directly supplemented to the growth media, taken up as a dipeptide precursor through a transporter,¹⁸ or biosynthesized by the host.¹⁹ An elegant example of the synthesis of all these components was the genetic encoding of *p*-aminophenylalanine in engineered bacteria that contained the orthogonal aaRS/tRNA pair, an amber codon inserted into the myoglobin sequence, and a gene cluster from *S. venezuelae* that enables the bacterial biosynthesis of *p*-aminophenylalanine to afford a 21 amino acid synthetic organism.¹⁹

Aminoacyl-tRNA Synthase Evolution. Recognition of a desired ncAA by the aaRS is often engineered using structure-based approaches in combination with powerful double-sieve selection schemes (Figure 2).^{3,21–27} This process typically begins with analysis of the crystal structure of the aaRS active site to identify those residues that contribute to specificity, followed by random mutagenesis of these sites to generate large aaRS libraries. The library is then subjected to positive selection in the presence of the desired ncAA (typically linked to suppression of an amber codon at a permissive site in a protein conferring antibiotic resistance) to ensure that only aaRSs that recognize and aminoacylate the desired ncAA proffer survival.²³ The survivors are then subjected to a negative selection in the absence of the ncAA (e.g., suppression of amber codons at permissive sites in a lethal gene product) to ensure survival in the positive selection step is not a result of aminoacylation with any natural host amino acid.²³ This process is repeated for several rounds to afford an aaRS that is highly selective for the ncAA over any endogenous amino acids.²⁸ A number of other selection and screening systems have also been developed, but all typically require both positive and negative screening or selection steps.^{29,30} One recent exciting example involves the

Received: November 13, 2017

Accepted: January 18, 2018

Published: January 18, 2018

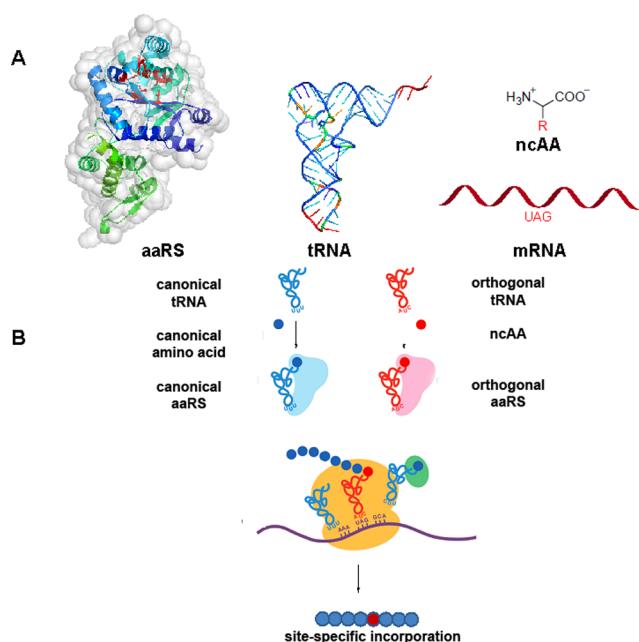


Figure 1. Expansion of the genetic code. (A) Requisite components for the site-specific incorporation of ncAAs into proteins. Red coloring indicates primary sites of mutation required to enable the efficient incorporation of the desired ncAA. (B) Site-specific incorporation of ncAAs through engineering of the translational machinery. Standard translation occurs with endogenous tRNA/aaRS pairs and canonical amino acids (blue) on the ribosome (yellow). An orthogonal tRNA/aaRS pair and ncAA are added (red) and encoded by a nonsense or frameshift codon in the mRNA (purple). Adapted with permission from ref 20. Copyright 2013 Elsevier.

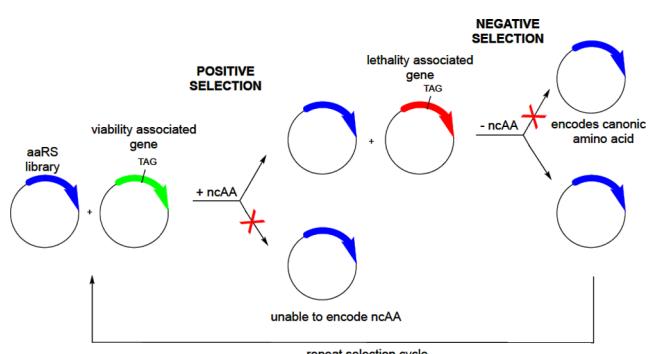


Figure 2. Standard protocol for generation of an aaRS to encode ncAAs.

use of rapid phage-assisted continuous evolution (PACE) methodology to evolve highly efficient and selective aaRSs.³¹ The X-ray crystal structures of a number of these evolved aaRSs reveal remarkable plasticity in the amino acid binding site such that a limited number of mutations can substantially alter the active site structure and specificity.^{32,33} More recently, it has been discovered that as a consequence of other ncAAs being absent from the negative selection, some aaRSs exhibit a degree of polyspecificity toward structurally related ncAAs.^{34–41} One such aaRS, evolved to encode *p*-cyanophenylalanine in *E. coli*, is capable of incorporating over 20 aromatic amino acid derivatives.^{34,37} Another such pair is the PylRS/tRNA_{CUA} pair from certain methanogens (notably *Methanosarcina barkeri* (*Mb*) and *Methanosarcina mazei* (*Mm*)).^{35,42,43}

Using the above methods, a large number (>150) of ncAAs with distinct structures and functions have been genetically encoded in prokaryotes and eukaryotes.^{44–46} These include metal chelating^{47–49} and fluorescent amino acids,^{50–53} a variety of photo-cross-linkers,^{54–58} amino acids with altered pK_a's and redox properties,^{36,59–62} amino acids with bioorthogonal chemical reactivity,^{12,59,63–71} and post-translationally modified amino acids and their stable analogues.^{18,26,72–77} Whereas many of the reported ncAAs are structural variants of canonical amino acid side chains, it has also been demonstrated that the amino acid backbone can be modified as well. These modifications include the incorporation of α -hydroxy acids,^{78–80} N-methyl amino acids,⁷⁸ and α - α -disubstituted amino acids, albeit the latter two amino acids involved an *in vitro* translation method.⁷⁸ Recently, β -amino acids were genetically encoded; specifically, β -phenylalanine analogs were introduced into a protein using mutant ribosomes and an orthogonal aaRS/tRNA pair.⁸¹

Ensuring Orthogonality. The selection of an appropriate aaRS/tRNA pair to encode an ncAA is dictated by the host organism requirements to achieve orthogonality. Early experiments relied on an engineered *Methanococcus jannaschii* (*Mj*) TyrRS/tRNA_{CUA} pair that is orthogonal in *E. coli* and other bacteria largely due to distinct acceptor stem recognition sequences.^{16,23,82} Since these initial studies, TyrRS/tRNA, LeuRS/tRNA, TrpRS/tRNA, SerRS/tRNA, AspRS/tRNA, GluRS/tRNA, LysRS/tRNA, and ProRS/tRNA pairs that are orthogonal in either prokaryotic and/or eukaryotic cells have been generated.^{8,9,15,22,46,83–88} PylRS/tRNA_{CUA} pairs from certain methanogens (notably *Methanosarcina barkeri* (*Mb*) and *Methanosarcina mazei* (*Mm*)) that are orthogonal in both bacteria and eukaryotic cells have been developed.^{24,42,43} These latter pairs are especially advantageous as they allow aaRSs to be evolved in *E. coli* prior to transfer of the machinery to more diverse eukaryotic hosts. Utilizing these orthogonal pairs, ncAAs have been encoded in *B. cereus*,⁸⁹ *P. pastoris*,⁹⁰ *C. elegans*,^{91,92} *D. melanogaster*,⁹³ *A. thaliana*,⁹⁴ Zebrafish embryos,⁹⁵ and the mouse.^{96–99} More recently, it has been demonstrated that a native Trp aaRS/tRNA pair in *E. coli* can be functionally replaced with a counterpart from yeast, and the liberated Trp pair can be used to encode ncAAs in bacteria.⁸⁶ Additionally, orthogonal aaRS/tRNA technologies have been used to incorporate ncAAs into proteins in mammalian cell lines on the grams per liter scale employing transient expression methods.^{100–102} Viral vectors have allowed the ncAA machinery to be delivered efficiently into primary cells, as well as tissues,^{96,103,104} where it was used among other applications to monitor voltage-sensitive changes in response to membrane depolarization events in neural cells.¹⁰⁰

Recent Advances. A variety of strategies have been reported to further improve the efficiency and specificity of ncAA incorporation into proteins, including mutations to the aaRS, tRNA, ribosomal peptidyl transferase, and elongation factor.^{13,17,104–110} Moreover, aaRS and tRNA expression levels have been modulated in order to facilitate high-level expressions of proteins containing ncAAs.^{13,104,105,111–113} These alterations have led to ncAA-incorporation on multigram per liter levels in large scale bacterial fermentation, and gram per liter scale in stable CHO cell lines, as demonstrated in the production of ncAA containing pegylated proteins and antibody-drug conjugates (ADCs).¹¹¹

An exciting recent advance is the ability to incorporate more than one ncAA into a protein sequence with the ultimate challenging goal of the mRNA template-directed biosynthesis

of monodisperse biopolymers made up of synthetic building blocks. Toward this end, several *E. coli* strains have been generated that either conditionally or constitutively remove release factors (RF1 in *E. coli* and eRF1 in eukaryotes) that terminate polypeptide synthesis in response to specific nonsense codons, in order to improve suppression efficiencies.^{75,114–116} Orthogonal bacterial ribosomes that are directed to an orthogonal message, by the incorporation of a mutant 16S rRNA into their small subunit (and therefore not essential to the cell), have also been created (Figure 3).^{117,118} One such

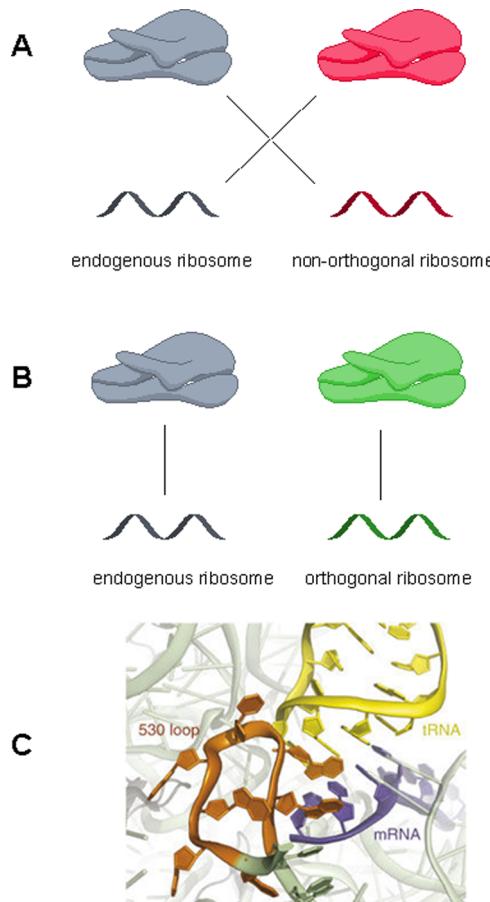


Figure 3. Generation of an orthogonal ribosome. (A) A non-orthogonal ribosome allows for cross talk between the two mRNAs, not providing efficient incorporation of ncAAs. (B) An orthogonal ribosome where the endogenous system (gray) and the engineered ribosome and mRNA (green) exhibit no cross-reactivity. (C) Crystal structure of the rRNA (orange), mRNA (purple), and tRNA (yellow), illustrating the key 530 loop within the ribosome that was subjected to mutagenesis to afford an orthogonal ribosome.¹¹⁹

orthogonal ribosome that no longer recognizes RF1 was discovered by directed evolution and enables the efficient incorporation of an ncAA in response to amber codons at multiple sites in a single polypeptide.¹¹⁹ Another approach involves recoding the genome such that some or all of the amber codons have been replaced by the ochre nonsense codon TAA in an effort to remove potential read-through of endogenous termination signals.^{120–122} These strains, which have TAG or TAGN (N = A, G, C, T) uniquely assigned to the ncAA, have been shown to enhance ncAA incorporation in response to the quadruplet codon TAGA, which is derived

from and competes with RF1 recognition of the amber codon (TAG).⁵

There is also interest in the incorporation of multiple distinct amino acids into a single protein, which requires aaRS/tRNA pairs that are mutually orthogonal and orthogonal to the host aaRS/tRNA pairs.⁹ Recently, a new expression cassette was engineered for bacterial expression that affords two aaRS/tRNA pairs (*M. jannaschii* and *M. barkeri*) that meet this requirement (recognizing TAG and TAA codons, respectively).⁶ This methodology allowed for the incorporation of two reactive ncAAs harboring either a ketone or azide, respectively, and also the generation of a fluorescence resonance energy transfer (FRET) pair within the same reporter protein.^{17,123} The synthesis of unnatural biopolymers with three or more ncAAs requires one new orthogonal aaRS/tRNA pair per ncAA and has catalyzed efforts to discover new orthogonal aaRS/tRNA pairs and repurpose additional codons.^{6,124–131} To this end, it has been shown that existing orthogonal pairs can be used as starting points to generate additional mutually orthogonal pairs by directed evolution.^{87,132} These new tRNA/aaRS pairs have allowed multiple, different amino acids to be incorporated in both *E. coli*^{123,126,133} and eukaryotic systems.^{134,135} These experiments suggest that the number of orthogonal pairs that can be discovered from natural sequence diversity does not place a limit on the number of distinct building blocks that may be simultaneously genetically encoded in cells. An active area of research now focuses on reassigning degenerate codons in the genetic code to encode additional ncAAs.^{124,136}

■ APPLICATIONS OF NON-CANONICAL AMINO ACIDS

While the technology development associated with ncAA incorporation has stimulated efforts to create other bioorthogonal cellular systems, the application of these methodologies illustrates the true utility of this technology.^{3,127,129,137–139} To date, ncAAs have been employed in a large number of studies of protein structure and function, and also to alter or enhance protein function. In addition, the recombinant introduction of ncAAs into proteins is allowing the development of new therapeutics and diagnostics.

Probing Protein Structure and Function. The ability to expand the genetic code beyond the 20 common amino acids facilitates the site-specific introduction of noncanonical amino acids with altered structures and functions to better understand protein function both *in vitro* and in living cells with minimal perturbation to protein structure. These ncAAs include residues with altered pK_a 's for mechanistic studies, isotopic labels for infrared and NMR studies, photo-cross-linkers for mapping biomolecular interactions in living cells, heavy atoms for X-ray crystallography, and spin labels and fluorescent side chains for EPR and optical applications, respectively. While ncAA probes have been used in numerous studies, below we highlight instructive examples of their use.

Altering pK_a and Redox Potential. Electron-withdrawing or donating substituents allow one to alter the acidity, basicity, and redox potential of canonical amino acids (Figure 4).^{36,61,62,140–146} For example, fluorinated tyrosine analogues served as effective EPR probes to monitor long-lived tyrosyl radicals in the complex mechanism of ribonucleotide reductase and better understand the role of conserved tyrosine residues in the prevention of undesirable radical chemistry.^{36,147} These studies complemented previous semisynthetic studies employing nitrotyrosine¹⁴⁰ and aminotyrosine,¹⁴³ which were used to

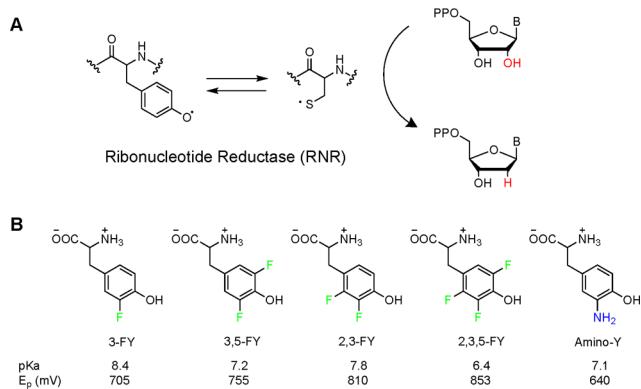


Figure 4. Modulation of pK_a and redox potential of tyrosine residues. (A) The ribonucleotide reductase reaction converting ribose to deoxyribose relies upon a catalytic cysteine radical. The generation of this radical is dependent on radical formation on several key tyrosine residues. Altering the pK_a 's and reduction potentials (E_p) of these residues affords key insights into the catalytic mechanism. (B) Examples of ncAAs conferring altered tyrosine pK_a 's and reduction potentials (E_p) that have been employed in the study of ribonucleotide reductase.

investigate the kinetics of radical intermediate formation within these ribonucleotide reductases.

Protein Cross-linking to Map Protein Interactions. An arena where ncAAs have found widespread use is in the field of protein cross-linking, both *in vitro* and inside living cells. Multiple cross-linking ncAAs have been genetically encoded including aryl azides,⁵⁹ aryl haloketones,^{148–150} benzophenones,⁵⁴ aryl carbamates,¹⁵¹ isothiocyanates,¹⁵² and diazirines^{56,58,153} to covalently cross-link interacting proteins and protein–nucleic acid complexes. This approach has been extensively used to map transient protein interactions that elude detection by other methods such as immunoprecipitation or are unstable outside the context of a living cell. Depending on the chemistry employed, cross-linking can occur via either direct chemical reaction with residues in close proximity or by photoactivation. One example using a benzophenone containing ncAA (which inserts relatively nonselectively into C–H bonds) was the identification of key protein–protein interactions necessary for lipopolysaccharide (LPS) transport to bacterial outer membranes (Figure 5).¹⁵⁴ This same ncAA was also employed to afford spatiotemporal control over the mapping of histone 2A interactions in yeast, elucidating key proteins involved in the histone modification cascade that induces mitosis.¹⁵⁵ The utility of this cross-linking ncAA has been further expanded via modification with an alkynyl handle for subsequent reaction, facilitating rapid isolation of cross-linked products.¹⁵⁶ More recently, new ncAA photo-cross-linkers have been developed that can be chemically cleaved following cross-linking, further simplifying the identification of a target protein by mass spectrometry.¹⁵³ In addition, chemical cross-linkers have also been generated that rely on the nascent chemical reactivity of lysine and cysteine residues with nearby electrophilic ncAAs to afford covalent cross-links between proteins.^{148,151,157} Other examples of the use of photo-cross-linking ncAAs include the identification of G-protein coupled receptor activating ligands,^{158,159} acid chaperone associated proteins in pathogens,¹⁶⁰ the binding site of antidepressant drugs in the serotonin transporter,¹⁶¹ and the interacting partners of short open reading-frame encoded peptides.¹⁶² An important new direction is the encoding of cross-linking agents in whole organisms to explore cell–cell interactions.

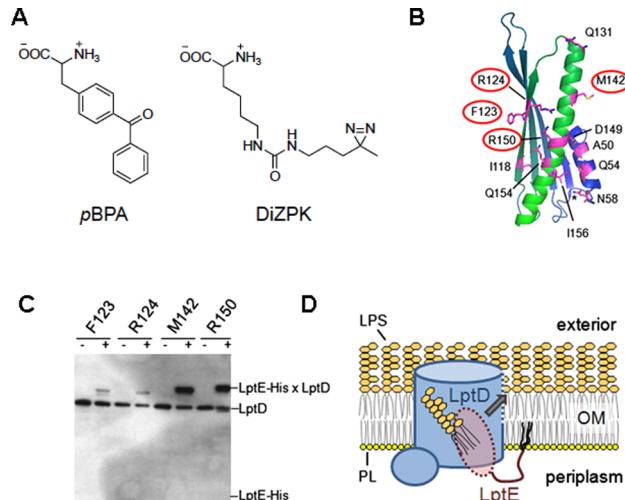


Figure 5. Applications of photo-cross-linking ncAAs. (A) Structures of two common photo-cross-linking ncAAs, *p*-benzoylphenylalanine and (3-(3-methyl-3H-diazirine-3-yl)-propaminocarbonyl- N^{ϵ} -lysine. (B) Structure of the lipopolysaccharide transport protein E (LptE) with the sites of ncAA incorporation circled in red. (C) Example of a photo-cross-linking gel employing the benzophenone ncAA at multiple residues of the LptE protein associated with lipopolysaccharide transport. Gel shifts observed in the irradiated (+) samples indicate a cross-linking event with LptD. (D) Proposed model of the LptD/LptE association for lipopolysaccharide transport established by the cross-linking experiments. Adapted with permission from ref 154. Copyright 2011 United States National Academy of Sciences.

Genetically Encoded Fluorescent Amino Acids. Genetically encoded fluorescent ncAAs have been used both *in vitro* and in living cells to label proteins with fluorescent probes. This approach has advantages over strategies that rely upon the reactivity of canonical amino acids such as Cys and Lys, as well as genetic fusions, due to the high level of control over the labeling site within a protein and minimal perturbation to protein structure and function. Several fluorescent ncAAs, including coumarin,^{50,163} dansyl,⁵³ naphthyl,⁵¹ terphenyl,¹⁶⁴ and prodan derivatives^{52,165} have been genetically encoded in both prokaryotes and eukaryotes (Figure 6). The fluorescence of some of these ncAAs is environmentally sensitive and can be used to probe cellular localization, biomolecular interactions, post-translational modifications, and residue exposure to solvent.⁵² For example, fluorescent ncAAs were used to detect the binding of glutamine to a glutamine binding protein¹⁶⁶ and the phosphorylation of STAT3.¹⁶⁷ The prodan derivative, ANAP, has also been used as a partner to generate a FRET pair within mammalian cells to investigate protein proteolysis and protein conformational changes.^{168,169} While the majority of genetically encoded fluorophores are excited at shorter wavelengths and provide a convenient way to site-specifically label proteins for *in vitro* studies, some directly encoded fluorophores have also been used for imaging proteins in live cells.^{52,163} An important direction for future research is the genetic encoding of longer wavelength, photostable fluorescent ncAAs.

Infrared Probes. The incorporation of deuterium,¹⁷⁰ cyano,¹⁷¹ nitro,¹⁷² and azido⁵⁹ IR active probes into proteins introduces unique vibrational modes that do not overlap those present in the canonical amino acids. These localized probes have facilitated investigations of enzymatic catalysis,¹⁷⁰ solvent exposure,^{173,174} and receptor activation using IR spectroscopy.

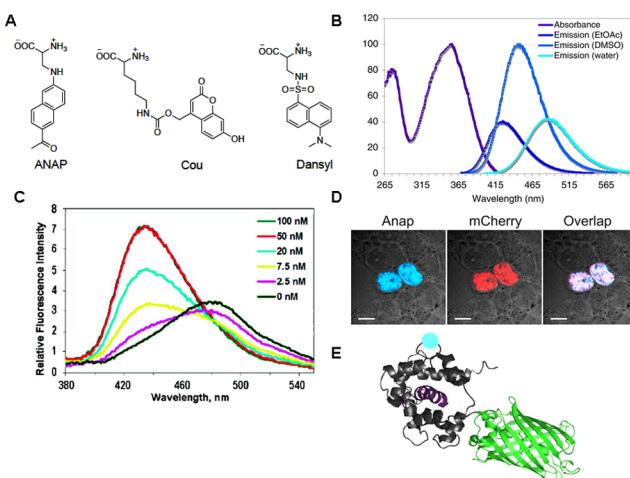


Figure 6. Genetically encoding fluorescent amino acids. (A) Structures of fluorescent ncAAs based on common fluorophores. (B) Solvent sensitivity of ANAP fluorescence.⁵² (C) Use of ANAP in a glutamine binding protein as a site-specific probe to detect glutamine binding by fluorescence.¹⁶⁶ (D) Demonstration of the use of an ANAP to track protein localization; the ncAA was incorporated into histones resulting in fluorescence only in nuclei.⁵² (E) A FRET pair prepared via site-specific incorporation of ANAP (blue circle) in a fusion construct with GFP. The presence of ANAP obviates the need for a second fluorescent fusion protein.¹⁶⁸

py.^{175,176} In one example, azidophenylalanine was introduced into rhodopsin at distinct sites within the transmembrane helices allowing for helix movements to be monitored when the receptor is activated by light.¹⁷⁵ In a second example, deuterium probes were introduced *via* a photocaged tyrosine derivative (which upon photodeprotection installed 2,3,4,5-D₄ tyrosine) at specific sites in the active site of dihydrofolate reductase. IR spectroscopy allowed one to monitor conformational changes in the presence of different ligands that correspond to different states along the catalytic pathway.¹⁷⁰ These studies have provided useful information regarding protein structure and function and all relied upon the novel functionality of ncAAs.

Spin-Label and NMR Probes. The introduction of spin labels into protein affords site-specific EPR probes and is traditionally accomplished via reaction of a chemically reactive spin label with sulfhydryl groups. This approach has limitations, as there can be multiple cysteine residues in a protein, and they often play a role in protein folding. Early examples involved the incorporation of *p*-acetylphenylalanine, a ketone containing ncAA with bioorthogonal chemical reactivity, which can be conjugated to a nitroxide group via an oxime ligation.¹⁷⁷ Recently, a spin label has been directly incorporated into proteins, facilitating in-cell EPR studies of endogenous proteins.^{178,179}

Two approaches have been developed to use ncAAs as site-specific NMR probes. The first involves genetically encoding a ¹⁵N-labeled tyrosine that is caged with a nitrobenzyl moiety, which upon uncaging by exposure to UV light results in the site-specific incorporation of ¹⁵N-tyrosine.¹⁸⁰ A second approach is based on the site specific incorporation of close analogues of canonical amino acids possessing isotopic labels, for example, O-methyl-phenylalanine containing ¹³C and/or ¹⁵N or a trifluoro derivative of the same amino acid possessing ¹⁹F.^{180–185} Both strategies were used to map the binding site of a fatty acid synthase with a model ligand, as well as active site conformational changes that occur upon ligand binding.¹⁸⁰

Fluorinated amino acid analogues, due to their increased signal sensitivity, have also been used to monitor protein conformational changes in living cells, for example, to monitor phosphorylation induced conformational shifts in arrestin-1 during cell signaling.¹⁸³

Environmental Sensors. Several ncAAs are capable of either specifically coordinating or directly reacting with analytes to act as sensors. One successful approach involved introducing ncAAs into green fluorescent protein (GFP) at sites that altered either the wavelength or intensity of protein fluorescence in the presence of an external stimulus (Figure 7).¹⁸⁶ Specifically, GFP

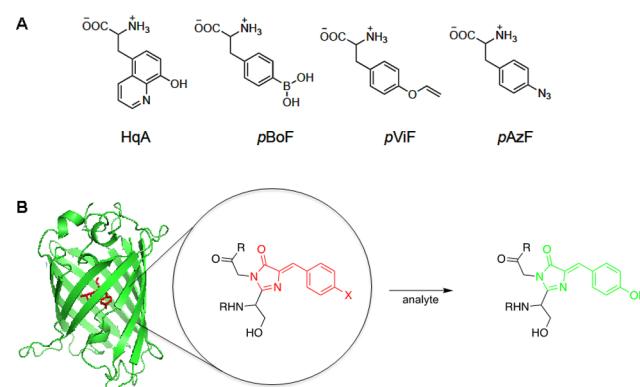


Figure 7. Using ncAAs as environmental sensors for biologically relevant analytes. (A) Structures of ncAAs commonly employed in environmental sensors, including metal binders, H₂O₂, H₂S, and ONOO[−] sensitive functionalities. (B) General method for the development of an “on” sensor. Incorporation of the ncAA within the GFP chromophore alters or quenches fluorescence; however, upon coordination or reaction with a desired analyte, the group is removed or fluorescence is shifted.

containing either L-3,4-dihydroxyphenylalanine (DOPA)¹⁸⁷ or 8-hydroxyquinoline alanine (HqA)¹⁸⁸ at residue 66 within the fluorophore was used to detect biologically relevant transition metals including Cu²⁺, Zn²⁺, Co²⁺, Fe²⁺, and Ni²⁺. More recently, a *p*-vinylphenylalanine was incorporated at this same site to act as a turn-on probe for the detection of mercuric ions in bacteria.¹⁸⁹ In addition to ions, this approach has been employed for the detection of H₂O₂ with a boronate ncAA,¹⁹⁰ H₂S with a azidophenylalanine ncAA,¹⁹¹ and sirtuins using an acetylated lysine which can be cleaved by these enzymes.¹⁹² Finally, the *p*-boronophenylalanine has also been exploited to detect peroxy nitrite in mammalian cells in order to monitor the production of this cell-signaling molecule at physiologically relevant concentrations.^{193,194} In addition ncAAs have been incorporated at residue 66 in GFP to alter the fluorescence properties of the protein,^{37,190,195–198} and into the coelenterazine binding site of aequorin to dramatically red-shift the bioluminescence of the interaction to afford *in vivo* bioluminescence reporters in mice.¹⁹⁷

Post-translational Modification. Protein post-translational modifications are ubiquitous in biology and control cellular processes ranging from signal transduction and transcription to cell division and protein degradation. It is relatively difficult to generate site-specific PTMs in proteins—the enzyme responsible for the modifications may be unknown or may not be site-specific, or the modifications may be removed in cells or lysates. Consequently, the ability to genetically encode post-translational modifications amino acids, or stable analogues thereof, provides a useful tool to better understand and characterize

individual PTMs in proteins. A number of post-translational modifications have been genetically encoded, including phosphoserine, phosphothreonine, phosphotyrosine, sulfotyrosine, nitrotyrosine, and numerous epigenetic lysine modifications.^{18,26,60,71,74–77,199–203} Stable mimicks of post-translational modifications, such as phosphotyrosine and carboxymethylphenylalanine, have also been genetically encoded.^{72,75,204,206} Incorporation of this latter phosphotyrosine mimetic into the arginine methyltransferase PRMT1 facilitated analysis of how phosphorylation alters the binding of this protein to its substrates.²⁰⁵ Phosphotyrosine was recently genetically incorporated into proteins by exploiting the polyspecificity of the previously evolved carboxymethyl-phenylalanine aaRS and a dipeptidyl transporter to transport this negatively charged amino acid into cells.²⁰⁶ Subsequent incorporation of phosphotyrosine into human Abl1 at specific residues allowed determination of the binding affinities of the various phosphorylated forms of the protein to its substrates. Additionally, 3-nitrotyrosine has been genetically encoded and used to gain new insights into post-translational modifications that modulate inflammatory responses and mimic disease-state processes.⁶⁰

Given the polyspecificity of the pyrrolysyl-aaRS toward a variety of lysine modifications, a number of laboratories have genetically encoded post-translationally modified lysines, including ϵ -N-2-hydroxyisobutyryl-lysine, ϵ -N-pivaloyl-lysine, ϵ -N-methyl-lysine, and ϵ -N-crotonyl-lysine, among others, as a means to study these post-translational modifications in a variety of proteins. For example, these ncAAs have also been exploited to characterize the effects of ubiquitination and histone methylation on gene regulation.^{71,207} Finally, *E. coli* has been engineered for the recombinant production of ribosomally synthesized post-translationally modified peptides (RIPPs) that contain ncAAs, and analogues of the lanthipeptide nisin with altered ring structures and sizes have been produced.^{208–212}

Site-Specific Protein Conjugation. A general strategy for site-specifically modifying proteins with synthetic moieties (e.g., biophysical probes, drugs, PEGs, oligonucleotides, etc.) involves installing a chemically (or enzymatically) reactive amino acid at the desired site and selectively conjugating the side chain to a reaction partner linked to the molecule of interest.²¹³ Cysteine and lysine residues are the most commonly targeted canonical amino acids and typically modified with electrophiles. However, there are often multiple such residues in a protein, resulting in heterogeneous labeling, and Cys residues may also be required for protein folding and function. The site-specific incorporation of ncAAs with unique chemical reactivity allows the selective modification of proteins with medicinal chemistry-like precision. This approach requires ncAAs with bioorthogonal chemical reactivity, i.e., unreactive with the canonical amino acids or metabolites in the cell. The most common bioorthogonal ncAAs that have been encoded have keto,²¹⁴ azido,⁵⁹ and acetylenic side chains (Figure 8).^{42,63,215,216} Azide and alkynyl ncAAs undergo highly selective copper-catalyzed 1,3-dipolar cycloaddition reactions and have been generally used for *in vitro* applications.^{217,218} However, strained alkynes have recently been genetically encoded to afford copper-free “click” reactions that can be used in living cells.^{219–221} Target proteins are efficiently and selectively labeled, and background nonspecific proteome labeling is minimal or undetectable, despite a significant number of endogenous genes that terminate in amber codons.²²² The use of Cys and an ncAA, or two ncAAs, with bio-orthogonal

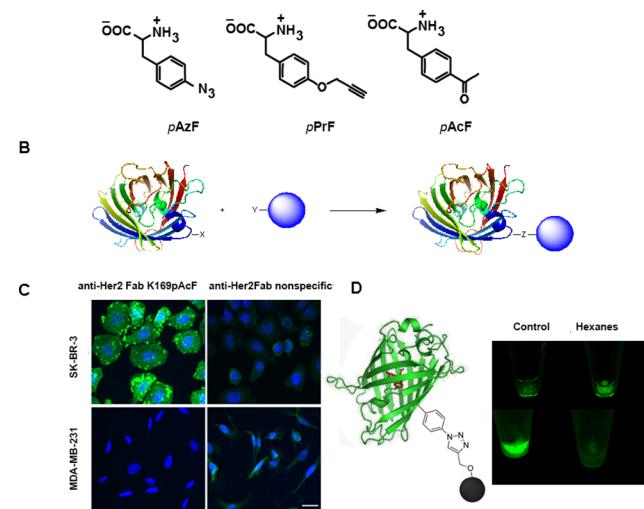


Figure 8. Bioorthogonal conjugation with ncAAs. (A) Common structures of ncAAs used in bioconjugations: *p*-azidophenylalanine (*p*AzF), *p*-propargyloxyphenylalanine (*p*PrF), and *p*-acetylphenylalanine (*p*AcF). (B) General scheme of a bioorthogonal conjugation. The protein harboring an ncAA with unique reactivity is subjected to appropriate reaction conditions with another molecule (protein, DNA/RNA, surface, small molecule, etc; blue sphere) that possesses a chemical functionality that will only react with the ncAA and no other biological molecules. (C) Incorporation of *p*-acetylphenylalanine into a Fab for Her2 followed by oxime ligation with DNA. This bioconjugate facilitates immuno-PCR with significantly higher levels of sensitivity than nonspecifically labeled antibodies.²²⁹ (D) By site-specifically incorporating *p*-azidophenylalanine, GFP can be site-specifically immobilized on a solid support, conferring a higher degree of protein stability in nonaqueous solvents.²⁴⁶

reactivity allows the site-specific labeling of proteins with two different moieties via sequential or one-pot bioorthogonal reactions for applications such as Förster resonance energy transfer (FRET) measurements^{17,133,136} with excellent dynamic range.²²³

Another commonly employed ncAA for bioorthogonal conjugations is *p*-acetylphenylalanine, which can be selectively modified at lower pH with highly efficient oxime formation reactions.²¹⁴ This reaction has been used to generate site-specific drug, polyethylene glycol (on kilogram commercial scale), peptide, and oligonucleotide-protein conjugates, and also to conjugate a large number of probes to proteins.^{20,111,224–232} The high selectivity of this reaction results in well-defined conjugates that do not significantly affect the intrinsic properties of a protein, in contrast to more nonspecific approaches. This advantage was demonstrated with the generation of a DNA-antibody conjugate for immuno-PCR: an anti-Her2 antibody/DNA conjugate was more sensitive with lower background when used to detect Her2 positive cells in complex blood mixtures relative to nonspecific lysine conjugates (Figure 8).²²⁹

Considerable effort has recently been devoted to expand the toolbox of genetically encoded ncAAs with bioorthogonal chemical reactivity. Alkenyl ncAAs have been used in bioorthogonal methathesis reactions to conjugate carbohydrates to proteins.^{233,234} hydroxytryptophan residues have been employed for azo-couplings,²³⁵ cyclopropanone and azide residues for reactions with phosphines;^{236,237} alkynyl amino acids have been exploited in Sonogashira,^{238,239} Glaser-Hay,²⁴⁰ and Cadiot-Chodkiewitz bioconjugation reactions;²⁴¹ and *p*-

boronophenylalanine and *p*-iodophenylalanine have been used in Suzuki and other transition metal-based couplings.^{242–245}

An important recent advance has been the encoding of new bioorthogonal reactive side chains that undergo rapid and selective labeling reactions in the cell. The importance of reaction rate has been increasingly appreciated, and reactions with rate constants that approach those of enzymatic mediated labeling reactions have been developed.²¹⁸ One such reaction is the inverse electron demand Diels–Alder reaction between strained alkenes or alkynes and tetrazines (Figure 9). The rate

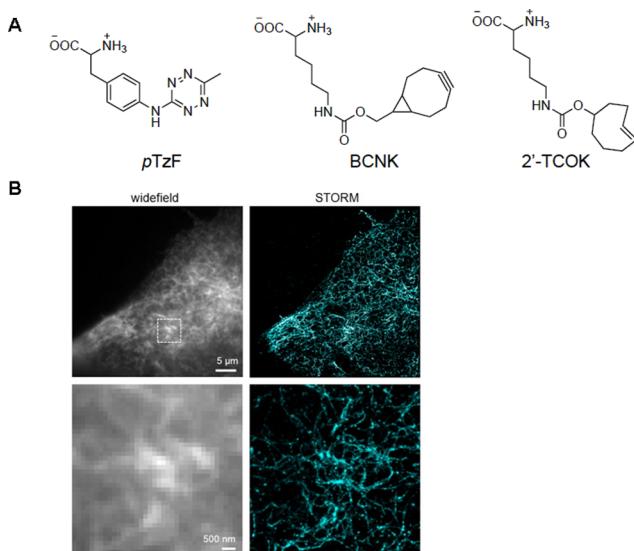


Figure 9. Live cell imaging using ncAAs. (A) Common structures of genetically encoded ncAAs for noncytotoxic and rapid live cell imaging. Using either standard 1,3-cycloadditions with strained alkynes or Diels–Alder type reactions with a tetrazine and a strained alkyne or alkene, reactions can be performed within living cells. (B) Genetic incorporation of the strained alkyne into vimentin for super-resolution, live cell imaging of proteins.²²²

constants for some tetrazine reactions with alkenes and alkynes can exceed $10^4 \text{ M}^{-1} \text{ s}^{-1}$, and several alkenes, alkynes, and tetrazines have now been recombinantly introduced into proteins and used for rapid, site-specific, live cell protein labeling.^{65,68,247–250} Live cell labeling and super resolution imaging have been further facilitated by the development of bright small molecule fluorophores that enter mammalian cells and have minimal background fluorescence.^{222,251} While the refinement and optimization of labeling and imaging approaches is still ongoing, early examples of applications that are not possible by other methods have already emerged.^{252–203}

Another application of ncAAs with bioorthogonal handles is their use in protein immobilization to a surface for a wide range of industrial and diagnostic uses. Immobilization typically leads to increased protein stability, tolerance to organic solvents, and recyclability of enzymes (Figure 8).²⁵⁴ ncAAs provide a means to both covalently link proteins to a surface as well as control the site of immobilization to prevent protein inactivation or heterogeneous immobilization. To date, ncAAs have been used for the immobilization of proteins to nanotubes, magnetic beads, and solid-supported resins and have been shown to confer increased stability to the immobilized protein.^{246,255–258} Further exploration and optimization of these immobilization

technologies has the potential to make significant advances in both materials chemistry and therapeutic diagnostics.

Control of Protein Function. The ability to control protein activity chemically or with light is useful for spatially and temporally activating biological processes in living cells. One such strategy involves the blocking of essential amino acid side chains with a removable protecting group (Figure 10). This moiety blocks normal side chain function until removed by an external stimulus. Many analogs of natural amino acids (especially lysine, serine, cysteine, and tyrosine) containing a photoremoveable protecting group have been genetically encoded in both prokaryotic and eukaryotic cells and organisms.^{259–262} These photocaged ncAAs have been used to control kinase, protease, intein, and other enzyme activities, nuclear localization, virus–host interactions, and cell signaling cascades.^{102,163,203,263–268}

One example of the use of this technology is the photoactivation of Cas9 using a photocaged lysine residue.²⁶⁹ The photocaged CRISPR/Cas9 system was used to both silence and activate exogenous reporter genes in mammalian cell culture, as well silence endogenous CD71, a receptor linked to leukemia and lymphoma. The regulation of this system in living cells has many potential applications in the rapidly developing field of gene editing. In another example, recombinant incorporation of a photocaged serine into the transcription factor Pho4 in yeast blocked phosphorylation and subsequent nuclear export until removal of the caging group with 400 nm light, allowing the kinetics of these processes to be followed.²⁷⁰ A final example of light-regulated *in vivo* protein function involved photocaging a lysine residue within an isocitrate dehydrogenase mutant known to produce the oncometabolite 2-hydroxyglutarate (2-HG).²⁷¹ Light activation of this oncogenic protein provided insights into the role of 2-HG in oncogenic activation.

The reversible photoregulation of protein activity has been accomplished through the genetic encoding of a photo-switchable azobenzene ncAA. This ncAA has been used to photoregulate enzyme activity by reversibly blocking the active site²⁶⁰ or by moving a tethered inhibitor in and out of the active site by *cis*–*trans* photoisomerization of the azobenzene moiety.²⁷²

While light is a useful external stimulus for protein activation, it has also been demonstrated that small molecule activators can be used in conjunction with ncAAs to modulate protein function in cells. A recent study incorporated a cyclooctene lysine derivative into a reporter luciferase protein in mammalian cells.²⁷³ Incubation with a tetrazine induced a Diels–Alder reaction that led to the removal of the cyclooctene cage, restoring wild-type lysine and consequently fluorescence. Palladium-catalyzed allene decaging²⁷⁴ and phosphine-mediated decaging of azides via a Staudinger reduction have also been employed to chemically modulate protein function *in vivo*.²⁷⁵

Therapeutic Applications. The ability to site-specifically modify proteins using bioorthogonal ncAAs greatly facilitates the generation of homogeneous therapeutic proteins whose structures are precisely controlled, much like the medicinal chemists' ability to precisely control the structures of small molecule therapeutics. Recent efforts have focused on introducing amino acids with bio-orthogonal reactivity into cytokines, growth factors, antibodies, and antibody domains, providing a route to their site-specific conjugation with diverse moieties.^{276,277} Pegylated proteins, antibody-drug conjugates

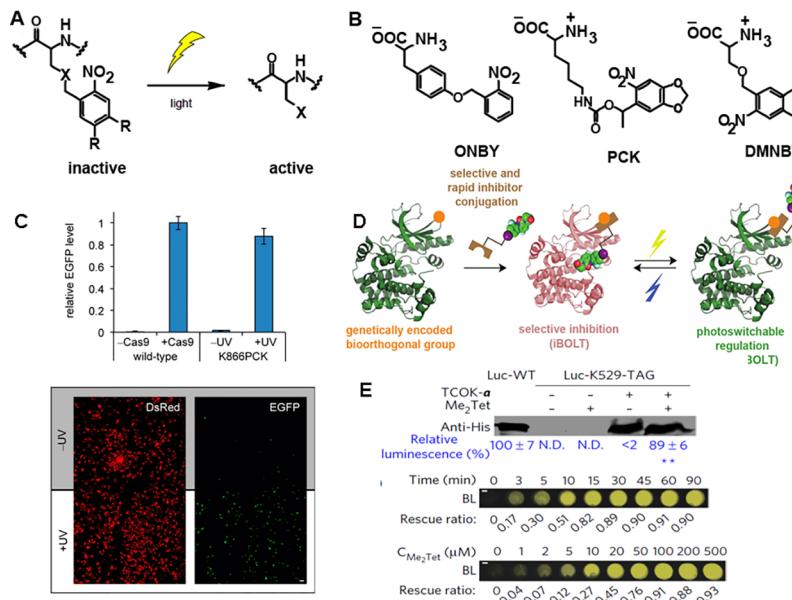


Figure 10. Photocontrol over protein function with ncAAs. (A) Photocaging strategy for activation of protein function with light. Typically, a key residue is substituted with a caged ncAA, rendering the protein of interest nonfunctional. Upon brief irradiation with noncytotoxic light, the caging group is removed to afford wild-type protein. X = O, N, S; R = H, O, OMe. (B) Photocaged tyrosine, lysine, and serine amino acids with variations of the common *o*-nitrobenzyl caging group. (C) Photocaging of the CRISPR/Cas9 system results in inactivity of the Cas9 protein, until brief irradiation with light restores its function close to WT levels, facilitating the expression of a reporter GFP plasmid.²⁶⁹ (D) Photoswitchable regulation of protein function *via* reaction of a specifically placed ncAA with a ligand harboring an azobenzene based linker. Light irradiation of different wavelengths results in *cis/trans* isomerization, blocking or exposing the active site.²⁷² (E) Genetically encoded decaging of a strained cycloalkene via reaction with a tetrazine, restoring a lysine residue and protein function. Proof-of-concept experiments were performed via caging luciferase, quenching luminescence until the tetrazine reagent is added.²⁷³

(ADCs), antibody-antisense oligonucleotide conjugates, and bispecific antibodies have been created for a variety of clinical indications,^{111,278,112,232,279–283} and a number of these new therapeutics have been approved or are showing positive results in clinical trials. Importantly, it has been shown that, by controlling both the stoichiometry and site of conjugation of a PEG or drug on a target protein, one can optimize both half-life and potency in a manner that is difficult to achieve with less specific chemistries. Moreover, it has been possible to conjugate small molecules that bind tumor antigens (e.g., DUPA and folate which selectively target PSMA and folate receptor, respectively^{281,284}) site-specifically to anti-CD3 antibodies to generate extremely potent semisynthetic bispecific antibodies, which in the case of the DUPA-anti-CD3 conjugate show impressive efficacy in preclinical models of metastatic prostate cancer (Figure 11).²⁸⁴ Recently, a modular strategy for chimeric antigen receptor (CAR)-T cell therapy was created in which a chimeric T cell receptor was created that binds FITC, and FITC was selectively conjugated to an aryl ketone containing ncAA selectively incorporated into an antibody specific for a tumor antigen. The resulting “switch” molecules bridge the CAR-T cell and the tumor cell, and their structures can again be precisely controlled to optimize formation of the immunological synapse. This approach enabled dose dependent tumor clearance²⁸⁵ with minimal cytokine release using a universal CAR-T cell that can be adapted to distinct tumor antigens by appropriate switch molecules.

Introducing ncAAs with immunogenic side chains, such as nitro- or sulfo-tyrosine, into proteins has been used to break immunological tolerance to the corresponding native protein. The immunogenic side chain forms a neoepitope that elicits a T cell response and leads to a cross-reactive antibody response to

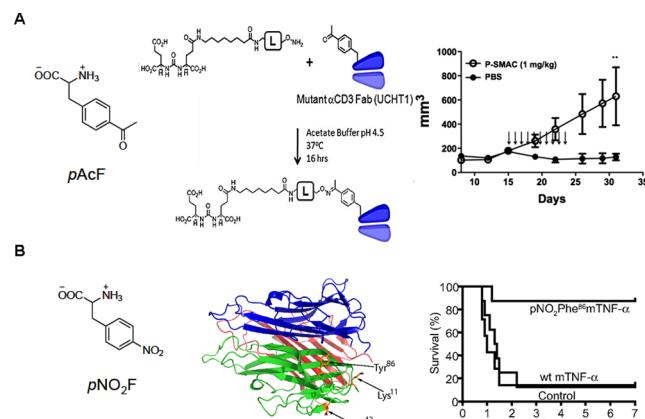


Figure 11. Therapeutic applications of ncAAs. (A) Incorporation of *p*-acetylphenylalanine into an anti-CD3 Fab followed by oxime ligation with an aminoxy-modified DUPA produces a bispecific agent for the treatment of prostate cancer. Treatment of mice in tumor xenografts resulted in complete tumor killing.²⁸⁸ (B) Example of using ncAAs to break immunological tolerance. The ncAA *p*-nitrophenylalanine was incorporated into mTNF- α at distinct sites and used to immunize mice, resulting in increased survival relative to both the WT protein and a control in an endotoxemia mouse model.²⁸⁷

the native proteins (Figure 10).²⁸⁶ Indeed, immunization of mice with a murine TNF that contains a single nitrotyrosine led to neutralizing antibodies to native TNF that were protective in an LPS mouse challenge model.²⁸⁷ These studies provide a route to breaking tolerance against specific native proteins and support the view that tyrosine nitration, which naturally results from viral infection and inflammation, could contribute to autoimmunity.

The introduction of amber codons into the genomes of viruses, including hepatitis D virus, HIV-1, and influenza A, has created viral genomes^{289–291} that can only be replicated in cells that contain orthogonal aaRS/tRNA_{CUA} pairs and their cognate ncAA. This strategy has enabled the creation of attenuated viruses that may be used for immunization. More recently, ncAAs have been used to create bacteria whose replication is strictly dependent on the presence of a specific ncAA at a functional site (e.g., catalytic, protein interface or metal ion binding site) in an essential protein. Such systems can be used for biological containment of engineered organisms. Immunization with these live conditional pathogens is expected to give robust immune responses, but the pathogen will die after the ncAA is depleted through rounds of replication in the host.²⁹⁰ Importantly, revertants for some sites of suppression have yet to be isolated (reversion rates less than 10e-11). Clearly, the use of ncAA technologies allows far more chemical control over the structures and, as a result, activities of therapeutic macromolecules.

Protein Design and Evolution with ncAAs. The addition of amino acids with novel properties to the genetic code presents new opportunities for the rational design of proteins with new or enhanced functions. Metal binding proteins are involved in a wide variety of cellular processes, but it has been challenging to design metal binding sites into proteins due, in part, to the need to precisely control the geometries of multiple amino acid side chains that chelate the metal.²⁹² The introduction of bipyridyl containing amino acids into proteins allows the design of metal ion binding sites that take advantage of the preorganization of the bidentate ligand for metal chelation. Computational design, using Rosetta, has enabled the generation and structural characterization of proteins and metallo-protein assemblies that bind metal ions with picomolar affinity.^{292,293} More recently, a protein was designed with Rosetta that binds the biphenyl side chain of an ncAA in a planar geometry which is stabilized by packing interactions in the hydrophobic core of a thermophilic protein (Figure 12).

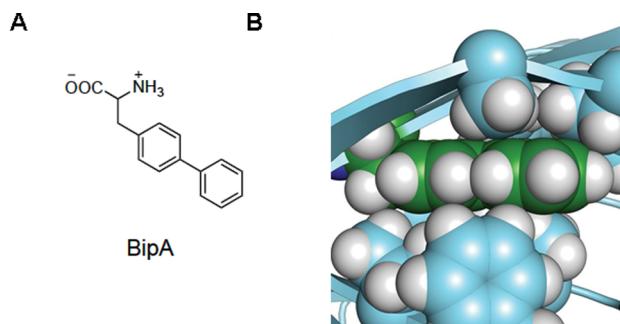


Figure 12. Computational design of ncAAs containing proteins. (A) Structure of *p*-phenylphenylalanine (BipA) used in the model study. (B) Space filling model of the interactions of the biphenyl (green) locked into its planar transition state with the designed pocket.²⁹⁴

Thus, the protein acts as a thermodynamic sink to make the transition state for biphenyl bond rotation kinetically persistent such that X-ray crystallography enabled direct observation of this transition state conformation.²⁹⁴

The ability to genetically encode noncanonical amino acids with chemically defined structures and properties has also allowed the *in vitro* evolution of proteins with novel or enhanced functions. For example, phage selections using genetically encoded sulfotyrosine and a germline antibody

library containing a randomized CDR3 loop led to the selection of high affinity sulfated antibodies that bind gp120, an HIV protein that naturally binds a sulfated receptor.²⁹⁵ Phage display was also used to evolve a zinc finger transcription factor that had an unusual high-spin-Fe(II) core containing a bipyridyl amino acid side chain,²⁹⁶ and which bound its operator sequence with high affinity and selectivity. A bacterial-based selection scheme has also been used to identify ncAA containing cyclic peptides that inhibit HIV protease in an antibiotic based selection (using a protease sensitive antibiotic antiporter).²⁹⁷ The highest affinity peptide to emerge incorporated a benzophenone containing ncAA that bound HIV protease through a Schiff base linkage to a surface Lys. More recently, beta-lactamase variants have been isolated using a growth-based selection from a library of mutants containing single amber codons at over half of the residues. One noncanonical amino acid substitution led to an enzyme with increased catalytic efficiency; X-ray crystallographic analysis suggested that the ncAA functioned by restricting the conformation of the active site to more efficiently stabilize the rate-limiting transition state.^{298,299} Similarly, a metA variant was isolated from a random ncAA library using a temperature dependent selection scheme that was stabilized by a keto-containing amino acid by a remarkable 23 °C and likely involves formation of a ketone adduct to a nucleophilic side chain at the homodimer interface. Similar *in vitro* evolution experiments have demonstrated that ncAAs with long chain thiols can form extended disulfide cross-links that lead to significant protein stabilization.³⁰⁰ Another example was the evolution of the T7 phage with an ncAA dependent growth advantage. The computational design and/or selection of variants of essential *E. coli* proteins that require a genetically encoded biphenyl or benzophenone amino acid for activity has provided an elegant strategy to make organism viability conditionally dependent on the presence of an ncAA.^{301–303} These experiments suggest that an expanded genetic code can indeed provide unique solutions to evolutionary challenges faced by living organisms, and that this is a rich area for further study.

CONCLUSIONS

Precise and highly tailored structural perturbations to proteins are now possible through the genetic encoding of noncanonical amino acids. The ability to incorporate ncAAs with diverse structures and properties into proteins in cells and living organisms has provided unique opportunities to probe, image, control, rationally engineer, and evolve protein structure and function, as well as to develop new therapeutics and approaches to biomaterials. The refinement of approaches for encoding and labeling amino acids containing bioorthogonal groups will be particularly important for coupling diverse functionalities at precise sites in proteins *in vitro* and *in vivo* as cell biological probes and precisely tailored therapeutic agents. *In vitro* evolution experiments will continue to reveal how life with additional genetically encoded amino acids may evolve new or enhanced protein functions. The development of additional strategies to create or repurpose codons, expand the substrate scope of translation, and create a suite of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs will further improve our ability to control protein structure at the building block level and potentially generate biopolymers in which all the building blocks are unnatural. Finally, these studies underscore the exciting opportunities that now exist to synthesize complex

molecular structures and functions through the rational manipulation of the cell's machinery using chemical and biological approaches synergistically.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: schultz@scripps.edu.

*E-mail: dyoung01@wm.edu.

ORCID

Douglas D. Young: 0000-0003-4433-4101

Peter G. Schultz: 0000-0003-3188-1202

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

P.G.S. would like to acknowledge the Department of Energy (DOE/DE-SC0011787) and the National Institute of Health (SR01 GM062159) for financial support. D.D.Y. would also like to acknowledge the National Institutes of Health (R15GM113203) and the Camille and Henry Dreyfus Foundation (TH-17-020) for financial support.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank H. Xiao and A. Chatterjee for their helpful comments. The authors would also like to thank K. Williams for her assistance with manuscript preparation and submission.

KEYWORDS

Noncanonical amino acid: any amino acid that is structurally different from the 20 endogenously encoded amino acids

Translation: process by which the genetic material in the form of RNA is converted into a protein context

Ribosome: cellular organelle that is the site of translation

Aminoacyl-tRNA Synthetase: protein responsible for recognizing an amino acid and cognate tRNA, resulting in the aminoacylation of the tRNA

Bioorthogonal: a process that occurs under physiological conditions but does not react with endogenous systems

Photocaging: the installation of a photoprotecting group often leading to the inactivation of a molecule until the group is removed via light irradiation

Biocoujugation: a process involving the linking of a biological molecule to another species (biomolecule, small molecule, surface, etc.)

Post-translational modification: alteration of natural amino acid structure after it has been incorporated into a protein

REFERENCES

- (1) Liu, C., Schultz, P., Kornberg, R., Raetz, C., Rothman, J., and Thorner, J. (2010) Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* 79, 413–444.
- (2) Young, T. S., and Schultz, P. G. (2010) Beyond the Canonical 20 Amino Acids: Expanding the Genetic Lexicon. *J. Biol. Chem.* 285, 11039–11044.
- (3) Wang, L., and Schultz, P. G. (2005) Expanding the genetic code. *Angew. Chem., Int. Ed.* 44, 34–66.

- (4) Chin, J. W. (2014) Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem.* 83, 379–408.
- (5) Chatterjee, A., Lajoie, M. J., Xiao, H., Church, G. M., and Schultz, P. G. (2014) A bacterial strain with a unique quadruplet codon specifying non-native amino acids. *ChemBioChem* 15, 1782–1786.
- (6) Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441–444.
- (7) Wang, N., Shang, X., Cerny, R., Niu, W., and Guo, J. (2016) Systematic Evolution and Study of UAGN Decoding tRNAs in a Genomically Recoded Bacteria. *Sci. Rep.* 6, 21898.
- (8) Anderson, J. C., and Schultz, P. G. (2003) Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry* 42, 9598–9608.
- (9) Anderson, J. C., Wu, N., Santoro, S. W., Lakshman, V., King, D. S., and Schultz, P. G. (2004) An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7566–7571.
- (10) Anderson, J. C., Magliery, T. J., and Schultz, P. G. (2002) Exploring the limits of codon and anticodon size. *Chem. Biol.* 9, 237–244.
- (11) Wang, F., Robbins, S., Guo, J., Shen, W., and Schultz, P. G. (2010) Genetic incorporation of unnatural amino acids into proteins in *Mycobacterium tuberculosis*. *PLoS One* 5, e9354.
- (12) Lang, K., Davis, L., and Chin, J. W. (2015) Genetic encoding of unnatural amino acids for labeling proteins. *Methods Mol. Biol.* 1266, 217–228.
- (13) Schmied, W. H., Elsässer, S. J., Uttamapinant, C., and Chin, J. W. (2014) Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J. Am. Chem. Soc.* 136, 15577–15583.
- (14) Sakamoto, K., Hayashi, A., and Sakamoto, A. (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res.* 30, 4692–4699.
- (15) Pastrnak, M., Magliery, T., and Schultz, P. (2000) A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code. *Helv. Chim. Acta* 83, 2277–2286.
- (16) Wang, L., Magliery, T., Liu, D., and Schultz, P. (2000) A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the *in vivo* incorporation of unnatural amino acids into proteins. *J. Am. Chem. Soc.* 122, 5010–5011.
- (17) Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in *Escherichia coli*. *Biochemistry* 52, 1828–1837.
- (18) Luo, X., Fu, G., Wang, R. E., et al. (2017) Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat. Chem. Biol.* 13, 845.
- (19) Mehl, R. A., Anderson, J. C., Santoro, S. W., et al. (2003) Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc.* 125, 935–939.
- (20) Kim, C. H., Axup, J. Y., and Schultz, P. G. (2013) Protein conjugation with genetically encoded unnatural amino acids. *Curr. Opin. Chem. Biol.* 17, 412–419.
- (21) Chin, J. W., Cropp, T. A., Chu, S., Meggers, E., and Schultz, P. G. (2003) Progress toward an expanded eukaryotic genetic code. *Chem. Biol.* 10, 511–519.
- (22) Chin, J. W., Cropp, T., Anderson, J., Mukherji, M., Zhang, Z., and Schultz, P. (2003) An Expanded Eukaryotic Genetic Code. *Science* 301, 964–967.
- (23) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*. *Science* 292, 498–500.
- (24) Chen, P. R., Groff, D., Guo, J., et al. (2009) A facile system for encoding unnatural amino acids in mammalian cells. *Angew. Chem., Int. Ed.* 48, 4052–4055.
- (25) Liu, D. R., Magliery, T. J., Pastrnak, M., and Schultz, P. G. (1997) Engineering a tRNA and aminoacyl-tRNA synthetase for the

- site-specific incorporation of unnatural amino acids into proteins *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10092–10097.
- (26) Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **4**, 232–234.
- (27) Melançon, C. E., and Schultz, P. G. (2009) One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases. *Bioorg. Med. Chem. Lett.* **19**, 3845–3847.
- (28) Xie, J., and Schultz, P. G. (2005) An expanding genetic code. *Methods* **36**, 227–238.
- (29) Santoro, S. W., Wang, L., Herberich, B., King, D. S., and Schultz, P. G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol.* **20**, 1044–1048.
- (30) Santoro, S. W., and Schultz, P. G. (2003) Directed evolution of the substrate specificities of a site-specific recombinase and an aminoacyl-tRNA synthetase using fluorescence-activated cell sorting (FACS). *Methods Mol. Biol.* **230**, 291–312.
- (31) 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.
- (32) Flügel, V., Vrabel, M., and Schneider, S. (2014) Structural basis for the site-specific incorporation of lysine derivatives into proteins. *PLoS One* **9**, e96198.
- (33) Turner, J. M., Graziano, J., Spraggon, G., and Schultz, P. G. (2006) Structural plasticity of an aminoacyl-tRNA synthetase active site. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6483–6488.
- (34) Young, D., Young, T., Jahnz, M., Ahmad, I., Spraggon, G., and Schultz, P. (2011) An Evolved Aminoacyl-tRNA Synthetase with Atypical Polysubstrate Specificity. *Biochemistry* **50**, 1894–1900.
- (35) Guo, L. T., Wang, Y. S., Nakamura, A., et al. (2014) Polyspecific pyrrolysyl-tRNA synthetases from directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16724–16729.
- (36) Minnihan, E. C., Young, D. D., Schultz, P. G., and Stubbe, J. (2011) Incorporation of fluorotyrosines into ribonucleotide reductase using an evolved, polyspecific aminoacyl-tRNA synthetase. *J. Am. Chem. Soc.* **133**, 15942–15945.
- (37) Young, D., Jockush, S., Turro, N., and Schultz, P. (2011) Synthetase polyspecificity as a tool to modulate protein function. *Bioorg. Med. Chem. Lett.* **21**, 7502–7504.
- (38) Brustad, E., Bushey, M. L., Brock, A., Chittuluru, J., and Schultz, P. G. (2008) A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins. *Bioorg. Med. Chem. Lett.* **18**, 6004–6006.
- (39) Miyake-Stoner, S. J., Refakis, C. A., Hammill, J. T., et al. (2010) Generating permissive site-specific unnatural aminoacyl-tRNA synthetases. *Biochemistry* **49**, 1667–1677.
- (40) Cooley, R. B., Karplus, P. A., and Mehl, R. A. (2014) Gleaning unexpected fruits from hard-won synthetases: probing principles of permissivity in non-canonical amino acid-tRNA synthetases. *ChemBioChem* **15**, 1810–1819.
- (41) Wang, Y. S., Fang, X., Wallace, A. L., Wu, B., and Liu, W. R. (2012) A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *J. Am. Chem. Soc.* **134**, 2950–2953.
- (42) Nguyen, D. P., Lusic, H., Neumann, H., Kapadnis, P. B., Deiters, A., and Chin, J. W. (2009) Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. *J. Am. Chem. Soc.* **131**, 8720–8721.
- (43) Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2008) Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem. Biophys. Res. Commun.* **371**, 818–822.
- (44) Liu, C. C., Mack, A. V., Tsao, M. L., et al. (2008) Protein evolution with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17688–17693.
- (45) Davis, L., and Chin, J. W. (2012) Designer proteins: applications of genetic code expansion in cell biology. *Nat. Rev. Mol. Cell Biol.* **13**, 168–182.
- (46) Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015) Designing logical codon reassignment - Expanding the chemistry in biology. *Chem. Sci.* **6**, 50–69.
- (47) Xie, J., Liu, W., and Schultz, P. (2007) A genetically encoded bidentate, metal-binding amino acid. *Angew. Chem., Int. Ed.* **46**, 9239–9242.
- (48) Lee, H. S., Spraggon, G., Schultz, P. G., and Wang, F. (2009) Genetic incorporation of a metal-ion chelating amino acid into proteins as a biophysical probe. *J. Am. Chem. Soc.* **131**, 2481–2483.
- (49) Tippmann, E., and Schultz, P. (2007) A genetically encoded metallocene containing amino acid. *Tetrahedron* **63**, 6182–6184.
- (50) Wang, J., Xie, J., and Schultz, P. (2006) A genetically encoded fluorescent amino acid. *J. Am. Chem. Soc.* **128**, 8738–8739.
- (51) Wang, L., Brock, A., and Schultz, P. G. (2002) Adding L-3-(2-Naphthyl)alanine to the genetic code of *E. coli*. *J. Am. Chem. Soc.* **124**, 1836–1837.
- (52) 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.
- (53) 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.
- (54) Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11020–11024.
- (55) Lin, S., He, D., Long, T., Zhang, S., Meng, R., and Chen, P. R. (2014) Genetically encoded cleavable protein photo-cross-linker. *J. Am. Chem. Soc.* **136**, 11860–11863.
- (56) Ai, H. W., Shen, W., Sagi, A., Chen, P. R., and Schultz, P. G. (2011) Probing protein-protein interactions with a genetically encoded photo-crosslinking amino acid. *ChemBioChem* **12**, 1854–1857.
- (57) Tippmann, E., Liu, W., Summerer, D., Mack, A., and Schultz, P. (2007) A genetically encoded diazirine photocrosslinker in *Escherichia coli*. *ChemBioChem* **8**, 2210–2214.
- (58) Chou, C., Upadhyay, R., Davis, L., Chin, J., and Deiters, A. (2011) Genetically encoding an aliphatic diazirine for protein photocrosslinking. *Chem. Sci.* **2**, 480–483.
- (59) Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **124**, 9026–9027.
- (60) Neumann, H., Hazen, J., Weinstein, J., Mehl, R., and Chin, J. (2008) Genetically encoding protein oxidative damage. *J. Am. Chem. Soc.* **130**, 4028–4033.
- (61) Seyed-Sayam-Dost, M., Xie, J., Chan, C., Schultz, P., and Stubbe, J. (2007) Site-specific insertion of 3-aminotyrosine into subunit alpha 2 of *E. coli* ribonucleotide reductase: Direct evidence for involvement of Y-730 and Y-731 in radical propagation. *J. Am. Chem. Soc.* **129**, 15060–15071.
- (62) Lv, X., Yu, Y., Zhou, M., et al. (2015) Ultrafast Photoinduced Electron Transfer in Green Fluorescent Protein Bearing a Genetically Encoded Electron Acceptor. *J. Am. Chem. Soc.* **137**, 7270–7273.
- (63) Deiters, A., and Schultz, P. G. (2005) In vivo incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg. Med. Chem. Lett.* **15**, 1521–1524.
- (64) Deiters, A., Cropp, T. A., Summerer, D., Mukherji, M., and Schultz, P. G. (2004) Site-specific PEGylation of proteins containing unnatural amino acids. *Bioorg. Med. Chem. Lett.* **14**, 5743–5745.
- (65) Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J. W. (2012) Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* **4**, 298–304.
- (66) Lee, Y. J., Kurra, Y., Yang, Y., Torres-Kolbus, J., Deiters, A., and Liu, W. R. (2014) Genetically encoded unstrained olefins for live cell labeling with tetrazine dyes. *Chem. Commun. (Cambridge, U. K.)* **50**, 13085–13088.

- (67) Lang, K., and Chin, J. (2014) Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* 114, 4764–4806.
- (68) Lang, K., Davis, L., Wallace, S., et al. (2012) 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.
- (69) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N(epsilon)-(o-Azidobenzoyloxycarbonyl) lysine for Site-Specific Protein Modification. *Chem. Biol.* 15, 1187–1197.
- (70) Xiao, H., Peters, F., Yang, P., Reed, S., Chittuluru, J., and Schultz, P. (2014) Genetic Incorporation of Histidine Derivatives Using an Engineered Pyrrolysyl-tRNA Synthetase. *ACS Chem. Biol.* 9, 1092–1096.
- (71) Xiao, H., Xuan, W., Shao, S., Liu, T., and Schultz, P. (2015) Genetic Incorporation of epsilon-N-2-Hydroxyisobutyryl-lysine into Recombinant Histones. *ACS Chem. Biol.* 10, 1599–1603.
- (72) Xie, J., Supekova, L., and Schultz, P. (2007) A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli. *ACS Chem. Biol.* 2, 474–478.
- (73) Xie, J., Wang, L., Wu, N., Brock, A., Spraggan, G., and Schultz, P. G. (2004) The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. *Nat. Biotechnol.* 22, 1297–1301.
- (74) Rogerson, D., Sachdeva, A., Wang, K., et al. (2015) Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* 11, 496.
- (75) Park, H., Hohn, M., Umebara, T., et al. (2011) Expanding the Genetic Code of Escherichia coli with Phosphoserine. *Science* 333, 1151–1154.
- (76) Groff, D., Chen, P., Peters, F., and Schultz, P. (2010) A Genetically Encoded epsilon-N-Methyl Lysine in Mammalian Cells. *ChemBioChem* 11, 1066–1068.
- (77) Kim, C., Kang, M., Kim, H., Chatterjee, A., and Schultz, P. (2012) Site-Specific Incorporation of epsilon-N-Crotonyllysine into Histones. *Angew. Chem., Int. Ed.* 51, 7246–7249.
- (78) Tan, Z., Forster, A., Blacklow, S., and Cornish, V. (2004) Amino acid backbone specificity of the Escherichia coli translation machinery. *J. Am. Chem. Soc.* 126, 12752–12753.
- (79) Guo, J., Wang, J., Anderson, J., and Schultz, P. (2008) Addition of an alpha-hydroxy acid to the genetic code of bacteria. *Angew. Chem., Int. Ed.* 47, 722–725.
- (80) Kobayashi, T., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2009) Recognition of Non-alpha-amino Substrates by Pyrrolysyl-tRNA Synthetase. *J. Mol. Biol.* 385, 1352–1360.
- (81) Melo Czekster, C., Robertson, W., Walker, A., Soll, D., and Schepartz, A. (2016) In Vivo Biosynthesis of a beta-Amino Acid-Containing Protein. *J. Am. Chem. Soc.* 138, 5194–5197.
- (82) Wang, L., and Schultz, P. G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol.* 8, 883–890.
- (83) Wu, N., Deiters, A., Cropp, T. A., King, D., and Schultz, P. G. (2004) A Genetically Encoded Photocaged Amino Acid. *J. Am. Chem. Soc.* 126, 14306–14307.
- (84) Sakamoto, K., Hayashi, A., and Sakamoto, A. (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res.* 30, 4692–4699.
- (85) Hughes, R. A., and Ellington, A. D. (2010) Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. *Nucleic Acids Res.* 38, 6813–6830.
- (86) Italia, J. S., Addy, P. S., Wrobel, C. J., et al. (2017) An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. *Nat. Chem. Biol.* 13, 446–450.
- (87) Chatterjee, A., Xiao, H., and Schultz, P. G. (2012) Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14841–14846.
- (88) Santoro, S. W., Anderson, J. C., Lakshman, V., and Schultz, P. G. (2003) An archaeabacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in Escherichia coli. *Nucleic Acids Res.* 31, 6700–6709.
- (89) Luo, X., Zambaldo, C., Liu, T., et al. (2016) Recombinant thiopeptides containing noncanonical amino acids. *Proc. Natl. Acad. Sci. U. S. A.* 113, 3615–3620.
- (90) Young, T. S., Ahmad, I., Brock, A., and Schultz, P. G. (2009) Expanding the genetic repertoire of the methylotrophic yeast *Pichia pastoris*. *Biochemistry* 48, 2643–2653.
- (91) Greiss, S., and Chin, J. W. (2011) Expanding the Genetic Code of an Animal. *J. Am. Chem. Soc.* 133, 14196–14199.
- (92) Parrish, A. R., She, X., Xiang, Z., et al. (2012) Expanding the genetic code of *Caenorhabditis elegans* using bacterial aminoacyl-tRNA synthetase/tRNA pairs. *ACS Chem. Biol.* 7, 1292–1302.
- (93) Bianco, A., Townsley, F. M., Greiss, S., Lang, K., and Chin, J. W. (2012) Expanding the genetic code of *Drosophila melanogaster*. *Nat. Chem. Biol.* 8, 748–750.
- (94) 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.* 52, 9700–9704.
- (95) Liu, J., Hemphill, J., Samanta, S., Tsang, M., and Deiters, A. (2017) Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling. *J. Am. Chem. Soc.* 139, 9100–9103.
- (96) Ernst, R. J., Krogager, T. P., Maywood, E. S., et al. (2016) Genetic code expansion in the mouse brain. *Nat. Chem. Biol.* 12, 776–778.
- (97) Elliott, T. S., Townsley, F. M., Bianco, A., et al. (2014) Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat. Biotechnol.* 32, 465–472.
- (98) Elliott, T. S., Bianco, A., Townsley, F. M., Fried, S. D., and Chin, J. W. (2016) Tagging and Enriching Proteins Enables Cell-Specific Proteomics. *Cell Chem. Biol.* 23, 805–815.
- (99) 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.
- (100) Shen, B., Xiang, Z., Miller, B., et al. (2011) Genetically encoding unnatural amino acids in neural stem cells and optically reporting voltage-sensitive domain changes in differentiated neurons. *Stem Cells* 29, 1231–1240.
- (101) Wang, W., Takimoto, J. K., Louie, G. V., et al. (2007) Genetically encoding unnatural amino acids for cellular and neuronal studies. *Nat. Neurosci.* 10, 1063–1072.
- (102) Kang, J.-Y., Kawaguchi, D., and Wang, L. (2016) Optical Control of a Neuronal Protein Using a Genetically Encoded Unnatural Amino Acid in Neurons. *J. Visualized Exp.*, e53818.
- (103) Chatterjee, A., Xiao, H., Bollong, M., Ai, H. W., and Schultz, P. G. (2013) Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11803–11808.
- (104) Zheng, Y., Lewis, T. L., Igo, P., Polleux, F., and Chatterjee, A. (2017) Virus-Enabled Optimization and Delivery of the Genetic Machinery for Efficient Unnatural Amino Acid Mutagenesis in Mammalian Cells and Tissues. *ACS Synth. Biol.* 6, 13–18.
- (105) Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* 395, 361–374.
- (106) Guo, J., Melançon, C. E., Lee, H. S., Groff, D., and Schultz, P. G. (2009) Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew. Chem., Int. Ed.* 48, 9148–9151.
- (107) Ryu, Y., and Schultz, P. G. (2006) Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods* 3, 263–265.
- (108) Amiram, M., Haimovich, A. D., Fan, C., et al. (2015) Evolution of translation machinery in recoded bacteria enables multi-site

- incorporation of nonstandard amino acids. *Nat. Biotechnol.* 33, 1272–1279.
- (109) Fan, C., Xiong, H., Reynolds, N. M., and Söll, D. (2015) Rationally evolving tRNAPyl for efficient incorporation of non-canonical amino acids. *Nucleic Acids Res.* 43, e156.
- (110) Takimoto, J. K., Adams, K. L., Xiang, Z., and Wang, L. (2009) Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells. *Mol. BioSyst.* 5, 931–934.
- (111) Tian, F., Lu, Y., Manibusan, A., et al. (2014) A general approach to site-specific antibody drug conjugates. *Proc. Natl. Acad. Sci. U. S. A.* 111, 1766–1771.
- (112) Cho, H., Daniel, T., Buechler, Y. J., et al. (2011) Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9060–9065.
- (113) Koehler, C., Sauter, P. F., Wawryszyn, M., et al. (2016) Genetic code expansion for multiprotein complex engineering. *Nat. Methods* 13, 997–1000.
- (114) Johnson, D. B., Xu, J., Shen, Z., et al. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.* 7, 779–786.
- (115) Johnson, D. B., Wang, C., Xu, J., et al. (2012) Release factor one is nonessential in Escherichia coli. *ACS Chem. Biol.* 7, 1337–1344.
- (116) Lajoie, M. J., Rovner, A. J., Goodman, D. B., et al. (2013) Genomically recoded organisms expand biological functions. *Science* 342, 357–360.
- (117) Rackham, O., and Chin, J. W. (2005) A network of orthogonal ribosome x mRNA pairs. *Nat. Chem. Biol.* 1, 159–166.
- (118) Orelle, C., Carlson, E. D., Szal, T., Florin, T., Jewett, M. C., and Mankin, A. S. (2015) Protein synthesis by ribosomes with tethered subunits. *Nature* 524, 119–124.
- (119) Wang, K., Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2007) Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol.* 25, 770–777.
- (120) Isaacs, F. J., Carr, P. A., Wang, H. H., et al. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* 333, 348–353.
- (121) Mukai, T., Hayashi, A., Iraha, F., et al. (2010) Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Res.* 38, 8188–8195.
- (122) Mukai, T., Hoshi, H., Ohtake, K., et al. (2015) Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. *Sci. Rep.* 5, 9699.
- (123) Wan, W., Huang, Y., Wang, Z., et al. (2010) A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli. *Angew. Chem., Int. Ed.* 49, 3211–3214.
- (124) Ostrov, N., Landon, M., Guell, M., et al. (2016) Design, synthesis, and testing toward a 57-codon genome. *Science* 353, 819–822.
- (125) Chin, J. W. (2012) Molecular biology. Reprogramming the genetic code. *Science* 336, 428–429.
- (126) Wang, K., Sachdeva, A., Cox, D. J., et al. (2014) Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* 6, 393–403.
- (127) Malyshev, D. A., Dhami, K., Lavergne, T., et al. (2014) A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509, 385–388.
- (128) Malyshev, D. A., and Romesberg, F. E. (2015) The expanded genetic alphabet. *Angew. Chem., Int. Ed.* 54, 11930–11944.
- (129) Zhang, Y., Lamb, B. M., Feldman, A. W., et al. (2017) A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc. Natl. Acad. Sci. U. S. A.* 114, 1317–1322.
- (130) Zeng, Y., Wang, W., and Liu, W. R. (2014) Towards reassigning the rare AGG codon in Escherichia coli. *ChemBioChem* 15, 1750–1754.
- (131) Rovner, A. J., Haimovich, A. D., Katz, S. R., et al. (2015) Recoded organisms engineered to depend on synthetic amino acids. *Nature* 518, 89–93.
- (132) Neumann, H., Slusarczyk, A. L., and Chin, J. W. (2010) De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J. Am. Chem. Soc.* 132, 2142–2144.
- (133) Sachdeva, A., Wang, K., Elliott, T., and Chin, J. W. (2014) Concerted, rapid, quantitative, and site-specific dual labeling of proteins. *J. Am. Chem. Soc.* 136, 7785–7788.
- (134) Zheng, Y., Addy, P. S., Mukherjee, R., and Chatterjee, A. (2017) Defining the current scope and limitations of dual non-canonical amino acid mutagenesis in mammalian cells. *Chem. Sci.* 8, 7211–7217.
- (135) Zheng, Y., Mukherjee, R., Chin, M. A., Igo, P., Gilgenast, M. J., and Chatterjee, A. (2017) Expanding the Scope of Single- and Double-Noncanonical Amino Acid Mutagenesis in Mammalian Cells Using Orthogonal Polyspecific Leucyl-tRNA Synthetases. *Biochemistry*, DOI: 10.1021/acs.biochem.7b00952.
- (136) Xiao, H., Chatterjee, A., Choi, S. H., Bajjuri, K. M., Sinha, S. C., and Schultz, P. G. (2013) Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew. Chem., Int. Ed.* 52, 14080–14083.
- (137) Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems. *Nat. Chem. Biol.* 1, 13–21.
- (138) Matsuda, S., Fillo, J. D., Henry, A. A., et al. (2007) Efforts toward expansion of the genetic alphabet: structure and replication of unnatural base pairs. *J. Am. Chem. Soc.* 129, 10466–10473.
- (139) Shogren-Knaak, M. A., Alaimo, P. J., and Shokat, K. M. (2001) Recent advances in chemical approaches to the study of biological systems. *Annu. Rev. Cell Dev. Biol.* 17, 405–433.
- (140) Yee, C. S., Seyedsayamost, M. R., Chang, M. C. Y., Nocera, D. G., and Stubbe, J. (2003) Generation of the R2 subunit of ribonucleotide reductase by intein chemistry: insertion of 3-nitrotyrosine at residue 356 as a probe of the radical initiation process. *Biochemistry* 42, 14541–14552.
- (141) Seyedsayamost, M. R., and Stubbe, J. (2006) Site-specific replacement of Y₃₅₆ with 3,4-dihydroxyphenylalanine in the β2 subunit of *E. coli* ribonucleotide reductase. *J. Am. Chem. Soc.* 128, 2522–2523.
- (142) Seyedsayamost, M. R., Yee, C. S., and Stubbe, J. (2007) Site-specific incorporation of fluorotyrosines into the R2 subunit of *E. coli* ribonucleotide reductase by expressed protein ligation. *Nat. Protoc.* 2, 1225–1235.
- (143) Minnihan, E. C., Seyedsayamost, M. R., Uhlin, U., and Stubbe, J. (2011) Kinetics of radical intermediate formation and deoxynucleotide production in 3-aminotyrosine-substituted *Escherichia coli* ribonucleotide reductases. *J. Am. Chem. Soc.* 133, 9430–9440.
- (144) Liu, X., Li, J., Dong, J., Hu, C., Gong, W., and Wang, J. (2012) Genetic incorporation of a metal-chelating amino acid as a probe for protein electron transfer. *Angew. Chem., Int. Ed.* 51, 10261–10265.
- (145) Xiao, H., Peters, F. B., Yang, P. Y., Reed, S., Chittuluru, J. R., and Schultz, P. G. (2014) Genetic incorporation of histidine derivatives using an engineered pyrrolysyl-tRNA synthetase. *ACS Chem. Biol.* 9, 1092–1096.
- (146) Wilkins, B., Marionni, S., Young, D., et al. (2010) Site-Specific Incorporation of Fluorotyrosines into Proteins in *Escherichia coli* by Photochemical Disguise. *Biochemistry* 49, 1557–1559.
- (147) Oyala, P. H., Ravichandran, K. R., Funk, M. A., et al. (2016) Biophysical Characterization of Fluorotyrosine Probes Site-Specifically Incorporated into Enzymes: *E. coli* Ribonucleotide Reductase As an Example. *J. Am. Chem. Soc.* 138, 7951–7964.
- (148) Xiang, Z., Lacey, V., Ren, H., et al. (2014) Proximity-Enabled Protein Crosslinking through Genetically Encoding Haloalkane Unnatural Amino Acids. *Angew. Chem., Int. Ed.* 53, 2190–2193.
- (149) Kobayashi, T., Hoppmann, C., Yang, B., and Wang, L. (2016) Using Protein-Confining Proximity To Determine Chemical Reactivity. *J. Am. Chem. Soc.* 138, 14832–14835.
- (150) Chen, X. H., Xiang, Z., Hu, Y. S., Lacey, V. K., Cang, H., and Wang, L. (2014) Genetically encoding an electrophilic amino acid for protein stapling and covalent binding to native receptors. *ACS Chem. Biol.* 9, 1956–1961.
- (151) Xuan, W., Shao, S., and Schultz, P. G. (2017) Protein Crosslinking by Genetically Encoded Noncanonical Amino Acids with

- Reactive Aryl Carbamate Side Chains. *Angew. Chem., Int. Ed.* 56, 5096–5100.
- (152) Xuan, W., Li, J., Luo, X., and Schultz, P. G. (2016) Genetic Incorporation of a Reactive Isothiocyanate Group into Proteins. *Angew. Chem., Int. Ed.* 55, 10065–10068.
- (153) Yang, Y., Song, H., He, D., et al. (2016) Genetically encoded protein photocrosslinker with a transferable mass spectrometry-identifiable label. *Nat. Commun.* 7, 12299.
- (154) Freinkman, E., Chng, S. S., and Kahne, D. (2011) The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. *Proc. Natl. Acad. Sci. U. S. A.* 108, 2486–2491.
- (155) Wilkins, B. J., Rall, N. A., Ostwal, Y., et al. (2014) A cascade of histone modifications induces chromatin condensation in mitosis. *Science* 343, 77–80.
- (156) Joiner, C. M., Breen, M. E., Clayton, J., and Mapp, A. K. (2017) A Bifunctional Amino Acid Enables Both Covalent Chemical Capture and Isolation of in Vivo Protein-Protein Interactions. *ChemBioChem* 18, 181–184.
- (157) Xiang, Z., Ren, H., Hu, Y. S., et al. (2013) Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. *Nat. Methods* 10, 885–888.
- (158) Coin, I., Katritch, V., Sun, T., et al. (2013) Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR. *Cell* 155, 1258–1269.
- (159) Grunbeck, A., Huber, T., Sachdev, P., and Sakmar, T. P. (2011) Mapping the ligand-binding site on a G protein-coupled receptor (GPCR) using genetically encoded photocrosslinkers. *Biochemistry* 50, 3411–3413.
- (160) Zhang, M., Lin, S., Song, X., et al. (2011) A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat. Chem. Biol.* 7, 671–677.
- (161) Rannversson, H., Andersen, J., Sørensen, L., et al. (2016) Genetically encoded photocrosslinkers locate the high-affinity binding site of antidepressant drugs in the human serotonin transporter. *Nat. Commun.* 7, 11261.
- (162) Hoppmann, C., and Wang, L. (2016) Proximity-enabled bioreactivity to generate covalent peptide inhibitors of p53-Mdm4. *Chem. Commun. (Cambridge, U. K.)* 52 (29), 5140–5143.
- (163) Luo, J., Upadhyay, R., Naro, Y., et al. (2014) Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. *J. Am. Chem. Soc.* 136, 15551–15558.
- (164) Lampkowski, J. S., Uthappa, D. M., and Young, D. D. (2015) Site-specific incorporation of a fluorescent terphenyl unnatural amino acid. *Bioorg. Med. Chem. Lett.* 25, 5277–5280.
- (165) Lee, H., Guo, J., Lemke, E., Dimla, R., and Schultz, P. (2009) Genetic Incorporation of a Small, Environmentally Sensitive, Fluorescent Probe into Proteins in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* 131, 12921.
- (166) 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.
- (167) Lacey, V. K., Parrish, A. R., Han, S., et al. (2011) A fluorescent reporter of the phosphorylation status of the substrate protein STAT3. *Angew. Chem., Int. Ed.* 50, 8692–8696.
- (168) Mitchell, A. L., Addy, P. S., Chin, M. A., and Chatterjee, A. (2017) A Unique Genetically Encoded FRET Pair in Mammalian Cells. *ChemBioChem* 18, 511–514.
- (169) Kalstrup, T., and Blunck, R. (2013) Dynamics of internal pore opening in K(V) channels probed by a fluorescent unnatural amino acid. *Proc. Natl. Acad. Sci. U. S. A.* 110, 8272–8277.
- (170) Groff, D., Thielges, M. C., Cellitti, S., Schultz, P. G., and Romesberg, F. E. (2009) Efforts toward the direct experimental characterization of enzyme microenvironments: tyrosine100 in dihydrofolate reductase. *Angew. Chem., Int. Ed.* 48, 3478–3481.
- (171) Schultz, K., Supekova, L., Ryu, Y., Xie, J., Perera, R., and Schultz, P. (2006) A genetically encoded infrared probe. *J. Am. Chem. Soc.* 128, 13984–13985.
- (172) Smith, E. E., Linderman, B. Y., Luskin, A. C., and Brewer, S. H. (2011) Probing local environments with the infrared probe: L-4-nitrophenylalanine. *J. Phys. Chem. B* 115, 2380–2385.
- (173) Taskent-Sezgin, H., Chung, J., Patsalo, V., et al. (2009) Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: a combined fluorescence, IR and molecular dynamics study. *Biochemistry* 48, 9040–9046.
- (174) Basom, E. J., Maj, M., Cho, M., and Thielges, M. C. (2016) Site-Specific Characterization of Cytochrome P450cam Conformations by Infrared Spectroscopy. *Anal. Chem.* 88, 6598–6606.
- (175) Ye, S., Huber, T., Vogel, R., and Sakmar, T. P. (2009) FTIR analysis of GPCR activation using azido probes. *Nat. Chem. Biol.* 5, 397–399.
- (176) Ye, S., Zaitseva, E., Caltabiano, G., et al. (2010) Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature* 464, 1386–1389.
- (177) Fleissner, M. R., Brustad, E. M., Kálai, T., et al. (2009) Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21637–21642.
- (178) Schmidt, M. J., Borbas, J., Drescher, M., and Summerer, D. (2014) A genetically encoded spin label for electron paramagnetic resonance distance measurements. *J. Am. Chem. Soc.* 136, 1238–1241.
- (179) Schmidt, M. J., Fedoseev, A., Bücker, D., et al. (2015) EPR Distance Measurements in Native Proteins with Genetically Encoded Spin Labels. *ACS Chem. Biol.* 10, 2764–2771.
- (180) Cellitti, S. E., Jones, D. H., Laggan, L., 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. *J. Am. Chem. Soc.* 130, 9268–9281.
- (181) Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* 129, 1160–1166.
- (182) Lampe, J. N., Floor, S. N., Gross, J. D., et al. (2008) Ligand-induced conformational heterogeneity of cytochrome P450 CYP119 identified by 2D NMR spectroscopy with the unnatural amino acid (13)C-p-methoxyphenylalanine. *J. Am. Chem. Soc.* 130, 16168–16169.
- (183) Yang, F., Yu, X., Liu, C., et al. (2015) Phospho-selective mechanisms of arrestin conformations and functions revealed by unnatural amino acid incorporation and (19)F-NMR. *Nat. Commun.* 6, 8202.
- (184) Shi, P., Wang, H., Xi, Z., Shi, C., Xiong, Y., and Tian, C. (2011) Site-specific F NMR chemical shift and side chain relaxation analysis of a membrane protein labeled with an unnatural amino acid. *Protein Sci.* 20, 224–228.
- (185) Shi, P., Xi, Z., Wang, H., Shi, C., Xiong, Y., and Tian, C. (2010) Site-specific protein backbone and side-chain NMR chemical shift and relaxation analysis of human vinexin SH3 domain using a genetically encoded 15N/19F-labeled unnatural amino acid. *Biochem. Biophys. Res. Commun.* 402, 461–466.
- (186) Niu, W., and Guo, J. (2013) Expanding the chemistry of fluorescent protein biosensors through genetic incorporation of unnatural amino acids. *Mol. BioSyst.* 9, 2961–2970.
- (187) Ayyadurai, N., Saravanan Prabhu, N., Deepankumar, K., et al. (2011) Development of a selective, sensitive, and reversible biosensor by the genetic incorporation of a metal-binding site into green fluorescent protein. *Angew. Chem., Int. Ed.* 50, 6534–6537.
- (188) Liu, X., Li, J., Hu, C., et al. (2013) Significant expansion of the fluorescent protein chromophore through the genetic incorporation of a metal-chelating unnatural amino acid. *Angew. Chem., Int. Ed.* 52, 4805–4809.
- (189) Shang, X., Wang, N., Cerny, R., Niu, W., and Guo, J. (2017) Fluorescent Protein-Based Turn-On Probe through a General Protection-Deprotection Design Strategy. *ACS Sens.* 2, 961–966.

- (190) Wang, F., Niu, W., Guo, J., and Schultz, P. G. (2012) Unnatural amino acid mutagenesis of fluorescent proteins. *Angew. Chem., Int. Ed.* 51, 10132–10135.
- (191) Chen, S., Chen, Z. J., Ren, W., and Ai, H. W. (2012) Reaction-based genetically encoded fluorescent hydrogen sulfide sensors. *J. Am. Chem. Soc.* 134, 9589–9592.
- (192) Wang, W. W., Zeng, Y., Wu, B., Deiters, A., and Liu, W. R. (2016) A Chemical Biology Approach to Reveal Sirt6-targeted Histone H3 Sites in Nucleosomes. *ACS Chem. Biol.* 11, 1973–1981.
- (193) Chen, Z. J., Tian, Z., Kallio, K., et al. (2016) The N-B Interaction through a Water Bridge: Understanding the Chemo-selectivity of a Fluorescent Protein Based Probe for Peroxynitrite. *J. Am. Chem. Soc.* 138, 4900–4907.
- (194) Chen, Z. J., Ren, W., Wright, Q. E., and Ai, H. W. (2013) Genetically encoded fluorescent probe for the selective detection of peroxynitrite. *J. Am. Chem. Soc.* 135, 14940–14943.
- (195) Maza, J., Howard, C., Vipani, M., Travis, C., and Young, D. (2017) Utilization of alkyne bioconjugations to modulate protein function. *Bioorg. Med. Chem. Lett.* 27, 30–33.
- (196) Padilla, M. S., and Young, D. D. (2015) Photosensitive GFP mutants containing an azobenzene unnatural amino acid. *Bioorg. Med. Chem. Lett.* 25, 470–473.
- (197) Grinstead, K. M., Rowe, L., Ensor, C. M., et al. (2016) Red-Shifted Aequorin Variants Incorporating Non-Canonical Amino Acids: Applications in In Vivo Imaging. *PLoS One* 11, e0158579.
- (198) Villa, J. K., Tran, H. A., and Vipani, M. (2017) Fluorescence Modulation of Green Fluorescent Protein Using Fluorinated Unnatural Amino Acids. *Molecules* 22, 1194.
- (199) Heinemann, I. U., Rovner, A. J., Aerni, H. R., et al. (2012) Enhanced phosphoserine insertion during *Escherichia coli* protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett.* 586, 3716–3722.
- (200) Liu, C. C., and Schultz, P. G. (2006) Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat. Biotechnol.* 24, 1436–1440.
- (201) Liu, C. C., Cellitti, S. E., Geierstanger, B. H., and Schultz, P. G. (2009) Efficient expression of tyrosine-sulfated proteins in *E. coli* using an expanded genetic code. *Nat. Protoc.* 4, 1784–1789.
- (202) Knight, W., and Cropp, T. (2015) Genetic encoding of the post-translational modification 2-hydroxyisobutyryl-lysine. *Org. Biomol. Chem.* 13, 6479–6481.
- (203) Erickson, S. B., Mukherjee, R., Kelemen, R. E., Wrobel, C. J., Cao, X., and Chatterjee, A. (2017) Precise Photoremovable Perturbation of a Virus-Host Interaction. *Angew. Chem., Int. Ed.* 56, 4234–4237.
- (204) Serwa, R., Wilkening, L., Del Signore, G., et al. (2009) Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. *Angew. Chem., Int. Ed.* 48, 8234–8239.
- (205) Rust, H. L., Subramanian, V., West, G. M., Young, D. D., Schultz, P. G., and Thompson, P. R. (2014) Using unnatural amino acid mutagenesis to probe the regulation of PRMT1. *ACS Chem. Biol.* 9, 649–655.
- (206) Luo, X., Fu, G., Wang, R., et al. (2017) Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat. Chem. Biol.* 13, 845.
- (207) Castañeda, C. A., Dixon, E. K., Walker, O., et al. (2016) Linkage via K27 Bestows Ubiquitin Chains with Unique Properties among Polyubiquitins. *Structure* 24, 423–436.
- (208) Zambaldo, C., Luo, X., Mehta, A. P., and Schultz, P. G. (2017) Recombinant Macroyclic Lanthipeptides Incorporating Non-Canonical Amino Acids. *J. Am. Chem. Soc.* 139, 11646–11649.
- (209) Bionda, N., Cryan, A. L., and Fasan, R. (2014) Bioinspired strategy for the ribosomal synthesis of thioether-bridged macrocyclic peptides in bacteria. *ACS Chem. Biol.* 9, 2008–2013.
- (210) Smith, J. M., Vitali, F., Archer, S. A., and Fasan, R. (2011) Modular assembly of macrocyclic organo-peptide hybrids using synthetic and genetically encoded precursors. *Angew. Chem., Int. Ed.* 50, 5075–5080.
- (211) Frost, J. R., Vitali, F., Jacob, N. T., Brown, M. D., and Fasan, R. (2013) Macrocyclization of organo-peptide hybrids through a dual bio-orthogonal ligation: insights from structure-reactivity studies. *ChemBioChem* 14, 147–160.
- (212) Bindman, N. A., Bobeica, S. C., Liu, W. R., and van der Donk, W. A. (2015) Facile Removal of Leader Peptides from Lanthipeptides by Incorporation of a Hydroxy Acid. *J. Am. Chem. Soc.* 137, 6975–6978.
- (213) Khare, P., Jain, A., Gulbake, A., Soni, V., Jain, N., and Jain, S. (2009) Bioconjugates: Harnessing Potential for Effective Therapeutics. *Crit. Rev. Ther. Drug Carrier Syst.* 26, 119–155.
- (214) Wang, L., Zhang, Z., Brock, A., and Schultz, P. G. (2003) Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 56–61.
- (215) Hancock, S. M., Upadhyay, 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.
- (216) Maza, J. C., McKenna, J. R., Raliski, B. K., Freedman, M. T., and Young, D. D. (2015) Synthesis and Incorporation of Unnatural Amino Acids To Probe and Optimize Protein Bioconjugations. *Bioconjugate Chem.* 26, 1884–1889.
- (217) Wang, Q., Chan, T. R., Hilgraf, R., Fokin, V. V., Sharpless, K. B., and Finn, M. G. (2003) Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. *J. Am. Chem. Soc.* 125, 3192–3193.
- (218) Lang, K., and Chin, J. W. (2014) Bioorthogonal reactions for labeling proteins. *ACS Chem. Biol.* 9, 16–20.
- (219) Plass, T., Milles, S., Koehler, C., Schultz, C., and Lemke, E. A. (2011) Genetically encoded copper-free click chemistry. *Angew. Chem., Int. Ed.* 50, 3878–3881.
- (220) Lang, K., Davis, L., Wallace, S., et al. (2012) Genetic Encoding of bicyclonynes 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.
- (221) Baskin, J. M., Prescher, J. A., Laughlin, S. T., et al. (2007) Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16793–16797.
- (222) Uttamapinant, C., Howe, J. D., Lang, K., et al. (2015) Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. *J. Am. Chem. Soc.* 137, 4602–4605.
- (223) Xue, L., Prifti, E., and Johnsson, K. (2016) A General Strategy for the Semisynthesis of Ratiometric Fluorescent Sensor Proteins with Increased Dynamic Range. *J. Am. Chem. Soc.* 138, 5258–5261.
- (224) Kumaresan, P., Luo, J., Song, A., Marik, J., and Lam, K. (2008) Evaluation of ketone-oxime method for developing therapeutic on-demand cleavable immunoconjugates. *Bioconjugate Chem.* 19, 1313–1318.
- (225) Brustad, E., Lemke, E., Schultz, P., and Deniz, A. (2008) A General and Efficient Method for the Site-Specific Dual-Labeling of Proteins for Single Molecule Fluorescence Resonance Energy Transfer. *J. Am. Chem. Soc.* 130, 17664.
- (226) Axup, J. Y., Bajjuri, K. M., Ritland, M., et al. (2012) Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proc. Natl. Acad. Sci. U. S. A.* 109, 16101–16106.
- (227) Hutchins, B., Kazane, S., Staflin, K., et al. (2011) Selective Formation of Covalent Protein Heterodimers with an Unnatural Amino Acid. *Chem. Biol.* 18, 299–303.
- (228) Hutchins, B., Kazane, S., Staflin, K., et al. (2011) Site-Specific Coupling and Sterically Controlled Formation of Multimeric Antibody Fab Fragments with Unnatural Amino Acids. *J. Mol. Biol.* 406, 595–603.
- (229) Kazane, S. A., Sok, D., Cho, E. H., et al. (2012) Site-specific DNA-antibody conjugates for specific and sensitive immuno-PCR. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3731–3736.
- (230) Kim, C. H., Axup, J. Y., Dubrovská, A., et al. (2012) Synthesis of bispecific antibodies using genetically encoded unnatural amino acids. *J. Am. Chem. Soc.* 134, 9918–9921.

- (231) Kazane, S. A., Axup, J. Y., Kim, C. H., et al. (2013) Self-assembled antibody multimers through peptide nucleic acid conjugation. *J. Am. Chem. Soc.* 135, 340–346.
- (232) Lu, H., Wang, D., Kazane, S., et al. (2013) Site-specific antibody-polymer conjugates for siRNA delivery. *J. Am. Chem. Soc.* 135, 13885–13891.
- (233) Ai, H. W., Shen, W., Brustad, E., and Schultz, P. G. (2010) Genetically encoded alkenes in yeast. *Angew. Chem., Int. Ed.* 49, 935–937.
- (234) Lin, Y. A., Chalker, J. M., Floyd, N., Bernardes, G. J., and Davis, B. G. (2008) Allyl sulfides are privileged substrates in aqueous cross-metathesis: application to site-selective protein modification. *J. Am. Chem. Soc.* 130, 9642–9643.
- (235) Addy, P. S., Erickson, S. B., Italia, J. S., and Chatterjee, A. (2017) A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for Unclickable Bioconjugation. *J. Am. Chem. Soc.* 139, 11670–11673.
- (236) Row, R. D., Shih, H. W., Alexander, A. T., Mehl, R. A., and Prescher, J. A. (2017) Cyclopropenones for Metabolic Targeting and Sequential Bioorthogonal Labeling. *J. Am. Chem. Soc.* 139, 7370–7375.
- (237) Wang, Z. A., Kurra, Y., Wang, X., et al. (2017) A Versatile Approach for Site-Specific Lysine Acylation in Proteins. *Angew. Chem., Int. Ed.* 56, 1643–1647.
- (238) Li, J., Lin, S., Wang, J., et al. (2013) Ligand-free palladium-mediated site-specific protein labeling inside gram-negative bacterial pathogens. *J. Am. Chem. Soc.* 135, 7330–7338.
- (239) Li, N., Lim, R., Edwardraja, S., and Lin, Q. (2011) Copper-Free Sonogashira Cross-Coupling for Functionalization of Alkyne-Encoded Proteins in Aqueous Medium and in Bacterial Cells. *J. Am. Chem. Soc.* 133, 15316–15319.
- (240) Lampkowski, J. S., Villa, J. K., Young, T. S., and Young, D. D. (2015) Development and Optimization of Glaser-Hay Bioconjugations. *Angew. Chem., Int. Ed.* 54, 9343–9346.
- (241) Maza, J. C., Nimmo, Z. M., and Young, D. D. (2016) Expanding the scope of alkyne-mediated bioconjugations utilizing unnatural amino acids. *Chem. Commun. (Cambridge, U. K.)* 52, 88–91.
- (242) Spicer, C. D., and Davis, B. G. (2011) Palladium-mediated site-selective Suzuki-Miyaura protein modification at genetically encoded aryl halides. *Chem. Commun. (Cambridge, U. K.)* 47, 1698–1700.
- (243) Spicer, C. D., Triemer, T., and Davis, B. G. (2012) Palladium-mediated cell-surface labeling. *J. Am. Chem. Soc.* 134, 800–803.
- (244) Dumas, A., Spicer, C. D., Gao, Z., et al. (2013) Self-ligated Suzuki-Miyaura coupling for site-selective protein PEGylation. *Angew. Chem., Int. Ed.* 52, 3916–3921.
- (245) Spicer, C. D., and Davis, B. G. (2013) Rewriting the bacterial glycocalyx via Suzuki-Miyaura cross-coupling. *Chem. Commun. (Cambridge, U. K.)* 49, 2747–2749.
- (246) Raliski, B. K., Howard, C. A., and Young, D. D. (2014) Site-specific protein immobilization using unnatural amino acids. *Bioconjugate Chem.* 25, 1916–1920.
- (247) Borrmann, A., Milles, S., Plass, T., et al. (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.
- (248) Plass, T., Milles, S., Koehler, C., et al. (2012) Amino Acids for Diels-Alder Reactions in Living Cells. *Angew. Chem., Int. Ed.* 51, 4166–4170.
- (249) Seitchik, J. L., Peeler, J. C., Taylor, M. T., et al. (2012) Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific in Vivo Bioorthogonal Ligation with trans-Cyclooctenes. *J. Am. Chem. Soc.* 134, 2898–2901.
- (250) Blizzard, R. J., Backus, D. R., Brown, W., Bazewicz, C. G., Li, Y., and Mehl, R. A. (2015) Ideal Bioorthogonal Reactions Using A Site-Specifically Encoded Tetrazine Amino Acid. *J. Am. Chem. Soc.* 137, 10044–10047.
- (251) Lukinavicius, G., Umezawa, K., Olivier, N., et al. (2013) A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat. Chem.* 5, 132–139.
- (252) 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.
- (253) Baumdick, M., Bruggemann, Y., and Schmick, M., et al. EGF-dependent re-routing of vesicular recycling switches spontaneous phosphorylation suppression to EGFR signaling. *eLife* 2015, 4, DOI: 10.7554/eLife.12223.
- (254) Rusmini, F., Zhong, Z., and Feijen, J. (2007) Protein immobilization strategies for protein biochips. *Biomacromolecules* 8, 1775–1789.
- (255) Deepankumar, K., Nadarajan, S., Mathew, S., et al. (2015) Engineering Transaminase for Stability Enhancement and Site-Specific Immobilization through Multiple Noncanonical Amino Acids Incorporation. *ChemCatChem* 7, 417–421.
- (256) Lim, S. I., Mizuta, Y., Takasu, A., Kim, Y. H., and Kwon, I. (2014) Site-specific bioconjugation of a murine dihydrofolate reductase enzyme by copper(I)-catalyzed azide-alkyne cycloaddition with retained activity. *PLoS One* 9, e98403.
- (257) Ikeda-Boku, A., Kondo, K., Ohno, S., et al. (2013) Protein fishing using magnetic nanobeads containing calmodulin site-specifically immobilized via an azido group. *J. Biochem.* 154, 159–165.
- (258) Yoshimura, S. H., Khan, S., Ohno, S., et al. (2012) Site-specific attachment of a protein to a carbon nanotube end without loss of protein function. *Bioconjugate Chem.* 23, 1488–1493.
- (259) Deiters, A., Groff, D., Ryu, Y., Xie, J., and Schultz, P. (2006) A genetically encoded photocaged tyrosine. *Angew. Chem., Int. Ed.* 45, 2728–2731.
- (260) Wu, N., Deiters, A., Cropp, T. A., King, D., and Schultz, P. G. (2004) A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* 126, 14306–14307.
- (261) Lemke, E., Summerer, D., Geierstanger, B., Brittain, S., and Schultz, P. (2007) Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat. Chem. Biol.* 3, 769–772.
- (262) Gautier, A., Nguyen, D., Lusic, H., An, W., Deiters, A., and Chin, J. (2010) Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells. *J. Am. Chem. Soc.* 132, 4086.
- (263) Edwards, W., Young, D., and Deiters, A. (2009) Light-Activated Cre Recombinase as a Tool for the Spatial and Temporal Control of Gene Function in Mammalian Cells. *ACS Chem. Biol.* 4, 441–445.
- (264) Gramespacher, J. A., Stevens, A. J., Nguyen, D. P., Chin, J. W., and Muir, T. W. (2017) Intein Zymogens: Conditional Assembly and Splicing of Split Inteins via Targeted Proteolysis. *J. Am. Chem. Soc.* 139, 8074–8077.
- (265) Ren, W., Ji, A., and Ai, H. W. (2015) Light activation of protein splicing with a photocaged fast intein. *J. Am. Chem. Soc.* 137, 2155–2158.
- (266) Arbely, E., Torres-Kolbus, J., Deiters, A., and Chin, J. W. (2012) Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. *J. Am. Chem. Soc.* 134, 11912–11915.
- (267) Nguyen, D., Mahesh, M., Elsasser, S., Hancock, S., Uttamapinant, C., and Chin, J. (2014) Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. *J. Am. Chem. Soc.* 136, 2240–2243.
- (268) Gautier, A., Nguyen, D. P., Lusic, H., An, W., Deiters, A., and Chin, J. W. (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* 132, 4086–4088.
- (269) Hemphill, J., Borchardt, E. K., Brown, K., Asokan, A., and Deiters, A. (2015) Optical Control of CRISPR/Cas9 Gene Editing. *J. Am. Chem. Soc.* 137, 5642–5645.
- (270) Lemke, E. A., Summerer, D., Geierstanger, B. H., Brittain, S. M., and Schultz, P. G. (2007) Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat. Chem. Biol.* 3, 769–772.
- (271) Walker, O. S., Elsässer, S. J., Mahesh, M., Bachman, M., Balasubramanian, S., and Chin, J. W. (2016) Photoactivation of

- Mutant Isocitrate Dehydrogenase 2 Reveals Rapid Cancer-Associated Metabolic and Epigenetic Changes. *J. Am. Chem. Soc.* **138**, 718–721.
- (272) Tsai, Y. H., Essig, S., James, J. R., Lang, K., and Chin, J. W. (2015) Selective, rapid and optically switchable regulation of protein function in live mammalian cells. *Nat. Chem.* **7**, 554–561.
- (273) Li, J., Jia, S., and Chen, P. R. (2014) Diels-Alder reaction-triggered bioorthogonal protein decaging in living cells. *Nat. Chem. Biol.* **10**, 1003–1005.
- (274) Li, J., Yu, J., Zhao, J., et al. (2014) Palladium-triggered deprotection chemistry for protein activation in living cells. *Nat. Chem.* **6**, 352–361.
- (275) Luo, J., Liu, Q., Morihiro, K., and Deiters, A. (2016) Small-molecule control of protein function through Staudinger reduction. *Nat. Chem.* **8**, 1027–1034.
- (276) Xiao, H., and Schultz, P. G. (2016) At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. *Cold Spring Harbor Perspect. Biol.* **8**, a023945.
- (277) Kelemen, R. E., Mukherjee, R., Cao, X., Erickson, S. B., Zheng, Y., and Chatterjee, A. (2016) A Precise Chemical Strategy To Alter the Receptor Specificity of the Adeno-Associated Virus. *Angew. Chem., Int. Ed.* **55**, 10645–10649.
- (278) VanBrunt, M. P., Shanebeck, K., Caldwell, Z., et al. (2015) Genetically Encoded Azide Containing Amino Acid in Mammalian Cells Enables Site-Specific Antibody-Drug Conjugates Using Click Cycloaddition Chemistry. *Bioconjugate Chem.* **26**, 2249–2260.
- (279) Mu, J., Pinkstaff, J., Li, Z., et al. (2012) FGF21 analogs of sustained action enabled by orthogonal biosynthesis demonstrate enhanced antidiabetic pharmacology in rodents. *Diabetes* **61**, 505–512.
- (280) Cho, H., Daniel, T., Buechler, Y. J., et al. (2011) Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9060–9065.
- (281) Kim, C. H., Axup, J. Y., and Schultz, P. G. (2013) Protein conjugation with genetically encoded unnatural amino acids. *Curr. Opin. Chem. Biol.* **17**, 412–419.
- (282) Sun, S. B., Schultz, P. G., and Kim, C. H. (2014) Therapeutic applications of an expanded genetic code. *ChemBioChem* **15**, 1721–1729.
- (283) Lu, H., Wang, D. L., Kazane, S., et al. (2013) Site-Specific Antibody-Polymer Conjugates for siRNA Delivery. *J. Am. Chem. Soc.* **135**, 13885–13891.
- (284) Kularatne, S. A., Deshmukh, V., Gymnopoulos, M., et al. (2013) Recruiting cytotoxic T cells to folate-receptor-positive cancer cells. *Angew. Chem., Int. Ed.* **52**, 12101–12104.
- (285) Ma, J. S., Kim, J. Y., Kazane, S. A., et al. (2016) Versatile strategy for controlling the specificity and activity of engineered T cells. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E450–E458.
- (286) Gauba, V., Grunewald, J., Gorney, V., et al. (2011) Loss of CD4 T-cell-dependent tolerance to proteins with modified amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 12821–12826.
- (287) Gruenewald, J., Tsao, M. L., Perera, R., et al. (2008) Immunochemical termination of self-tolerance. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 11276–11280.
- (288) Kim, C. H., Axup, J. Y., Lawson, B. R., et al. (2013) Bispecific small molecule-antibody conjugate targeting prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17796–17801.
- (289) Wang, N., Li, Y., Niu, W., et al. (2014) Construction of a live-attenuated HIV-1 vaccine through genetic code expansion. *Angew. Chem., Int. Ed.* **53**, 4867–4871.
- (290) Si, L., Xu, H., Zhou, X., et al. (2016) Generation of influenza A viruses as live but replication-incompetent virus vaccines. *Science* **354**, 1170–1173.
- (291) Lin, S., Yan, H., Li, L., et al. (2013) Site-specific engineering of chemical functionalities on the surface of live hepatitis D virus. *Angew. Chem., Int. Ed.* **52**, 13970–13974.
- (292) Mills, J. H., Khare, S. D., Bolduc, J. M., et al. (2013) Computational design of an unnatural amino acid dependent metalloprotein with atomic level accuracy. *J. Am. Chem. Soc.* **135**, 13393–13399.
- (293) Mills, J. H., Sheffler, W., Ener, M. E., et al. (2016) Computational design of a homotrimeric metalloprotein with a trisbipyridyl core. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 15012–15017.
- (294) Pearson, A. D., Mills, J. H., Song, Y., et al. (2015) Transition states. Trapping a transition state in a computationally designed protein bottle. *Science* **347**, 863–867.
- (295) Liu, C. C., Mack, A. V., Brustad, E. M., et al. (2009) Evolution of proteins with genetically encoded "chemical warheads". *J. Am. Chem. Soc.* **131**, 9616–9617.
- (296) Kang, M., Light, K., Ai, H. W., et al. (2014) Evolution of Iron(II)-Finger Peptides by Using a Bipyridyl Amino Acid. *ChemBioChem* **15**, 822–825.
- (297) Young, T. S., Young, D. D., Ahmad, I., Louis, J. M., Benkovic, S. J., and Schultz, P. G. (2011) Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 11052–11056.
- (298) Xiao, H., Nasertorabi, F., Choi, S. H., et al. (2015) Exploring the potential impact of an expanded genetic code on protein function. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6961–6966.
- (299) Tack, D. S., Ellefson, J. W., Thyer, R., et al. (2016) Addicting diverse bacteria to a noncanonical amino acid. *Nat. Chem. Biol.* **12**, 138–140.
- (300) Ohtake, K., Yamaguchi, A., Mukai, T., et al. (2015) Protein stabilization utilizing a redefined codon. *Sci. Rep.* **5**, 9762.
- (301) Hammerling, M. J., Ellefson, J. W., Boutz, D. R., Marcotte, E. M., Ellington, A. D., and Barrick, J. E. (2014) Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness. *Nat. Chem. Biol.* **10**, 178–180.
- (302) Mandell, D. J., Lajoie, M. J., Mee, M. T., et al. (2015) Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60.
- (303) Koh, M., Nasertorabi, F., Han, G. W., Stevens, R. C., and Schultz, P. G. (2017) Generation of an Orthogonal Protein-Protein Interface with a Noncanonical Amino Acid. *J. Am. Chem. Soc.* **139**, 5728–5731.