HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2018 July 26.

Published in final edited form as:

ACS Chem Biol. 2018 April 20; 13(4): 854–870. doi:10.1021/acschembio.7b00974.

Playing with the molecules of life

Douglas D. Young[b] and **Peter G. Schultz**[*,a]

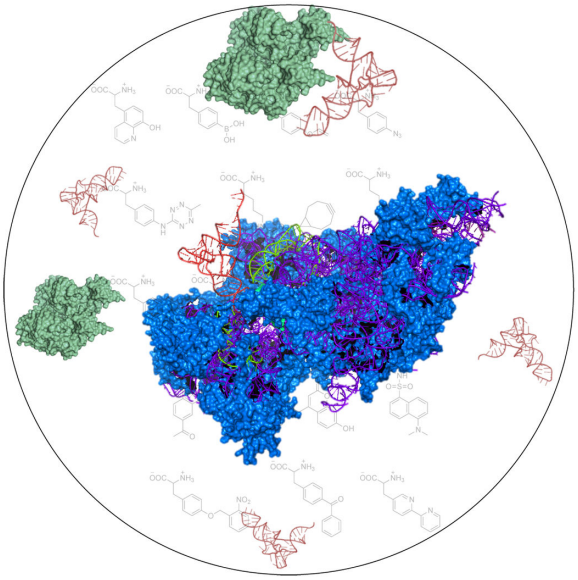
[a]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037 (USA), schultz@scripps.edu

[b]Department of Chemistry, College of William & Mary, P.O. Box 8795, Williamsburg, VA 23187 (USA)

Abstract

Our understanding of the complex 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. 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 in ways previously unimaginable

Graphical abstract



[*]**Corresponding Author** schultz@scripps.edu.

Author Contributions

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

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; Bioothogonal: a process that occurs under physiological conditions by does not react with endogenous systems; Photocaging: the installation of a photo-protecting group often leading to the inactivation of a molecule until the group is remove via light irradiation; Bioconjugation: 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

1. Introduction

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.

2. Expanding the Genetic Code

2.1. 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]

2.1.1. 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 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 towards 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-crosslinkers,[54–58] amino acids with altered pKas 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]

2.1.2. 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 at gm/L 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]

2.1.3. 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/L levels in large scale bacterial fermentation, and gram/L 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 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 3 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]

3. 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 illustrate 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.

3.1. 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 pKas for mechanistic studies, isotopic labels for infrared and NMR studies, photocrosslinkers 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 ncAAs probes have been used in numerous studies, below we highlight instructive examples of their use.

3.1.1 Altering pKa 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 nitrotyrosine[140] and aminotyrosine,[143] which were used to investigate the kinetics of radical intermediate formation within these ribonucleotide reductases.

3.1.2 Protein crosslinking to map protein interactions—An arena where ncAAs have found widespread use is in the field of protein crosslinking, both in vitro and inside living cells. Multiple crosslinking 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 crosslink 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, crosslinking 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 photocrosslinkers have been developed that can be chemically cleaved following crosslinking, further simplifying the identification of a target protein by mass spectrometry.[153] In addition, chemical crosslinkers have also been generated that rely on the nascent chemical reactivity of lysine and cysteine residues with nearby electrophilic ncAAs to afford covalent crosslinks between proteins.[148, 151, 157] Other examples of the use of photocrosslinking 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 crosslinking agents in whole organisms to explore cell-cell interactions.

3.1.3. 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, posttranslational modifications, and residue exposure to solvent.[52] For example,

fluorescent ncAAs were used to detect the binding of glutamine to glutamine binding protein,[166] 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 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.

3.1.4. Infrared probes—The incorporation of deuterium,[170] cyano,[171] nitro,[172] and azido[59] 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.[175, 176] In one example, azidophenylalanine was introduce ed 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-D_4 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.

3.1.5. 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 ^{15}N-labeled tyrosine that is caged with a nitrobenzyl moiety, which upon uncaging by exposure to UV light results in the site-specific incorporation of ^{15}N-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 ^{13}C and/or ^{15}N, or a trifluoro derivative of the same amino acid possessing ^{19}F.[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]

3.1.6 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 containing either L-3,4-dihydroxyphenylalainine (DOPA)[187] or 8-hydroxyquinoline alanine (HqA)_[188] at residue 66 within the fluorophore was used to detect biologically relevant transition metals including Cu^{2+}, Zn^{2+}, Co^{2+}, Fe^{2+}, and Ni^{2+}. 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 H_2O_2 with a boronate ncAA,[190] H_2S with a azidophenylalanine ncAA,[191] and sirtuins using an actylated lysine which can be cleaved by these enzymes.[192] Finally, the *p*-boronophenylalanine has also been exploited to detect peroxynitrite 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]

3.2. 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. [199] 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,[200] azido[59] and acetylenic side chains (Figure 8).[42, 63, 201, 202] Azide and alkynyl ncAAs undergo highly selective copper-catalyzed 1,3-dipolar cycloaddition reactions and have been generally used for in vitro applications.[203, 204] However, strained alkynes have recently been genetically encoded to afford copper-free “click” reactions that can be used in living cells.[205–207] 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[208]. The use of Cys and an ncAA, or two ncAAs, with bio-orthogonal reactivity allows the site-specific labeling of proteins with two different moieties via sequential or one-pot bioorthogonal reactions for applications such as Forster resonance energy transfer (FRET) measurements[17, 133, 136] with excellent dynamic range[209].

Another commonly employed ncAA for bioorthogonal conjugations is *p*-acetylphenylalanine, which can be selectively modified at lower pH with highly efficient

oxime formation reactions.[200] 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, 210–218] 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).[215]

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:[219, 220] hydroxytryptophan residues have been employed for azo-couplings,[221] cyclopropenone and azide residues for reactions with phosphines,[222, 223] alkynyl amino acids have been exploited in Sonogoshira,[224, 225] Glaser-Hay,[226] and Cadiot-Chodkewitz bioconjugation reactions;[227] and *p*-boronophenylalanine and *p*-iodophenylalanine have been used in Suzuki and other transition metal-based couplings.[228–231]

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.[204] One such reaction is the inverse electron demand Diels-Alder reaction between strained alkenes or alkynes and tetrazines (Figure 9). The rate constants for some tetrazine reactions with alkenes and alkynes can exceed 10^4 $M^{-1}s^{-1}$, 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, 233–236] 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.[208, 237] While the refinement and optimization of labeling and imaging approaches is still on going, early examples of applications that are not possible by other methods have already emerged.[238–240]

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).[241] 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.[232, 242–245] Further exploration and optimization of these immobilization technologies has the potential to make significant advances to both materials chemistry and therapeutic diagnostics.

3.3 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 photo-removable protecting group have been genetically encoded in both prokaryotic and eukaryotic cells and organisms.[246–249] 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, 240, 250–255]

One example of the use of this technology is the photoactivation of Cas9 using a photocaged lysine residue.[256] 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 400nm light, allowing the kinetics of these processes to be followed.[257] A final example of light-regulated in vivo protein function involved photocaging a lysine residue within an isocitrate dehydrogenase mutant known to produce the onco-metabolite 2-hydroxygluterate (2-HG).[258] 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 site[247] or by moving a tethered inhibitor in and out of the active site by *cis-trans* photoisomerization of the azobenzene moiety.[259]

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.[260] 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 decaging[261] and phosphine-mediated decaging of azides via a Staudinger reduction have also been employed to chemically modulate protein function in vivo.[262]

3.4. 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, may not be site-specific, or the modifications may be removed in cells or lysates. Consequently, the ability to genetically

encode post-translationally modified 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, 240, 263–266] Stable mimicks of posttranslational modifications, such as phosphonotyrosine and carboxymethyl-phenylalanine have also been genetically encoded.[72, 75, 267] 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.[268] Phosphonotyrosine 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.[269] Subsequent incorporation of phosphonotyrosine 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 towards a variety of lysine modifications, a number of labs have genetically encoded post-translationally modified lysines, including ε-*N*-2-hydroxyisobutylyl-lysine, ε-*N*-pivaloyl-lysine, ε-*N*-methyl-lysine, and ε-*N*-crotonlyl-lysine amongst 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, 270] Finally, *E. coli* has been engineered for the recombinant production of ribosomally synthesized posttranslationally modified peptides (RIPPs) that contain ncAAs, and analogues of the lanthipeptide nisin with altered ring structures and sizes have been produced.[271–275]

3.5 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 (ADCs), antibody-antisense oligonucleotide conjugates and bispecific antibodies have been created for a variety of clinical indications[111, 278112, 218, 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 clearance[285] 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 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 a 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.

3.6. 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, for 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). 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 (Figure 13). 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 crosslinks that lead to significant protein stabilization.[300] Another example was the evolution of 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 a 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.

4. 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 approached synergistically.

ACKNOWLEDGMENT

The authors would like to thank H. Xiao and A. Chatterjee for their helpful comments

Funding Sources

PGS would like to acknowledge the Department of Energy (DOE/DE-SC0011787) and the National Institute of Health (5R01 GM062159) for financial support. DDY would also like to acknowledge the National Institute of Health (R15GM113203) and the Camille and Henry Dreyfus Foundation (TH-17–020) for financial support.

REFERENCES

1. Liu C, Schultz P, Kornberg R, Raetz C, Rothman J, Thorner J. Adding New Chemistries to the Genetic Code. Ann Rev Biochem 2010;79: 413–444.20307192
2. Young TS, Schultz PG. Beyond the Canonical 20 Amino Acids: Expanding the Genetic Lexicon. J Biol Chem 2010;285: 11039–11044.20147747
3. Wang L, Schultz PG. Expanding the genetic code. Angew Chem Int Ed Engl 2004;44: 34–66.15599909
4. Chin JW. Expanding and reprogramming the genetic code of cells and animals. Annu Rev Biochem 2014;83: 379–408.24555827
5. Chatterjee A, Lajoie MJ, Xiao H, Church GM, Schultz PG. A bacterial strain with a unique quadruplet codon specifying non-native amino acids. Chembiochem 2014;15: 1782–1786.24867343
6. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. Nature 2010;464: 441–444.20154731
7. Wang N, Shang X, Cerny R, Niu W, Guo J. Systematic Evolution and Study of UAGN Decoding tRNAs in a Genomically Recoded Bacteria. Sci Rep 2016;6: 21898.26906548
8. Anderson JC, Schultz PG. Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. Biochem 2003;42: 9598–9608.12911301
9. Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, Schultz PG. An expanded genetic code with a functional quadruplet codon. Proc Natl Acad Sci U S A 2004;101: 7566–7571.15138302
10. Anderson JC, Magliery TJ, Schultz PG. Exploring the limits of codon and anticodon size. Chem Biol 2002;9: 237–244.11880038
11. Wang F, Robbins S, Guo J, Shen W, Schultz PG. Genetic incorporation of unnatural amino acids into proteins in Mycobacterium tuberculosis. PLoS One 2010;5: e9354.20179771
12. Lang K, Davis L, Chin JW. Genetic encoding of unnatural amino acids for labeling proteins. Methods Mol Biol 2015;1266: 217–228.25560078
13. Schmied WH, Elsässer SJ, Uttamapinant C, Chin JW. Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. J Am Chem Soc 2014;136: 15577–15583.25350841
14. Sakamoto K, Hayashi A, Sakamoto A, Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. Nucleic Acids Res 2002;30: 4692–4699.12409460

15. Pastrnak M, Magliery T, Schultz P. A new orthogonal suppressor tRNA/aminoacyl-tRNA synthetase pair for evolving an organism with an expanded genetic code. Helvetica Chimica Acta 2000;83: 2277–2286.
16. Wang L, Magliery T, Liu D, Schultz P. A new functional suppressor tRNA/aminoacyl-tRNA synthetase pair for the in vivo incorporation of unnatural amino acids into proteins. J Am Chem Soc 2000;122: 5010–5011.
17. Chatterjee A, Sun SB, Furman JL, Xiao H, Schultz PG. A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. Biochem 2013;52: 1828–1837.23379331
18. Luo X, Fu G, Wang RE, Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. Nat Chem Biol 2017.
19. Mehl RA, Anderson JC, Santoro SW, Generation of a bacterium with a 21 amino acid genetic code. J Am Chem Soc 2003;125: 935–939.12537491
20. Kim CH, Axup JY, Schultz PG. Protein conjugation with genetically encoded unnatural amino acids. Curr Opin Chem Biol 2013;17: 412–419.23664497
21. Chin JW, Cropp TA, Chu S, Meggers E, Schultz PG. Progress toward an expanded eukaryotic genetic code. Chem Biol 2003;10: 511–519.12837384
22. Chin JW, Cropp T, Anderson J, Mukherji M, Zhang Z, Schultz P. An Expanded Eukaryotic Genetic Code. Science 2003;301: 964–967.12920298
23. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science 2001;292: 498–500.11313494
24. Chen PR, Groff D, Guo J, A facile system for encoding unnatural amino acids in mammalian cells. Angew Chem Int Ed Engl 2009;48: 4052–4055.19378306
25. Liu DR, Magliery TJ, Pastrnak M, Schultz PG. 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 1997;94: 10092–10097.9294168
26. Neumann H, Peak-Chew SY, Chin JW. Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. Nat Chem Biol 2008;4: 232–234.18278036
27. Melançon CE, Schultz PG. One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases. Bioorg Med Chem Lett 2009;19: 3845–3847.19398201
28. Xie J, Schultz PG. An expanding genetic code. Methods 2005;36: 227–238.16076448
29. Santoro SW, Wang L, Herberich B, King DS, Schultz PG. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. Nat Biotechnol 2002;20: 1044–1048.12244330
30. Santoro SW, Schultz PG. 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 2003;230: 291–312.12824591
31. Bryson DI, Fan C, Guo LT, Miller C, Söll D, Liu DR. Continuous directed evolution of aminoacyl-tRNA synthetases. Nat Chem Biol 2017;13: 1253–1260.29035361
32. Flügel V, Vrabel M, Schneider S. Structural basis for the site-specific incorporation of lysine derivatives into proteins. PLoS One 2014;9: e96198.24760130
33. Turner JM, Graziano J, Spraggon G, Schultz PG. Structural plasticity of an aminoacyl-tRNA synthetase active site. Proc Natl Acad Sci U S A 2006;103: 6483–6488.16618920
34. Young D, Young T, Jahnz M, Ahmad I, Spraggon G, Schultz P. An Evolved Aminoacyl-tRNA Synthetase with Atypical Polysubstrate Specificity. Biochem 2011;50: 1894–1900.21280675
35. Guo LT, Wang YS, Nakamura A, Polyspecific pyrrolysyl-tRNA synthetases from directed evolution. Proc Natl Acad Sci U S A 2014;111: 16724–16729.25385624
36. Minnihan EC, Young DD, Schultz PG, Stubbe J. Incorporation of fluorotyrosines into ribonucleotide reductase using an evolved, polyspecific aminoacyl-tRNA synthetase. J Am Chem Soc 2011;133: 15942–15945.21913683
37. Young D, Jockush S, Turro N, Schultz P. Synthetase polyspecificity as a tool to modulate protein function. Bioorg Med Chem Lett 2011;21: 7502–7504.22041062

38. Brustad E, Bushey ML, Brock A, Chittuluru J, Schultz PG. A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins. Bioorg Med Chem Lett 2008;18: 6004–6006.18845434
39. Miyake-Stoner SJ, Refakis CA, Hammill JT, Generating permissive site-specific unnatural aminoacyl-tRNA synthetases. Biochemistry 2010;49: 1667–1677.20082521
40. Cooley RB, Karplus PA, Mehl RA. Gleaning unexpected fruits from hard-won synthetases: probing principles of permissivity in non-canonical amino acid-tRNA synthetases. Chembiochem 2014;15: 1810–1819.25044993
41. Wang YS, Fang X, Wallace AL, Wu B, Liu WR. A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. J Am Chem Soc 2012;134: 2950–2953.22289053
42. Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin JW. Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. J Am Chem Soc 2009;131: 8720–8721.19514718
43. Mukai T, Kobayashi T, Hino N, Yanagisawa T, Sakamoto K, Yokoyama S. Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. Biochem Biophys Res Commun 2008;371: 818–822.18471995
44. Liu CC, Mack AV, Tsao ML, Protein evolution with an expanded genetic code. Proc Natl Acad Sci U S A 2008;105: 17688–17693.19004806
45. Davis L, Chin JW. Designer proteins: applications of genetic code expansion in cell biology. Nat Rev Mol Cell Biol 2012;13: 168–182.22334143
46. Dumas A, Lercher L, Spicer CD, Davis BG. Designing logical codon reassignment - Expanding the chemistry in biology. Chem Sci 2015;6: 50–69.28553457
47. Xie J, Liu W, Schultz P. A genetically encoded bidentate, metal-binding amino acid. Angew Chem Int Ed 2007;46: 9239–9242.
48. Lee HS, Spraggon G, Schultz PG, Wang F. Genetic incorporation of a metal-ion chelating amino acid into proteins as a biophysical probe. J Am Chem Soc 2009;131: 2481–2483.19193005
49. Tippmann E, Schultz P. A genetically encoded metallocene containing amino acid. Tetrahedron 2007;63: 6182–6184.
50. Wang J, Xie J, Schultz P. A genetically encoded fluorescent amino acid. J Am Chem Soc 2006;128: 8738–8739.16819861
51. Wang L, Brock A, Schultz PG. Adding L-3-(2-Naphthyl)alanine to the genetic code of E. coli. J Am Chem Soc 2002;124: 1836–1837.11866580
52. Chatterjee A, Guo J, Lee HS, Schultz PG. A genetically encoded fluorescent probe in mammalian cells. J Am Chem Society 2013;135: 12540–12543.
53. Summerer D, Chen S, Wu N, Deiters A, Chin JW, Schultz PG. A genetically encoded fluorescent amino acid. Proc Natl Acad Sci U S A 2006;103: 9785–9789.16785423
54. Chin JW, Martin AB, King DS, Wang L, Schultz PG. Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. Proc Natl Acad Sci U S A 2002;99: 11020–11024.12154230
55. Lin S, He D, Long T, Zhang S, Meng R, Chen PR. Genetically encoded cleavable protein photo-cross-linker. J Am Chem Soc 2014;136: 11860–11863.25084056
56. Ai HW, Shen W, Sagi A, Chen PR, Schultz PG. Probing protein-protein interactions with a genetically encoded photo-crosslinking amino acid. Chembiochem 2011;12: 1854–1857.21678540
57. Tippmann E, Liu W, Summerer D, Mack A, Schultz P. A genetically encoded diazirine photocrosslinker in Escherichia coli. Chembiochem 2007;8: 2210–2214.18000916
58. Chou C, Uprety R, Davis L, Chin J, Deiters A. Genetically encoding an aliphatic diazirine for protein photocrosslinking. Chem Sci 2011;2: 480–483.
59. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. J Am Chem Soc 2002;124: 9026–9027.12148987
60. Neumann H, Hazen J, Weinstein J, Mehl R, Chin J. Genetically encoding protein oxidative damage. J Am Chem Soc 2008;130: 4028–4033.18321101

61. Seyedsayamdost M, Xie J, Chan C, Schultz P, Stubbe J. 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 2007;129: 15060–15071.17990884
62. Lv X, Yu Y, Zhou M, Ultrafast Photoinduced Electron Transfer in Green Fluorescent Protein Bearing a Genetically Encoded Electron Acceptor. J Am Chem Soc 2015;137: 7270–7273.26020364
63. Deiters A, Schultz PG. In vivo incorporation of an alkyne into proteins in Escherichia coli. Bioorg Med Chem Lett 2005;15: 1521–1524.15713420
64. Deiters A, Cropp TA, Summerer D, Mukherji M, Schultz PG. Site-specific PEGylation of proteins containing unnatural amino acids. Bioorg Med Chem Lett 2004;14: 5743–5745.15501033
65. Lang K, Davis L, Torres-Kolbus J, Chou C, Deiters A, Chin JW. Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. Nat Chem 2012;4: 298–304.22437715
66. Lee YJ, Kurra Y, Yang Y, Torres-Kolbus J, Deiters A, Liu WR. Genetically encoded unstrained olefins for live cell labeling with tetrazine dyes. Chem Commun (Camb) 2014;50(: 13085–13088.25224663
67. Lang K, Chin J. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. Chem Rev 2014;114: 4764–4806.24655057
68. Lang K, Davis L, Wallace S, 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 2012;134: 10317–10320.22694658
69. Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N(epsilon)-(o-Azidobenzyloxycarbonyl) lysine for Site-Specific Protein Modification. Chemi Biol 2008;15: 1187–1197.
70. Xiao H, Peters F, Yang P, Reed S, Chittuluru J, Schultz P. Genetic Incorporation of Histidine Derivatives Using an Engineered Pyrrolysyl-tRNA Synthetase. ACS Chem Biol 2014;9: 1092–1096.24506189
71. Xiao H, Xuan W, Shao S, Liu T, Schultz P. Genetic Incorporation of epsilon-N-2-Hydroxyisobutyryl-lysine into Recombinant Histones. ACS Chem Biol 2015;10: 1599–1603.25909834
72. Xie J, Supekova L, Schultz P. A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli. ACS Chem Biol 2007;2: 474–478.17622177
73. Xie J, Wang L, Wu N, Brock A, Spraggon G, Schultz PG. The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination. Nat Biotechnol 2004;22: 1297–1301.15378068
74. Rogerson D, Sachdeva A, Wang K. Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. Nat Chem Biol 2015;11: 496.26030730
75. Park H, Hohn M, Umehara T, Expanding the Genetic Code of Escherichia coli with Phosphoserine. Science 2011;333: 1151–1154.21868676
76. Groff D, Chen P, Peters F, Schultz P. A Genetically Encoded epsilon-N-Methyl Lysine in Mammalian Cells. Chembiochem 2010;11: 1066–1068.20422671
77. Kim C, Kang M, Kim H, Chatterjee A, Schultz P. Site-Specific Incorporation of epsilon-N-Crotonyllysine into Histones. Angew Chem Int Ed 2012;51: 7246–7249.
78. Tan Z, Forster A, Blacklow S, Cornish V. Amino acid backbone specificity of the Escherichia coli translation machinery. J Am Chem Soc 2004;126: 12752–12753.15469251
79. Guo J, Wang J, Anderson J, Schultz P. Addition of an alpha-hydroxy acid to the genetic code of bacteria. Angew Chem Int Ed 2008;47: 722–725.
80. Kobayashi T, Yanagisawa T, Sakamoto K, Yokoyama S. Recognition of Non-alpha-amino Substrates by Pyrrolysyl-tRNA Synthetase. J Mol Biol 2009;385: 1352–1360.19100747
81. Czekster C, Robertson W, Walker A, Soll D, Schepartz A. In Vivo Biosynthesis of a beta-Amino Acid-Containing Protein. J Am Chem Soc 2016;138: 5194–5197.27086674
82. Wang L, Schultz PG. A general approach for the generation of orthogonal tRNAs. Chem Biol 2001;8: 883–890.11564556

83. Wu N, Deiters A, Cropp TA, King D, Schultz PG. A Genetically Encoded Photocaged Amino Acid. J Am Chem Soc 2004;126: 14306–14307.15521721
84. Sakamoto K, Hayashi A, Sakamoto A, Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. Nucleic Acids Res 2002;30: 4692–4699.12409460
85. Hughes RA, Ellington AD. Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. Nucleic Acids Res 2010;38: 6813–6830.20571084
86. Italia JS, Addy PS, Wrobel CJ, An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. Nat Chem Biol 2017;13: 446–450.28192410
87. Chatterjee A, Xiao H, Schultz PG. 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 2012;109: 14841–14846.22927411
88. Santoro SW, Anderson JC, Lakshman V, Schultz PG. An archaebacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in Escherichia coli. Nucleic Acids Res 2003;31: 6700–6709.14627803
89. Luo X, Zambaldo C, Liu T, Recombinant thiopeptides containing noncanonical amino acids. Proc Natl Acad Sci U S A 2016;113: 3615–3620.26976568
90. Young TS, Ahmad I, Brock A, Schultz PG. Expanding the genetic repertoire of the methylotrophic yeast Pichia pastoris. Biochem 2009;48: 2643–2653.19265424
91. Greiss S, Chin JW. Expanding the Genetic Code of an Animal. J Am Chem Soc 2011;133: 14196–14199.21819153
92. Parrish AR, She X, Xiang Z, Expanding the genetic code of Caenorhabditis elegans using bacterial aminoacyl-tRNA synthetase/tRNA pairs. ACS Chem Biol 2012;7: 1292–1302.22554080
93. Bianco A, Townsley FM, Greiss S, Lang K, Chin JW. Expanding the genetic code of Drosophila melanogaster. Nat Chem Biol 2012;8: 748–750.22864544
94. Li F, Zhang H, Sun Y, Pan Y, Zhou J, Wang J. Expanding the genetic code for photoclick chemistry in E. coli, mammalian cells, and A. thaliana. Angew Chem Int Ed Engl 2013;52: 9700–9704.23873613
95. Liu J, Hemphill J, Samanta S, Tsang M, Deiters A. Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling. J Am Chem Soc 2017;139: 9100–9103.28657738
96. Ernst RJ, Krogager TP, Maywood ES, Genetic code expansion in the mouse brain. Nat Chem Biol 2016;12: 776–778.27571478
97. Elliott TS, Townsley FM, Bianco A, Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. Nat Biotechnol 2014;32: 465–472.24727715
98. Elliott TS, Bianco A, Townsley FM, Fried SD, Chin JW. Tagging and Enriching Proteins Enables Cell-Specific Proteomics. Cell Chem Biol 2016;23: 805–815.27447048
99. Stone SE, Glenn WS, Hamblin GD, Tirrell DA. Cell-selective proteomics for biological discovery. Curr Opin Chem Biol 2017;36: 50–57.28088696
100. Shen B, Xiang Z, Miller B, Genetically encoding unnatural amino acids in neural stem cells and optically reporting voltage-sensitive domain changes in differentiated neurons. Stem Cells 2011;29: 1231–1240.21681861
101. Wang W, Takimoto JK, Louie GV, Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat Neurosci 2007;10: 1063–1072.17603477
102. Kang JY, Kawaguchi D, Wang L. Optical Control of a Neuronal Protein Using a Genetically Encoded Unnatural Amino Acid in Neurons. J Vis Exp 2016: e53818.27078635
103. Chatterjee A, Xiao H, Bollong M, Ai HW, Schultz PG. Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells. Proc Natl Acad Sci U S A 2013;110: 11803–11808.23818609
104. Zheng Y, Lewis TL, Igo P, Polleux F, Chatterjee A. Virus-Enabled Optimization and Delivery of the Genetic Machinery for Efficient Unnatural Amino Acid Mutagenesis in Mammalian Cells and Tissues. ACS Synth Biol 2017;6: 13–18.27482719
105. Young TS, Ahmad I, Yin JA, Schultz PG. An enhanced system for unnatural amino acid mutagenesis in E. coli. J Mol Biol 2010;395: 361–374.19852970

106. Guo J, Melançon CE, Lee HS, Groff D, Schultz PG. Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. Angew Chem Int Ed Engl 2009;48: 9148–9151.19856359

107. Ryu Y, Schultz PG. Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. Nat Methods 2006;3: 263–265.16554830

108. Amiram M, Haimovich AD, Fan C, Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. Nat Biotechnol 2015;33: 1272–1279.26571098

109. Fan C, Xiong H, Reynolds NM, Söll D. Rationally evolving tRNAPyl for efficient incorporation of noncanonical amino acids. Nucleic Acids Res 2015;43: e156.26250114

110. Takimoto JK, Adams KL, Xiang Z, Wang L. Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells. Mol Biosyst 2009;5: 931–934.19668857

111. Tian F, Lu Y, Manibusan A, A general approach to site-specific antibody drug conjugates. Proc Natl Acad Sci U S A 2014;111: 1766–1771.24443552

112. Cho H, Daniel T, Buechler YJ, Optimized clinical performance of growth hormone with an expanded genetic code. Proc Natl Acad Sci U S A 2011;108: 9060–9065.21576502

113. Koehler C, Sauter PF, Wawryszyn M, Genetic code expansion for multiprotein complex engineering. Nat Methods 2016;13: 997–1000.27749839

114. Johnson DB, Xu J, Shen Z, RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol 2011;7: 779–786.21926996

115. Johnson DB, Wang C, Xu J, Release factor one is nonessential in Escherichia coli. ACS Chem Biol 2012;7: 1337–1344.22662873

116. Lajoie MJ, Rovner AJ, Goodman DB, Genomically recoded organisms expand biological functions. Science 2013;342: 357–360.24136966

117. Rackham O, Chin JW. A network of orthogonal ribosome x mRNA pairs. Nat Chem Biol 2005;1: 159–166.16408021

118. Orelle C, Carlson ED, Szal T, Florin T, Jewett MC, Mankin AS. Protein synthesis by ribosomes with tethered subunits. Nature 2015;524: 119–124.26222032

119. Wang K, Neumann H, Peak-Chew SY, Chin JW. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. Nat Biotechnol 2007;25: 770–777.17592474

120. Isaacs FJ, Carr PA, Wang HH, Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. Science 2011;333: 348–353.21764749

121. Mukai T, Hayashi A, Iraha F, Codon reassignment in the Escherichia coli genetic code. Nucleic Acids Res 2010;38: 8188–8195.20702426

122. Mukai T, Hoshi H, Ohtake K, Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. Sci Rep 2015;5: 9699.25982672

123. Wan W, Huang Y, Wang Z, A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli. Angew Chem Int Ed Engl 2010;49: 3211–3214.20340150

124. Ostrov N, Landon M, Guell M, Design, synthesis, and testing toward a 57-codon genome. Science 2016;353: 819–822.27540174

125. Chin JW. Molecular biology. Reprogramming the genetic code. Science 2012;336: 428–429.22539711

126. Wang K, Sachdeva A, Cox DJ, Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. Nat Chem 2014;6: 393–403.24755590

127. Malyshev DA, Dhami K, Lavergne T, A semi-synthetic organism with an expanded genetic alphabet. Nature 2014;509: 385–388.24805238

128. Malyshev DA, Romesberg FE. The expanded genetic alphabet. Angew Chem Int Ed Engl 2015;54: 11930–11944.26304162

129. Zhang Y, Lamb BM, Feldman AW, A semisynthetic organism engineered for the stable expansion of the genetic alphabet. Proc Natl Acad Sci U S A 2017;114: 1317–1322.28115716

130. Zeng Y, Wang W, Liu WR. Towards reassigning the rare AGG codon in Escherichia coli. Chembiochem 2014;15: 1750–1754.25044341
131. Rovner AJ, Haimovich AD, Katz SR, Recoded organisms engineered to depend on synthetic amino acids. Nature 2015;518: 89–93.25607356
132. Neumann H, Slusarczyk AL, Chin JW. De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. J Am Chem Soc 2010;132: 2142–2144.20121121
133. Sachdeva A, Wang K, Elliott T, Chin JW. Concerted, rapid, quantitative, and site-specific dual labeling of proteins. J Am Chem Soc 2014;136: 7785–7788.24857040
134. Zheng Y, Addy PS, Mukherjee R, Chatterjee A. Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cells. Chem Sci 2017;8: 7211–7217.29081953
135. Zheng Y, Mukherjee R, Chin MA, Igo P, Gilgenast MJ, Chatterjee A. Expanding the Scope of Single- and Double-Noncanonical Amino Acid Mutagenesis in Mammalian Cells Using Orthogonal Polyspecific Leucyl-tRNA Synthetases. Biochem 2017.
136. Xiao H, Chatterjee A, Choi SH, Bajjuri KM, Sinha SC, Schultz PG. Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. Angew Chem Int Ed Engl 2013;52: 14080–14083.24353230
137. Prescher JA, Bertozzi CR. Chemistry in living systems. Nat Chem Biol 2005;1: 13–21.16407987
138. Matsuda S, Fillo JD, Henry AA, Efforts toward expansion of the genetic alphabet: structure and replication of unnatural base pairs. J Am Chem Soc 2007;129: 10466–10473.17685517
139. Shogren-Knaak MA, Alaimo PJ, Shokat KM. Recent advances in chemical approaches to the study of biological systems. Annu Rev Cell Dev Biol 2001;17: 405–433.11687495
140. Yee CS, Seyedsayamdost MR, Chang MCY, Nocera DG, Stubbe J. 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. Biochem 2003;42: 14541–14552.14661967
141. Seyedsayamdost MR, Stubbe J. Site-specific replacement of Y_{356} with 3,4-dihydroxyphenylalanine in the β2 subunit of E. *coli* ribonucleotide reductase. J Am Chem Soc 2006;128: 2522–2523.
142. Seyedsayamdost MR, Yee CS, Stubbe J. Site-specific incorporation of fluorotyrosines into the R2 subunit of *E. coli* ribonucleotide reductase by expressed protein ligation. Nat Protoc 2007;2: 1225–1235.17546018
143. Minnihan EC, Seyedsayamdost MR, Uhlin U, Stubbe J. Kinetics of radical intermediate formation and deoxynucleotide production in 3-aminotyrosine-substituted Escherichia coli ribonucleotide reductases. J Am Chem Soc 2011;133: 9430–9440.21612216
144. Liu X, Li J, Dong J, Hu C, Gong W, Wang J. Genetic incorporation of a metal-chelating amino acid as a probe for protein electron transfer. Angew Chem Int Ed Engl 2012;51: 10261–10265.22936654
145. Xiao H, Peters FB, Yang PY, Reed S, Chittuluru JR, Schultz PG. Genetic incorporation of histidine derivatives using an engineered pyrrolysyl-tRNA synthetase. ACS Chem Biol 2014;9: 1092–1096.24506189
146. Wilkins B, Marionni S, Young D, Site-Specific Incorporation of Fluorotyrosines into Proteins in Escherichia coli by Photochemical Disguise. Biochem 2010;49: 1557–1559.20136096
147. Oyala PH, Ravichandran KR, Funk MA, Biophysical Characterization of Fluorotyrosine Probes Site-Specifically Incorporated into Enzymes: E. coli Ribonucleotide Reductase As an Example. J Am Chem Soc 2016;138: 7951–7964.27276098
148. Xiang Z, Lacey V, Ren H, Proximity-Enabled Protein Crosslinking through Genetically Encoding Haloalkane Unnatural Amino Acids. Angew Chem Int Ed 2014;53: 2190–2193.
149. Kobayashi T, Hoppmann C, Yang B, Wang L. Using Protein-Confined Proximity To Determine Chemical Reactivity. J Am Chem Soc 2016;138: 14832–14835.27797495
150. Chen XH, Xiang Z, Hu YS, Lacey VK, Cang H, Wang L. Genetically encoding an electrophilic amino acid for protein stapling and covalent binding to native receptors. ACS Chem Biol 2014;9: 1956–1961.25010185

151. Xuan W, Shao S, Schultz PG. Protein Crosslinking by Genetically Encoded Noncanonical Amino Acids with Reactive Aryl Carbamate Side Chains. Angew Chem Int Ed Engl 2017;56: 5096–5100.28371162
152. Xuan W, Li J, Luo X, Schultz PG. Genetic Incorporation of a Reactive Isothiocyanate Group into Proteins. Angew Chem Int Ed Engl 2016;55: 10065–10068.27418387
153. Yang Y, Song H, He D, Genetically encoded protein photocrosslinker with a transferable mass spectrometry-identifiable label. Nat Commun 2016;7: 12299.27460181
154. Freinkman E, Chng SS, Kahne D. The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. Proc Natl Acad Sci U S A 2011;108: 2486–2491.21257904
155. Wilkins BJ, Rall NA, Ostwal Y, A cascade of histone modifications induces chromatin condensation in mitosis. Science 2014;343: 77–80.24385627
156. Joiner CM, Breen ME, Clayton J, Mapp AK. A Bifunctional Amino Acid Enables Both Covalent Chemical Capture and Isolation of in Vivo Protein-Protein Interactions. Chembiochem 2017;18: 181–184.27966261
157. Xiang Z, Ren H, Hu YS, Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. Nat Methods 2013;10: 885–888.23913257
158. Coin I, Katritch V, Sun T, Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR. Cell 2013;155: 1258–1269.24290358
159. Grunbeck A, Huber T, Sachdev P, Sakmar TP. Mapping the ligand-binding site on a G protein-coupled receptor (GPCR) using genetically encoded photocrosslinkers. Biochem 2011;50: 3411–3413.21417335
160. Zhang M, Lin S, Song X, A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. Nat Chem Biol 2011;7: 671–677.21892184
161. Rannversson H, Andersen J, Sørensen L, Genetically encoded photocrosslinkers locate the high-affinity binding site of antidepressant drugs in the human serotonin transporter. Nat Commun 2016;7: 11261.27089947
162. Hoppmann C, Wang L. Proximity-enabled bioreactivity to generate covalent peptide inhibitors of p53-Mdm4. Chem Commun (Camb) 2016;52(29): 5140–5143.26996321
163. Luo J, Uprety R, Naro Y, Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. J Am Chem Soc 2014;136: 15551–15558.25341086
164. Lampkowski JS, Uthappa DM, Young DD. Site-specific incorporation of a fluorescent terphenyl unnatural amino acid. Bioorg Med Chem Lett 2015;25: 5277–5280.26421994
165. Lee H, Guo J, Lemke E, Dimla R, Schultz P. Genetic Incorporation of a Small, Environmentally Sensitive, Fluorescent Probe into Proteins in Saccharomyces cerevisiae. J Am Chem Soc 2009;131: 12921.19702307
166. Lee HS, Guo J, Lemke EA, Dimla RD, Schultz PG. Genetic incorporation of a small, environmentally sensitive, fluorescent probe into proteins in Saccharomyces cerevisiae. J Am Chem Soc 2009;131: 12921–12923.19702307
167. Lacey VK, Parrish AR, Han S, A fluorescent reporter of the phosphorylation status of the substrate protein STAT3. Angew Chem Int Ed Engl 2011;50: 8692–8696.21805546
168. Mitchell AL, Addy PS, Chin MA, Chatterjee A. A Unique Genetically Encoded FRET Pair in Mammalian Cells. Chembiochem 2017;18: 511–514.28093840
169. Kalstrup T, Blunck R. Dynamics of internal pore opening in K(V) channels probed by a fluorescent unnatural amino acid. Proc Natl Acad Sci U S A 2013;110: 8272–8277.23630265
170. Groff D, Thielges MC, Cellitti S, Schultz PG, Romesberg FE. Efforts toward the direct experimental characterization of enzyme microenvironments: tyrosine100 in dihydrofolate reductase. Angew Chem Int Ed Engl 2009;48: 3478–3481.19347910
171. Schultz K, Supekova L, Ryu Y, Xie J, Perera R, Schultz P. A genetically encoded infrared probe. J Am Chem Soc 2006;128: 13984–13985.17061854
172. Smith EE, Linderman BY, Luskin AC, Brewer SH. Probing local environments with the infrared probe: L-4-nitrophenylalanine. J Phys Chem B 2011;115: 2380–2385.21344930

173. Taskent-Sezgin H, Chung J, Patsalo V, 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. Biochem 2009;48: 9040–9046.19658436
174. Basom EJ, Maj M, Cho M, Thielges MC. Site-Specific Characterization of Cytochrome P450cam Conformations by Infrared Spectroscopy. Anal Chem 2016;88: 6598–6606.27185328
175. Ye S, Huber T, Vogel R, Sakmar TP. FTIR analysis of GPCR activation using azido probes. Nat Chem Biol 2009;5: 397–399.19396177
176. Ye S, Zaitseva E, Caltabiano G, Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. Nature 2010;464: 1386–1389.20383122
177. Fleissner MR, Brustad EM, Kálai T, Site-directed spin labeling of a genetically encoded unnatural amino acid. Proc Natl Acad Sci U S A 2009;106: 21637–21642.19995976
178. Schmidt MJ, Borbas J, Drescher M, Summerer D. A genetically encoded spin label for electron paramagnetic resonance distance measurements. J Am Chem Soc 2014;136: 1238–1241.24428347
179. Schmidt MJ, Fedoseev A, Bücker D, EPR Distance Measurements in Native Proteins with Genetically Encoded Spin Labels. ACS Chem Biol 2015;10: 2764–2771.26421438
180. Cellitti SE, Jones DH, Lagpacan L, 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 2008;130: 9268–9281.18576636
181. Jackson JC, Hammill JT, Mehl RA. 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 2007;129: 1160–1166.17263397
182. Lampe JN, Floor SN, Gross JD, 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 2008;130: 16168–16169.18998650
183. Yang F, Yu X, Liu C, Phospho-selective mechanisms of arrestin conformations and functions revealed by unnatural amino acid incorporation and (19)F-NMR. Nat Commun 2015;6: 8202.26347956
184. Shi P, Wang H, Xi Z, Shi C, Xiong Y, Tian C. Site-specific F NMR chemical shift and side chain relaxation analysis of a membrane protein labeled with an unnatural amino acid. Protein Sci 2011;20: 224–228.21080424
185. Shi P, Xi Z, Wang H, Shi C, Xiong Y, Tian C. 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 2010;402: 461–466.20946873
186. Niu W, Guo J. Expanding the chemistry of fluorescent protein biosensors through genetic incorporation of unnatural amino acids. Mol Biosys 2013;9: 2961–2970.
187. Ayyadurai N, Saravanan Prabhu N, Deepankumar K,. 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 Engl 2011;50: 6534–6537.21656613
188. Liu X, Li J, Hu C, Significant expansion of the fluorescent protein chromophore through the genetic incorporation of a metal-chelating unnatural amino acid. Angew Chem Int Ed Engl 2013;52: 4805–4809.23554162
189. Shang X, Wang N, Cerny R, Niu W, Guo J. Fluorescent Protein-Based Turn-On Probe through a General Protection-Deprotection Design Strategy. ACS Sens 2017;2: 961–966.28750537
190. Wang F, Niu W, Guo J, Schultz PG. Unnatural amino acid mutagenesis of fluorescent proteins. Angew Chem Int Ed Engl 2012;51: 10132–10135.22951916
191. Chen S, Chen ZJ, Ren W, Ai HW. Reaction-based genetically encoded fluorescent hydrogen sulfide sensors. J Am Chem Soc 2012;134: 9589–9592.22642566
192. Wang WW, Zeng Y, Wu B, Deiters A, Liu WR. A Chemical Biology Approach to Reveal Sirt6-targeted Histone H3 Sites in Nucleosomes. ACS Chem Biol 2016;11: 1973–1981.27152839
193. Chen ZJ, Tian Z, Kallio K, The N-B Interaction through a Water Bridge: Understanding the Chemoselectivity of a Fluorescent Protein Based Probe for Peroxynitrite. J Am Chem Soc 2016;138: 4900–4907.27019313

194. Chen ZJ, Ren W, Wright QE, Ai HW. Genetically encoded fluorescent probe for the selective detection of peroxynitrite. J Am Chem Soc 2013;135: 14940–14943.24059533

195. Maza J, Howard C, Vipani M, Travis C, Young D. Utilization of alkyne bioconjugations to modulate protein function. Bioorg Med Chem Lett 2017;27: 30–33.27894869

196. Padilla MS, Young DD. Photosensitive GFP mutants containing an azobenzene unnatural amino acid. Bioorg Med Chem Lett 2015;25: 470–473.25563892

197. Grinstead KM, Rowe L, Ensor CM, Red-Shifted Aequorin Variants Incorporating Non-Canonical Amino Acids: Applications in In Vivo Imaging. PLoS One 2016;11: e0158579.27367859

198. Villa JK, Tran HA, Vipani M, Fluorescence Modulation of Green Fluorescent Protein Using Fluorinated Unnatural Amino Acids. Molecules 2017;22.

199. Khare P, Jain A, Gulbake A, Soni V, Jain N, Jain S. Bioconjugates: Harnessing Potential for Effective Therapeutics. Crit Rev Therap Drug Carr Sys 2009;26: 119–155.

200. Wang L, Zhang Z, Brock A, Schultz PG. Addition of the keto functional group to the genetic code of Escherichia coli. Proc Natl Acad Sci U S A 2003;100: 56–61.12518054

201. Hancock SM, Uprety R, Deiters A, Chin JW. Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. J Am Chem Soc 2010;132: 14819–14824.20925334

202. Maza JC, McKenna JR, Raliski BK, Freedman MT, Young DD. Synthesis and Incorporation of Unnatural Amino Acids To Probe and Optimize Protein Bioconjugations. Bioconjug Chem 2015; 26: 1884–1889.26287719

203. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG. Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. J Am Chem Soc. 2003;125: 3192–3193.12630856

204. Lang K, Chin JW. Bioorthogonal reactions for labeling proteins. ACS Chem Biol 2014;9: 16–20.24432752

205. Plass T, Milles S, Koehler C, Schultz C, Lemke EA. Genetically encoded copper-free click chemistry. Angew Chem Int Ed Engl 2011;50: 3878–3881.21433234

206. Lang K, Davis L, Wallace S, 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 2012;134: 10317–10320.22694658

207. Baskin JM, Prescher JA, Laughlin ST, Copper-free click chemistry for dynamic in vivo imaging. Proc Natl Acad Sci U S A 2007;104: 16793–16797.17942682

208. Uttamapinant C, Howe JD, Lang K, Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. J Am Chem Soc 2015;137: 4602–4605.25831022

209. Xue L, Prifti E, Johnsson K. A General Strategy for the Semisynthesis of Ratiometric Fluorescent Sensor Proteins with Increased Dynamic Range. J Am Chem Soc 2016;138: 5258–5261.27071001

210. Kumaresan P, Luo J, Song A, Marik J, Lam K. Evaluation of ketone-oxime method for developing therapeutic on-demand cleavable immunoconjugates. Bioconjug Chem 2008;19: 1313–1318.18494516

211. Brustad E, Lemke E, Schultz P, Deniz A. A General and Efficient Method for the Site-Specific Dual-Labeling of Proteins for Single Molecule Fluorescence Resonance Energy Transfer. J Am Chem Soc 2008;130: 17664.19108697

212. Axup JY, Bajjuri KM, Ritland M, Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. Proc Natl Acad Sci U S A 2012;109: 16101–16106.22988081

213. Hutchins B, Kazane S, Staflin K, Selective Formation of Covalent Protein Heterodimers with an Unnatural Amino Acid. Chem Biol 2011;18: 299–303.21439474

214. Hutchins B, Kazane S, Staflin K, Site-Specific Coupling and Sterically Controlled Formation of Multimeric Antibody Fab Fragments with Unnatural Amino Acids. J Mol Biol 2011;406: 595–603.21237172

215. Kazane SA, Sok D, Cho EH, Site-specific DNA-antibody conjugates for specific and sensitive immuno-PCR. Proc Natl Acad Sci U S A 2012;109: 3731–3736.22345566

216. Kim CH, Axup JY, Dubrovska A, Synthesis of bispecific antibodies using genetically encoded unnatural amino acids. J Am Chem Soc 2012;134: 9918–9921.22642368
217. Kazane SA, Axup JY, Kim CH, Self-assembled antibody multimers through peptide nucleic acid conjugation. J Am Chem Soc 2013;135: 340–346.23210862
218. Lu H, Wang D, Kazane S, Site-specific antibody-polymer conjugates for siRNA delivery. J Am Chem Soc 2013;135: 13885–13891.23924037
219. Ai HW, Shen W, Brustad E, Schultz PG. Genetically encoded alkenes in yeast. Angew Chem Int Ed Engl 2010;49: 935–937.20025009
220. Lin YA, Chalker JM, Floyd N, Bernardes GJ, Davis BG. Allyl sulfides are privileged substrates in aqueous cross-metathesis: application to site-selective protein modification. J Am Chem Soc 2008;130: 9642–9643.18593118
221. Addy PS, Erickson SB, Italia JS, Chatterjee A. A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for Unclickable Bioconjugation. J Am Chem Soc 2017;139: 11670–11673.28787141
222. Row RD, Shih HW, Alexander AT, Mehl RA, Prescher JA. Cyclopropenones for Metabolic Targeting and Sequential Bioorthogonal Labeling. J Am Chem Soc 2017;139: 7370–7375.28478678
223. Wang ZA, Kurra Y, Wang X, A Versatile Approach for Site-Specific Lysine Acylation in Proteins. Angew Chem Int Ed Engl 2017;56: 1643–1647.28042700
224. Li J, Lin S, Wang J, Ligand-free palladium-mediated site-specific protein labeling inside gram-negative bacterial pathogens. J Am Chem Soc 2013;135: 7330–7338.23641876
225. Li N, Lim R, Edwardraja S, Lin Q. Copper-Free Sonogashira Cross-Coupling for Functionalization of Alkyne-Encoded Proteins in Aqueous Medium and in Bacterial Cells. J Am Chem Soc 2011;133: 15316–15319.21899368
226. Lampkowski JS, Villa JK, Young TS, Young DD. Development and Optimization of Glaser-Hay Bioconjugations. Angew Chem Int Ed Engl 2015;54: 9343–9346.26088803
227. Maza JC, Nimmo ZM, Young DD. Expanding the scope of alkyne-mediated bioconjugations utilizing unnatural amino acids. Chem Commun (Camb) 2016;52: 88–91.26499242
228. Spicer CD, Davis BG. Palladium-mediated site-selective Suzuki-Miyaura protein modification at genetically encoded aryl halides. Chem Commun (Camb) 2011;47: 1698–1700.21206952
229. Spicer CD, Triemer T, Davis BG. Palladium-mediated cell-surface labeling. J Am Chem Soc 2012;134: 800–803.22175226
230. Dumas A, Spicer CD, Gao Z, Self-liganded Suzuki-Miyaura coupling for site-selective protein PEGylation. Angew Chem Int Ed Engl 2013;52: 3916–3921.23440916
231. Spicer CD, Davis BG. Rewriting the bacterial glycocalyx via Suzuki-Miyaura cross-coupling. Chem Commun (Camb) 2013;49: 2747–2749.23338477
232. Raliski BK, Howard CA, Young DD. Site-specific protein immobilization using unnatural amino acids. Bioconjug Chem 2014;25: 1916–1920.25340826
233. Borrmann A, Milles S, Plass T, Genetic Encoding of a Bicyclo[6.1.0]nonyne-Charged Amino Acid Enables Fast Cellular Protein Imaging by Metal-Free Ligation. Chembiochem 2012;13: 2094–2099.22945333
234. Plass T, Milles S, Koehler C, Amino Acids for Diels-Alder Reactions in Living Cells. Angew Chem Int Ed 2012;51: 4166–4170.
235. Seitchik JL, Peeler JC, Taylor MT, Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific in Vivo Bioorthogonal Ligation with trans-Cyclooctenes. J Am Chem Soc 2012;134: 2898–2901.22283158
236. Blizzard RJ, Backus DR, Brown W, Bazewicz CG, Li Y, Mehl RA. Ideal Bioorthogonal Reactions Using A Site-Specifically Encoded Tetrazine Amino Acid. J Am Chem Soc 2015;137: 10044–10047.26237426
237. Lukinavicius G, Umezawa K, Olivier N, A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. Nat Chem 2013;5: 132–139.23344448
238. Peng T, Hang HC. Site-Specific Bioorthogonal Labeling for Fluorescence Imaging of Intracellular Proteins in Living Cells. J Am Chem Soc 2016;138: 14423–14433.27768298

239. Baumdick M, Bruggemann Y, Schmick M, EGF-dependent re-routing of vesicular recycling switches spontaneous phosphorylation suppression to EGFR signaling. Elife 2015;4.

240. Erickson SB, Mukherjee R, Kelemen RE, Wrobel CJ, Cao X, Chatterjee A. Precise Photoremovable Perturbation of a Virus-Host Interaction. Angew Chem Int Ed Engl 2017;56: 4234–4237.28294501

241. Rusmini F, Zhong Z, Feijen J. Protein immobilization strategies for protein biochips. Biomacromol 2007;8: 1775–1789.

242. Deepankumar K, Nadarajan S, Mathew S, Engineering Transaminase for Stability Enhancement and Site-Specific Immobilization through Multiple Noncanonical Amino Acids Incorporation. Chemcatchem 2015;7: 417–421.

243. Lim SI, Mizuta Y, Takasu A, Kim YH, Kwon I. Site-specific bioconjugation of a murine dihydrofolate reductase enzyme by copper(I)-catalyzed azide-alkyne cycloaddition with retained activity. PLoS One 2014;9: e98403.24887377

244. Ikeda-Boku A, Kondo K, Ohno S, Protein fishing using magnetic nanobeads containing calmodulin site-specifically immobilized via an azido group. J Biochem 2013;154: 159–165.23653405

245. Yoshimura SH, Khan S, Ohno S, Site-specific attachment of a protein to a carbon nanotube end without loss of protein function. Bioconjug Chem 2012;23: 1488–1493.22693944

246. Deiters A, Groff D, Ryu Y, Xie J, Schultz P. A genetically encoded photocaged tyrosine. Angew Chem Int Ed 2006;45: 2728–2731.

247. Wu N, Deiters A, Cropp TA, King D, Schultz PG. A genetically encoded photocaged amino acid. J Am Chem Soc 2004;126: 14306–14307.15521721

248. Lemke E, Summerer D, Geierstanger B, Brittain S, Schultz P. Control of protein phosphorylation with a genetically encoded photocaged amino acid. Nat Chem Biol 2007;3: 769–772.17965709

249. Gautier A, Nguyen D, Lusic H, An W, Deiters A, Chin J. Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells. J Am Chem Soc 2010;132: 4086.20218600

250. Edwards W, Young D, Deiters A. Light-Activated Cre Recombinase as a Tool for the Spatial and Temporal Control of Gene Function in Mammalian Cells. ACS Chem Biol 2009;4: 441–445.19413301

251. Gramespacher JA, Stevens AJ, Nguyen DP, Chin JW, Muir TW. Intein Zymogens: Conditional Assembly and Splicing of Split Inteins via Targeted Proteolysis. J Am Chem Soc 2017;139: 8074–8077.28562027

252. Ren W, Ji A, Ai HW. Light activation of protein splicing with a photocaged fast intein. J Am Chem Soc 2015;137: 2155–2158.25647354

253. Arbely E, Torres-Kolbus J, Deiters A, Chin JW. Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. J Am Chem Soc 2012;134: 11912–11915.22758385

254. Nguyen D, Mahesh M, Elsasser S, Hancock S, Uttamapinant C, Chin J. Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. J Am Chem Soc 2014;136: 2240–2243.24479649

255. Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW. Genetically encoded photocontrol of protein localization in mammalian cells. J Am Chem Soc 2010;132: 4086–4088.20218600

256. Hemphill J, Borchardt EK, Brown K, Asokan A, Deiters A. Optical Control of CRISPR/Cas9 Gene Editing. J Am Chem Soc 2015;137: 5642–5645.25905628

257. Lemke EA, Summerer D, Geierstanger BH, Brittain SM, Schultz PG. Control of protein phosphorylation with a genetically encoded photocaged amino acid. Nat Chem Biol 2007;3: 769–772.17965709

258. Walker OS, Elsässer SJ, Mahesh M, Bachman M, Balasubramanian S, Chin JW. Photoactivation of Mutant Isocitrate Dehydrogenase 2 Reveals Rapid Cancer-Associated Metabolic and Epigenetic Changes. J Am Chem Soc 2016;138: 718–721.26761588

259. Tsai YH, Essig S, James JR, Lang K, Chin JW. Selective, rapid and optically switchable regulation of protein function in live mammalian cells. Nat Chem 2015;7: 554–561.26100803

260. Li J, Jia S, Chen PR. Diels-Alder reaction-triggered bioorthogonal protein decaging in living cells. Nat Chem Biol 2014;10: 1003–1005.25362360

261. Li J, Yu J, Zhao J, Palladium-triggered deprotection chemistry for protein activation in living cells. Nat Chem 2014;6: 352–361.24651204
262. Luo J, Liu Q, Morihiro K, Deiters A. Small-molecule control of protein function through Staudinger reduction. Nat Chem 2016;8: 1027–1034.27768095
263. Heinemann IU, Rovner AJ, Aerni HR, Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. FEBS Lett 2012;586: 3716–3722.22982858
264. Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in Escherichia coli. Nat Biotechnol 2006;24: 1436–1440.17072302
265. Liu CC, Cellitti SE, Geierstanger BH, Schultz PG. Efficient expression of tyrosine-sulfated proteins in E. coli using an expanded genetic code. Nat Protoc 2009;4: 1784–1789.20010929
266. Knight W, Cropp T. Genetic encoding of the post-translational modification 2-hydroxyisobutyryl-lysine. Org Biomol Chem 2015;13: 6479–6481.25999185
267. Serwa R, Wilkening I, Del Signore G, Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. Angew Chem Int Ed Engl 2009;48: 8234–8239.19637176
268. Rust HL, Subramanian V, West GM, Young DD, Schultz PG, Thompson PR. Using unnatural amino acid mutagenesis to probe the regulation of PRMT1. ACS Chem Biol 2014;9: 649–655.24358983
269. Luo X, Fu G, Wang R, Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. Nat Chem Biol 2017;13: 845.28604693
270. Castañeda CA, Dixon EK, Walker O, Linkage via K27 Bestows Ubiquitin Chains with Unique Properties among Polyubiquitins. Structure 2016;24: 423–436.26876099
271. Zambaldo C, Luo X, Mehta AP, Schultz PG. Recombinant Macrocyclic Lanthipeptides Incorporating Non-Canonical Amino Acids. J Am Chem Soc 2017;139: 11646–11649.28809560
272. Bionda N, Cryan AL, Fasan R. Bioinspired strategy for the ribosomal synthesis of thioether-bridged macrocyclic peptides in bacteria. ACS Chem Biol 2014;9: 2008–2013.25079213
273. Smith JM, Vitali F, Archer SA, Fasan R. Modular assembly of macrocyclic organo-peptide hybrids using synthetic and genetically encoded precursors. Angew Chem Int Ed Engl 2011;50: 5075–5080.21506227
274. Frost JR, Vitali F, Jacob NT, Brown MD, Fasan R. Macrocyclization of organo-peptide hybrids through a dual bio-orthogonal ligation: insights from structure-reactivity studies. Chembiochem 2013;14: 147–160.23203912
275. Bindman NA, Bobeica SC, Liu WR, van der Donk WA. Facile Removal of Leader Peptides from Lanthipeptides by Incorporation of a Hydroxy Acid. J Am Chem Soc 2015;137: 6975–6978.26006047
276. Xiao H, Schultz PG. At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. Cold Spring Harb Perspect Biol 2016;8.
277. Kelemen RE, Mukherjee R, Cao X, Erickson SB, Zheng Y, Chatterjee A. A Precise Chemical Strategy To Alter the Receptor Specificity of the Adeno-Associated Virus. Angew Chem Int Ed Engl 2016;55: 10645–10649.27483453
278. VanBrunt MP, Shanebeck K, Caldwell Z, Genetically Encoded Azide Containing Amino Acid in Mammalian Cells Enables Site-Specific Antibody-Drug Conjugates Using Click Cycloaddition Chemistry. Bioconjug Chem 2015;26: 2249–2260.26332743
279. Mu J, Pinkstaff J, Li Z, FGF21 analogs of sustained action enabled by orthogonal biosynthesis demonstrate enhanced antidiabetic pharmacology in rodents. Diabetes 2012;61: 505–512.22210323
280. Cho H, Daniel T, Buechler YJ, Optimized clinical performance of growth hormone with an expanded genetic code. Proc Natl Acad Sci U S A 2011;108: 9060–9065.21576502
281. Kim CH, Axup JY, Schultz PG. Protein conjugation with genetically encoded unnatural amino acids. Curr Opin Chem Biol 2013;17: 412–419.23664497
282. Sun SB, Schultz PG, Kim CH. Therapeutic applications of an expanded genetic code. Chembiochem 2014;15: 1721–1729.25044800

283. Lu H, Wang DL, Kazane S, Site-Specific Antibody-Polymer Conjugates for siRNA Delivery. J Am Chem Soc 2013;135: 13885–13891.23924037

284. Kularatne SA, Deshmukh V, Gymnopoulos M, Recruiting cytotoxic T cells to folate-receptor-positive cancer cells. Angew Chem Int Ed Engl 2013;52: 12101–12104.24573789

285. Ma JS, Kim JY, Kazane SA, Versatile strategy for controlling the specificity and activity of engineered T cells. Proc Natl Acad Sci U S A 2016;113: E450–458.26759368

286. Gauba V, Grunewald J, Gorney V, Loss of CD4 T-cell-dependent tolerance to proteins with modified amino acids. Proc Natl Acad Sci U S A 2011;108: 12821–12826.21768354

287. Gruenewald J, Tsao ML, Perera R, Immunochemical termination of self-tolerance. Proc Natl Acad Sci U S A 2008;105: 11276–11280.18685087

288. Kim CH, Axup JY, Lawson BR, Bispecific small molecule-antibody conjugate targeting prostate cancer. Proc Natl Acad Sci U S A 2013;110: 17796–17801.24127589

289. Wang N, Li Y, Niu W, Construction of a live-attenuated HIV-1 vaccine through genetic code expansion. Angew Chem Int Ed Engl 2014;53: 4867–4871.24715496

290. Si L, Xu H, Zhou X, Generation of influenza A viruses as live but replication-incompetent virus vaccines. Science 2016;354: 1170–1173.27934767

291. Lin S, Yan H, Li L, Site-specific engineering of chemical functionalities on the surface of live hepatitis D virus. Angew Chem Int Ed Engl 2013;52: 13970–13974.24222614

292. Mills JH, Khare SD, Bolduc JM, Computational design of an unnatural amino acid dependent metalloprotein with atomic level accuracy. J Am Chem Soc 2013;135: 13393–13399.23924187

293. Mills JH, Sheffler W, Ener ME, Computational design of a homotrimeric metalloprotein with a trisbipyridyl core. Proc Natl Acad Sci U S A 2016;113: 15012–15017.27940918

294. Pearson AD, Mills JH, Song Y, Transition states. Trapping a transition state in a computationally designed protein bottle. Science 2015;347: 863–867.25700516

295. Liu CC, Mack AV, Brustad EM, Evolution of proteins with genetically encoded “chemical warheads”. J Am Chem Soc 2009;131: 9616–9617.19555063

296. Kang M, Light K, Ai HW, Evolution of Iron(II)-Finger Peptides by Using a Bipyridyl Amino Acid. Chembiochem 2014;15: 822–825.24591102

297. Young TS, Young DD, Ahmad I, Louis JM, Benkovic SJ, Schultz PG. Evolution of cyclic peptide protease inhibitors. Proc Natl Acad Sci U S A 2011;108: 11052–11056.21690365

298. Xiao H, Nasertorabi F, Choi SH, Exploring the potential impact of an expanded genetic code on protein function. Proc Natl Acad Sci U S A 2015;112: 6961–6966.26038548

299. Tack DS, Ellefson JW, Thyer R, Addicting diverse bacteria to a noncanonical amino acid. Nat Chem Biol 2016;12: 138–140.26780407

300. Ohtake K, Yamaguchi A, Mukai T, Protein stabilization utilizing a redefined codon. Sci Rep 2015;5: 9762.25985257

301. Hammerling MJ, Ellefson JW, Boutz DR, Marcotte EM, Ellington AD, Barrick JE. Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness. Nat Chem Biol 2014;10: 178–180.24487692

302. Mandell DJ, Lajoie MJ, Mee MT, Biocontainment of genetically modified organisms by synthetic protein design. Nature 2015;518: 55–60.25607366

303. Koh M, Nasertorabi F, Han GW, Stevens RC, Schultz PG. Generation of an Orthogonal Protein-Protein Interface with a Noncanonical Amino Acid. J Am Chem Soc 2017;139: 5728–5731.28413876

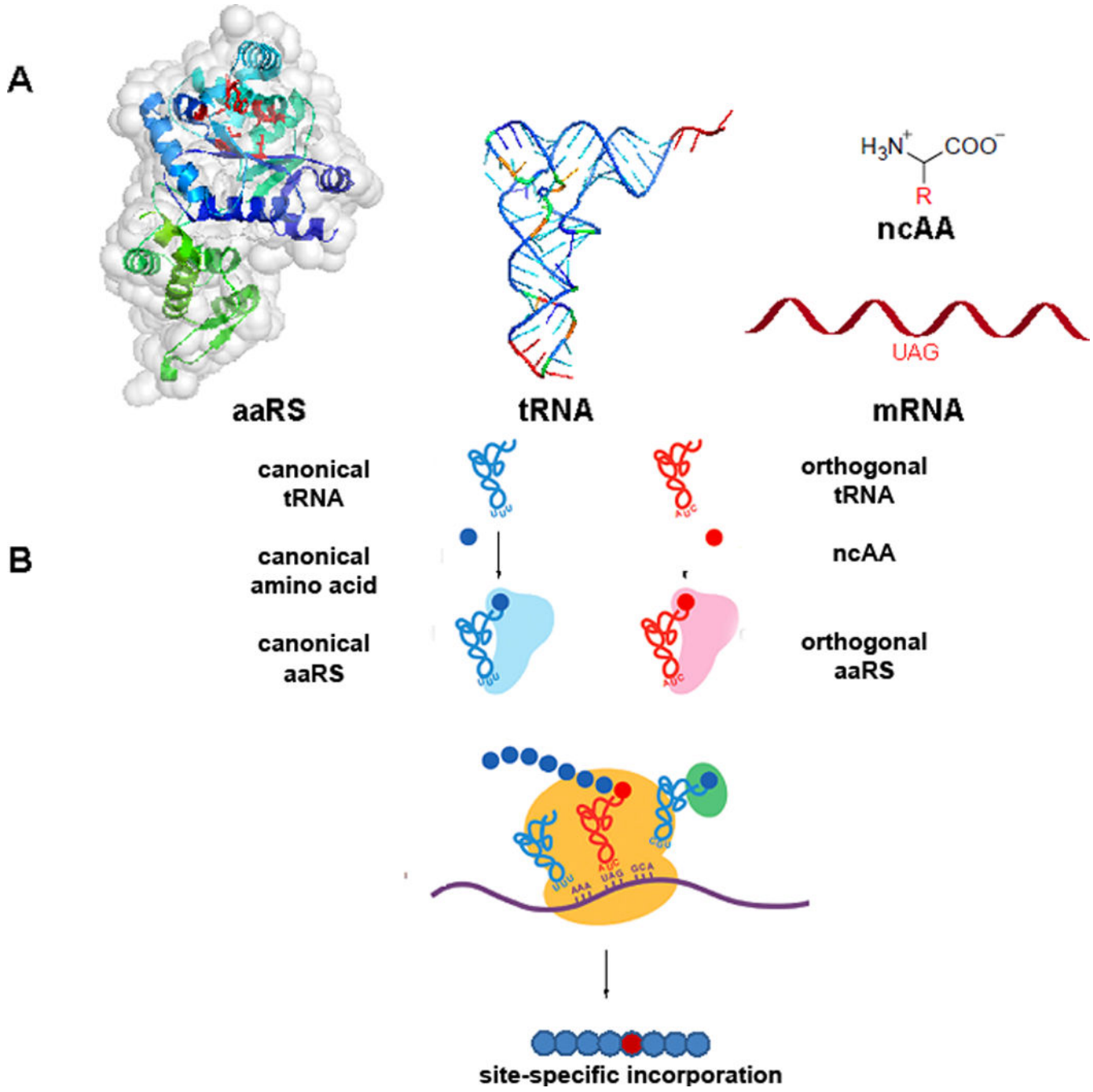


Figure 1.

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 on the mRNA (purple). Adapted from Kim, C.H. *et. al* 2013.[20]

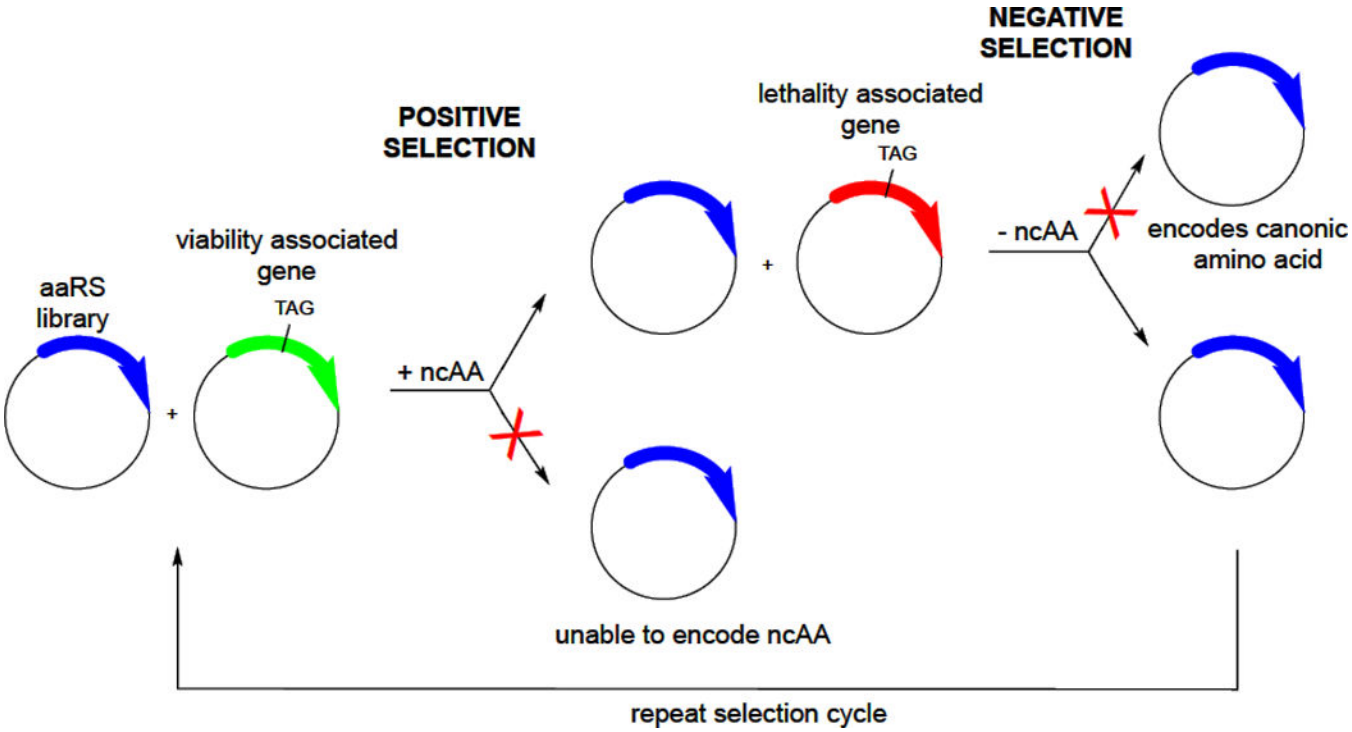


Figure 2.
Standard protocol for generation of an aaRS to encode ncAAs.

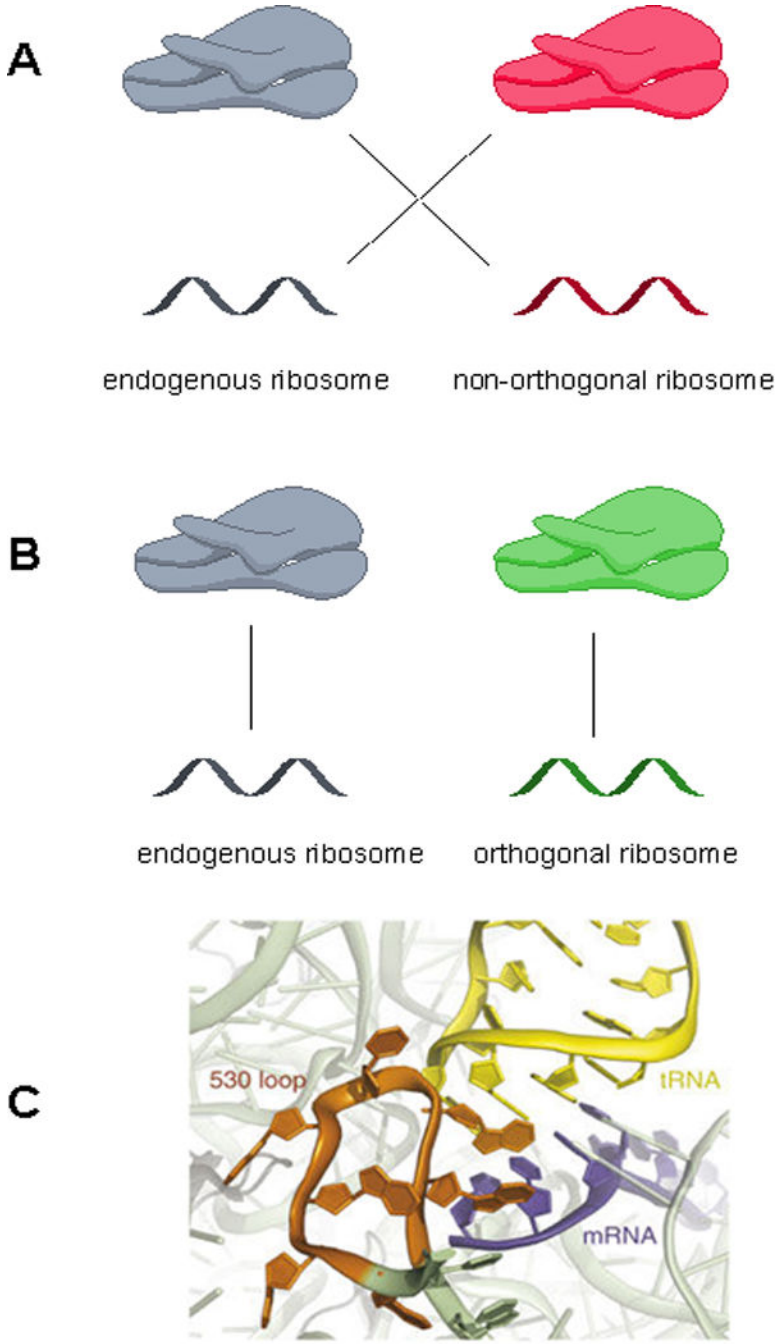


Figure 3.

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 (grey) 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]

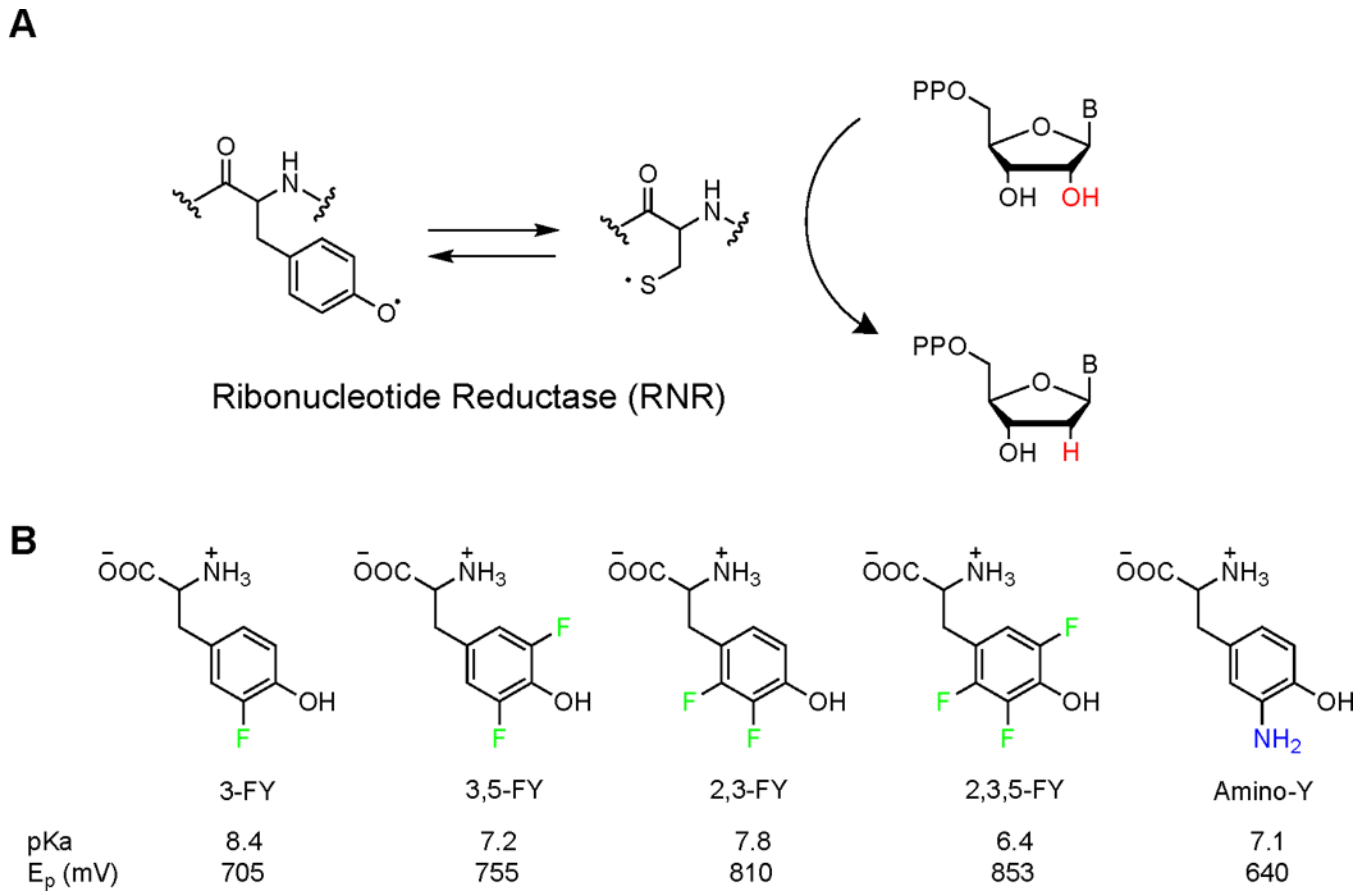


Figure 4.

Modulation of pKa 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 pKas and redox potentials of these residues affords key insights into the catalytic mechanism. B) Examples of ncAAs conferring altered tyrosine pKas and reduction potentials (E_p) that have been employed in the study of ribonucleotide reductase.

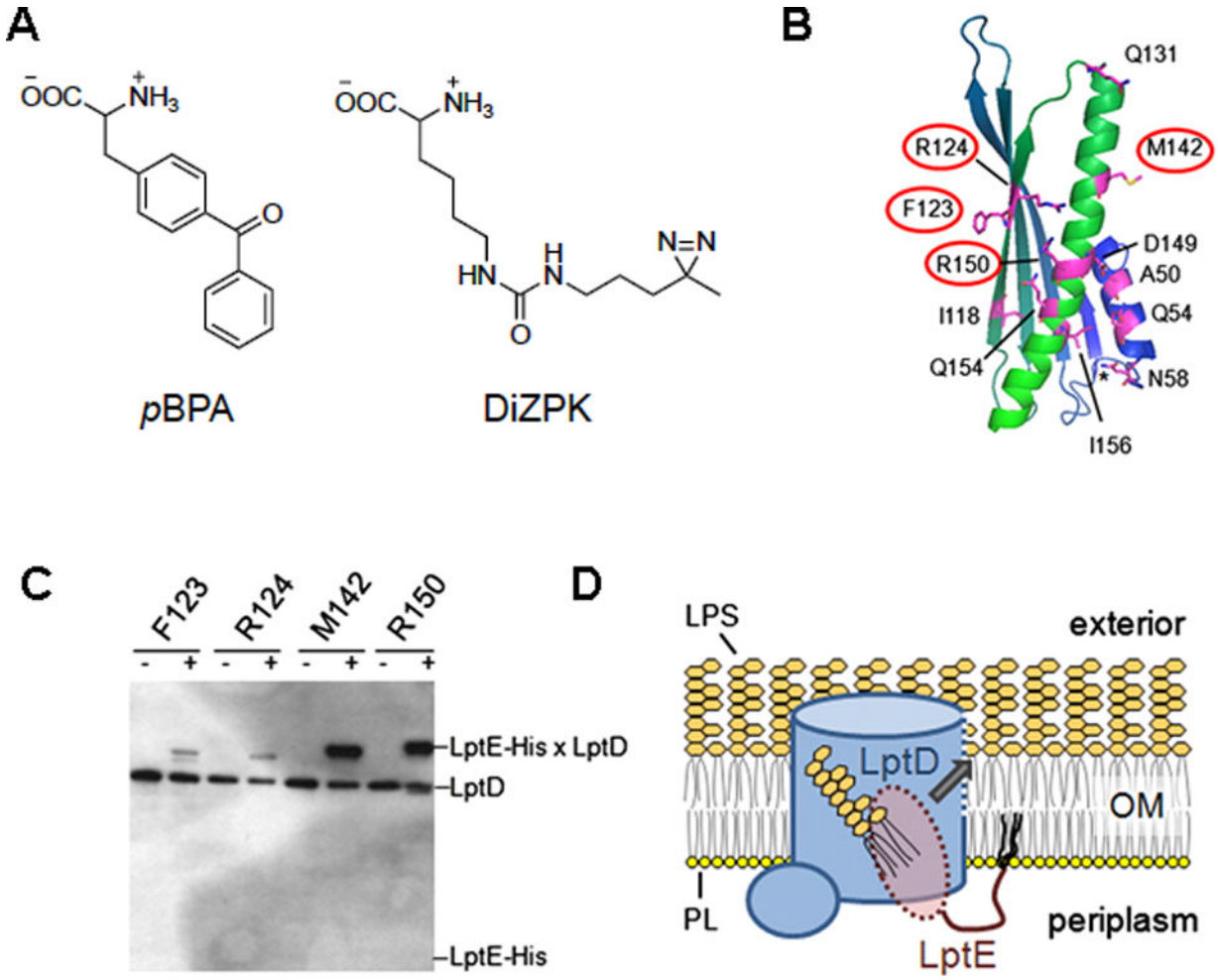


Figure 5.

Applications of photocrosslinking ncAAs. A) Structures of two common photocrosslinking 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 photocrosslinking 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 crosslinking event with LptD. Proposed model of the LptD/LptE association for lipopolysaccharide transport established by the crosslinking experiments. Adapted from Freinkman, E. et al. 2011.[154]

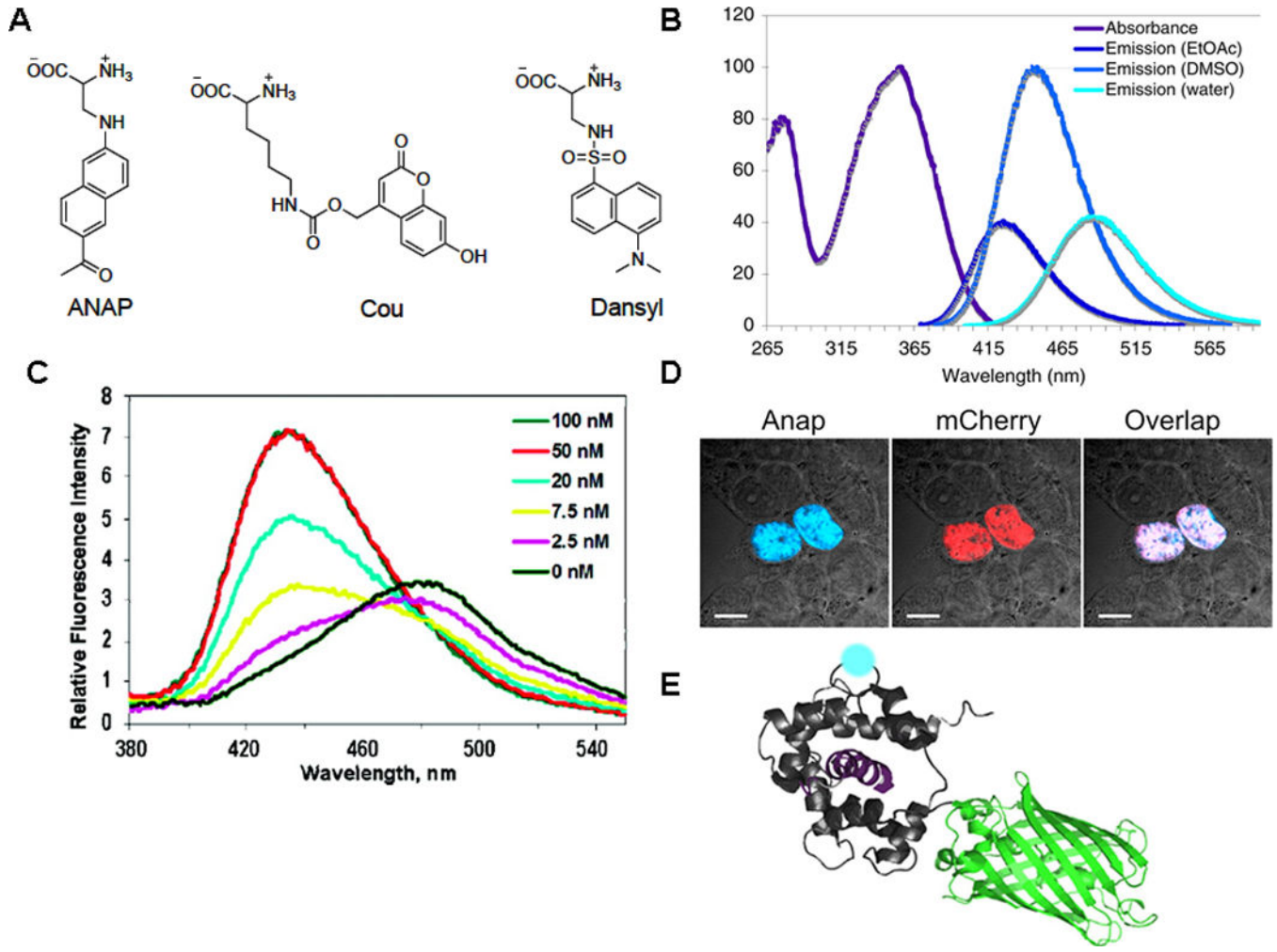


Figure 6.

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]

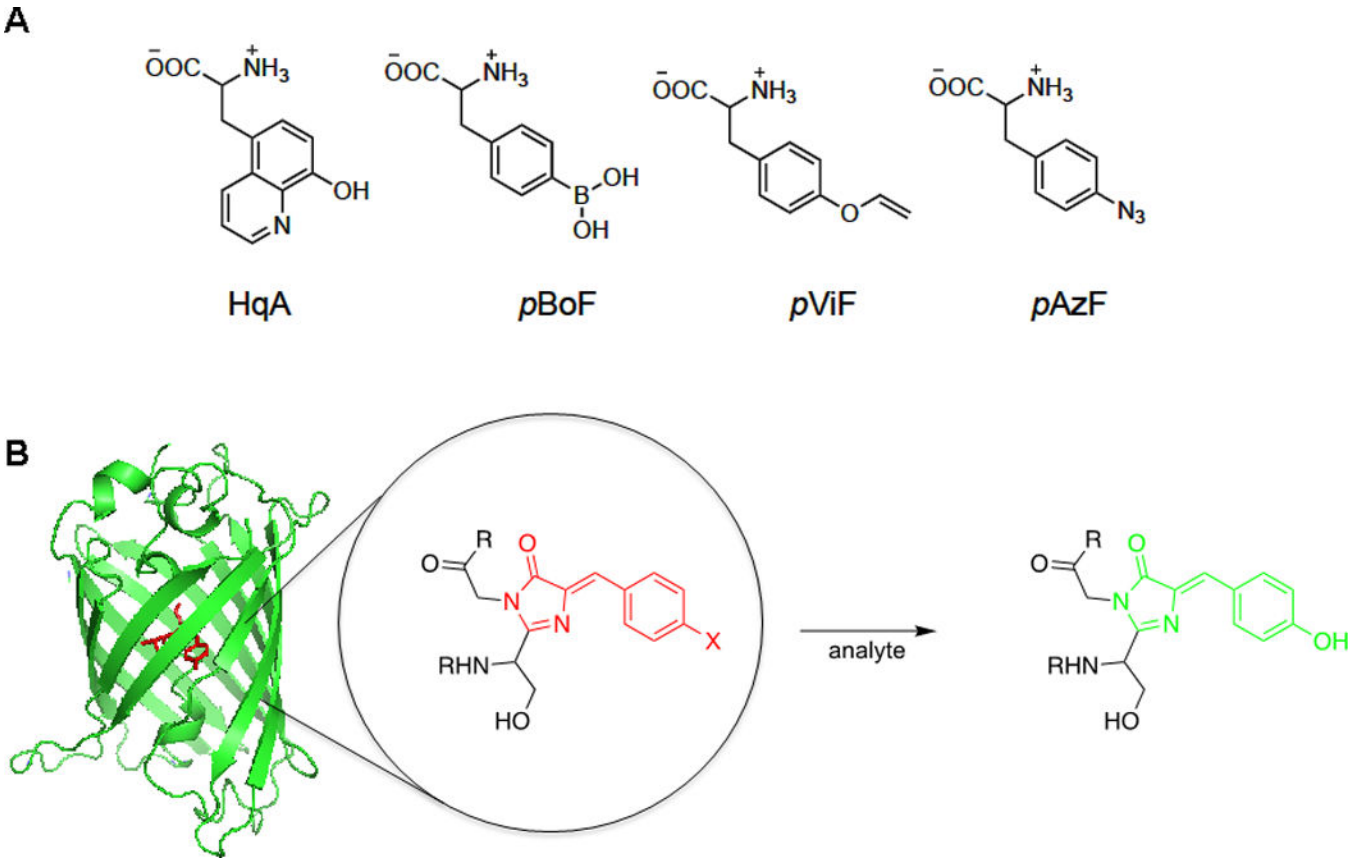


Figure 7.

Using ncAAs as environmental sensors for biologically relevant analytes. A) Structures of ncAAs commonly employed in environmental sensors, including metal binders, H_20_2, H_2S, 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.

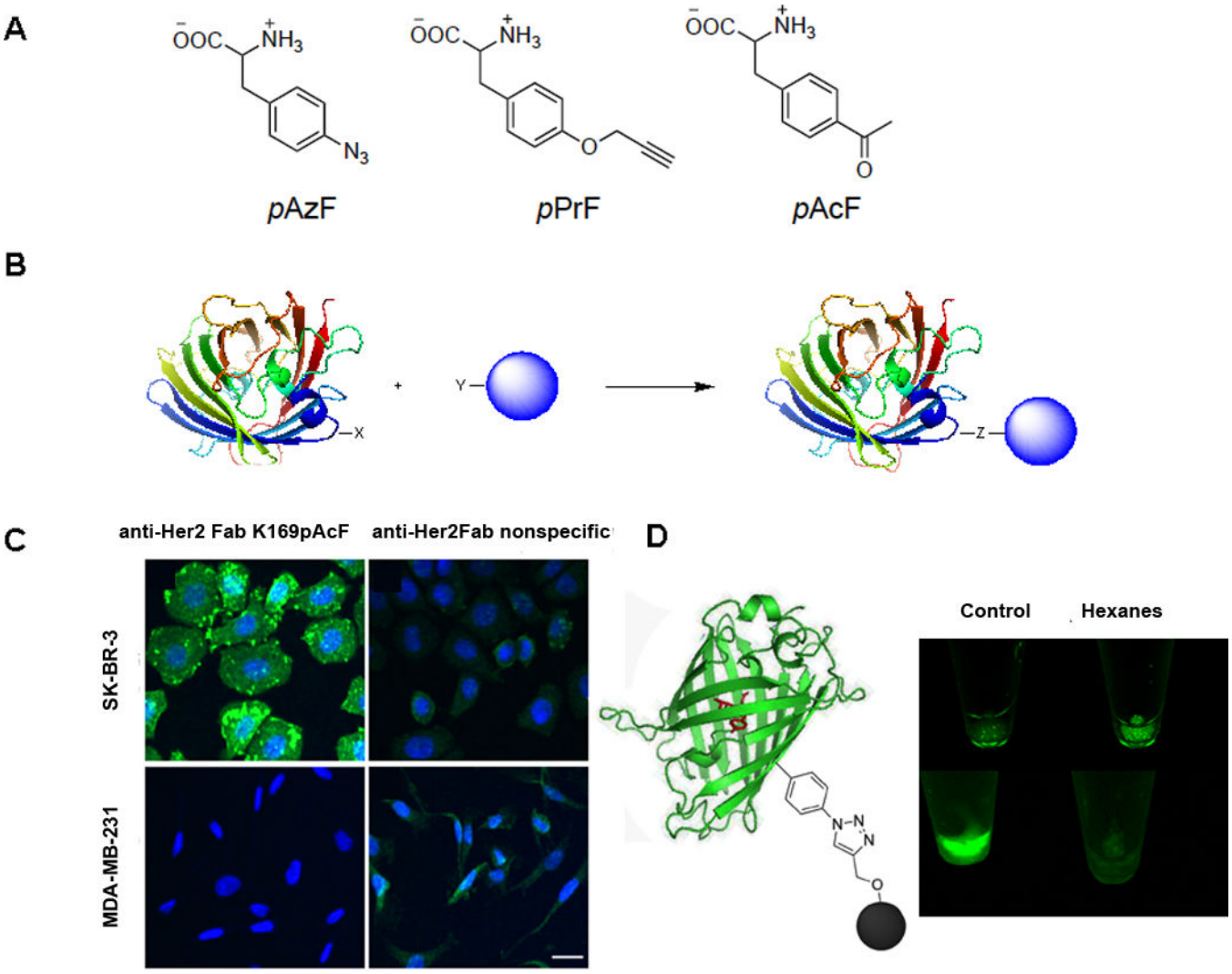


Figure 8.

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.[215] D) By site-specifically incorporating *p*-azidophenylalalanine GFP can be site-specifically immobilized on a solid support, conferring a higher degree of protein stability in non-aqueous solvents.[232]

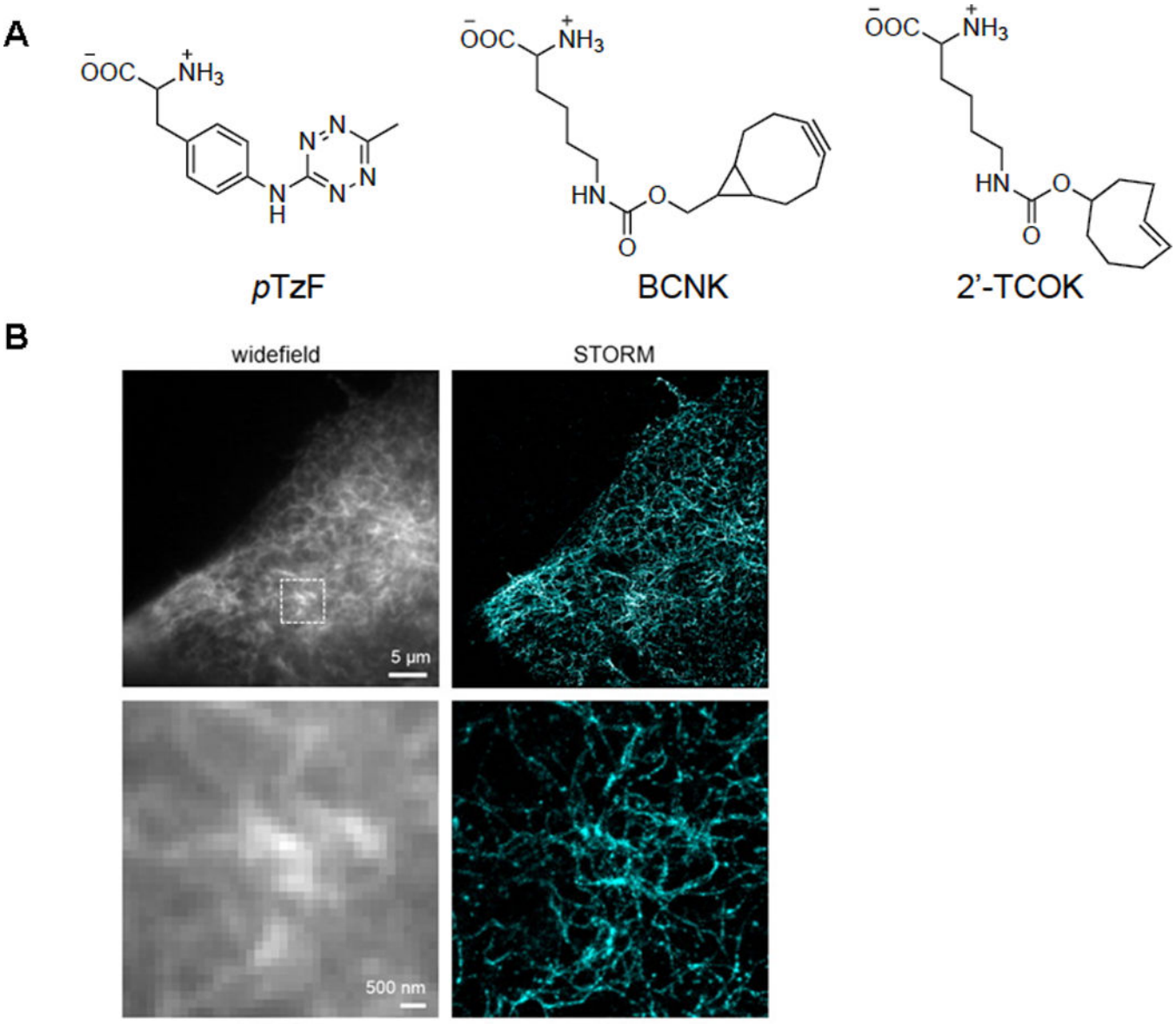


Figure 9.

Live cell imaging using ncAAs. A) Common structures of genetically encoded ncAAs for non-cytotoxic 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.[208]

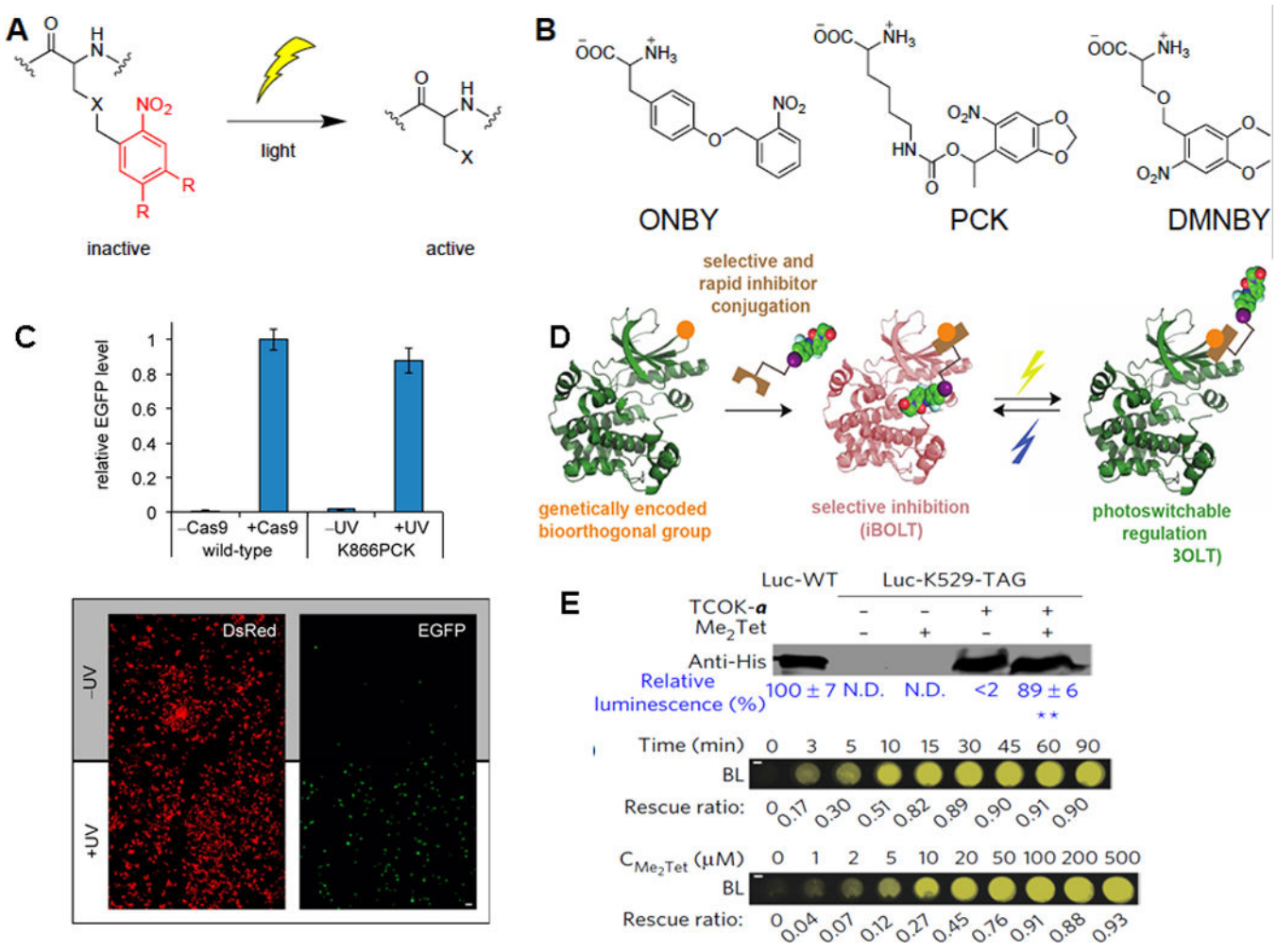


Figure 10.

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 non-functional. Upon brief irradiation with non-cytotoxic 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.[256] 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.[259] 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.[260]

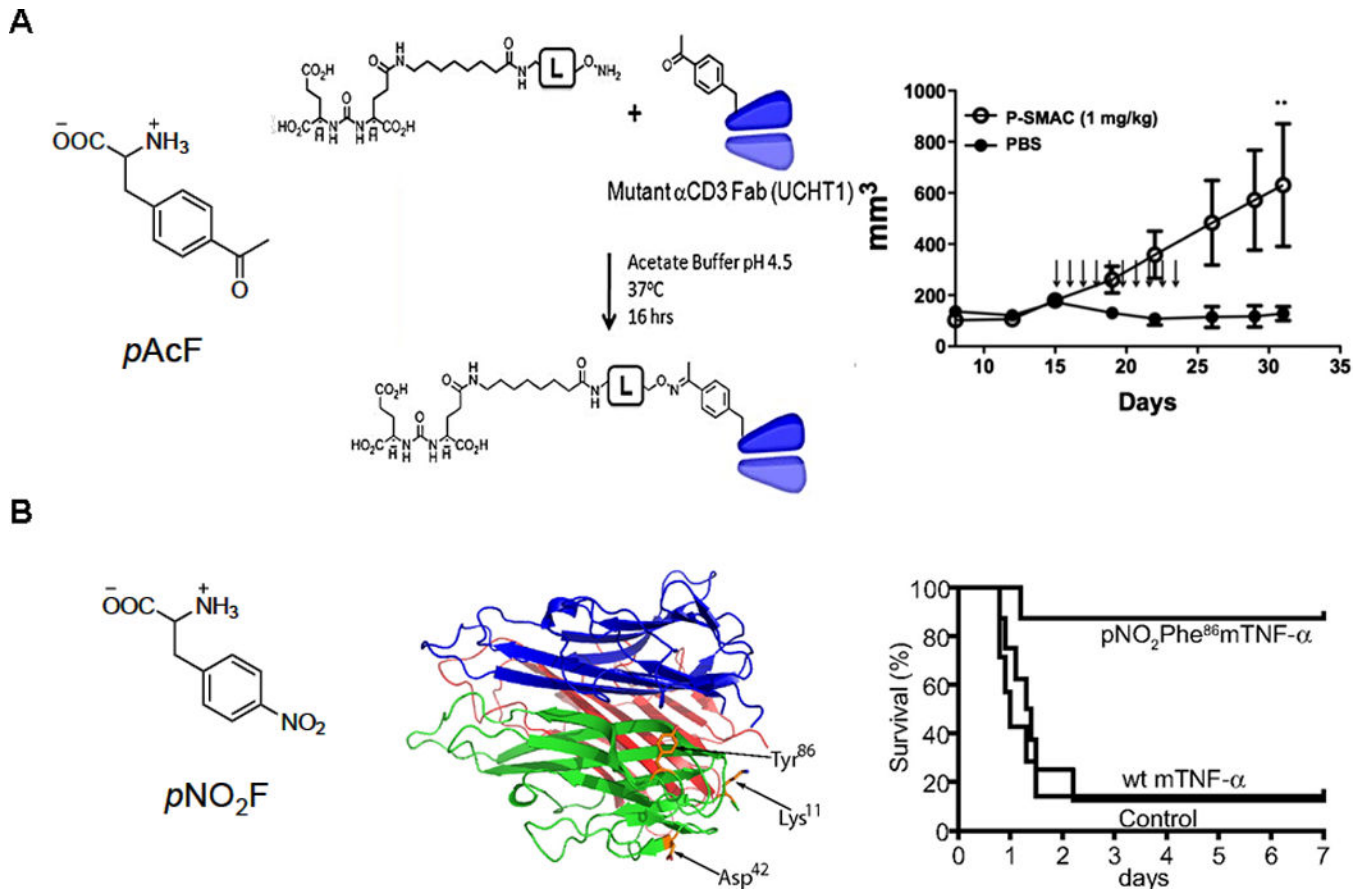


Figure 11.

Therapeutic applications of ncAAs. A) Incorporation of *p*-acetylphenylalanine into an anti-CD3 Fab followed by oxime ligation with an aminooxy-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 a endotoxemia mouse model.[287]

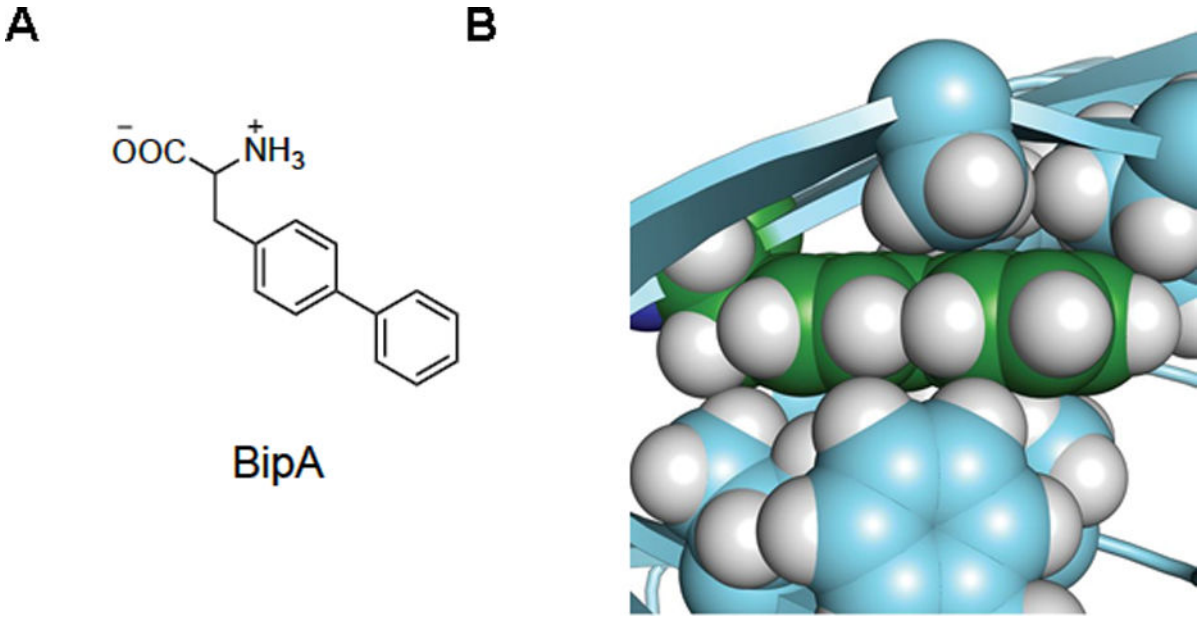


Figure 12.

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]