

Published in final edited form as:

*Curr Opin Chem Biol.* 2013 June ; 17(3): 412–419. doi:10.1016/j.cbpa.2013.04.017.

## Protein conjugation with genetically encoded unnatural amino acids

Chan Hyuk Kim<sup>1,\*</sup>, Jun Y. Axup<sup>2,\*</sup>, and Peter G. Schultz<sup>1,2</sup>

<sup>1</sup>California Institute for Biomedical Research, 11119 N. Torrey Pines Road, La Jolla, CA 92037

<sup>2</sup>Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Rd, La Jolla, CA 92037

### Abstract

The site-specific incorporation of unnatural amino acids with orthogonal chemical reactivity into proteins enables the synthesis of structurally-defined protein conjugates. Amino acids containing ketone, azide, alkyne, alkene, and tetrazine side chains can be genetically encoded in response to nonsense and frameshift codons. These bio-orthogonal chemical handles allow precise control over the site and stoichiometry of conjugation, and have enabled medicinal chemistry-like optimization of the physical and biological properties of protein conjugates, especially the next-generation protein therapeutics.

### Introduction

Methods for the site-specific modification of proteins have become powerful tools for probing protein structure and function, as well as generating proteins with new or enhanced properties. Historically, protein conjugation reactions involved the modification of lysine, cysteine, or nucleophilic serine side chains with electrophilic agents [1]. However, these reactions generally lead to heterogeneous mixtures of protein conjugates with distinct properties, especially when modifying large proteins such as antibodies. The introduction of chemically orthogonal functional groups into proteins, on the other hand, has allowed for precise control of conjugation site and stoichiometry. Common methods for this purpose include enzymatic or chemical reactions that modify specific peptide sequences with chemical tags [2], and semi-synthetic methods such as expressed protein ligation (EPL) [3]. Alternatively, recombinant methods have also been used to incorporate unnatural amino acids (UAA) into proteins as chemical handles for bio-orthogonal conjugation reactions [4]. The latter approach is particularly attractive because, in principle, the UAA can be incorporated at any desired position in any protein; the structure of the wild-type protein is minimally perturbed by UAA incorporation since it does not require any particular sequence

© 2013 Published by Elsevier Ltd.

\*Corresponding author: Schultz, Peter (schultz@scripps.edu).

\* authors contributed equally

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

context to specify the UAA site, and site-specifically modified proteins are expressed recombinantly in high yields in bacteria, yeast, or mammalian cells, and require little additional manipulation for further conjugation. In this review, we discuss several methods for the incorporation of UAAs with orthogonal chemical reactivity into proteins in living cells and recent examples of their application, especially with regards to protein therapeutics.

## Methods for Genetically Incorporating Unnatural Amino Acids into Proteins

Among the early methods for the genetic incorporation of UAAs into proteins was the use of bacterial strains auxotrophic for a structurally-related canonical amino acid. UAAs that are structurally similar to a common amino acid are recognized by that wild-type or engineered aminoacyl-tRNA synthetase (aaRS) and incorporated in the absence of the cognate amino acid. This method has been applied to analogues of Pro, Tyr, Phe, Leu, Val, and most commonly to Met (the latter has low abundance in the proteosome and hence less toxicity when replaced globally) [5]. For example, the incorporation of selenomethionine into proteins using methionyl-tRNA synthetase (MetRS) has been extensively used for the determination of protein structures by X-ray crystallography [6]. In another example, an engineered phenylalanyl-tRNA synthetase (PheRS) was used to replace Phe with *p*-acetylphenylalanine to introduce a reactive ketone functional group [7]. The low efficiency of UAA incorporation in this system was improved by over-expressing the cognate aaRS. However, this general approach has a number of drawbacks: the global replacement of canonical amino acids with structural analogues is often too toxic to support exponential growth of cells; wild-type aaRSs typically only mis-acylate close structural analogues of the canonical amino acid; and in larger proteins this method typically leads to the incorporation of UAAs at multiple sites.

In some archaeobacteria, nature has used nonsense codons to encode selenocysteine (Sec) and pyrrolysine (Pyl), which can be used for selective protein conjugation [8]. Selenol is relatively acidic (pKa ~5.2) compared to thiol (pKa ~8.3), and mostly exists as the selenate anion at physiological pH. Thus one can selectively modify proteins by reacting Sec with electrophiles (e.g., maleimide or  $\alpha$ -haloketones) in the presence of free Cys [9]. Unlike selenocysteine, pyrrolysine is unreactive. However, pyrrolysine analogues bearing reactive groups have been genetically encoded using the corresponding wild-type aaRS (or mutants thereof) and used in protein conjugation reactions [10].

To expand the genetic code beyond the canonical amino acids, we and others have reprogrammed the cell's translational machinery by evolving additional tRNA/aaRS pairs that site-specifically incorporate UAAs in response to either nonsense or 4-base frameshift codons [4] (Figure 1). Importantly, these tRNAs and aaRSs are orthogonal to their counterparts in the host cell (i.e., they do not cross-react with their natural counterparts) in order to maintain high translational fidelity. The most successful strategy for generating orthogonal tRNA/aaRS pairs is to import them from phylogenetically distant organisms (e.g., from archaea to *E. coli*, from *E. coli* to yeast, from archaea to mammalian cells, etc.). For example, an engineered tRNA<sup>Tyr</sup>/TyrRS pair derived from *Methanococcus jannaschii* has been used to incorporate a large number of UAAs into proteins in *E. coli* in response to the

amber codon TAG [11]. To genetically encode a specific unnatural amino acid, an active site library of aaRS mutants is subjected to multiple rounds of positive selection (to select mutants that incorporate any amino acid) and negative selection (to select against mutants that incorporate endogenous host amino acids). This method has been used to incorporate over 70 unnatural amino acids in bacteria, yeast, and mammalian cells (with yields up to 5 g/L in a 40,000 L bacterial fermentation) [4]. More recent advancements have expanded the technology to include orthogonal tRNA/aaRS pairs derived from leucyl [12], tryptophanyl [13], lysyl [14], pyrrolysyl [15], and prolyl [16] pairs to allow increased diversity in UAA structure. To incorporate two or more UAAs, tRNA/aaRS pairs have been developed that recognize other nonsense codons (TGA or TAA) or frameshift codons [14]. In addition, an orthogonal ribosome [17] that only recognizes orthogonal mRNAs (by means of a modified Shine-Dalgarno sequence) has also been engineered to efficiently decode quadruplet codons with minimal toxicity [18]. Moreover, it has been shown that a reengineered EF-Tu [19] with improved binding properties allows the incorporation of phosphoserine into proteins, and that deletion of RF1 with compensating mutations in RF2 can improve suppression efficiency [20]. More recently, the TAG nonsense codon has been more rigorously reassigned to only encode unnatural amino acids by genome-wide substitution of TAG with other synonymous stop codons (Church *et al.* submitted). Hence, one can very effectively reprogram much of the translational machinery to efficiently incorporate UAAs with novel chemical functionality not found in the canonical 20 amino acids.

## Unnatural Amino acids with Orthogonal Chemical Reactivity

Several chemically reactive UAAs have been genetically encoded to facilitate the selective conjugation of proteins to small molecules, oligonucleotides, synthetic polymers, and other proteins. Inspired by recent advances in bio-orthogonal or protein-compatible reactions, the repertoire of genetically encoded chemically reactive amino acids has grown considerably. Among them, the most widely used reactions include oxime formation [21], Cu(I)-catalyzed and strain-promoted Huisgen 1,3-dipolar cycloadditions (“Click” reaction) [22], and, most recently, inverse electron demand hetero Diels-Alder (HDA) reactions [23]. In addition, other chemically reactive amino acids have also been encoded that allow the selective modification of proteins by Michael reactions [24], metathesis reactions [25], transition metal catalyzed cross-couplings [26], radical polymerizations [27], oxidative couplings [28], acyl-transfer reactions [29], and photo click reactions [30].

Ketone-containing UAAs react with alkoxy-amine derivatives to form a stable oxime bond (whereas adducts with lysine and cysteine side chains are unstable in aqueous buffer) [31] (Figure 2A). Although the reaction rate is relatively slow ( $k \sim <10^{-2} \text{ M}^{-1}\text{s}^{-1}$ ), this drawback can be easily overcome by the addition of nucleophilic catalysts or excess conjugation partners. This strategy has a number of attractive features: the oxime bond is stable under physiological conditions; keto groups can be introduced into proteins with excellent efficiency and fidelity using genetically encoded p-acetylphenylalanine (pAcF) [32]; alkoxy-amine derivatives of small molecules and macromolecules are relatively easy to synthesize; pAcF causes minimal perturbation to protein function and stability; and the relatively mild conjugation conditions provide homogeneous reaction products in almost quantitative yields. Indeed this conjugation method has been the most extensively used method to date,

however, the acidic reaction conditions (pH <~5) preclude its use with pH-sensitive proteins and live cells, and it cannot be applied to molecules containing functional groups that react with alkoxy-amines.

Since the first description of click chemistry in 2001 [33], the Cu(I)-catalyzed cycloaddition of an azide and alkyne has found many uses in the biological sciences due to its high selectivity, yield and relatively mild reaction conditions. In the presence of Cu(I), an azide and alkyne rapidly form a 1,2,3-triazole linkage, which is stable under physiological conditions (Figure 2B). The reaction components, other than alkyne and azide, typically include Cu(II), reducing agent (e.g. ascorbic acid), and ligands to stabilize *in situ*-generated Cu(I) species. Although the reaction can take place at physiological pH, its applications are significantly limited due to Cu(I)-dependent side reactions with proteins and cellular toxicity. To overcome the limitations associated with Cu(I) catalysts, “activated” alkynes have been developed that react with good kinetics in the absence of catalyst [22,34,35] (Figure 2C). Typically, a cyclooctyne moiety has increased reactivity as a result of ring strain; other modifications such as fluorination of the vicinal carbons and fusion with a biaryl/cyclopropane ring further increase the reactivity of alkynes. Some of the cyclooctyne variants have even been used to label proteins in whole cells [35]. However, the more complex synthesis and short shelf-life (especially at ambient temperatures) of cyclooctyne compounds can be a drawback, and their increased bulk and hydrophobicity may affect protein structure and/or solubility.

Recently, there has been considerable interest in conjugation strategies that employ electron-deficient tetrazines, which selectively react with strained alkenes (norbornene, *trans*-cyclooctene, and cyclopropene) or alkynes to form stable bicyclic linkages [36] (Figure 2D). These reactions can typically be carried out under physiological conditions and do not require any catalyst. Their reaction rates ( $k = 10 \sim 10000 \text{ M}^{-1}\text{s}^{-1}$ ) are orders of magnitude faster than oxime formation or [3+2] cycloadditions, which makes this reaction more attractive for *in vivo* applications. Indeed a number of these reactive groups have been successfully genetically encoded [34,37–39].

## Applications to Protein Therapeutics

The ability to engineer chemically-defined macromolecules using unnatural amino acids provides for the first time medicinal chemistry-like control over the pharmacological and physical properties of protein therapeutics. More traditional methods like lysine or cysteine conjugation often form heterogeneous products with limited control over conjugation site and stoichiometry. Furthermore, it is likely that individual product species have distinct efficacy, safety, and pharmacokinetic properties that complicate the optimization of a therapeutic candidate. Some recent applications of genetically incorporated unnatural amino acids to the development of protein therapeutics are highlighted below (Figure 3).

One of the earliest applications involved site-specific polyethylene glycol (PEG) conjugation to extend the serum half-life of proteins (Figure 3A). The orthogonal chemical handle allows for the control of conjugation site and stoichiometry to maximize biological activity, serum half-life and stability/solubility. For example, human growth hormone [40]

and fibroblast growth factor 21 [41] were expressed with pAcF at defined positions and conjugated to an alkoxy-amine derivatized PEG in 80–97% yield. The resulting conjugate showed excellent pharmacokinetics, minimal loss of biological activity, and no apparent immunogenicity *in vivo*.

Genetically encoded UAAs have also been used to synthesize antibody-drug conjugates (ADCs). ADCs preferentially deliver cytotoxic drugs to cells presenting tumor-associated antigens to achieve improved drug efficacy and safety. For example, brentuximab vedotin (Adcetris/SGN-35; an anti-CD30 antibody conjugated with the microtubule-disrupting agent monomethyl auristatin E) has been recently approved for Hodgkin's lymphoma [42], and several other ADCs are currently in late-stage clinical trials. However, the first approved ADC, gemtuzumab ozogamicin (Mylotarg; an anti-CD33 antibody conjugated with the DNA cleaving agent calicheamicin) [43], was pulled from the market in 2010 because of toxicity and lack of efficacy in larger trials, showing the need for further optimization of this class of drugs.

Typical synthetic methods for most ADCs involve modification of lysine or cysteine residues which leads to a distribution of zero to eight toxins per antibody at various sites. The individual species within this mixture are likely to have distinct affinities, stabilities, pharmacokinetics, efficacies and safety profiles. A site-specific ADC with only two conjugation sites (generated by Thiomab technology in which additional cysteines are introduced into the antibody for maleimide conjugation [44]) showed similar efficacy to randomly labeled ADCs, but had an improved therapeutic index and better pharmacokinetic properties in rodents. This methodology, however, requires complicated reduction/oxidization steps and still has inherent cysteine-maleimide instability [45]. An anti-Her2 IgG expressed with two genetically encoded pAcF residue was conjugated to an auristatin derivative via a stable oxime-linkage in excellent yield [46] (Figure 3B). This construct showed selective *in vitro* cytotoxicity (EC<sub>50</sub> 100–400 pM) against Her2<sup>+</sup> breast cancer cell lines as well as complete regression of Her2<sup>+</sup> mammary fat pad tumors in mouse xenograft models. Full-length IgGs containing pAcF can be expressed in mammalian cells at yields over 1 g/L, making this a commercially viable technology. The ability to synthesize homogeneous ADCs enables medicinal chemistry-like optimization of ADCs, and is currently being applied to the synthesis of antibodies conjugated to kinase and phosphatase inhibitors, nuclear hormone receptor agonists and antagonists, and other drug classes for treatment of autoimmune, cardiovascular, and metabolic disease.

Bispecific antibodies that can bind two different antigens simultaneously (on the same or distinct proteins and/or cells) also have gained considerable attention as next-generation immunotherapeutics for cancer. Bispecific antibodies consisting of an antibody (or fragments thereof) specific for a tumor associated antigen linked to an anti-human CD3 (cluster of differentiation 3) antibody have demonstrated impressive efficacy in the clinic [47]. These engineered antibodies recruit CD8<sup>+</sup>/CD3<sup>+</sup> cytotoxic T lymphocytes to tumor cells and form a pseudo-immunological synapse that results in the activation of T lymphocytes and subsequent lysis of the target cells. A number of recombinant methods (including single chain variable fragment (scFv) and IgG-based formats) and some chemical crosslinking methods have been developed to generate bispecific antibodies [48–51].

However, genetically fused scFvs can have short serum half-lives and relatively poor stability, and some IgG-based quadroma formats, which consist of chimeric mouse IgG2a and rat IgG2b, can generate human anti-mouse (HAMA) or human anti-rat (HARA) antibody responses in patients [52]. Moreover, these methods often lead to restricted (and less optimal) orientations of the fusion proteins (e.g., C- to N-terminal fusions). Synthesis of bispecific antibodies using traditional chemical methods is challenging since one generates many different orientations of the two antigen binding sites, which may or may not form productive immunological synapses.

We have developed a two-step method to synthesize homogeneous bispecific antibodies using genetically encoded UAAs and flexible bifunctional linkers [53]. This approach involves conjugating the two pAcF-containing Fabs to bifunctional linkers (a short PEG with an alkoxy-amine on one end and either an azide or cyclooctyne on the other). In a second step, the antibody-linker conjugates are coupled to each other via copper-free [3+2] Huisgen cycloaddition in good yields. The affinity of anti-Her2 Fab homodimers generated by this method toward Her2 was comparable to full-length IgG. Moreover, heterodimers of anti-Her2/anti-CD3 Fabs showed sub-nanomolar killing ( $EC_{50} \sim 20$  pM) of Her2<sup>+</sup> cancer cells in the presence of purified peripheral blood mononuclear cells (PBMCs) *in vitro*, which is comparable to the most potent bispecific antibodies reported to date. Furthermore, the conjugate has a half-life of  $\sim 7$  hours in rodents and has shown effective tumor suppression in mouse xenograft models. Similar bispecific antibodies against other targets, e.g., CD20, CD33, and EGFR, are being developed.

Proteins have also been conjugated to single-stranded DNAs or peptide nucleic acids (PNAs) and can be subsequently hybridized to form homodimers, heterodimers, tetramers or higher order complexes [54] (Figure 3D). For example, PNA was conjugated to a pAcF-containing Fab via oxime ligation in excellent yields. The PNA-linked anti-Her2/anti-CD3 Fabs showed potent and specific cytotoxicity *in vitro*, similar to that of the PEG-linked bispecific Fabs. A tetramer of anti-CD20 Fab was generated utilizing PNAs that form a Holiday junction, and potently induced apoptosis of CD20<sup>+</sup> Ramos cells. Additionally, an anti-Her2 Fab conjugated to single-stranded DNA was used to detect Her2<sup>+</sup> cells by immuno-PCR with exquisite sensitivity [55].

Other applications of selective conjugation with genetically encoded unnatural amino acids include the conjugation of fluorescent dyes and other spectroscopic probes to proteins. Fluorescent dyes have been conjugated both *in vitro*, such as for FRET experiments to study protein dynamics [56] and *in vivo* for specific labeling on cell surfaces [35,57,58]. Spin labels also have been conjugated to probe protein structure by electron paramagnetic resonance (EPR) [59].

## Conclusion

An advantage of using genetically encoded UAAs for bio-orthogonal protein conjugation is the ability to generate chemically-defined, homogeneous protein conjugates that can be characterized and optimized with respect to their efficacy, safety and pharmacokinetics. There are clearly many directions in which the technology and its applications can be further

developed. For example, currently stoichiometry is largely limited to one site per protein (or two sites on a full-length antibody); the efficient incorporation of two or more identical or different UAAs into a single polypeptide will allow the synthesis of more-complex conjugates. Moreover, the development of stable cell lines each encoding a distinct amino acid will simplify mammalian expression of mutant proteins. Different payloads for antibody-drug conjugates, such as kinase inhibitors, nuclear hormone receptor agonists, or HDAC inhibitors will further expand the therapeutic applications of ADCs. Also, novel bi- and multi-specific protein conjugates can be created to bind specific antigens on cells in defined orientations with respect to each other. These and other advances promise to enable a new generation of chemically tailored protein therapeutics.

## Acknowledgments

The authors thank the National Institutes of Health (NIH) for support (R01GM062159). This manuscript is number 22069 of The Scripps Research Institute.

## References

1. Basle E, Joubert N, Pucheault M. Protein chemical modification on endogenous amino acids. *Chemistry & biology*. 2010; 17(3):213–227. [PubMed: 20338513]
2. O'Hare HM, Johnsson K, Gautier A. Chemical probes shed light on protein function. *Current opinion in structural biology*. 2007; 17(4):488–494. [PubMed: 17851069]
3. Muir TW. Semisynthesis of proteins by expressed protein ligation. *Annual review of biochemistry*. 2003; 72:249–289.
- 4. Liu CC, Schultz PG. Adding new chemistries to the genetic code. *Annual review of biochemistry*. 2010; 79:413–444. This comprehensive review describes genetic incorporation and function of a wide range of unnatural amino acids.
5. Hendrickson TL, de Crecy-Lagard V, Schimmel P. Incorporation of nonnatural amino acids into proteins. *Annual review of biochemistry*. 2004; 73:147–176.
6. Yang W, Hendrickson WA, Crouch RJ, Satow Y. Structure of ribonuclease H phased at 2 Å resolution by mad analysis of the selenomethionyl protein. *Science*. 1990; 249(4975):1398–1405. [PubMed: 2169648]
7. Datta D, Wang P, Carrico IS, Mayo SL, Tirrell DA. A designed phenylalanyl-trna synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins. *Journal of the American Chemical Society*. 2002; 124(20):5652–5653. [PubMed: 12010034]
8. Ambrogelly A, Palioura S, Soll D. Natural expansion of the genetic code. *Nature chemical biology*. 2007; 3(1):29–35.
9. Hofer T, Skeffington LR, Chapman CM, Rader C. Molecularly defined antibody conjugation through a selenocysteine interface. *Biochemistry*. 2009; 48(50):12047–12057. [PubMed: 19894757]
10. Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. Multistep engineering of pyrrolysyl-trna synthetase to genetically encode N(epsilon)-(o-azidobenzoyloxycarbonyl) lysine for site-specific protein modification. *Chemistry & biology*. 2008; 15(11):1187–1197. [PubMed: 19022179]
11. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of *Escherichia coli*. *Science*. 2001; 292(5516):498–500. [PubMed: 11313494]
12. Wu N, Deiters A, Cropp TA, King D, Schultz PG. A genetically encoded photocaged amino acid. *Journal of the American Chemical Society*. 2004; 126(44):14306–14307. [PubMed: 15521721]
13. Zhang ZW, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, Schultz PG. Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(24):8882–8887. [PubMed: 15187228]

14. Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, Schultz PG. An expanded genetic code with a functional quadruplet codon. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(20):7566–7571. [PubMed: 15138302]
15. Neumann H, Peak-Chew SY, Chin JW. Genetically encoding N-epsilon-acetyllysine in recombinant proteins. *Nature chemical biology*. 2008; 4(4):232–234.
16. Chatterjee A, Xiao H, Schultz PG. Evolution of multiple, mutually orthogonal prolyl-trna synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(37):14841–14846. [PubMed: 22927411]
17. Rackham O, Chin JW. A network of orthogonal ribosome x mRNA pairs. *Nature chemical biology*. 2005; 1(3):159–166.
- 18. 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(7287):441–444. This paper describes the use of an orthogonal ribosome-mRNA pair to encode quadruplet codons. [PubMed: 20154731]
19. Park HS, Hohn MJ, Umehara T, Guo LT, Osborne EM, Benner J, Noren CJ, Rinehart J, Soll D. Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science*. 2011; 333(6046):1151–1154. [PubMed: 21868676]
20. Johnson DB, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, Xiang Z, Ecker JR, Briggs SP, Wang L. RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nature chemical biology*. 2011; 7(11):779–786.
21. Dirksen A, Dawson PE. Rapid oxime and hydrazone ligations with aromatic aldehydes for biomolecular labeling. *Bioconjugate chemistry*. 2008; 19(12):2543–2548. [PubMed: 19053314]
- 22. Jewett JC, Bertozzi CR. Cu-free click cycloaddition reactions in chemical biology. *Chemical Society reviews*. 2010; 39(4):1272–1279. This review discusses advances in Cu-free click cycloaddition for protein conjugation *in vitro* and *in vivo*. [PubMed: 20349533]
23. Devaraj NK, Weissleder R, Hilderbrand SA. Tetrazine-based cycloadditions: Application to pretargeted live cell imaging. *Bioconjugate chemistry*. 2008; 19(12):2297–2299. [PubMed: 19053305]
24. Guo J, Wang J, Lee JS, Schultz PG. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angewandte Chemie*. 2008; 47(34):6399–6401. [PubMed: 18624319]
25. Ai HW, Shen W, Brustad E, Schultz PG. Genetically encoded alkenes in yeast. *Angewandte Chemie*. 2010; 49(5):935–937. [PubMed: 20025009]
26. Brustad E, Bushey ML, Lee JW, Groff D, Liu W, Schultz PG. A genetically encoded boronate-containing amino acid. *Angewandte Chemie*. 2008; 47(43):8220–8223. [PubMed: 18816552]
27. Peeler JC, Woodman BF, Averick S, Miyake-Stoner SJ, Stokes AL, Hess KR, Matyjaszewski K, Mehl RA. Genetically encoded initiator for polymer growth from proteins. *Journal of the American Chemical Society*. 2010; 132(39):13575–13577. [PubMed: 20839808]
28. Behrens CR, Hooker JM, Obermeyer AC, Romanini DW, Katz EM, Francis MB. Rapid chemoselective bioconjugation through oxidative coupling of anilines and aminophenols. *Journal of the American Chemical Society*. 2011; 133(41):16398–16401. [PubMed: 21919497]
29. Nguyen DP, Elliott T, Holt M, Muir TW, Chin JW. Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *Journal of the American Chemical Society*. 2011; 133(30):11418–11421. [PubMed: 21736333]
30. Yu Z, Pan Y, Wang Z, Wang J, Lin Q. Genetically encoded cyclopropene directs rapid, photoclick-chemistry-mediated protein labeling in mammalian cells. *Angewandte Chemie*. 2012; 51(42):10600–10604. [PubMed: 22997015]
31. Kalia J, Raines RT. Hydrolytic stability of hydrazones and oximes. *Angewandte Chemie*. 2008; 47(39):7523–7526. [PubMed: 18712739]
32. Wang L, Zhang Z, Brock A, Schultz PG. Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100(1):56–61. [PubMed: 12518054]

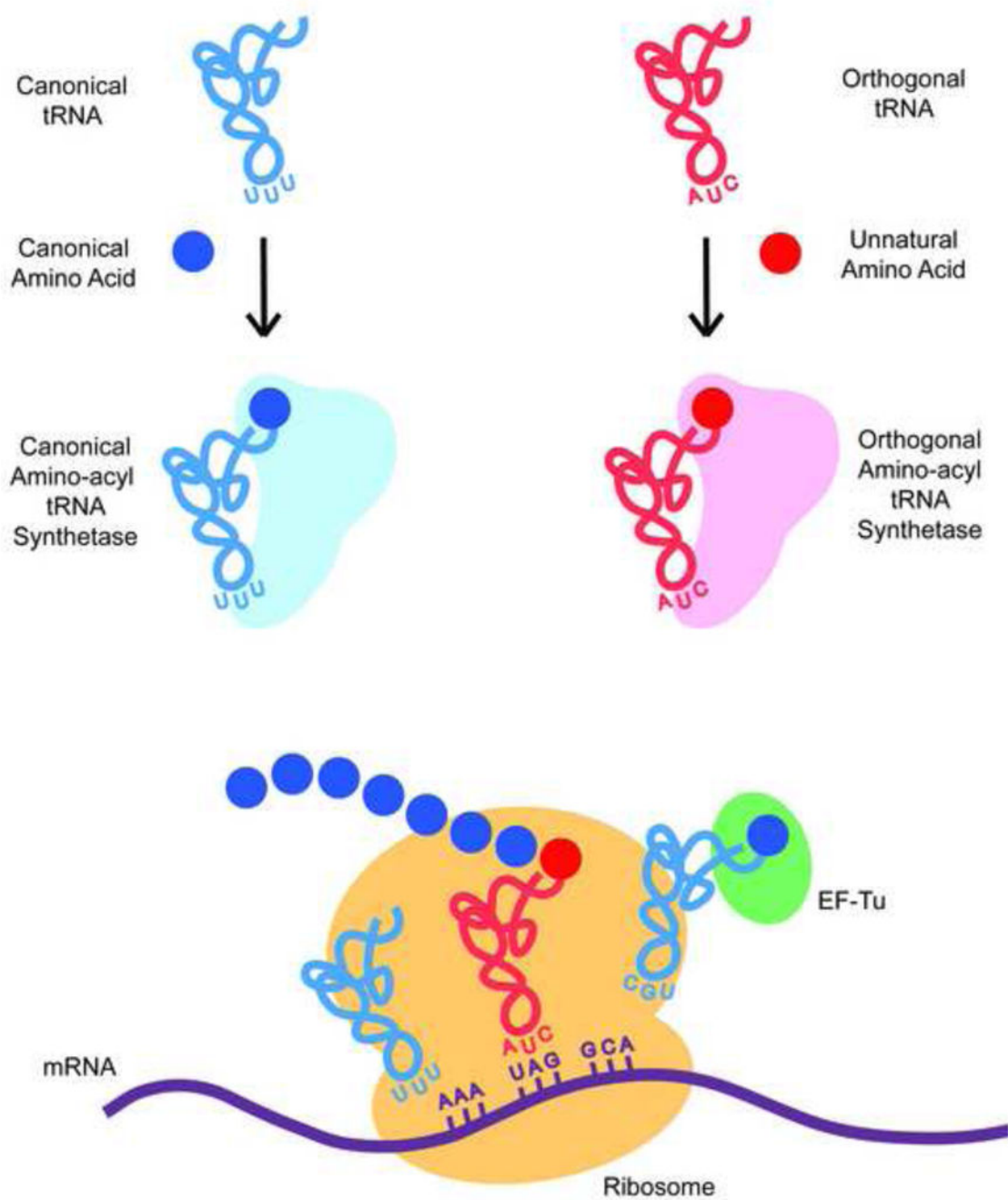


33. Kolb HC, Finn MG, Sharpless KB. Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie*. 2001; 40(11):2004–2021. [PubMed: 11433435]
34. Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic diels-alder reactions. *Journal of the American Chemical Society*. 2012; 134(25):10317–10320. [PubMed: 22694658]
35. Plass T, Milles S, Koehler C, Schultz C, Lemke EA. Genetically encoded copper-free click chemistry. *Angewandte Chemie*. 2011; 50(17):3878–3881. [PubMed: 21433234]
36. Blackman ML, Royzen M, Fox JM. Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *Journal of the American Chemical Society*. 2008; 130(41):13518–13519. [PubMed: 18798613]
37. Seitchik JL, Peeler JC, Taylor MT, Blackman ML, Rhoads TW, Cooley RB, Refakis C, Fox JM, Mehl RA. Genetically encoded tetrazine amino acid directs rapid site-specific in vivo bioorthogonal ligation with trans-cyclooctenes. *Journal of the American Chemical Society*. 2012; 134(6):2898–2901. [PubMed: 22283158]
38. Kaya E, Vrabel M, Deiml C, Prill S, Fluxa VS, Carell T. A genetically encoded norbornene amino acid for the mild and selective modification of proteins in a copper-free click reaction. *Angew Chem Int Edit*. 2012; 51(22):5277–5277.
39. 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. *Nature chemistry*. 2012; 4(4):298–304.
- 40. Cho H, Daniel T, Buechler YJ, Litzinger DC, Maio Z, Putnam AM, Kraynov VS, Sim BC, Bussell S, Javahishvili T, Kaphle S, et al. Optimized clinical performance of growth hormone with an expanded genetic code. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(22):9060–9065. This paper introduces the use of genetically encoded unnatural amino acids for site-specific PEGylation that enhances the pharmacokinetic properties of therapeutic proteins. [PubMed: 21576502]
41. Mu J, Pinkstaff J, Li Z, Skidmore L, Li N, Myler H, Dallas-Yang Q, Putnam AM, Yao J, Bussell S, Wu M, et al. FGF21 analogs of sustained action enabled by orthogonal biosynthesis demonstrate enhanced antidiabetic pharmacology in rodents. *Diabetes*. 2012; 61(2):505–512. [PubMed: 22210323]
42. Doronina SO, Mendelsohn BA, Bovee TD, Cervený CG, Alley SC, Meyer DL, Oflazoglu E, Toki BE, Sanderson RJ, Zabinski RF, Wahl AF, et al. Enhanced activity of monomethylauristatin F through monoclonal antibody delivery: Effects of linker technology on efficacy and toxicity. *Bioconjugate chemistry*. 2006; 17(1):114–124. [PubMed: 16417259]
43. Hamann PR, Hinman LM, Hollander I, Beyer CF, Lindh D, Holcomb R, Hallett W, Tsou HR, Upešlaciš J, Shochat D, Mountain A, et al. Gemtuzumab ozogamicin, a potent and selective anti-CD33 antibody-calicheamicin conjugate for treatment of acute myeloid leukemia. *Bioconjugate chemistry*. 2002; 13(1):47–58. [PubMed: 11792178]
- 44. Junutula JR, Raab H, Clark S, Bhakta S, Leipold DD, Weir S, Chen Y, Simpson M, Tsai SP, Dennis MS, Lu Y, et al. Site-specific conjugation of a cytotoxic drug to an antibody improves the therapeutic index. *Nature biotechnology*. 2008; 26(8):925–932. This paper describes THIOMAB technology that utilizes reactive cysteines for site-specific conjugation of drugs to an antibody.
45. Shen BQ, Xu K, Liu L, Raab H, Bhakta S, Kenrick M, Parsons-Repointe KL, Tien J, Yu SF, Mai E, Li D, et al. Conjugation site modulates the in vivo stability and therapeutic activity of antibody-drug conjugates. *Nature biotechnology*. 2012; 30(2):184–189.
- 46. Axup JY, Bajjuri KM, Ritland M, Hutchins BM, Kim CH, Kazane SA, Halder R, Forsyth JS, Santidrian AF, Stafin K, Lu Y, et al. Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(40):16101–16106. This paper introduces the use of genetically encoded unnatural amino acids for the synthesis of site-specific homogeneous antibody-drug conjugates. [PubMed: 22988081]
47. Baeuerle PA, Reinhardt C. Bispecific t-cell engaging antibodies for cancer therapy. *Cancer research*. 2009; 69(12):4941–4944. [PubMed: 19509221]

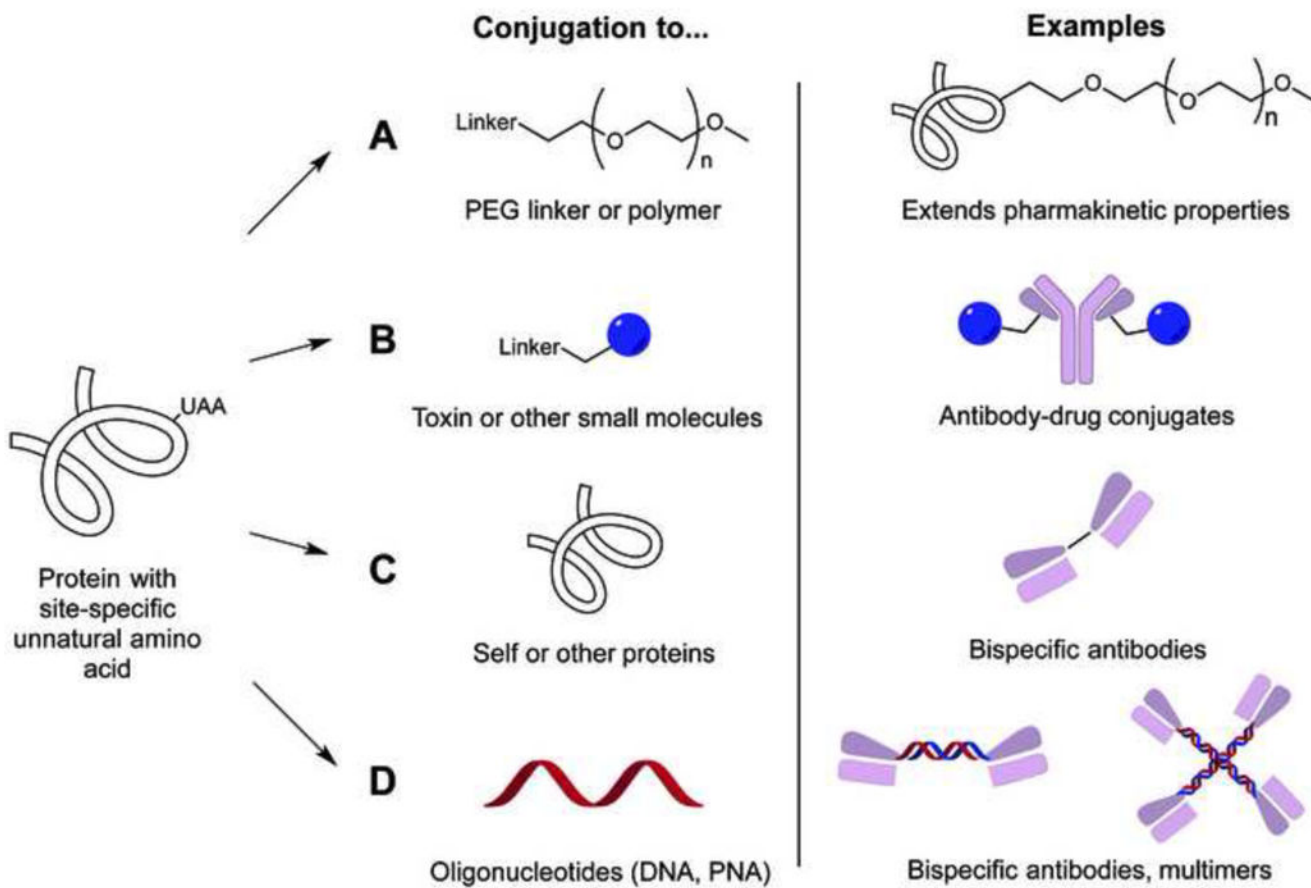
48. Moore PA, Zhang W, Rainey GJ, Burke S, Li H, Huang L, Gorlatov S, Veri MC, Aggarwal S, Yang Y, Shah K, et al. Application of dual affinity retargeting molecules to achieve optimal redirected T-cell killing of B-cell lymphoma. *Blood*. 2011; 117(17):4542–4551. [PubMed: 21300981]
49. Jager M, Schoberth A, Ruf P, Hess J, Lindhofer H. The trifunctional antibody ertumaxomab destroys tumor cells that express low levels of human epidermal growth factor receptor 2. *Cancer research*. 2009; 69(10):4270–4276. [PubMed: 19435924]
50. Wu C, Ying H, Grinnell C, Bryant S, Miller R, Clabbers A, Bose S, McCarthy D, Zhu RR, Santora L, Davis-Taber R, et al. Simultaneous targeting of multiple disease mediators by a dual-variable-domain immunoglobulin. *Nature biotechnology*. 2007; 25(11):1290–1297.
51. Graziano RF, Guptill P. Chemical production of bispecific antibodies. *Methods in molecular biology*. 2004; 283:71–85. [PubMed: 15197303]
52. Chames P, Baty D. Bispecific antibodies for cancer therapy: The light at the end of the tunnel? *mAbs*. 2009; 1(6):539–547. [PubMed: 20073127]
- 53. Kim CH, Axup JY, Dubrovskaya A, Kazane SA, Hutchins BA, Wold ED, Smider VV, Schultz PG. Synthesis of bispecific antibodies using genetically encoded unnatural amino acids. *Journal of the American Chemical Society*. 2012; 134(24):9918–9921. This paper describes the generation of bispecific antibodies using unnatural amino acids and bifunctional linkers. [PubMed: 22642368]
54. Kazane SA, Axup JY, Kim CH, Ciobanu M, Wold ED, Barluenga S, Hutchins BA, Schultz PG, Winssinger N, Smider VV. Self-assembled antibody multimers through peptide nucleic acid conjugation. *Journal of the American Chemical Society*. 2012
55. Kazane SA, Sok D, Cho EH, Uson ML, Kuhn P, Schultz PG, Smider VV. Site-specific DNA-antibody conjugates for specific and sensitive immuno-PCR. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(10):3731–3736. [PubMed: 22345566]
56. Brustad EM, Lemke EA, Schultz PG, Deniz AA. A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *Journal of the American Chemical Society*. 2008; 130(52):17664–17665. [PubMed: 19108697]
57. Ye S, Köhrer C, Huber T, Kazmi M, Sachdev P, Yan ECY, Bhagat A, RajBhandary UL, Sakmar TP. Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. *Journal of Biological Chemistry*. 2008; 283(3):1525–1533. [PubMed: 17993461]
58. Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels–Alder reactions. *Journal of the American Chemical Society*. 2012; 134(25):10317–10320. [PubMed: 22694658]
59. Fleissner MR, Brustad EM, Kálai T, Altenbach C, Cascio D, Peters FB, Hideg K, Peuker S, Schultz PG, Hubbell WL. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proceedings of the National Academy of Sciences*. 2009; 106(51):21637–21642.

### Highlights

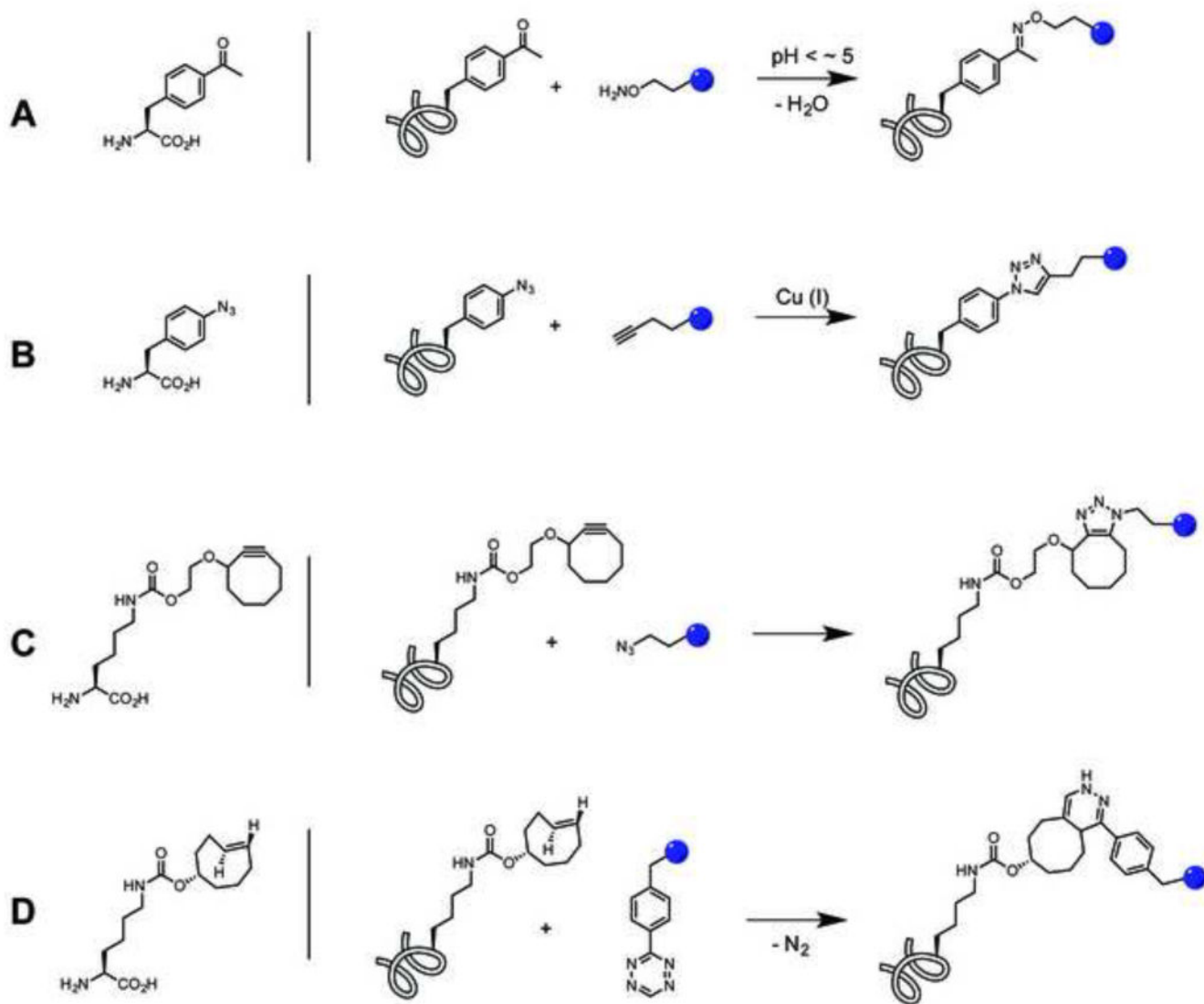
1. Unnatural amino acids with orthogonal chemical reactivity can be genetically encoded.
2. Efficient methods for the selective incorporation of unnatural amino acids into proteins have been developed.
3. Small molecules and macromolecules can be site-specifically conjugated to proteins.
4. Recent applications of this method include the generation of protein therapeutics.



**Figure 1.**  
Scheme for the genetic incorporation of unnatural amino acids.



**Figure 2.**  
UAAs used for site-specific protein conjugation.



**Figure 3.** Protein conjugates synthesized with unnatural amino acids and examples of their applications.