BIOLOGY REPORTS

Published: 26 November 2009
© 2009 Faculty of 1000 Ltd

# Unnatural amino acids: better than the real things?

Ellen C Minnihan[1], Kenichi Yokoyama[1] and JoAnne Stubbe[1,2]*

Addresses: [1]Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA; [2]Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA

* Corresponding author: JoAnne Stubbe (stubbe@mit.edu)

*F1000 Biology Reports* 2009, **1**:88 (doi:10.3410/B1-88)

The electronic version of this article is the complete one and can be found at: http://F1000.com/Reports/Biology/content/1/88

## Abstract

Considerable effort has been dedicated to the development of technology for the site-specific incorporation of unnatural amino acids into proteins, with nonsense codon suppression and expressed protein ligation emerging as two of the most promising methods. Recent research advances in which these methods have been applied to study protein function and mechanism are briefly highlighted, and the potential of the methods for efficient, widespread future use *in vitro* and *in vivo* is critically evaluated.

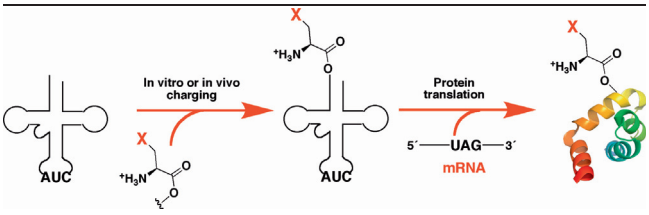
## Introduction and context

Robust technology has the potential to revolutionize scientific inquiry in ways often unimaginable at the time of the method's infancy. In the fields of protein chemistry and enzymology, an example of such a technology is site-directed mutagenesis, which allows researchers to site-specifically replace any naturally occurring amino acid with another amino acid in the genetic code [1]. The technology of site-directed mutagenesis evolved over several decades and is now used routinely by scientists across many disciplines to address diverse questions related to protein structure and function. The ability to site-specifically replace any natural amino acid with an unnatural amino acid (UAA) has even greater potential for protein biochemistry, as it allows the introduction of unique chemical properties and reactivities into the protein environment. Recent advances in this field suggest that the tools to achieve this goal, once developed to their full potential, will also facilitate solving problems that cannot be addressed by current methods [2]. This technology will be transforming.

Two general approaches are currently available to replace natural amino acids with unnatural analogues: *in vitro* or *in vivo* nonsense-codon suppressor tRNA technology and expressed protein ligation (EPL) technology. Recent advances in these two methods have been reviewed thoroughly by pioneers in the respective fields, Peter Schultz and Tom Muir [3-6]. In the former technique (Figure 1), a tRNA with a suppressor anticodon is charged with a UAA and incorporates it site-specifically during mRNA translation in response to a nonsense codon placed site-specifically in a gene of interest. Exploitation of Nature's translational machinery, the ribosome, in this way has allowed for incorporation of more than 100 UAAs into a wide variety of proteins in a number of model organisms. Incorporation by this method is astonishing, as Nature has evolved many mechanisms to ensure fidelity of the translational process [7,8]. The most efficient means of charging the tRNA with the UAA is an area of intense research. The method pioneered by Schultz allows for *in vivo* tRNA charging via the evolution of an orthogonal tRNA synthetase (RS) [9] and has proven successful in many proof-of-concept studies [4]. *In vitro* attachment of the UAA to the tRNA by chemical or enzymatic synthesis has also proven to be a viable alternative in select systems [10].

The foundation of EPL lies in intein-mediated protein self-splicing [11]. Coupling this naturally occurring phenomenon with the synthetic technique of native chemical ligation [12] has allowed for the semisynthesis of proteins containing UAAs by a combination of

**Figure 1. Unnatural amino acid incorporation though nonsense codon suppression**

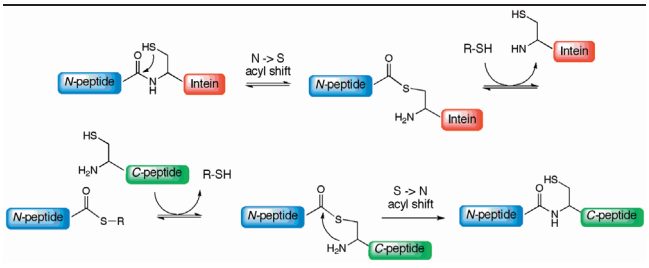



An evolved tRNA containing a nonsense anticodon is charged *in vitro* or *in vivo* with an unnatural amino acid. The machinery of the ribosome recognizes a nonsense codon within the mRNA, site-specifically engineered into the gene of interest, and uses the charged tRNA to site-specifically incorporate the unnatural amino acid into the protein.

molecular biology and solid-phase peptide synthesis (Figure 2). While the original rendition of this method necessitated additional point mutations beyond the UAA of interest, this problem has been addressed by recent efforts in the development of traceless ligations [13,14].

It is important to highlight that, in addition to the two well-established methods described herein, there exist many emerging technologies for UAA incorporation. Noteworthy examples include reconstituted, cell-free translation systems [15] and flexizyme technology [16,17]. While both show promise for the simultaneous incorporation of multiple different UAAs with high fidelity, these technologies are presently limited to synthesis of peptides and are not yet suitable for incorporation of UAAs into large proteins. Thus, these methods, while promising, are not discussed at length in this brief review.

**Figure 2. Expressed protein ligation (EPL)**




The amino-terminal peptide of the target protein is expressed *in vivo* as a fusion protein with an intein that has been modified to eliminate the branched ligation. Thiols (R = benzyl, phenyl, alkyl, $CH_2CH_2SO_3Na$) are used to cleave the intein and generate a reactive carboxy-terminal thioester. This construct is then reacted with a peptide generated by solid phase peptide synthesis that contains an amino-terminal cysteine and an unnatural amino acid(s) in the position of interest. Reaction between the two constructs results in thiotransesterification, followed by S to N acyl shift to regenerate a peptide. There are many variations on this general theme.

## Major recent advances

As the methods described above are evolving, a rapidly increasing number of protein studies involving UAAs are being reported. Below, three studies are highlighted in which UAA incorporation has provided a unique opportunity to address mechanistic questions.

Over the past 10 years, ion channels and receptors have been a particularly interesting target for UAA incorporation. By microinjection of *in vitro*-charged tRNAs and mRNA bearing a nonsense codon, Dougherty, Lester and coworkers have conducted whole-cell electrophysiological studies in *Xenopus* oocytes probing the activation of ion channels of the Cys-loop superfamily [10]. Site-specific incorporation of a series of fluorinated tyrosine or tryptophan analogues at conserved aromatic amino acids in the agonist binding sites of $GABA_c$ ($\gamma$-aminobutyric acid receptor C) and nicotinic acetylcholine receptors, respectively, revealed the importance of cation-$\pi$ interactions in agonist binding [18,19]. The success of UAA studies in ion channels/receptors is intricately linked to the sensitivity of electrophysiology techniques, which has allowed measurable readout in cells containing only small amounts of the mutant proteins. Recently, nonsense codon suppression has been extended to study the activation of G-protein-coupled receptors (GPCRs). Thirteen UAAs were incorporated into the M2 muscarinic acetylcholine and D2 dopamine GPCRs via microinjection of charged tRNAs into *Xenopus* cells, and the role of specific residues in ligand binding was assessed by an optimized assay in which activation of a downstream G-protein-coupled $K^+$ channel was monitored electrophysiologically [20]. Interestingly, GPCRs in yeast have also been targeted recently for UAA studies. Incorporation of the photoactivatable amino acid *p*-benzoyl-L-phenylalanine (Bpa) into the ligand-binding site of the *Saccharomyces cerevisiae* GPCR Ste2p was demonstrated using an evolved, orthogonal tRNA-RS pair specific for Bpa [21]. Photolysis of the mutant protein resulted in the capture of Ste2p with its $\alpha$ factor ligand.

The use of EPL has contributed significantly to understanding structure-function relationships in ion channels as well. High-resolution crystal structures of the potassium channel KcsA led to the hypothesis that the channel's ion selectivity was derived from backbone carbonyl groups that lined the pore, or selectivity filter, region of this integral membrane protein [22]. By EPL, Muir and colleagues [23] reported the semisynthesis of a truncated KcsA in which the amino-terminal 73 amino acids were expressed recombinantly and the carboxy-terminal 52 residues, which encompass the entire selectivity filter, were generated by solid-phase peptide

synthesis. Semisynthetic KscA channels were generated in which D-amino acids were introduced in the selectivity filter to report on the role of a left-handed helix in mediating ion conductivity. Likewise, semisynthetic channels with amide-to-ester backbone mutations provided insight into how the energetics of different binding configurations contribute to maintaining an optimized ion current [24].

Both nonsense suppression and EPL methods have been utilized in mechanistic studies of the class Ia ribonucleotide reductases (RNRs). RNRs catalyze the reduction of all four nucleotides to their 2′-deoxy analogues, and thus play a key role in modulating the relative ratios and absolute amounts of deoxynucleotide pools necessary for DNA synthesis and repair. Class Ia RNRs utilize a unique mechanism of long-range proton-coupled electron transfer (PCET) through transient aromatic amino acid radicals to generate an active site protein radical, which then initiates substrate reduction. This oxidation occurs over 35 Å and through two protein subunits, and is proposed to involve the participation of three tyrosines that act as oxidation 'stepping stones' [25,26]. Initial studies focused on the participation of $Y_{356}$, located in the conformationally flexible carboxy-terminal tail of one of the subunits ($\beta$). Because of its location in the protein (residue 356 of 375 amino acids), this tyrosine was targeted for UAA replacement via EPL. To date, more than a half-dozen UAAs with varying reduction potentials and p$K_a$s have been incorporated into the $\beta$ subunit made semisynthetically to examine the role of $Y_{356}$ in long-range PCET [27-29]. Two additional tyrosines implicated in this mechanism are $Y_{730}$ and $Y_{731}$ of the second subunit ($\alpha$). These have been the successful target of site-specific UAA incorporation via the *in vivo* orthogonal tRNA-RS nonsense suppression method [2,30]. Two UAAs, the redox trap 3-aminotyrosine and the p$K_a$ probe 3-nitrotyrosine, have been incorporated at positions 730 and 731, with isolation of the mutant proteins occurring on the 100 mg scale. The incorporation of these UAAs will allow for the unprecedented study of the mechanism of PCET, previously inaccessible due to a rate-limiting conformational change in wild-type RNR.

## Future directions

While all methods have certain niches in which they prove most useful for problem solving, in our opinion, *in vivo* nonsense suppression is the method that demonstrates the most potential for future development and widespread applicability. Methodology development to date has been most extensively focused on incorporation of UAAs into proteins expressed in *Escherichia coli*. This is no coincidence, as the bacterial translational machinery is simpler and the details better understood than in eukaryotic systems. Undoubtedly, as our understanding of translation continues to expand in both prokaryotic and eukaryotic systems, new methods will evolve to take advantage of this knowledge.

Recent methods developed by Wang and Wang [31] to increase expression levels of proteins containing UAAs in *S. cerevisiae* provide a case in point. In this organism, reported expression levels of UAA-containing proteins are substantially lower than in *E. coli* expression systems. Protein yields in the yeast-based expression systems were increased significantly by improvements in the expression efficiency of the orthogonal prokaryotic tRNA and the intracellular stability of the nonsense codon-containing mRNA. Similarly, order-of-magnitude improvements in protein yields were reported for an expression system in *Pichia pastoris* in which the genes encoding the orthogonal tRNA-RS were incorporated into the organism's genome under the control of optimized promoters [32].

This technology would also benefit from a more efficient means of evolving orthogonal tRNA-RS pairs for unique UAAs. At present, selections using *E. coli* as the host organism occur on the month timescale, typically requiring four to six rounds of positive and negative selection and a battery of control reactions. Recently, Melançon and Schultz [33] described a single plasmid selection system in which positive and negative selections were conducted via expression of one fusion gene, thus eliminating the need for plasmid isolation and re-transformation in between selection rounds and greatly reducing the time and labor involved in RS evolution.

Improving the suppression efficiency of nonsense codons has been the focus of recent work by the Chin lab. One problem inherent to the *in vivo* nonsense suppression method is the competition between mutant, orthogonal tRNAs and endogenous release factors (RFs) for interaction with the nonsense codon during mRNA translation. This competition often results in high levels of truncated proteins that can complicate purification of proteins and biochemical analyses. Chin and coworkers have reported the evolution of an orthogonal *E. coli* ribosome optimized in the 530 loop of the 16s rRNA for preferential interaction with the orthogonal tRNA over RF-1 [34]. The ribosome translates an orthogonal, nonsense codon-containing mRNA, identified by its modified ribosomal binding site [35], with high specificity and fidelity. This technology has been demonstrated to increase suppression efficiency in genes containing single nonsense codons by up to 3-fold, and genes containing two nonsense codons by up to 20-fold. These orthogonal ribosome-mRNA pairs

serve as a starting point for the engineering of complete orthogonal gene expression systems in *E. coli* [35,36].

Finally, evolution of a mutant orthogonal tRNA-RS pair for expression in mammalian systems must be modified relative to the *E. coli* selection scheme due to difficulty in generating large, stable synthetase libraries and inefficient survival rates relative to prokaryotes in consecutive selection rounds. The emerging strategy of a bacterial-mammalian 'shuttle' shows promise in addressing this issue. By this method, a pyrrolysyl tRNA-RS pair from the mutually orthogonal organism *Methanosarcina mazei* was evolved in the *E. coli* host and then transfected into mammalian cells for protein expression [37].

Given the excitement expressed by the research community for this technology and the increasing number of young scientists dedicated to its development, the methodology for robust, efficient incorporation of UAAs into proteins is primed to evolve very rapidly in the immediate future. We anticipate the day when ready-to-use tRNA-RS pairs are at our disposal to tackle biological problems both *in vitro* and *in vivo*.

## Abbreviations

Bpa, *p*-benzoyl-L-phenylalanine; EPL, expressed protein ligation; $GABA_C$; $\gamma$-aminobutyric acid receptor C; GPCR, G-protein-coupled receptor; PCET, proton-coupled electron transfer; RNR, ribonucleotide reductase; RF, release factor; RS, tRNA synthetase; UAA, unnatural amino acid.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgments

This work was supported by NIH grant GM29595 (to JS).

## References

1. Zoller MJ, Smith M: **Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template.** *Methods Enzymol* 1987, **154:**329-50.
2. Muir TW, Abelson JN (Eds): **Non-natural amino acids.** In *Methods Enzymol* 2009, **462:**1-264.
3. Xie J, Schultz PG: **An expanding genetic code.** *Methods* 2005, **36:**227-38.
4. Wang L, Xie J, Schultz PG: **Expanding the genetic code.** *Annu Rev Biophys Biomol Struct* 2006, **35:**225-49.
5. Muralidharan V, Muir TW: **Protein ligation: an enabling technology for the biophysical analysis of proteins.** *Nat Methods* 2006, **3:**429-38.
6. Flavell RR, Muir TW: **Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology.** *Acc Chem Res* 2009, **42:**107-16.
7. Rodnina MV, Wintermeyer W: **Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms.** *Annu Rev Biochem* 2001, **70:**415-35.
8. Zaher HS, Green R: **Quality control by the ribosome following peptide bond formation.** *Nature* 2009, **457:**161-6.

   F1000 Factor 6.0 *Must Read*
   Evaluated by Tom Rapoport 04 Feb 2009
9. 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 USA* 1997, **94:**10092-7.
10. Nowak MW, Gallivan JP, Silverman SK, Labarca CG, Dougherty DA, Lester HA: ***In vivo* incorporation of unnatural amino acids into ion channels in *Xenopus* oocyte expression system.** *Methods Enzymol* 1998, **293:**504-29.
11. Paulus H: **Protein splicing and related forms of protein autoprocessing.** *Annu Rev Biochem* 2000, **69:**447-96.
12. Dawson PE, Muir TW, Clark-Lewis I, Kent SB: **Synthesis of proteins by native chemical ligation.** *Science* 1994, **266:**776-9.
13. Tam A, Soellner MB, Raines RT: **Water-soluble phosphinothiols for traceless staudinger ligation and integration with expressed protein ligation.** *J Am Chem Soc* 2007, **129:**11421-30.
14. Lockless SW, Muir TW: **Traceless protein splicing utilizing evolved split inteins.** *Proc Natl Acad Sci USA* 2009, **106:**10999-1004.
15. Shimizu Y, Kanamori T, Ueda T: **Protein synthesis by pure translation systems.** *Methods* 2005, **36:**299-304.
16. Murakami H, Ohta A, Ashigai H, Suga H: **A highly flexible tRNA acylation method for non-natural polypeptide synthesis.** *Nat Methods* 2006, **3:**357-9.

   F1000 Factor 6.4 *Must Read*
   Evaluated by Gottfried Otting 02 Jun 2006, Virginia Cornish 23 Jan 2007
17. Ohta A, Yamagishi Y, Suga H: **Synthesis of biopolymers using genetic code reprogramming.** *Curr Opin Chem Biol* 2008, **12:**159-67.
18. Lummis SC, Beene DL, Harrison NJ, Lester HA, Dougherty DA: **A cation-$\pi$ binding interaction with a tyrosine in the binding site of the GABAc receptor.** *Chem Biol* 2005, **12:**993-7.
19. Xiu X, Puskar NL, Shanata JA, Lester HA, Dougherty DA: **Nicotine binding to brain receptors requires a strong cation-pi interaction.** *Nature* 2009, **458:**534-7.
20. Torrice MM, Bower KS, Lester HA, Dougherty DA: **Probing the role of the cation-pi interaction in the binding sites of GPCRs using unnatural amino acids.** *Proc Natl Acad Sci USA* 2009, **106:**11919-24.
21. Huang LY, Umanah G, Hauser M, Son C, Arshava B, Naider F, Becker JM: **Unnatural amino acid replacement in a yeast G protein-coupled receptor in its native environment.** *Biochemistry* 2008, **47:**5638-48.
22. Zhou Y, Morais-Cabral JH, Kaufman A, MacKinnon R: **Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 Å resolution.** *Nature* 2001, **414:**43-8.

   F1000 Factor 10.0 *Exceptional*
   Evaluated by Gary Yellen 07 Nov 2001, Tom Rapoport 15 Nov 2001, Christopher Miller 16 Nov 2001, Eaton E Lattman 04 Dec 2001
23. Valiyaveetil FI, Leonetti M, Muir TW, MacKinnon R: **Ion selectivity in a semisynthetic K+ channel locked in the conductive conformation.** *Science* 2006, **314:**1004-7.
24. Valiyaveetil FI, Sekedat M, MacKinnon R, Muir TW: **Structural and functional consequences of an amide-to-ester substitution in the selectivity filter of a potassium channel.** *J Am Chem Soc* 2006, **128:**11591-9.
25. Uhlin U, Eklund H: **Structure of ribonucleotide reductase protein R1.** *Nature* 1994, **370:**533-9.
26. Stubbe J, Nocera DG, Yee CS, Chang MC: **Radical initiation in the class I ribonucleotide reductase: long-range proton-coupled electron transfer?** *Chem Rev* 2003, **103:**2167-201.

27. Yee CS, Seyedsayamdost MR, Chang MC, 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.** *Biochemistry* 2003, **42:**14541-52.

28. 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:**2522-3.

F1000 Factor 3.0 *Recommended*
Evaluated by Mark Nelson 01 Jun 2006

29. 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:**1225-35.

30. Seyedsayamdost MR, Xie J, Chan CT, Schultz PG, Stubbe J: **Site-specific insertion of 3-aminotyrosine into subunit α2 of *E. coli* ribonucleotide reductase: direct evidence for involvement of Y730 and Y731 in radical propagation.** *J Am Chem Soc* 2007, **129:**15060-71.

F1000 Factor 6.0 *Must Read*
Evaluated by Marc Fontecave 21 Feb 2008

31. Wang Q, Wang L: **New methods enabling efficient incorporation of unnatural amino acids in yeast.** *J Am Chem Soc* 2008, **130:**6066-7.

32. Young TS, Ahmad I, Brock A, Schultz PG: **Expanding the genetic repertoire of the methylotrophic yeast *Pichia pastoris*.** *Biochemistry* 2009, **48:**2643-53.

33. Melançon CE 3rd, Schultz PG: **One plasmid selection system for the rapid evolution of aminoacyl-tRNA synthetases.** *Bioorg Med Chem Lett* 2009, **19:**3845-7.

34. Wang K, Neumann H, Peak-Chew SY, Chin JW: **Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion.** *Nat Biotechnol* 2007, **25:**770-7.

35. Rackham O, Chin JW: **A network of orthogonal ribosome x mRNA pairs.** *Nat Chem Biol* 2005, **1:**159-66.

36. An W, Chin JW: **Synthesis of orthogonal transcription-translation networks.** *Proc Natl Acad Sci USA* 2009, **106:**8477-82.

37. Chen PR, Groff D, Guo J, Ou W, Cellitti S, Geierstanger BH, Schultz PG: **A facile system for encoding unnatural amino acids in mammalian cells.** *Angew Chem Int Ed Engl* 2009, **48:**4052-5.

F1000 Factor 3.0 *Recommended*
Evaluated by Jan Grunewald 23 Apr 2009