Review Article

Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells

Alexander R. Nödling, Luke A. Spear, Thomas L. Williams, Louis Y.P. Luk and Yu-Hsuan Tsai

School of Chemistry, Cardiff University, Cardiff, Wales, United Kingdom

Correspondence: Yu-Hsuan Tsai (tsaiy5@cardiff.ac.uk)

Genetic code expansion allows unnatural (non-canonical) amino acid incorporation into proteins of interest by repurposing the cellular translation machinery. The development of this technique has enabled site-specific incorporation of many structurally and chemically diverse amino acids, facilitating a plethora of applications, including protein imaging, engineering, mechanistic and structural investigations, and functional regulation. Particularly, genetic code expansion provides great tools to study mammalian proteins, of which dysregulations often have important implications in health. In recent years, a series of methods has been developed to modulate protein function through genetically incorporated unnatural amino acids. In this review, we will first discuss the basic concept of genetic code expansion and give an up-to-date list of amino acids that can be incorporated into proteins in mammalian cells. We then focus on the use of unnatural amino acids to activate, inhibit, or reversibly modulate protein function by translational, optical or chemical control. The features of each approach will also be highlighted.

Introduction

Knowledge of protein function is of pivotal importance to life science research. It can guide conventional drug development programmes and lead to novel strategies to address currently non-targetable systems [1–3]. In order to understand the precise role and interacting network of a protein, it is essential to analyse it within its native environment. For a mammalian protein, its function often also depends on its host cell (e.g. cell type and cell cycle stage), specific subcellular location and post-translational modifications. In addition, a protein of interest often exists in the presence of other closely related homologues (e.g. proteins within the same family), making it difficult to decipher the precise function of a specific protein in cells. Targeting the protein by small-molecule inhibition is often not possible in these cases, as protein homologues will also be affected. To tackle this problem, over the last two decades there has been a drive to develop and refine the technique of genetic code expansion which allows researchers to exploit the cellular translation machinery for site-specific incorporation of unnatural (non-canonical) amino acids into target proteins [4–14]. Consequently, this enables the use of building blocks beyond the 20 canonical amino acids and incorporation of unnatural amino acids with unprecedented functionality into target proteins in live cells. The repurposing of the translational machinery by this approach has paved the way for revealing the functions of proteins under physiological conditions [15–19]. For example, the technique can be used to site-specifically introduce an unnatural amino acid into the homologue of interest, whereby unique functionality (on the unnatural amino acid) can be used for selective activation, inhibition, or reversible regulation of the target homologue [7].

At the molecular level, the mechanism of protein translation is highly conserved in all organisms, where the cellular machinery 'translates' every nucleotide triplet as a codon consecutively on the mRNA
Figure 1. Mechanism of genetic code expansion for site-specific incorporation of an unnatural amino acid by amber suppression

Figure 2. Allowed and not allowed reactivities between the orthogonal and endogenous aaRS/tRNA pairs
(A) Matching amino acid and aaRS/tRNA pairs; (B) mismatched amino acids; (C) mismatched aaRS/tRNA pairs.

into the corresponding amino acid. In nature, the endogenous aminoacyl-tRNA synthetase (aaRS)/tRNA pairs within the cell decode 61 of the total 64 codons to 20 canonical amino acids. The remaining three codons (UAG, UGA and UAA) are used for translation termination, and hence they are also known as ‘stop’ codons. In order to achieve site-specific incorporation of an unnatural amino acid, an orthogonal aaRS/tRNA pair is needed, which must decode a codon that does not correspond to any canonical amino acid, a so-called blank codon (Figure 1). Stop codons are most commonly used as a blank codon in genetic code expansion, and decoding of a stop codon is known as ‘suppression’ because it suppresses the translation termination. The amber stop codon (UAG) is often used as the blank codon due to its minimal occurrence in most organisms.

Within the concept of genetic code expansion, ‘orthogonality’ refers to the non-reactivity of the orthogonal aaRS/tRNA pair with the endogenous pair and canonical amino acids in the host cell. The orthogonal synthetase must only acylate the orthogonal tRNA with the designated unnatural amino acid; neither canonical amino acids nor endogenous tRNAs are substrates of the orthogonal synthetase; similarly, neither the unnatural amino acid nor orthogonal tRNA is a substrate of the endogenous synthetases (Figure 2).

Besides the amber codon, other stop codons [20–26] and different four-nucleotide codons [27,28] have been used as a blank codon. The use of four-nucleotide codons expands the theoretical codon numbers from $4^3$ (64) to $4^4$ (256) so that multiple different unnatural amino acids can be incorporated at the same time. However, decoding a four-nucleotide codon by the ribosome is less efficient than decoding the normal three-nucleotide codons. Although this issue has been addressed in Escherichia coli through ribosome engineering [29–31], the lower efficiency in decoding four-nucleotide codons remains an issue in mammalian systems [27,28].

To date, many unnatural amino acids (1–110, Table 1) can be site-specifically incorporated into proteins produced by mammalian cells using genetic code expansion [5,32]. While the amino acids are structurally diverse, the majority of them can be incorporated through only a few orthogonal synthetases and their mutants. The Pyrrolysyl-tRNA synthetase (PylRS)/tRNA<sub>Pyl</sub> pairs from archaea species Methanosarcina barkeri (Mb) and
Figure 3. Use of amber suppression to switch on protein production

(A) Absence of the unnatural amino acid leads to recognition of the UAG codon for translation termination. (B) Addition of the unnatural amino acid leads to amber suppression and successful production of the full-length and functional protein.

Methanosarcina mazei (Mm) have proven to be extraordinarily useful pairs [4]. The tRNA^Pyl^ naturally recognises the UAG codon and thus engineering of this tRNA is not needed. In addition, this pair is orthogonal in both E. coli and mammalian cells; hence, it facilitates the engineering of PylRS in E. coli and subsequently using the engineered PylRS mutant for incorporation of the designated unnatural amino acid in mammalian systems. As shown in Table 1, a wide range of amino acids has been incorporated into proteins in mammalian cells through only a few point mutations on the PylRS gene.

Many engineered E. coli aaRS/tRNA pairs have also been used as orthogonal pairs in mammalian cells. The most successful ones are the E. coli tyrosine, leucine and tryptophan pairs [33]. However, as all these synthetases naturally recognise a canonical amino acid, it is necessary to abolish their natural activity towards the canonical amino acid and to recognise only the designated unnatural amino acid. As it is technically difficult to perform directed evolution in mammalian cells due to low efficiency in transfection and screening, synthetase engineering is normally carried out in E. coli [34,35] or yeast [15,36,37] so that large mutant libraries can be easily screened. It is also necessary to modify the E. coli tRNA so that it decodes a blank codon instead of a codon corresponding to a canonical amino acid.

Based on the simplicity of the established methodology [38–40] and the promiscuity of many orthogonal synthetases towards different unnatural amino acids (vide infra) [41], the number of genetically incorporable unnatural amino acids has steadily increased. In addition, some orthogonal aaRS/tRNA pairs are mutually orthogonal [21,24–26,42] and can be used at the same time to incorporate multiple different unnatural amino acids into a protein of interest.

As recent reviews cover fundamental aspects of genetic code expansion [4,6,9,13], the engineering of new orthogonal synthetases [8], and general [5,10,11,14] or specific [7,12] applications of genetic code expansion in eukaryotic systems, we will focus on recent advances and applications of genetic code expansion for controlling protein function in mammalian cells through translational, optical or chemical means.

Translational control by amber suppression

Genetic code expansion by unnatural amino acid incorporation in response to an amber stop codon provides the simplest way to ‘switch on’ protein production. In this case, an amber stop codon is placed into the gene of interest (Figure 3) for incorporation of the unnatural amino acid into a permissive site of the target protein [43]. In the absence of the designated unnatural amino acid, protein translation stops prematurely at the amber stop codon, generating truncated and non-functional protein product and thus giving an effect as a nonsense mutation (Figure 3A). In the presence of the unnatural amino acid, the orthogonal tRNA is acylated and decodes the amber codon, leading to amber suppression and generation of full-length, functional protein product (Figure 3B). Thus, simple addition of the unnatural amino acid into the growth medium ‘switches on’ the protein production and function [44]. In contrast with commonly used systems for inducible mammalian protein expression (e.g. the tetracycline transcriptional transactivation) [45], the lag time is shorter in the translational control by amber suppression (i.e. time for translation and folding) than the gene activation approaches (i.e. time for transcription, mRNA processing, translation and folding). Amber suppression also allows a more stringent control, as the background activity (if any) can be further
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date

| Amino acid     | aaRS                           | Mutations | tRNA        | Application          |
|----------------|--------------------------------|-----------|-------------|----------------------|
| Cysteine and selenocysteine derivatives |                                |           |             |                      |
| ![Cysteine](image) | EcLeuRS [24] M40I T252A Y499I Y527A H529G | Ec tRNA<sub>Leu</sub><sup>CUA</sup> | Method development |
| ![Selenocysteine](image) | EcLeuRS [24] M40I T252A Y499I Y527A H529G | Ec tRNA<sub>Leu</sub><sup>CUA</sup> | Method development |
| ![Modified Cysteine](image) | EcLeuRS [24] M40I T252A Y499I Y527A H529G | Ec tRNA<sub>Leu</sub><sup>CUA</sup> | Method development |
| ![Modified Selenocysteine](image) | EcLeuRS [24] M40I T252A Y499I Y527A H529G | Ec tRNA<sub>Leu</sub><sup>CUA</sup> | Method development |

Continued over
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Structure](image1) | MbPylRS [73,74] | N311Q C313A V366M | MbtRNA<sub>CUA</sub> | Photoactivation |
| ![Structure](image2) | MbPylRS [75] | M241F A267S Y271C L274M | MbtRNA<sub>CUA</sub> | Photoactivation |
| ![Structure](image3) | MbPylRS [76] | C313W W382T | MbPyltRNA | Method development |
| ![Structure](image4) | MbPylRS [40] | L274A C313S Y349F | MbtRNA<sub>CUA</sub> | Photocrosslinking |
| ![Structure](image5) | MbPylRS [40] | L274A C313S Y349F | MbtRNA<sub>CUA</sub> | Photocrosslinking |
| ![Structure](image6) | EcTyrRS [77] | Y37V D182S F183M D265R | Ec<sub>Tr</sub>|Method development |

Phenylalanine derivatives
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![26](image) | EcTyrRS | Y37V D182S F183M D265R [77] | EcRNA\textsubscript{Tyr}[\textsubscript{CUA}] [35, 37, 43, 77] | Method development [35, 37, 77] |
| ![27](image) | EcTyrRS [77] | Y37V D182S F183M D265R | EcRNA\textsubscript{Tyr}[\textsubscript{CUA}] [35, 77, 78, 81, 86, 90] | Protein engineering [43, 49, 55, 83] |
| ![28](image) | EcTyrRS | Y37L D182S F183A L186A D265R [78, 81, 84, 85] | EcRNA\textsubscript{Tyr}[\textsubscript{CUA}] [16,17,37,38,43,49,55,68,77,79,80,82,83,86–89,91–96] | Bioorthogonal labelling [38, 79, 83, 87–89, 96] |
| ![29](image) | EcTyrRS\textsubscript{CUA} [25] | Y37L D182S F183M L186A [16,17,37,38,43,49,55,68,79,80,82,83,86–89,91–96] | EcRNA\textsubscript{Tyr}[\textsubscript{CUA}] [35, 77, 78, 81, 86, 90] | Photocrosslinking [38, 68, 81, 84, 85, 91–95] |
| ![30](image) | EcTyrRS\textsubscript{CUA} [35, 97] | Y37I D182S F183I L186E D265R [35] | EcRNA\textsubscript{Tyr}[\textsubscript{CUA}] [35, 77, 78, 81, 86, 90] | Protein engineering [43, 49, 55, 83] |
| ![31](image) | MmPylRS [50] | L301M Y306L L309A C348F | Mm\textsubscript{PylRNA}[\textsubscript{CUA}] | Method development [35] |
| ![32](image) | MmPylRS [98] | N346A C348A | Mm\textsubscript{PylRNA}[\textsubscript{CUA}] | Method development |
| ![33](image) | MmPylRS [98] | N346A C348A | Mm\textsubscript{PylRNA}[\textsubscript{CUA}] | Method development |

Continued over
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Cl](image) | MmPyRS [98] | N346A C348A | MmRNA<sub>CUA</sub> | Method development |
| ![Br](image) | MmPyRS [98] | N346A C348A | MmRNA<sub>CUA</sub> | Method development |
| ![I](image) | MmPyRS [98] | N346A C348A | MmRNA<sub>CUA</sub> | Method development |
| ![NC](image) | MmPyRS [98] | N346A C348A | MmRNA<sub>CUA</sub> | Method development |
| ![NO2](image) | MmPyRS [98] | N346A C348A | MmRNA<sub>CUA</sub> | Method development |
| ![O](image) | EcTyrRS<sub>UCA</sub> [21,35,37,77,83,87,90,99,100] | Y37I N165G D182G F183M L186A D265R [37,43,87,100] | EcRNA<sub>CUA</sub> [21,35,77,90] | Bioorthogonal labelling [83,87,100] Method development [25,36,37,77,83,90,99] Protein engineering [21,43] |
| ![O](image) | EcTyrRS<sub>UCA</sub> [21,35,37,77,83,87,90,99,100] | Y37I N165G D182G F183M L186A D265R [37,43,87,100] | EcRNA<sub>CUA</sub> [21,35,77,90] | Bioorthogonal labelling [83,87,100] Method development [25,36,37,77,83,90,99] Protein engineering [21,43] |
| ![O](image) | EcTyrRS<sub>UCA</sub> [25] | Y37I N165G D182G F183M L186A D265R [35] | EcRNA<sub>CUA</sub> [25] | Chemical crosslinking [84,101] Method development [101] |
| ![O](image) | EcTyrRS [84,101] | Y37I N165G D182G F183M L186A D265R | EcRNA<sub>UCA</sub> [84,101] | Chemical crosslinking [84,101] Method development [101] |
| ![O](image) | EcTyrRS [15,78,103] | Y37G N165G D182G L186A D265R [15,78,103] | EcRNA<sub>UCA</sub> [15,78,103] | Mechanistic studies [15] Method development [37,78,95,99,103,106] Photocrosslinking [85,92,102,104,106] Photosensitization [58] |
| ![O](image) | EcTyrRS [15,78,103] | Y37G N165G D182G L186A D265R [15,78,103] | EcRNA<sub>UCA</sub> [15,78,103] | Mechanistic studies [15] Method development [37,78,95,99,103,106] Photocrosslinking [85,92,102,104,106] Photosensitization [58] |
| ![O](image) | EcTyrRS [15,78,103] | Y37G N165G D182G L186A D265R [15,78,103] | EcRNA<sub>UCA</sub> [15,78,103] | Mechanistic studies [15] Method development [37,78,95,99,103,106] Photocrosslinking [85,92,102,104,106] Photosensitization [58] |
| ![O](image) | EcTyrRS [15,78,103] | Y37G N165G D182G L186A D265R [15,78,103] | EcRNA<sub>UCA</sub> [15,78,103] | Mechanistic studies [15] Method development [37,78,95,99,103,106] Photocrosslinking [85,92,102,104,106] Photosensitization [58] |
| ![O](image) | MmPyRS [106] | A302T N346T C348T W417C [106] | MmRNA<sub>UCA</sub> [106] | Method development |

Continued over
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Histidine derivative](image) | MnPyRS [106] | A302T N346G C348T V401I W417Y | Mm tRNA<sub>Pyl</sub> <sub>CUA</sub> | Method development |
| ![Histidine derivative](image) | EcLeuRS [107–109] | L38F M40G L41P Y499V Y500L Y527A H537E L538S F541C A560V | Ec tRNA<sub>Leu</sub> <sub>CUA</sub> | Method development [107,108] Spectroscopic probe [109] |
| ![Histidine derivative](image) | MnPyRS [110] | N346Q C348S V401G W417T | Mm tRNA<sub>Pyl</sub> <sub>CUA</sub> | Spectroscopic Probe |
| ![Histidine derivative](image) | MbPyRS [111] | L270I Y271F L274G C313F Y349F | Mb tRNA<sub>Pyl</sub> <sub>CUA</sub> | Photoswitching |
| ![Histidine derivative](image) | MbPyRS [111] | L270F L274M N311G C313G Y349F | Mb tRNA<sub>Pyl</sub> <sub>CUA</sub> | Photoswitching |
| ![Histidine derivative](image) | MbPyRS [111] | L270F L274M N311G C313G Y349F | Mb tRNA<sub>Pyl</sub> <sub>CUA</sub> | Photoswitching |
| ![Histidine derivative](image) | MnPyRS [56,112] | A302T L309S N346V C348G | Mm tRNA<sub>Pyl</sub> <sub>CUA</sub> | Method development [112] Photoswitching [56] |
| ![Histidine derivative](image) | MaPyRS [28] | L121M L125I Y126F M129A V168F | Ma tRNA<sub>Pyl</sub> <sub>CUA</sub> | Method development |
| ![Histidine derivative](image) | MbPyRS [113] | L270I Y271F L274G C313F Y349F | Mb tRNA<sub>Pyl</sub> <sub>CUA</sub> | Method development |

Histidine derivatives

© 2019 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Amino acid](image) | MbPylRS [113] | L270I Y271F L274G C313F Y349F | MbtRNA<sub>CUA</sub> | Method development |

**Lysine derivatives**

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Amino acid](image) | MbPylRS [114] | L266M L270I Y271F L274A C313F | MbtRNA<sub>CUA</sub> | Method development |
| ![Amino acid](image) | MbPylRS [115–117] | D76G L266V L270I Y271F L274A C313F [115] D76G L266M L270I Y271F L274A C313F [116,117] | MbtRNA<sub>CUA</sub> | Method development [44,50,115,117,118] Spectroscopic probe [116] |
| ![Amino acid](image) | MmPylRS [44,50,118] | L301M Y306L L309A C348F [44,50] | MmrRNA<sub>CUA</sub> [50,118] | Method development |
| ![Amino acid](image) | MmPylRS [50] | L301M Y306L L309A C348F | MmrRNA<sub>CUA</sub> [44] | Method development |
| ![Amino acid](image) | MbPylRS [119] | D76G L266M L270I Y271F L274A C313F | MbtRNA<sub>CUA</sub> | Method development |
| ![Amino acid](image) | MbPylRS [77,120] | L274A C313F Y349F [120] wt [77] | MbtRNA<sub>CUA</sub> | Method development |

Continued over
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| Mb | PylRS [121] | Y271M L274A C313A | Mb tRNAPyl | Photocrosslinking |
| Mm | PylRS [122] | Y306V L309A C348F Y384F | Mm tRNAPyl | Photocrosslinking |
| Ma | PylRS [26] | wt | Mm tRNAPyl | Method development |
| Mb | PylRS [21,25,44,125,129,130] | wt | Mm tRNAPyl | Method development |
| Mb | PylRS [21,25,48,77,132–136] | L274A C313S Y349F [131] | Mb tRNAPyl | Bioorthogonal labelling |
| Mb | PylRS [57,69,77,138,139] | wt | Mb tRNAPyl | Bioorthogonal labelling |
| Mm | PylRS [25,130] | wt | Mm tRNAPyl | Bioorthogonal labelling |
| Mm | PylRS [130] | wt | Mm tRNAPyl | Bioorthogonal labelling |
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Diagram](63) | MbPyRS [131,140] | L274A C313S Y349F | *MbtRNA*<sup>CUA</sup> | Bioorthogonal labelling |
| ![Diagram](64) | MnPyRS [141] | wt | *MmntRNA*<sup>CUA</sup> | Method development |
| ![Diagram](65) | MbPyRS [140] | wt | *MbtRNA*<sup>CUA</sup> | Method development |
| ![Diagram](66) | MnPyRS [118,142,143] | R61K G131E L309A C348V Y384F [118] Y306A Y384F [142] R61K G131E Y306A Y384F [143] | *MmntRNA*<sup>CUA</sup> [118,142] *MmntRNA*<sup>CUA</sup> BU25C [143] | Method development [118,142,143] |
| ![Diagram](67) | MbPyRS [140] | Y271L L274A C313A Y349F | *MbtRNA*<sup>CUA</sup> | Method development [140,141] Photoactivation [61,144] |
| ![Diagram](68) | MbPyRS [145] | Y271M L274T C313A Y349F | *MbtRNA*<sup>CUA</sup> | Method development |

Continued over...
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [146] | Y271I274M C313A | MbtrNA<sub>Pyl</sub> | Method development |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [146] | Y271I274M C313A | MbtrNA<sub>Pyl</sub> | Method development |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [63] | Y271A Y349F | MbtrNA<sub>Pyl</sub> | Chemical decaging |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [62] | L274A C313S Y349F | MbtrNA<sub>Pyl</sub> | Bioorthogonal labelling |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [66,67,69,75,125,147–152] | M241F A267S Y271C L274M | MbtrNA<sub>Pyl</sub> | Method development |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [66,67,69,75,125,147–152] | M241F A267S Y271C L274M | MbtrNA<sub>Pyl</sub> | Photoactivation |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [153] | Y271A L274M | MbtrNA<sub>Pyl</sub> | Photoactivation |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [153] | Y271A L274M | MbtrNA<sub>Pyl</sub> | Photoactivation |
| ![Image](https://portlandpress.com/essaysbiochem/article-pdf/63/2/237/843204/ebc-2018-0042c.pdf) | MbPylRS [153] | Y271A L274M | MbtrNA<sub>Pyl</sub> | Method development |

Continued over
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| MbPylRS    | [154]| L266M L270I Y271L L274A C313 | MbtRNA\(^{PyL}\)\(_{CUA}\) | Method development |
| MbPylRS    | [24,25]| wt | MbtRNA\(^{PyL}\)\(_{CUA}\) [25] | |
| MmPylRS    | [22,69,115,123,155,156]| Wt [22,69,115,123,155,156] Y306A Y384F [155] | MmtRNA\(^{PyL}\)\(_{CUA}\) [22,155] MmtRNA\(^{PyL}\)\(_{CUA}\) [69,115,123,156] Mx1201tRNAPyl\(^{CUA}\) [69,115,123,156] | Imaging [123] Method development [22,24,25,69,115,155,156] |
| MmPylRS    | [124]| wt | MmtRNA\(^{PyL}\)\(_{CUA}\) | Bioorthogonal labelling |
| MbPylRS    | [77,157,158]| Wt [77,158] L274M 313A Y349F [157] | MbtRNA\(^{PyL}\)\(_{CUA}\) | Method development [77,155,159] Photocrosslinking [157,158] |
| MmPylRS    | [155,159]| Y306A Y384F | MmtRNA\(^{PyL}\)\(_{CUA}\) [159] MmtRNA\(^{PyL}\)\(_{CUA}\) [155] | |
| MmPylRS    | [159]| Y306A Y384F | MmtRNA\(^{PyL}\)\(_{CUA}\) | Method development |
| MbPylRS    | [132,140,160]| L274A C313S Y349F | MbtRNA\(^{PyL}\)\(_{CUA}\) | Method development [140] Photocrosslinking [132,160] Protein engineering [132] |
| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| ![Amino acid 82](image) | MmPyRS [159] | Y306A Y384F | Mm\text{tRNA}_{\text{CUA}}\text{P}_{\text{Y}} | Method development |
| ![Amino acid 83](image) | MmPyRS [142] | Y306A Y384F | Mm\text{tRNA}_{\text{CUA}}\text{P}_{\text{Y}} | Photocrosslinking |
| ![Amino acid 84](image) | MmPyRS [143] | R61K G131E Y306A Y384F | Mm\text{tRNA}_{\text{CUA}}\text{B}_{25C} | Photocrosslinking |
| ![Amino acid 85](image) | MmPyRS [18,39,59–61,123,155,161–166] | Y306A Y384F | Mm\text{tRNA}_{\text{CUA}}\text{B}_{25C} | Imaging [123,161,162,164,166] Chemical decaging [18,59–61] Chemical crosslinking [163] Method development [155,165] Protein labelling [59] |
| ![Amino acid 86](image) | MbPyRS [123,167,168] | Y271A L274M C313A | Mb\text{tRNA}_{\text{CUA}}\text{P}_{\text{Y}} | Bioorthogonal labelling [124,167] Imaging [123,168] Method development [169] |
| ![Amino acid 87](image) | MmPyRS [124,169] | Y306A Y384F | Mm\text{tRNA}_{\text{CUA}}\text{P}_{\text{Y}} | Method development |

Continued over
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| MbPylRS [167,168] | Y271A L274M C313A | MbrRNA<sub>CUA</sub> | Bioorthogonal labelling [167] |
| MbPylRS [169] | | MmmRNA<sub>CUA</sub> | Method development [167,169] |
| MmPylRS [24] | wt | MbrRNA<sub>CUA</sub> | Bioorthogonal labelling [39,127,131] |
| MmPylRS [39,127,131,169] | Y306A Y384F | MmmRNA<sub>CUA</sub> | Method development [24,169] |
| MmPylRS [39,127,131,169] | Y306A Y384F | MmmRNA<sub>CUA</sub> | Method development |
| MmPylRS [39,127,131,169] | Y306A Y384F | MmmRNA<sub>CUA</sub> | Method development |
| MbPylRS [19,64,140,168,171,172] | Y271M L274G C313A | MbrRNA<sub>CUA</sub> | Bioorthogonal labelling [39,131,167] |
| | [19,64,168,171,172] | MmmRNA<sub>CUA</sub> | Imaging [161,166,170] |
| | | | Method development [169] |
| | | | Chemical inhibition [64] |
| | | | Bioorthogonal labelling [39,131,167] |
| | | | Imaging [161,166,170] |
| | | | Method development [169] |
| | | | Protein engineering [143] |
| | | | Spectroscopic probe [19] |
| Tryptophan derivatives | | | |
| EcTrpRS [34] | S8A V144S V146A | EctRNA<sub>Tip</sub> | Method development |
| | S8A V144G V146C | | |

© 2019 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| Ec | TrpRS [34] | S8A V144S V146A | EctRNA • CUA | Method development |
| Ec | TrpRS [34] | S8A V144G V146C | EctRNA • CUA | Method development |
| Ec | TrpRS [34] | S8A V144G V146C | EctRNA • CUA | Method development |
| Ec | TyrRS [46] | Y37V Q195C | BstRNA • CUA | Method development |
| Ec | TyrRS [35,37] | Y37V D182S F183M D265R | EctRNA • CUA | Mechanistic studies [15] |
| Ec | TyrRSUCA [25] | Y37V D182S F183M | BstRNA • CUA | Method development [21,25,35,77,78,90] |
| Ec | TyrRS [77] | Y37V D182S F183M D265R | EctRNA • CUA | Method development |

Tyrosine derivatives

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| Ec | TyrRS [46] | Y37V Q195C | BstRNA • CUA | Method development |
| Ec | TyrRS [35,37] | Y37V D182S F183M D265R | EctRNA • CUA | Mechanistic studies [15] |
| Ec | TyrRSUCA [25] | Y37V D182S F183M | BstRNA • CUA | Method development [21,25,35,77,78,90] |
| Ec | TyrRS [77] | Y37V D182S F183M D265R | EctRNA • CUA | Method development |
Table 1 Overview of unnatural amino acids that have been successfully incorporated into proteins in mammalian cells and used for a variety of applications to date (Continued)

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| EcTyrRS  | [25,35,77] | Y37V D182S F183M D265R [77] | Ec tRNATyr CUA [25,35,77] | Method development |
|            |       | Y37S D182T F183M L186V [37] | * | |
|            |       | Y37V D182S F183M [25] | * | |
|            |       | Y37V D165G D182S F183M L186A D265R [35] | * | |
| MbPylRS   | [126,148,173] | L270F L274M N311G C313G Y349F [173] | Mb tRNAPyl CUA | Photoactivation |
|            |       | L270F L274M N311G C313G [126,148] | * | |
| MbPylRS   | [173] | L270F L274M N311G C313G Y349F | Mb tRNAPyl CUA | Photoactivation |
| MbPylRS   | [173] | L270F L274M N311G C313G Y349F | Mb tRNAPyl CUA | Photoactivation |
| MbPylRS   | [174] | N346T C348I Y384L W417K | Mmt tRNAPyl CUA | Bioorthogonal labelling |
| EcTyrRS   | [35] | Y37V D182S F183M D265R | Ec tRNATyr CUA | Method development |
|            |       | Y37V D165G D182S F183M L186A D265R | * | |

Miscellaneous unnatural amino acids

| Amino acid | aaRS | Mutations | tRNA | Application |
|------------|------|-----------|------|-------------|
| EcLeuRS   | [24,25] | M40I T252A Y499I Y527A H529G [24] | Ec tRNALeu CUA | Method development |
|            |       | E20K M40V L41S T252R Y499S Y527L H529G H537G [25] | * | |
minimised by including multiple amber codons into the gene of interest. The translational control by amber suppression approach is fully complementary to conventional genetic approaches (e.g. knockout, knockdown) that deplete a protein in cells to ‘switch off’ its function. In addition, the unnatural amino acid approach is reversible, as removing the unnatural amino acid in the growth medium will ‘switch off’ the translation of the protein of interest.

The translational switch-on process has been widely employed as a reporter system to test incorporation of new unnatural amino acids by using luminescent proteins like green fluorescent protein [46] or luciferase [20]. Upon successful incorporation, cells can emit light, whose intensity directly correlates to the unnatural amino acid incorporation efficiency. Apart from the reporter strategy, this approach has also been used to regulate function of other proteins, such as Cas9 for controllable gene editing in mouse embryos [47].

Besides the general use of the ‘translational activation’ approach to study protein function, this principle has been proven to be powerful in controlling virus replication [43,48,49]. By introducing TAG codons within the virus genes, viruses can only be generated using cell lines containing an orthogonal synthetase/tRNA$_{CUA}$ pair, and the resulting viruses are replication-incompetent in normal cells due to the lack of amber suppressor tRNA (Figure 4A) [43,48]. Such replication-incompetent viruses offer an additional tier of control for live-attenuated vaccines and significantly increase their safety. This concept has been further developed by including the genes encoding the orthogonal aaRS/tRNA pair into the viral genome (Figure 4B) [49]. In this case, viruses can be replicated in wild-type cells and the native hosts, as long as the unnatural amino acid is supplemented. Here, spatial control can also be achieved by local administration of the unnatural amino acid as demonstrated in examples of mice with an expanded genetic code [50,51]. Thus, the approach can be used for controlling viral vectors in gene therapy, where spatiotemporal virus replication and gene editing are highly desirable.

While the translational control approach is quite simple, the response is not instantaneous. There is always a lag time from when the unnatural amino acid is administered into the culture medium until the full-length protein is produced and folded. Similarly, depleting the unnatural amino acid in the growth medium will stop production of new proteins, but the protein function will only be completely switched off when all previously produced proteins are degraded in the cells. Thus, the kinetics of the switching off process largely depend on the half-life of the protein, so the response rate is the same as with genetic knockdown.

**Light-induced activation or inhibition**

The slower kinetics of the translational control approach limit its applicability to study biological processes where fast response is needed. This can be addressed by using light to unmask or modify unnatural amino acids and subsequently regulate protein function. Depending on the nature of the light-responsive group, it is possible to either activate, inhibit, or reversibly switch on/off protein function (Table 2). Unnatural amino acids containing a photocage (i.e. a photolabile protecting group) [52] have been widely used for protein activation. When a functional critical amino acid residue with the corresponding photocaged amino acid, the target protein becomes inactive; upon
Table 2 Overview of photocaged unnatural amino acids that have been incorporated in mammalian cell systems for activating protein function upon irradiation with light of specific wavelength $\lambda$

| Amino acid | Photocaging group (R) | System | Proteins | $\lambda_{\text{decag}}$ (nm) |
|------------|-----------------------|--------|----------|-----------------------------|
| ![Image of amino acid 17](image1.png) | ![Image of photocaging group](image2.png) | HEK293T | eGFP, potassium channel Kir2.1 [71] | 385 |
| ![Image of amino acid 18](image3.png) | ![Image of photocaging group](image4.png) | HEK293T | eGFP, mCherry, Npu DnaE intein, Src kinase [72] | 385 |
| ![Image of amino acid 19](image5.png) | ![Image of photocaging group](image6.png) | HEK293T | TEV protease [73,74], Npu DnaE intein [74] | 385 |
| ![Image of amino acid 20](image7.png) | ![Image of photocaging group](image8.png) | HEK293T | sfGFP, luciferase [75] | 355 |
| ![Image of amino acid 21](image9.png) | ![Image of photocaging group](image10.png) | HEK 293T | Nuclear localisation peptide for subcellular localisation of SATB1 and FOXO3 transcription factors, and TEV protease [149] | 350 |
| ![Image of amino acid 72](image11.png) | ![Image of photocaging group](image12.png) | HEK293T [67,125], HEK293T [66,75,147–152], HEK293ET [151], HeLa [150,152] | Cas9 endonuclease [150], Cre recombinase [148], Capsid of adeno-associated virus 2 [147], T7RNA polymerase [152], MEK1 kinase [151], LCK kinase [66] | 365 |
| ![Image of amino acid 67](image13.png) | ![Image of photocaging group](image14.png) | HEK293T | Luciferase [144] | 365 |
| ![Image of amino acid 73](image15.png) | ![Image of photocaging group](image16.png) | HEK293T or CHO K1 | eGFP and luciferase [153] | 365 |
| ![Image of amino acid 74](image17.png) | ![Image of photocaging group](image18.png) | HEK293T or CHO K1 | eGFP and luciferase [153] | 365 |

Continued over...
light irradiation, the photocage is removed, thereby restoring the protein's function. To date, photocaged cysteines (17–21), lysines (67, 72–74) and tyrosines (102–105) have been used to control enzyme function, intein splicing, protein subcellular localisation, virus–host interactions, and cell signalling cascades [13]. The light-activation approach is particularly useful for kinetic studies as it provides extreme spatiotemporal resolution. Spatial control can be achieved to even subcellular locations using focused light beams, which is virtually impossible when using the translational control approach. Theoretically, it is also possible to incorporate photocaged serine in mammalian cells as it has been demonstrated in yeast [53]. Therefore, the light-activation approach is applicable to regulate any protein that has a functionally critical cysteine, lysine, tyrosine, or serine residue in mammalian cells, including but not limited to kinases, DNA- and RNA-binding proteins, proteases, phosphatases, oxidoreductases, isomerases, and ubiquitin-modifying enzymes [54].

In contrast, the incorporation of a photocrosslinking amino acid can be used to inhibit protein function upon light irradiation [55]. In this case, a photocrosslinking amino acid is placed in the interior of the target protein. Upon light irradiation, a highly reactive functionality (e.g. radical, nitrene, carbene) is generated and reacted non-specifically with a nearby amino acid residue, causing cross-linking of the protein and subsequent abolishment of the protein's activity. The feasibility of this approach has been demonstrated with the use of \( p \)-benzoylphenylalanine (41) in the study of glutamate receptors, GluA1 and GluA2 [55]. When compared with the use of a photocaged amino acid, inhibiting a protein by photocrosslinking does not rely on the existence of a functionally critical residue, and thus theoretically, it can be used to investigate any protein in mammalian cells. Nevertheless, for each protein target it is necessary to screen a suitable site for placing the photocrosslinking amino acid. Protein variants containing the photocrosslinking amino acid must (i) behave in the same way as the wild-type protein (i.e. phenotypically silent) before light irradiation; and (ii) be fully inhibited after light irradiation causing the photocrosslinking. Due to these criteria, the screening process can be laborious and time-consuming.

In addition to light-induced activation and inhibition, reversible regulation of a protein function can be achieved through incorporation of a photoswitchable amino acid (Table 3). For example, 48, containing an azobenzene functionality which undergoes reversible \( cis \)-\( trans \) isomerisation upon irradiation with blue and UV light, has been used to control the activity of a glutamate receptor [56]. However, the general applicability of this approach suffers from similar constraints as inhibition by photocrosslinking. Extensive screening is often needed to identify a suitable site for incorporation, such that the resulting protein variant is fully active or inactive upon irradiation with light of a specific wavelength. At the current state of the art, there is no guarantee that such a site can be found in the target protein.

Table 2 Overview of photocaged unnatural amino acids that have been incorporated in mammalian cell systems for activating protein function upon irradiation with light of specific wavelength \( \lambda \) (Continued)

| Amino acid | Photocaging group (R) | System | Proteins | \( \lambda_{\text{decag}} \text{ (nm)} \) |
|------------|-----------------------|--------|----------|-----------------|
|            |                      | HEK293 [126] | Cre recombinase | 365 |
|            |                      | HEK293T [148,173] | [148]STAT transcription factor |  |
|            |                      | HEK293T luciferase, TEV protease [173] | Luciferase |  |

1 Decaging is only achieved at this wavelength via a two-photon activation using a specialised multiphoton laser setup. Abbreviation: Npu, Nostoc punctiforme.
Table 3 Overview of photoswitchable unnatural amino acids that have been incorporated in mammalian cell systems used to modulate protein function upon irradiation with the given wavelengths $\lambda$

| Amino acid | System | Proteins | $\lambda_{\text{trans-cis}}$ (nm) | $\lambda_{\text{cis-trans}}$ (nm) |
|------------|--------|----------|-------------------------------|-------------------------------|
| ![Image](45.png) | HEK 293T | Luciferase [111] | 355 | 450 |
| ![Image](46.png) | HEK 293T |  | 530 | 406 |
| ![Image](47.png) |  | NMDAR glutamate receptor [56] | 365 | 460 |

Overall, the use of light-responsive amino acids offers superior temporal control of protein function as the response is significantly faster (seconds) than the translational control by amber suppression (minutes to hours). Additionally, spatial control can be achieved at subcellular level, which is not possible with the translational approach. Generally, UV light at approximately 360 nm (i.e. UVA) is required (Tables 2 and 3) to induce the change (i.e. decaging, cross-linking, or isomerisation). However, UVA light has been shown to alter cellular signalling processes [57] or influence proper cellular function, if high intensity irradiation is applied (i.e. 50 J.cm$^{-2}$) [58]. Though not necessarily problematic, this has to be considered when planning to apply light-responsive unnatural amino acids. Thus, there is a trend to develop new functionalities that can be modulated by light of higher wavelengths [52]. In particular, red and near-infrared light (650–750 nm) are appealing because they cause no harm to cells even under excessive exposure, and they can penetrate tissues for \textit{in vivo} applications. To date, coumarin-caged lysines (73 and 74) are the only genetically incorporable unnatural amino acids that can be decaged within these wavelengths, although by two-photon approach that requires a specialised multiphoton laser setup [52]. Nevertheless, with the continuous advances in light-responsive chemical functionalities and orthogonal aaRS engineering, it is expected that more unnatural amino acids with the desired photophysical properties can be incorporated through genetic code expansion.

Small-molecule induced activation or inhibition

In addition to light, small molecules can also be used to unmask or modify unnatural amino acids and subsequently regulate protein function with prompt response. For example, several protecting groups can be removed bioorthogonally inside live mammalian cells, and these chemistries have been used to switch on protein function by genetic code expansion. Intracellular bioorthogonal reactions that have been used in this purpose include inverse electron demand Diels–Alder reactions [18,59–61], 1,3-dipolar cycloadditions [62], Staudinger reactions [63], and palladium-catalysed propargyl removal (Table 4) [57]. Currently, all of these have only been demonstrated in caged lysine molecules (61, 70, 71, 85) through a number of examples, including activation of luciferases, kinases, nucleases etc. Theoretically, all these protecting groups can be applied to other nucleophilic amino acids (e.g. cysteine, serine, threonine, tyrosine) subjected to successful engineering of the corresponding orthogonal synthetases.

© 2019 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC By).
Figure 4. Translational activation approaches to control virus replication
(A) Use of genetic code expansion to control replication of an amber codon tagged virus within transgenic host cells containing the orthogonal tRNA/synthetase pairs [43,48].
(B) Use of genetic code expansion to control replication of an amber codon tagged virus within normal host cells with the orthogonal tRNA/synthetase pair gene encoded by the viral genome [49].

Table 4 Overview of bioorthogonally protected unnatural amino acids tested in mammalian cell systems and their deprotection conditions

| Amino acid | System | Proteins | Reaction | Reagent |
|------------|--------|----------|----------|---------|
| ![Amino Acid 61](image) | HeLa, CHO, HEK293T, IH3T3, Caco-2, A549, HeLa | GFP, OspF phosphothreonine lyase | Pd-catalysed Tsuji–Trost-like reaction | Pd(II) complexes [57] |
| ![Amino Acid 70](image) | HEK293T | eGFP, SATB1 transcription factor, Cre recombinase, Cas9 endonuclease | Staudinger | Various phosphines [63] |
| ![Amino Acid 71](image) | HEK293T | eGFP, luciferase, OspF phosphothreonine lyase, Src kinase | 1,3-dipolar cycloaddition | trans-cyclooctenes [62] |
| ![Amino Acid 85](image) | HEK293T | GFP [61], Luciferase [59,61], MEK1 [18,60] and MEK2 [18] and FAK [60] and Src [60] kinases | Inverse electron demand Diels–Alder reactions | Tetrazines [18,59–61] |

On the other hand, bioorthogonal amino acids (e.g. 77 and 92, Table 5) have been used for rapid and selective inhibition of a specific enzyme in live mammalian cells [64]. In this case, a bioorthogonal amino acid is placed into the target enzyme without affecting the enzyme function. Upon contact with an inhibitor conjugate bearing the complementary bioorthogonal group, the enzyme variant is tethered to the conjugate and thus the enzyme activity is inhibited (Figure 5). The inhibition is exquisitely selective and can even discriminate between isoforms that differ by a single amino acid residue. Using this approach, selective inhibition of an intracellular kinase for which no selective small-molecule inhibitor exists was achieved. In addition, placing a photoswitchable moiety (i.e. azobenzene) into the inhibitor conjugate enables reversible modulation of enzyme activity by light.
Figure 5. Example of the small-molecule approach to protein inhibition
Use of unnatural amino acid incorporation for selective inhibition of protein function by bioorthogonal tethering [64].

Table 5 Overview of bioorthogonally reactive unnatural amino acids that have been incorporated in mammalian cell systems used to deactivate protein function upon reaction with the specified reagents

| Amino acid | System | Proteins | Reaction | Reagent |
|------------|--------|----------|----------|---------|
|            | HEK293T| MEK1 and LCK kinases | inverse electron demand Diels–Alder reactions | Inhibitor–tetrazine conjugates [64] |
|            | MEK1 and MEK2 kinases | | | |

Table 6 Comparison of protein function control approaches currently enabled by genetic code expansion

| Approach          | Temporal Control | Spatial Control | Reversibility |
|-------------------|------------------|-----------------|--------------|
| Translational control | Yes, slow       | Only by local administration | Yes, but high lag time |
| Optical control   | Yes, fast        | Yes, very high (to subcellular levels) | For photoswitchable amino acids |
| Chemical control  | Yes, medium      | Only by local administration | Yet to be established |

In comparison with light-induced activation or inhibition, small molecules can be used to activate or inhibit the target protein in deep animal tissue or intact animals which are not easily accessible by light. However, as mentioned above, only a few reactions have so far been shown to be bioorthogonal with high reaction rates to allow fast response [65]. The extension of this methodology is therefore tied to the development of novel bioorthogonal reactions. In contrast to the use of light for activation or inhibition, the small-molecule approach, similar to the translational approach, only allows spatial control by using cell-compartment selective compounds or reactions, or local injections.

Conclusion

Genetic code expansion has matured into a technique that can be routinely used in mammalian systems. Controlling protein function is currently mostly achieved by translation, light, and small molecules. These methods have been summarised and discussed, and their features have been compared (Table 6). Translational control is arguably the easiest to perform but suffers from the longer response time (up to several hours). On the other hand, both light and small-molecule induced methods have a faster response (seconds to minutes). The light approach is particularly appealing where subcellular spatial resolution is needed. While all three approaches have shown promise, most of the reviewed applications are so far proof-of-principle studies. The dissemination of this technique could be enhanced by the community simplifying access to plasmids (e.g. through plasmid repository), standardising the reporting format of aaRS mutants with full sequencing information, and providing protocols with extensive details. Only general implementation of protein control by genetic code expansion in the wider scientific community to unravel new biological insights will truly demonstrate the power of these techniques [55,56,66–68]. Since genetic code expansion has also been recently demonstrated in mice [47,50,51,60,69,70], we foresee that optimisation will be tailored to target cells, tissues, and mammalian models and many of the aforementioned approaches will be applied in vivo, providing the complete native environment to study function of mammalian proteins. With further promotion and adaptation of genetic code expansion, this is to be expected.
Summary

- Genetic code expansion can now be routinely used for incorporation of more than 100 different unnatural amino acids in mammalian cells using only four orthogonal aaRS/tRNA pairs and their mutants as shown in Table 1.

- Protein function can be temporally regulated (activation or inhibition) by simple translational control, i.e. supplementation of the desired unnatural amino acid to allow full-length, functional protein production.

- More rapid control can be achieved by incorporating stimuli responsive amino acids which allow activation, inhibition, or reversible regulation of protein function by light or small molecules.

- Most of the approaches have been demonstrated in proof-of-principle studies, but are ready for adaptation by the broader scientific community.

Funding

This work was supported by the BBSRC [grant number BB/P009506/1 (to Y.-H.T.);] and the Wellcome Trust [grant numbers 202056/Z/16/Z (to L.Y.P.L.), 200730/Z/16/Z (to Y.-H.T.). The funders play no role in design, decision to publish, or preparation of this manuscript.

Competing interests

The authors declare that there are no competing interests associated with the manuscript.

Abbreviations

aaRS, aminoacyl-tRNA synthetase; Ec, Escherichia coli; Mx, Methanomethylophilus alvus Mx1201; PylRS, Pyrrolysyl-tRNA synthetase.

References

1 McDonnell, D.P., Wardell, S.E. and Norris, J.D. (2015) Oral selective estrogen receptor downregulators (SERDs), a breakthrough endocrine therapy for breast cancer. J. Med. Chem. 58, 4883–4887, https://doi.org/10.1021/acs.jmedchem.5b00760
2 Petta, I., Lievens, S., Libert, C., Tavernier, J. and De Bosscher, K. (2016) Modulation of protein-protein interactions for the development of novel therapeutics. Mol. Ther. 24, 707–718, https://doi.org/10.1038/mt.2015.214
3 van Montfort, R.L.M. and Workman, P. (2017) Structure-based drug design: aiming for a perfect fit. Essays Biochem 61, 431–437, https://doi.org/10.1042/EBC20170052
4 Wan, W., Tharp, J.M. and Liu, W.R. (2014) Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. Biochim. Biophys. Acta 1844, 1059–1070, https://doi.org/10.1016/j.bbapap.2014.03.002
5 Dumas, A., Lercher, L., Spicer, C.D. and Davis, B.G. (2015) Designing logical codon reassignment - Expanding the chemistry in biology. Chem. Sci. 6, 50–69, https://doi.org/10.1039/C4SC01534G
6 Leisle, L., Vaiyaveetil, F., Meh, R.A. and Ahern, C.A. (2015) Incorporation of non-canonical amino acids. Adv. Exp. Med. Biol. 869, 119–151, https://doi.org/10.1007/978-1-4939-2845-3
7 Neumann-Staubitz, P. and Neumann, H. (2016) The use of unnatural amino acids to study and engineer protein function. Curr. Opin. Chem. Biol. 38, 119–128
8 Wang, L. (2017) Engineering the genetic code in cells and animals: biological considerations and impacts. Acc. Chem. Res. 50, 2767–2775, https://doi.org/10.1021/acs.accounts.7b00376
9 Itàlia, J.S., Zheng, Y., Kelemen, R.E., Erickson, S.B., Addy, P.S. and Chatterjee, A. (2017) Expanding the genetic code of mammalian cells. Biochem. Soc. Trans. 45, 555–562, https://doi.org/10.1042/BST20160336
10 Debelouchina, G.T. and Muir, T.W. (2017) A molecular engineering toolbox for the structural biologist. Q. Rev. Biophys. 50, e7, https://doi.org/10.1017/S0033583517000051
11 Chin, J.W. (2017) Expanding and reprogramming the genetic code. Nature 550, 53–60, https://doi.org/10.1038/nature24031
12 Zhang, S., He, D., Lin, Z., Yang, Y., Song, H. and Chen, P.R. (2017) Conditional chaperone-client interactions revealed by genetically encoded photo-cross-linkers. Acc. Chem. Res. 50, 1184–1192, https://doi.org/10.1021/acs.accounts.6b00647
13 Young, D.D. and Schultz, P.G. (2018) Playing with the molecules of life. ACS Chem. Biol. 13, 854–870, https://doi.org/10.1021/acschembio.7b00974
Brown, W., Liu, J. and Deiters, A. (2018) Genetic code expansion in animals. ACS Chem. Biol. 13, 2375–2386, https://doi.org/10.1021/acscambio.8b00520

Wang, W., Takimoto, J.K., Louie, G.V., Baiga, T.J., Noel, J.P., Lee, K.F. et al. (2007) Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat. Neurosci. 10, 1063–1072, https://doi.org/10.1038/nn1932

Ye, S., Zaitseva, E., Cattabiani, G., Scherfier, G.F.X., Säkmä, T.P., Deupi, X. et al. (2010) Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. Nature 464, 1386–1390, https://doi.org/10.1038/nature08946

Zhu, S., Rieu, M., Yao, C.A., Carvalho, S., Rodríguez, P.C., Bensaude, D. et al. (2014) Genetically encoding a light switch in an ionotropic glutamate receptor reveals subunit-specific interfaces. Proc. Natl. Acad. Sci. U.S.A. 111, 6081–6086, https://doi.org/10.1073/pnas.1318801111

Zheng, S., Fan, X., Wang, J., Zhao, J. and Chen, P.R. (2017) Dissection of kinase isoforms through orthogonal and chemical inducible signaling cascades. ChemBioChem 18, 1593–1598, https://doi.org/10.1002/cbic.201700255

Baumdick, M., Gelleri, M., Uttamapinant, C., Beránek, V., Chin, J.W. and Bastaens, P.I.H. (2018) A conformational sensor based on genetic code expansion reveals an autocatalytic component in EGFR activation. Nat. Commun. 9, 3847, https://doi.org/10.1038/s41467-018-06299-7

Köhre, C., Sullivan, E.L. and RajBhandary, U.L. (2004) Complete set of orthogonal 21st aminoacyl-tRNA synthetase-amber, ochre and opal suppressor tRNAs: concomitant suppression of three different termination codons in an mRNA in mammalian cells. Nucleic Acids Res. 32, 6200–6211, https://doi.org/10.1093/nar/gkh959

Xiao, H., Chatterjee, A., Choi, S.H., Bajuri, K.M., Sinha, S.C. and Schultz, P.G. (2013) Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. Angew. Chem. Int. Ed. 52, 14080–14083, https://doi.org/10.1002/anie.201308137

Schmied, W.H., Elssässer, S.J., Uttamapinant, C. and Chin, J.W. (2014) Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. J. Am. Chem. Soc. 136, 15577–15583, https://doi.org/10.1021/ja5069728

Serfling, R., Lorenz, C., Etzel, M., Schicht, G., Böttke, T., Mörl, M. et al. (2018) Designer tRNAEs for efficient incorporation of non-canonical amino acids by the pyrrolysin system in mammalian cells. Nucleic Acids Res. 46, 1–10, https://doi.org/10.1093/nar/gkx1156

Zheng, Y., Mukherjee, R., Chin, M.A., Igo, P., Gilgenast, M.J. and Chatterjee, A. (2018) Expanding the scope of single- and double-noncanonical amino acid mutagenesis in mammalian cells using orthogonal polyspecific leucyl-tRNA synthetases. Biochemistry 57, 441–445, https://doi.org/10.1021/acs.biochem.7b00952

Zheng, Y., Addy, P.S., Mukherjee, R. and Chatterjee, A. (2017) Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cells. Chem. Sci. 8, 7211–7217, https://doi.org/10.1039/C7SC02560B

Beranek, V., Willis, J.C.W. and Chin, J.W. (2019) An evolved Methanomethylphilus albus pyrrolysyl-tRNA Synthetase/tRNA pair is highly active and orthogonal in mammalian cells. Biochemistry 58, 387–390, https://doi.org/10.1021/acs.biochem.9b00808

Niu, W., Schultz, P.G. and Guo, J. (2013) An expanded genetic code in mammalian cells with a functional quadruplet codon. ACS Chem. Biol. 8, 1640–1645, https://doi.org/10.1021/cb4001662

Taki, M., Matsushita, J. and Sisido, M. (2008) Expanding the genetic code in a mammalian cell line by the introduction of four-base codon/anticodon pairs. ChemBioChem 9, 425–428, https://doi.org/10.1002/cbic.200800360

Neumann, H., Wang, K., Davis, L., García-Alai, M. and Chin, J.W. (2016) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. Nature 547, 441–444, https://doi.org/10.1038/nature22757

Wang, K., Sachdeva, A., Cox, D.J., Will, N.M., Lang, K., Wallacoe, S. et al. (2016) Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. Nat. Chem. 8, 393–403, https://doi.org/10.1038/nchem.2199

Schmied, W.H., Trinov, Z., Uttamapinant, C., Rae, C.D., Fried, S.D. and Chin, J.W. (2018) Controlling orthogonal ribosome subunit interactions enables evolution of new function. Nature 564, 444–448, https://doi.org/10.1038/s41586-018-0773-z

Xiao, H. and Schultz, P.G. (2016) At the interface of chemical and biological synthesis: an expanded genetic code. Cold Spring Harbor Perspect. Biol. 8, a023945, https://doi.org/10.1101/cshperspect.a023945

Vargas-Rodriguez, O., Sevostyanova, A., Söll, D. and Crnković, A. (2018) Upgrading aminoacyl-tRNA synthetases for genetic code expansion. Curr. Opin. Chem. Biol. 46, 115–122, https://doi.org/10.1016/j.cbpa.2018.07.014

Italia, J.S., Addy, P.S., Wrobel, C.J., Crawford, L.A., Lajoie, M.J., Zheng, Y. et al. (2017) An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. Nat. Chem. Biol. 13, 446–450, https://doi.org/10.1038/nchembio.2312

Italia, J.S., Latour, C., Wrobel, C.J.J. and Chatterjee, A. (2018) Resurrecting the bacterial turocy1-tRNA Synthetase/tRNA pair for expanding the genetic code of both E. coli and mammalian cells. Cell Chem. Biol. 25, 1304–1312, https://doi.org/10.1016/j.chembiol.2018.07.002

Chin, J.W., Crompton, T.A., Anderson, J.C., Mukherji, M., Zhang, Z. and Schultz, P.G. (2003) An expanded eukaryotic genetic code. Science 301, 964–967, https://doi.org/10.1126/science.1084772

Liu, W., Brock, A., Chen, S., Chen, S. and Schultz, P.G. (2007) Genetic incorporation of unnatural amino acids into proteins in mammalian cells. Nat. Methods 4, 239–244, https://doi.org/10.1038/nmeth1016

Naganathan, S., Grünbeck, A., Tian, H., Huber, T. and Sakmar, T.P. (2013) Genetically-encoded molecular probes to study G protein-coupled receptors. J. Vis. Exp., https://doi.org/10.3791/50588

Nikić, I., Kang, J.H., Girona, G.E., Aramburu, I.V. and Lemke, E.A. (2015) Labeling proteins on live mammalian cells using click chemistry. Nat. Protoc. 10, 780–791, https://doi.org/10.1038/nprot.2015.045

Yang, Y., Song, H., He, D., Zhang, S., Dai, S., Xie, X. et al. (2017) Genetically encoded releasable photo-cross-linking strategies for studying protein-protein interactions in living cells. Nat. Protoc. 12, 2147–2168, https://doi.org/10.1038/nprot.2017.090

Young, D.D., Young, T.S., Jahnz, M., Ahmad, I., Spraggon, G. and Schultz, P.G. (2011) An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. Biochemistry 50, 1894–1900, https://doi.org/10.1021/bi101929e
72 Ren, W., Ji, A. and Ai, H.-W. (2015) Light activation of protein splicing with a photocaged fast intein. J. Am. Chem. Soc. 137, 2155–2158, https://doi.org/10.1021/ja508597d

73 Nguyen, D.P., Mahesh, M., Elsässer, S.J., Hancock, S.M., Uttamapinant, C. and Chin, J.W. (2014) Genetic encoding of photocaged cysteine allows photoactivation of TEV protease in live mammalian cells. J. Am. Chem. Soc. 136, 2240–2243, https://doi.org/10.1021/ja412191m

74 Gramespacher, J.A., Stevens, A.J., Nguyen, D.P., Chin, J.W. and Muir, T.W. (2017) Inteinzymogens: conditional assembly and splicing of split inteins via targeted proteolysis. J. Am. Chem. Soc. 139, 8074–8077, https://doi.org/10.1021/jacs.7b02618

75 Uprety, R., Luo, J., Liu, J., Naro, Y., Samanta, S. and Deiters, A. (2014) Genetic encoding of caged cysteine and caged homocysteine in bacterial and mammalian cells. ChemBioChem 15, 1793–1799, https://doi.org/10.1002/cbic.201400073

76 Liu, J., Zheng, F., Cheng, R., Li, S., Rozovsky, S., Wang, Q. et al. (2018) Site-specific incorporation of selenocysteine using an expanded genetic code and palladium-mediated chemical deprotection. J. Am. Chem. Soc. 140, 8807–8816, https://doi.org/10.1021/jacs.8b04603

77 Chatterjee, A., Xiao, H., Bollong, M., Ai, H.-W. and Schultz, P.G. (2013) Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells. Proc. Natl. Acad. Sci. U.S.A. 110, 11803–11808, https://doi.org/10.1073/pnas.1309584110

78 Takimoto, J.K., Adams, K.L., Xiang, Z. and Wang, L. (2009) Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells. Mol. Biol. Syst. 5, 931–934, https://doi.org/10.1039/b904228h

79 Smits, A.H., Bornmann, A., Roosjen, M., van Hest, J.C. and Vermeulen, M. (2016) Click-MIS: tagless protein enrichment using bioorthogonal chemistry for quantitative proteomics. ACS Chem. Biol. 11, 3245–3250, https://doi.org/10.1021/acschembio.6b00520

80 Ye, S., Huber, T., Vogel, R. and Sakmar, T.P. (2009) FTIR analysis of GPCR activation using azido probes. Nat. Chem. Biol. 5, 397–399, https://doi.org/10.1038/nchembio.167

81 Cion, I., Perrin, M.H., Vale, W.W. and Wang, L. (2011) Photo-cross-linkers incorporated into G-protein-coupled receptors in mammalian cells: a ligand comparison. Angew. Chem. Int. Ed. 50, 8077–8081, https://doi.org/10.1002/anie.201102646

82 Chen, S., Chen, Z.-j., Ren, W. and Ai, H.-W. (2012) Reaction-based genetically encoded fluorescent hydrogen sulfide sensors. J. Am. Chem. Soc. 134, 9589–9592, https://doi.org/10.1021/ja303261d

83 Naganathan, S., Ye, S., Sakmar, T.P. and Huber, T. (2013) Site-specific epitope tagging of G protein-coupled receptors by bioorthogonal modification of a genetically encoded unnatural amino acid. Biochemistry 52, 1028–1036, https://doi.org/10.1021/bi301292h

84 Cion, I., Katritch, V., Sun, T., Xiang, Z., Siu, F.Y., Beyermann, M. et al. (2013) Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR. Cell 155, 1258–1269, https://doi.org/10.1016/j.cell.2013.11.008

85 Wang, W., Li, T., Felsovalyi, K., Chen, C., Cardozo, T. and Krogsgaard, M. (2014) Quantitative analysis of T cell receptor complex interaction sites using genetically encoded photo-cross-linkers. ACS Chem. Biol. 9, 2165–2172, https://doi.org/10.1021/cb500351s

86 Chen, Z.-j. and Ai, H.-W. (2014) A highly responsive and selective fluorescent probe for imaging physiological hydrogen sulfide. Biochemistry 53, 5966–5974, https://doi.org/10.1021/bi500830d

87 Tian, H., Naganathan, S., Kazmi, M.A., Schwartz, T.W., Sakmar, T.P. and Huber, T. (2014) Bioorthogonal fluorescent labeling of functional G-protein-coupled receptors. ChemBioChem 15, 1820–1829, https://doi.org/10.1002/cbic.201402193

88 Naganathan, S., Ray-Saha, S., Park, M., Tian, H., Sakmar, T.P. and Huber, T. (2015) Multiplex detection of functional G-protein-coupled receptors harboring site-specifically modified unnatural amino acids. Biochemistry 54, 776–786, https://doi.org/10.1021/bi501267x

89 Tian, H., Sakmar, T.P. and Huber, T. (2015) Micelle-enhanced bioorthogonal labeling of genetically encoded azido groups on the lipid-embedded surface of a GPCR. ChemBioChem 16, 1314–1322, https://doi.org/10.1002/cbic.201500030

90 Zheng, Y., Lewis, Jr, T.L., Igo, P., Polletux, F. and Chatterjee, A. (2017) Virus-enabled optimization and delivery of the genetic machinery for efficient unnatural amino acid mutagenesis in mammalian cells and tissues. ACS Synth. Biol. 6, 13–18, https://doi.org/10.1021/acsbio.6b00092

91 Koole, C., Reynolds, C.A., Mobarec, J.C., Hick, C., Sexton, P.M. and Sakmar, T.P. (2017) Genetically encoded photocross-linkers determine the biological binding site of exendin-4 peptide in the N-terminal domain of the intact human glucagon-like peptide-1 receptor (GLP-1R). J. Biol. Chem. 292, 7131–7144, https://doi.org/10.1074/jbc.M117.779496

92 Grunbeck, A., Huber', T., Sachdev, P. and Sakmar, T.P. (2011) Mapping the ligand-binding site on a G protein-coupled receptor (GPCR) using genetically encoded photocrosslinkers. Biochemistry 50, 3411–3413, https://doi.org/10.1021/bi200214r

93 Ray-Saha, S., Huber, T. and Sakmar, T.P. (2014) Antibody epitopes on g protein-coupled receptors mapped with genetically encoded photoactivatable cross-linkers. Biochemistry 53, 1302–1310, https://doi.org/10.1021/bi401298p

94 Grunbeck, A., Huber, T., Abrol, R., Trzaskowski, B., Goddard, III, W.A. and Sakmar, T.P. (2012) Genetically encoded photo-cross-linkers map the binding site of an allosteric drug on a G protein-coupled receptor. ACS Chem. Biol. 7, 967–972, https://doi.org/10.1021/cb300592g

95 Rannoverson, H., Andersen, J., Sorensen, L., Bang-Andersen, B., Park, M., Huber, T. et al. (2016) Genetically encoded photocrosslinkers locate the high-affinity binding site of antidepressant drugs in the human serotonin transporter. Nat. Commun. 7, 11261, https://doi.org/10.1038/ncomms11261

96 Park, M., Sivertsen, B.B., Eils-Heindl, S., Holst, B., Beck-Sickinger, A.G. et al. (2015) Bioorthogonal labeling of ghrelin receptor to facilitate studies of ligand-dependent conformational dynamics. Chem. Biol. 22, 1431–1436, https://doi.org/10.1016/j.chembiol.2015.09.014

97 Chen, Z.-j., Ren, W., Wright, O.E. and Ai, H.-W. (2013) Genetically encoded fluorescent probe for the selective detection of peroxynitrite. J. Am. Chem. Soc. 135, 14940–14943, https://doi.org/10.1021/ja408011q

98 Tharp, J.M., Wang, Y.S., Lee, Y.J., Yang, Y. and Liu, W.R. (2014) Genetic incorporation of seven ortho-substituted phenylalanine derivatives. ACS Chem. Biol. 9, 884–890, https://doi.org/10.1021/cb400917a

99 Ye, S., Köhner, C., Huber, T., Kazmi, M., Sachdev, P., Yan, E.C. et al. (2008) Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. J. Biol. Chem. 283, 1525–1533, https://doi.org/10.1074/jbc.M707355200

100 Axup, J.Y., Bajjuri, K.M., Rftland, M., Hutchins, B.M., Kim, C.H., Kazane, S.A. et al. (2012) Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. Proc. Natl. Acad. Sci. U.S.A. 109, 16101–16106, https://doi.org/10.1073/pnas.1211023109
Xiang, Z., Ren, H., Hu, Y.S., Coin, I., Wei, J., Cang, H. et al. (2013) Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. *Nat. Methods* **10**, 885–888, https://doi.org/10.1038/nmeth.2595

Hino, N., Okazaki, Y., Kobayashi, T., Hayashi, A., Sakamoto, K. and Yokoyama, S. (2005) Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photocross-linkable amino acid. *Nat. Methods* **2**, 201–206, https://doi.org/10.1038/nmeth.379

Shen, B., Xiang, Z., Miller, B., Louie, G., Wang, W., Noel, J.P. et al. (2011) Genetically encoding unnatural amino acids in neural stem cells and optically reporting voltage-sensitive domain changes in differentiated neurons. *Stem Cells* **29**, 1231–1240, https://doi.org/10.1002/stem.679

Murray, C.I., Westhoff, M., Eldstrom, J., Thompson, E., Ernes, R. and Fedida, D. (2016) Unnatural amino acid photo-cross-linking of the Iks channel complex demonstrates a KCNE1-KCNQ1 stoichiometry of up to 4:4. *Elite 5*, e11815, https://doi.org/10.7554/eLife.11815

Valentin-Hansen, L., Park, M., Huber, T., Grunbeck, A., Naganathan, S., Schwartz, T.W. et al. (2014) Mapping substance P binding sites on the neurokinin-1 receptor using genetic incorporation of a photoactive amino acid. *J. Biol. Chem.* **289**, 18045–18054, https://doi.org/10.1074/jbc.M113.527085

Lacey, V.K., Louie, G.V., Noel, J.P. and Wang, L. (2013) Expanding the library and substrate diversity of the pyrrolysyl-tRNA synthetase to incorporate unnatural amino acids containing conjugated rings. *ChemBioChem* **14**, 2100–2105, https://doi.org/10.1002/cbic.201300400

Chatterjee, A., Guo, J., Lee, H.S. and Schultz, P.G. (2013) A genetically encoded fluorescent probe in mammalian cells. *J. Am. Chem. Soc.* **135**, 12540–12543, https://doi.org/10.1021/ja405955x

Mitchell, A.L., Addy, P.S., Chin, M.A. and Chatterjee, A. (2017) A unique genetically encoded FRET pair in mammalian cells. *ChemBioChem* **18**, 511–514, https://doi.org/10.1002/cbic.201600668

Park, S.-H., Ko, W., Lee, H.S. and Shin, I. (2019) Analysis of protein–protein interaction in a single live cell by using a FRET system based on genetic code expansion technology. *J. Am. Chem. Soc.* **141**, 4273–4281, https://doi.org/10.1021/jacs.8b10098

Fu, C., Kobayashi, T., Wang, N., Hoppmann, C., Yang, B., Irnanejad, R. et al. (2018) Genetically encoding quinoline reverses chromophore charge and enables fluorescent protein brightening in acidic vesicles. *J. Am. Chem. Soc.* **140**, 11058–11066, https://doi.org/10.1021/jacs.8b05814

Luo, J., Samanta, S., Convertino, M., Dokholyan, N.V. and Deiters, A. (2018) Reversible and tunable photoswitching of protein function through genetic encoding of azobenzene amino acids in mammalian cells. *ChemBioChem* **19**, 2178–2185, https://doi.org/10.1002/cbic.201800226

Hopmann, C., Lacey, V.K., Louie, G.V., Wei, J., Noel, J.P. and Wang, L. (2014) Genetically encoding photoswitchable click amino acids in Escherichia coli and mammalian cells. *Angew. Chem. Int. Ed.* **53**, 3932–3936, https://doi.org/10.1002/anie.201400001

Xiao, H., Peters, F.B., Yang, P.Y., Reed, S., Chittluru, J.R. and Schultz, P.G. (2014) Genetic incorporation of histidine derivatives using an engineered pyrrolysyl-tRNA synthetase. *ACS Chem. Biol.* **9**, 1092–1096, https://doi.org/10.1021/cb500032c

Wang, T., Zhou, Q., Li, F., Yu, Y., Yin, X. and Wang, J. (2015) Genetic incorporation of N[1-π(ε(p)-π)]-formyllysine, a new histone post-translational modification. *ChemBioChem* **16**, 1440–1442, https://doi.org/10.1002/cbic.201500170

Elssäser, S.J., Ernst, R.J., Walker, O.S. and Chin, J.W. (2016) Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat. Methods* **13**, 158–164, https://doi.org/10.1038/nmeth.3701

Xuan, W., Yao, A. and Schultz, P.G. (2017) Genetically encoded fluorescent probe for detecting sirtuins in living cells. *J. Am. Chem. Soc.* **139**, 12350–12353, https://doi.org/10.1021/jacs.7b05725

Cohen, S. and Arbely, E. (2016) Single-plasmid-based system for efficient noncanonical amino acid mutagenesis in cultured mammalian cells. *ChemBioChem* **17**, 1008–1011, https://doi.org/10.1002/cbic.201500681

Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K. and Yokoyama, S. (2008) Adding l-lysine derivatives to the genetic code of *Arabidopsis thaliana*. *Angew. Chem. Int. Ed.* **56**, 298–304, https://doi.org/10.1002/anie.200702839

Li, T., Zhang, H., Sun, Y., Pan, Y., Zhou, J. and Wang, J. (2013) Expanding the genetic code for photoclick chemistry in E. coli, mammalian cells, and A. thaliana. *Angew. Chem. Int. Ed.* **52**, 9700–9704, https://doi.org/10.1002/anie.201305477

Kim, C.H., Kang, M., Kim, H.J., Chatterjee, A. and Schultz, P.G. (2012) Site-specific incorporation of epsilon-N-crotonyllysine into histones. *Angew. Chem. Int. Ed.* **51**, 7246–7249, https://doi.org/10.1002/anie.201203349

Cigler, M., Müller, T.G., Horn-Geltho, D., von Wrisberg, M.K., Fottner, M., Goody, R.S. et al. (2017) Proximity-triggered covalent stabilization of low-affinity protein complexes in vitro and in vivo. *Angew. Chem. Int. Ed.* **56**, 15737–15741, https://doi.org/10.1002/anie.201706927

Tian, Y., Jacinto, M.P., Zang, Y., Yu, Z., Qu, J., Liu, W.R. et al. (2017) Genetically encoded 2-aryl-5-carboxytetrazoles for site-selective protein photo-cross-linking. *J. Am. Chem. Soc.* **139**, 6078–6081, https://doi.org/10.1021/jacs.7b02615

Peng, T. and Hang, H.C. (2016) Site-specific bioorthogonal labeling for fluorescence imaging of intracellular proteins in living cells. *J. Am. Chem. Soc.* **138**, 14423–14433, https://doi.org/10.1021/jacs.6b08733

Ramil, C.P., Dong, M., An, P., Lewandowski, T.M., Yu, Z., Miller, L.J. et al. (2017) Spirohexene-tetrazine ligation enables bioorthogonal labeling of class B G protein-coupled receptors in live cells. *J. Am. Chem. Soc.* **139**, 13376–13386, https://doi.org/10.1021/jacs.7b05674

Gautier, A., Nguyen, D.P., Lusic, H., An, W., Deiters, A. and Chin, J.W. (2010) Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* **132**, 4086–4088, https://doi.org/10.1021/ja910688s

Arbely, E., Torres-Kolbus, J., Deiters, A. and Chin, J.W. (2012) Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine. *J. Am. Chem. Soc.* **134**, 11912–11915, https://doi.org/10.1021/ja3046958

Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A. and Chin, J.W. (2012) Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* **4**, 298–304, https://doi.org/10.1038/nchem.1250

Atoukh, N., Schwartz, T., König, A.I., Cohen, S., Brozgel, E., Tam, B. et al. (2018) Live cell imaging of bioorthogonally labelled proteins generated with a single pyrrolysine tRNA gene. *Sci. Rep.* **8**, 14527, https://doi.org/10.1038/s41598-018-32824-1

Schneider, N., Gabelein, C., Wiener, J., Georgiev, T., Gobet, N., Weber, W. et al. (2018) Genetic code expansion method for temporal labeling of endogenously expressed proteins. *ACS Chem. Biol.* **13**, 3049–3053, https://doi.org/10.1021/acschembio.8b00594
130 Li, N., Ramil, C.P., Lim, R.K. and Lin, Q. (2015) A genetically encoded alkyne directs palladium-mediated protein labeling on live mammalian cell surface. ACS Chem. Biol. 10, 379–384, https://doi.org/10.1021/acschembio.5b00649
131 Yang, Y., Lin, S., Lin, W. and Chen, P.R. (2014) Ligand-assisted dual-site click labeling of EGFR on living cells. ChemBioChem 15, 1738–1743, https://doi.org/10.1002/cbic.201400057
132 Zheng, Y., Yu, F., Wu, Y., Si, L., Xu, H., Zhang, C. et al. (2015) Broadening the versatility of lentiviral vectors as a tool in nucleic acid research via genetic code expansion. Nucleic Acids Res. 43, e73, https://doi.org/10.1093/nar/gkv202
133 Zhang, C., Yao, T., Zheng, Y., Li, Z., Zhang, Q., Zhang, L. et al. (2016) Development of next generation adeno-associated viral vectors capable of selective tropism and efficient gene delivery. Biomaterials 80, 134–145, https://doi.org/10.1016/j.biomaterials.2015.11.066
134 Zhang, Z., Xu, H., Si, L., Chen, Y., Zhang, B., Wang, Y. et al. (2017) Construction of an inducible stable cell line for efficient incorporation of unnatural amino acids in mammalian cells. Biochim. Biophys. Res. Commun. 469, 490–496, https://doi.org/10.1016/j.bbrc.2017.05.178
135 Yao, T., Zhou, X., Zhang, C., Yu, X., Tian, Z., Zhang, L. et al. (2017) Site-specific PEGylated adeno-associated viruses with increased serum stability and reduced immunogenicity. Molecules 22, 1155, https://doi.org/10.3390/molecules22071155
136 Katrekar, D., Moreno, A.M., Chen, G., Worlak, A. and Mali, P. (2018) Oligonucleotide conjugated multi-functional adeno-associated viruses. Sci. Rep. 8, 3589, https://doi.org/10.1038/s41598-018-21742-x
137 Kelemen, R.E., Mukherjee, R., Cao, X., Erickson, S.B., Zheng, Y. and Chatterjee, A. (2016) A precise chemical strategy to alter the receptor specificity of the adeno-associated virus. Angew. Chem. Int. Ed. 55, 10645–10649, https://doi.org/10.1002/anie.201604067
138 Zhang, J., Yan, S., He, Z., Ding, C., Zhai, T., Chen, Y. et al. (2018) Small unnatural amino acid carried raman tag for molecular imaging of genetically targeted proteins. J. Phys. Chem. Lett. 9, 4679–4685, https://doi.org/10.1021/acs.jpclett.8b01991
139 Lu¨ohmann, T., Jones, G., Gutmann, M., Rybak, J.-C., Nickell, J., Rubini, M. et al. (2015) Bio-orthogonal immobilization of fibroblast growth factor 2 for spatial controlled cell proliferation. ACS Biomater. Sci. Eng. 1, 740–746, https://doi.org/10.1021/acsbiomaterials.5b00236
140 Lin, S., Yan, H., Li, L., Yang, M., Peng, B., Chen, S. et al. (2013) Site-specific engineering of chemical functionalities on the surface of live hepatitis D virus. Angew. Chem. Ed. 52, 13970–13974, https://doi.org/10.1002/ange.201305787
141 Chen, P.R., Groff, D., Guo, J., Ou, W., Cellitti, S., Geierstanger, B.H. et al. (2009) A facile system for encoding unnatural amino acids in mammalian cells. Angew. Chem. Int. Ed. 48, 4052–4055, https://doi.org/10.1002/anie.200900683
142 Yanagisawa, T., Hino, N., Iraha, F., Mukai, T., Sakamoto, K. and Yokoyama, S. (2012) Wide-range protein photo-crosslinking achieved by a genetically encoded N(epsilon)-(benzyloxycarbonyl)lysine derivative with a diazirinyl moiety. Mol. Biosyst. 8, 1131–1135, https://doi.org/10.1039/c2mb05321g
143 Kita, A., Hino, N., Higashi, S., Hirota, K., Narumi, R., Adachi, J. et al. (2016) Adenovirus vector-based incorporation of a photo-cross-linkable amino acid into proteins in human primary cells and cancerous cell lines. Sci. Rep. 6, 36946, https://doi.org/10.1038/srep36946
144 Zhao, J., Lin, S., Huang, Y., Zhao, J. and Chen, P.R. (2013) Mechanism-based design of a photoactivatable firefly luciferase. J. Am. Chem. Soc. 135, 7410–7413, https://doi.org/10.1021/ja4013535
145 Ren, W., Ji, A., Wang, M.X. and Ai, H.-W. (2015) Expanding the Genetic Code for a Dinitrophenyl Hapten. ChemBioChem 16, 2007–2010, https://doi.org/10.1002/cbic.201500204
146 Groff, D., Chen, F.R., Peters, F.B. and Schultz, P.G. (2010) A genetically encoded epsilon-N-methyl lysine in mammalian cells. ChemBioChem 11, 1066–1068, https://doi.org/10.1002/cbic.200900690
147 Erickson, S.B., Mukherjee, R., Kelemen, R.E., Wrobel, C.J.J., Cao, X. and Chatterjee, A. (2017) Precise photoremovable perturbation of a virus–host interaction. Angew. Chem. Int. Ed. 56, 4233–4237, https://doi.org/10.1002/anie.201700683
148 Luo, J., Arbely, E., Zhang, J., Chou, C., Uprety, R., Chin, J.W. et al. (2016) Genetically encoded optical activation of DNA recombination in human cells. Chem. Commun. 52, 8529–8532, https://doi.org/10.1039/c6cc03934k
149 Engelke, H., Chou, C., Uprety, R., Jess, P. and Deiters, A. (2014) Control of protein function through optochemical translocation. ACS Synth. Biol. 3, 731–736, https://doi.org/10.1021/sb400192a
150 Hemphill, J., Borchardt, E.K., Brown, K., Asokan, A. and Deiters, A. (2015) Optical control of CRISPR/Cas9 gene editing. J. Am. Chem. Soc. 137, 5642–5645, https://doi.org/10.1021/jacs.5b0192a
151 Gautier, A., Deiters, A. and Chin, J.W. (2011) Light-activated kinases enable temporal dissection of signaling networks in living cells. J. Am. Chem. Soc. 133, 2124–2127, https://doi.org/10.1021/ja110997x
152 Hemphill, J., Chou, C., Chin, J.W. and Deiters, A. (2013) Genetically encoded light-activated transcription for spatiotemporal control of gene expression and gene silencing in mammalian cells. J. Am. Chem. Soc. 135, 13433–13439, https://doi.org/10.1021/ja4051026
153 Luo, J., Uprety, R., Naro, Y., Chou, C., Nguyen, D.P., Chin, J.W. et al. (2014) Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation. J. Am. Chem. Soc. 136, 15551–15558, https://doi.org/10.1021/ja505586
154 Yu, Z., Pan, Y., Wang, Z., Wang, J. and Lin, Q. (2012) Genetically encoded cyclopropene directs rapid, photoclick-chemistry-mediated protein labeling in mammalian cells. Angew. Chem. Int. Ed. 51, 10600–10604, https://doi.org/10.1002/anie.201205352
155 Meineke, B., Heimgartner, J., Lefrancli, L. and Els¨asser, S.J. (2018) Methanomethylophilus alvus Mx1201 provides basis for mutual orthogonal pyrrolysyl tRNA/Aminoacyl-tRNA synthetase pairs in mammalian cells. ACS Chem. Biol. 13, 3087–3096, https://doi.org/10.1021/acschembio.8b00571
156 Elliott, T.S., Townsley, F.M., Bianco, A., Ernst, R.J., Sachdeva, A., Els¨asser, S.J. et al. (2014) Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. Nat. Biotechnol. 32, 465–472, https://doi.org/10.1038/nbt.2860
157 Al, H.-W., Shen, W., Sagi, A., Chen, P.R. and Schultz, P.G. (2011) Probing protein-protein interactions with a genetically encoded photo-crosslinking amino acid. ChemBioChem 12, 1854–1857, https://doi.org/10.1002/cbic.201100194
158 Chou, C., Uprety, R., Davis, L., Chin, J.W. and Deiters, A. (2011) Genetically encoding an aliphatic diazirine for protein photocrosslinking. Chem. Sci. 2, 480–483, https://doi.org/10.1039/C0SC00373E
Hoffmann, J.E., Dziuba, D., Stein, F. and Schultz, C. (2018) A bifunctional noncanonical amino acid: synthesis, expression, and residue-specific proteome-wide incorporation. Biochemistry 57, 4747–4752, https://doi.org/10.1021/acs.biochem.8b00397.

Zhang, C., Lu, J., Su, H., Yang, J. and Zhou, D. (2017) Fatty acid synthase cooperates with protrudin to facilitate membrane outgrowth of cellular protrusions. Sci. Rep. 7, 46569, https://doi.org/10.1038/srep46569.

Nikić, I., Plass, T., Schraidt, O., Szymański, J., Briggs, J.A., Schultz, C. et al. (2014) Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. Angew. Chem. Int. Ed. 53, 2245–2249, https://doi.org/10.1002/anie.201309847.

Nikić, I., Estrada Girona, G., Kang, J.H., Pacci, G., Mikhaleva, S., Koehler, C., Smyrniotis, N.V. et al. (2016) Debugging eukaryotic genetic code expansion for site-specific Click-PAINT super-resolution microscopy. Angew. Chem. Int. Ed. 55, 16172–16176, https://doi.org/10.1002/anie.201608284.

Rutkowska, A., Plass, T., Hoffmann, J.E., Yushchenko, D.A., Feng, S. and Schultz, C. (2014) T-CrAsH: a heterologous chemical crosslinker. ChemBioChem 15, 1765–1768, https://doi.org/10.1002/cbic.201401218.

Hoffmann, J.E., Plass, T., Nikić, I., Aramburu, I.V., Koehler, C., Gillandt, H. et al. (2015) Highly stable trans-cyclooctene amino acids for live-cell labeling. Chem. Eur. J. 21, 12266–12270, https://doi.org/10.1002/chem.201501647.

Kozma, E., Nikić, I., Varga, B.R., Aramburu, I.V., Kang, J.H., Fackler, O.T. et al. (2016) Hydrophilic trans-cyclooctenylated noncanonical amino acids for fast intracellular protein labeling. ChemBioChem 17, 1518–1524, https://doi.org/10.1002/chem.201600284.

Sakin, V., Hanne, J., Dunder, J., Anders-Össwein, M., Laketa, V., Nikić, I. et al. (2017) A versatile tool for live-cell imaging and super-resolution nanoscopy studies of hiv-1 env distribution and mobility. Cell Chem. Biol. 24, 635–645, https://doi.org/10.1016/j.chembiol.2017.04.007.

Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D.J., Blackman, M.L. et al. (2012) Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. J. Am. Chem. Soc. 134, 10317–10320, https://doi.org/10.1021/ja302832g.

Uttamapinant, C., Howe, J.D., Lang, K., Beránek, V., Davis, L., Mahesh, M. et al. (2015) Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. J. Am. Chem. Soc. 137, 4602–4605, https://doi.org/10.1021/jacs.a512838z.

Plass, T., Milles, S., Koehler, C., Szymarski, J., Mueller, R., Wießler, M. et al. (2012) Amino acids for Diels-Alder reactions in living cells. ChemBioChem 13, 4166–4170, https://doi.org/10.1002/cbic.201108231.

Xue, L., Prifti, E. and Johnsson, K. (2016) A general strategy for the semisynthesis of ratiometric fluorescent sensor proteins with increased dynamic range. J. Am. Chem. Soc. 138, 5258–5261, https://doi.org/10.1021/jacs.6b03034.

Bormann, A., Milles, S., Plass, T., Dommerholt, J., Verkade, J.M., Wissler, M. et al. (2012) Genetic encoding of a bicyclo[6.1.0]nonyne-charged amino acid enables fast cellular protein imaging by metal-free ligation. ChemBioChem 13, 2094–2099, https://doi.org/10.1002/cbic.201200407.

Schwartz, T., Aloush, N., Golandi, I., Segal, I., Nachmias, D., Arbel, E. et al. (2017) Direct fluorescent-dye labeling of alpha-tubulin in mammalian cells for fast intracellular protein labeling. Mol. Biol. Cell 28, 2747–2756, https://doi.org/10.1091/mbc.e17-03-0161.

Luo, J., Torres-Kolbus, J., Liu, J. and Deiters, A. (2017) Genetic encoding of photocaged tyrosines with improved light-activation properties for the optical control of protease function. ChemBioChem 18, 1442–1447, https://doi.org/10.1002/cbic.201700147.

Wang, N., Yang, B., Fu, C., Zhu, H., Zheng, F., Kobayashi, T. et al. (2018) Genetically encoding fluorosulfate-l-tyrosine to react with lysine, histidine, and tyrosine via SuFEx in proteins in vivo. J. Am. Chem. Soc. 140, 4995–4999, https://doi.org/10.1021/jacs.8b01087.

Takimoto, J.K., Xiang, Z., Kang, J.Y. and Wang, L. (2010) Esterification of an unnatural amino acid structurally deviating from canonical amino acids promotes its uptake and incorporation into proteins in mammalian cells. ChemBioChem 11, 2268–2272, https://doi.org/10.1002/cbic.201000436.

Beránek, V., Reinkemeier, C.D., Zhang, M.S., Liang, A.D., Kym, G. and Chin, J.W. (2018) Genetically encoded protein phosphorylation in mammalian cells. Cell Chem. Biol. 25, 1067–1074, https://doi.org/10.1016/j.chembiol.2018.05.013.