Published in final edited form as:

Curr Opin Chem Biol. 2011 June ; 15(3): 387–391. doi:10.1016/j.cbpa.2011.03.007.

The Pyrrolysine Translational Machinery as a Genetic-Code Expansion Tool

Tomasz Fekner^a and Michael K. Chan^b

Tomasz Fekner: tfekner@chemistry.ohio-state.edu; Michael K. Chan: chan@chemistry.ohio-state.edu

^a Department of Chemistry, The Ohio State University, 100 W 18th Ave., Columbus, OH 43210, USA

^b Departments of Chemistry and Biochemistry, The Ohio State University, 484 W 12th Ave., Columbus, OH 43210, USA., Phone: (+1) 614 292 8375., Fax: (+1) 614 292 6773., Homepage: <http://www.chemistry.ohio-state.edu/~chan/>

Abstract

The discovery of pyrrolysine not only expanded the set of the known proteinogenic amino acids but also revealed unusual features of its encoding mechanism. The engagement of a canonical stop codon and a unique aminoacyl-tRNA synthetase-tRNA pair that can be used to accommodate a broad range of unnatural amino acids while maintaining strict orthogonality in a variety of prokaryotic and eukaryotic expression systems has proven an invaluable combination. Within a few years since its unique properties were elucidated, the pyrrolysine translational machinery has become a popular choice for the synthesis of recombinant proteins bearing a wide variety of otherwise hard-to-introduce functional groups. It is also central to the development of new synthetic strategies that rely on stop-codon suppression.

Introduction

In 2002, the internal UAG (amber) codon present in the monomethylamine methyltransferase gene of *Methanosarcina barkeri* [1], a methanogenic archaeon, was shown to encode pyrrolysine (1, Pyl, Figure 1) [2,3], establishing it as the 22nd proteinogenic amino acid [4,5]. Apart from the engagement of UAG which is usually a nonsense (stop) codon, pyrrolysine behaves like a typical canonical amino acid and is charged directly into its cognate amber suppressor tRNA (tRNA^{Pyl}) by its own pyrrolysyl-tRNA synthetase (PylRS) [6,7].

Pyrrolysine is present in a very limited number of organisms including some other members of the *Methanosarcinaceae* family and, curiously, several bacteria [2,8,9]. However, because many successful synthetic strategies engage stop codons in translational incorporation of noncanonical amino acids (NAAs) [10,11], the PylRS-tRNA^{Pyl} system can be harnessed for the genetic-code expansion in unrelated organisms. This notion is supported by the discovery that pyrrolysine can still be translationally incorporated when a vector containing the PylRS-tRNA^{Pyl} genes is transformed into *Escherichia coli*, providing that either an

© 2011 Elsevier Ltd. All rights reserved.

Correspondence to: Michael K. Chan, chan@chemistry.ohio-state.edu.

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

exogenous source of the synthetic amino acid is made available [6,12] or additional genes associated with its biosynthesis are also included [13]. Moreover, it has been shown that while a specific pyrrolysine insertion sequence (PYLIS) element [14] within the coding mRNA promotes efficient amber suppression in *Methanosarcina* [15], it is not essential in *E. coli* [16]. The PylRS–tRNA^{Pyl} pair also retains excellent orthogonality in the new host, that is, tRNA^{Pyl} and *E. coli* tRNAs are not charged with any of the 20 canonical amino acids and pyrrolysine, respectively [17]. Perhaps most importantly from the standpoint of future synthetic applications, it was demonstrated that the lysine derivatives (**2–4**) are acceptable substrates for the *M. barkeri* PylRS–tRNA^{Pyl} (*Mb*PylRS–tRNA^{Pyl}) pair and can be efficiently and site-specifically incorporated into recombinant proteins in vivo [18,19].

These pioneering discoveries set the stage for the recent explosion of interest in the PylRS–tRNA^{Pyl} system for the synthesis of recombinant proteins containing NAAs and our review aims to provide a concise account of key developments, primarily from 2009 onwards, in three main areas: (1) the substrate specificity of PylRS–tRNA^{Pyl}, (2) the use of the PylRS–tRNA^{Pyl} system in *E. coli*, and (3) the transfer and application of PylRS–tRNA^{Pyl} in eukaryotic organisms. We hope that our review will contribute to the general awareness of benefits associated with the use of the PylRS–tRNA^{Pyl} pairs as a highly flexible synthetic tool.

Substrate specificity of the PylRS–tRNA^{Pyl} system

Methanosarcina mazei PylRS–tRNA^{Pyl} (*Mm*PylRS–tRNA^{Pyl}) and *Mb*PylRS–tRNA^{Pyl} are the systems of choice for the translational incorporation of pyrrolysine analogs. To date, no systematic comparative study has been performed that could unequivocally ascertain superiority of either pair in terms of stability, substrate specificity, readthrough efficiency, amenability to directed evolution, etc. However, thanks to a significant number of diverse pyrrolysine surrogates reported to date [20,21], it was possible to identify key structural features of potentially viable PylRS substrates [20]. Typically, an α -amino acid needs to possess a carbonyl group separated from its carboxylate by six atoms. As no canonical amino acid contains this structural element, it no doubt contributes to the overall orthogonality of the PylRS–tRNA^{Pyl} systems. An additional heteroatom located close to the carbonyl on the side opposite to the carbonyl–carboxylate linker is also beneficial. The exact position of this extra heteroatom and the stereochemical arrangement of substituents in the vicinity of the carbonyl group are also important factors. For example, an epimer of **2** containing natural proline is not a viable PylRS substrate [18]. Changing the position of the oxygen atom within the tetrahydrofuran ring in **4** has a similar effect [19]. Although the wild-type PylRSs are more convenient to use, their evolved versions have often been developed in order to accommodate a desired substrate. While most analogs are acylated lysines, exceptions such as **5** [22] and **6** [23] (Figure 2) are known.

The PylRS–tRNA^{Pyl} pairs in the *E. coli* expression system

Following the discovery of the pyrrolysine surrogates **2–4**, it appeared that their steric and electronic similarity to the parent amino acid was a key factor in their recognition by PylRS. An analogous reasoning led to the design of **7** for the synthesis of recombinant proteins amenable for further functionalization via copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) [24]. This method was used to prepare calmodulin site-specifically labeled with two different dyes. Conformational changes of the protein were subsequently studied by Förster resonance energy transfer (FRET) measurements. Recently, synthetically more accessible clickable analogs **8** [25] and **9** [26] were identified. The former was used to suppress up to three amber codons within the same mRNA [27], albeit with modest efficiency, during the studies on the synthesis of glycosylated proteins.

The PylRS–tRNA^{Pyl} system has the greatest potential in the synthesis and study of proteins containing modified lysine residues. Acetylation, methylation, ubiquitylation, and sumoylation are common post-translational lysine modifications and strategies to generate proteins containing them have been reported. For example, the pyrrolysine surrogate **10**, a structural analog of **9**, made it possible to prepare recombinant proteins suitable for native chemical ligation [28], which was central to the development of a protocol for nonenzymatic ubiquitylation [29].

The substrate specificity and readthrough efficiency of the wild-type PylRS–tRNA^{Pyl} systems can be augmented by modifying the synthetase. Yokoyama et al. showed that a structure-based one-point mutation in *Mm*PylRS can significantly improve the level of the in vivo amber suppression with **11** and **12** [30]. By introducing two rational mutations, even the bulkier analogs **13** and **14** become viable substrates for the enzyme. Translational incorporation of **14** provided a recombinant protein suitable for labeling with a phosphine-bearing fluorescein derivative via Staudinger ligation.

Translational incorporation of *N*^ε-acetyllysine (**15**) [31,32,33], which is not a viable substrate for wild-type PylRS, was accomplished using a modified enzyme identified by subjecting a large library (~10⁸) of *Mm*PylRS mutants to directed evolution involving three rounds of selection [31]. Additional randomization and selection involving a different set of PylRS residues increased the efficiency of the enzyme even further [32].

It enabled the preparation and study of recombinant histone H3 [32] and Cyclophilin A [33] in which a specific lysine residue was acetylated. Interestingly, the originally evolved PylRS [31] was also used to translationally incorporate ketone **6** [23]. Thanks to increased nucleophilicity of the keto group relative to its amide counterpart, recombinant proteins containing **6** can be tagged with hydrazide- and alkoxyamine-bearing probes under mild conditions.

Although all the known lysine-based pyrrolysine surrogates contain an *N*^ε-acyl substituent that is a required structural element for effective recognition by PylRS, a two-step strategy for the synthesis of recombinant proteins containing NAAs devoid of this key feature was developed. As such, while *N*^ε-methyllysine (**16**) is not a suitable substrate for the PylRS–tRNA^{Pyl} system, its *N*^ε-substituted derivatives **17–19** are [34,35,36]. Consequently, their translational incorporation, followed by selective removal of the carbamate substituents under mild conditions, results in the overall insertion of **16**. This strategy was used to site-specifically install **16** in full-length histones H3 [34] and H2B [36] and the green fluorescent protein (GFPuv) [35].

In a similar manner, the strategy of the genetically-encoded orthogonal protection developed by Chin et al. also relies on the use of a lysine derivative bearing an easily removable *N*^ε-substituent [37,38]. Thus, the PylRS–tRNA^{Pyl}-mediated translational incorporation of **11**, followed by orthogonal protection of both the remaining lysine residues and the N terminal end of the resulting recombinant protein, makes it possible to selectively remove the Boc protection which, in turn, exposes just one lysine residue to further functionalization. This method was used to prepare a homogenous diubiquitin bearing an atypical linkage [37] and site-specifically install *N*^ε,*N*^ε-dimethyllysine (**20**) in a recombinant histone [38].

As a testament to the yet untapped potential of the pyrrolysine incorporation system, Liu et al. reported an evolved *Mm*PylRS–tRNA^{Pyl} pair that enables incorporation of a first canonical amino acid, phenylalanine (**21**) [39]. Moreover, yet another engineered pair was shown to accept *p*-iodo- and *p*-bromophenylalanine (**22** and **23**, respectively) as competent substrates. A recombinant protein containing **22** was subsequently site-specifically labeled

via Suzuki–Miyaura cross-coupling with a boronic acid derived from a fluorescent dye. These results are significant as they demonstrate that the PylRS–tRNA^{Pyl} system is remarkably flexible and can be made to accommodate amino acids bearing even less steric and electronic resemblance to pyrrolysine than previously thought.

PylRS–tRNA^{Pyl} pairs are also key components in two systems for simultaneous translational incorporation of two different NAAs into a single protein [40,41]. In one of such systems, an orthogonal ribosome was evolved which efficiently decodes both a series of quadruplet codons (in preference to their unexpanded triplet counterparts) and the UAG codon on an orthogonal mRNA. In the other system [40], to avoid a conflict with another amber suppressor, tRNA^{Pyl} was successfully reassigned to a different nonsense codon (UAA, ochre) by taking advantage of the previously reported insignificance of the tRNA^{Pyl} anticodon for PylRS recognition [42]. By encoding alkyne **8** and an azide-containing amino acid into a single protein, it was then possible to perform intramolecular cyclization via CuAAC [40]. Intriguingly, the same amino acid residues in a different protein were also engaged in intermolecular CuAAC reactions with fluorescent dyes containing either an azide or a terminal alkyne group [41].

The PylRS–tRNA^{Pyl} pairs in eukaryotic expression systems

There is a growing body of evidence that PylRS–tRNA^{Pyl} pairs are also fully functional and orthogonal in a range of eukaryotic expression systems. Yokoyama et al. identified an external promoter that enables transcription of the tRNA^{Pyl} gene in mammalian cells [43]. It was also demonstrated that PylRS specific for an NAA can be either rationally designed [43] or evolved [44,45] in *E. coli* and then transferred into mammalian expression systems without any loss of orthogonality. Interestingly, a range of photocaged amino acids, including **19** [46], **24** [44], and **25** [47] were incorporated into proteins in human cells using evolved PylRS–tRNA^{Pyl} pairs. Additionally, Chin et al. overcame a series of problems related to transcription of the tRNA^{Pyl} gene in yeast and expanded its genetic code by a range of amino acid including, among others, **25** and **26** [48].

Conclusions

Within a few years since its discovery, the PylRS–tRNA^{Pyl} system has been established as a very powerful tool for the synthesis of recombinant proteins containing NAAs adorned with a wide variety of functional groups. Although a series of empirical rules make it possible to predict with some level of confidence if an amino acid is a viable pyrrolysine surrogate for the PylRS–tRNA^{Pyl} system, the level of tolerance for some structural features has so far not been explored. For example, it is unknown if lysine analogs **27–29** (Figure 3) bearing additional substituents on the side chain or heteroatoms/multiple bonds within it are acceptable by the synthetase. Because the studies of the PylRS–tRNA^{Pyl} system are primarily application driven, the above questions will no doubt be addressed if a need for a recombinant protein containing any of these amino acids emerges. Given that the amber-codon suppression is frequently exploited in the development of advanced protein-engineering strategies, the unique and beneficial features of the PylRS–tRNA^{Pyl} pair will almost certainly elevate it to the status currently occupied by its *Methanococcus jannaschii* TyrRS–tRNA^{Tyr} counterpart. It should be noted that the two systems generally complement each other as, with the notable exception of the phenylalanine derivatives **22–23**, they can accommodate different sets of substrates.

Acknowledgments

Financial support by the National Institute of General Medical Sciences, the National Institutes of Health (grant GM061796 supplemented by American Recovery and Reinvestment Act funds) is gratefully acknowledged.

References

1. James CM, Ferguson TK, Leykam JF, Krzycki JA. The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. *J Biol Chem*. 2001; 276:34252–34258. [PubMed: 11435424]
2. Srinivasan G, James CM, Krzycki JA. Pyrrolysine encoded by UAG in Archaea. Charging of a UAG-decoding specialized tRNA. *Science*. 2002; 296:1459–1462. [PubMed: 12029131]
3. Hao B, Gong W, Ferguson TK, James CM, Krzycki JA, Chan MK. A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science*. 2002; 296:1462–1466. [PubMed: 12029132]
4. Krzycki, JA. Translation of UAG as pyrrolysine. In: Atkins, JF.; Gesteland, RF., editors. *Recoding. Expansion of Decoding Rules Enriches Gene Expression*. Springer; 2010. p. 53-78.
5. Rother M, Krzycki JA. Selenocysteine, pyrrolysine, and the unique energy metabolism of methanogenic Archaea. *Archaea*. 2010; 10:1155/2010/453642
6. Blight SK, Larue RC, Mahapatra A, Longstaff DG, Chang E, Zhao G, Kang PT, Green-Church KB, Chan MK, Krzycki JA. Direct charging of tRNA^{CUA} with pyrrolysine *in vitro* and *in vivo*. *Nature*. 2004; 431:333–335. [PubMed: 15329732]
7. Polycarpo C, Ambrogelly A, Bérubé A, Winbush SAM, McCloskey JA, Crain PF, Wood JL, Söll D. An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proc Natl Acad Sci U S A*. 2004; 101:12450–12454. [PubMed: 15314242]
8. Herring S, Ambrogelly A, Polycarpo CR, Söll D. Recognition of pyrrolysine tRNA by the *Desulfitobacterium hafniense* pyrrolysyl-tRNA synthetase. *Nucleic Acids Res*. 2007; 35:1270–1278. [PubMed: 17267409]
9. Zhang Y, Gladyshev VN. High content of proteins containing 21st and 22nd amino acids, selenocysteine and pyrrolysine, in a symbiotic deltaproteobacterium of gutless worm *Olavius algarvensis*. *Nucleic Acid Res*. 2007; 35:4952–4963. [PubMed: 17626042]
10. Liu CC, Schultz PG. Adding new chemistries to the genetic code. *Annu Rev Biochem*. 2010; 79:413–444. [PubMed: 20307192]
11. Wang Q, Parrish AR, Wang L. Expanding the genetic code for biological studies. *Chem Biol*. 2009; 16:323–336. [PubMed: 19318213]
12. Hao B, Zhao G, Kang PT, Soares JA, Ferguson TK, Gallucci J, Krzycki JA, Chan MK. Reactivity and chemical synthesis of L-pyrrolysine – the 22nd genetically encoded amino acid. *Chem Biol*. 2004; 11:1317–1324. [PubMed: 15380192]
13. Longstaff DG, Larue RC, Faust JE, Mahapatra A, Zhang L, Green-Church KB, Krzycki JA. A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine. *Proc Natl Acad Sci U S A*. 2007; 104:1021–1026. [PubMed: 17204561]
14. Théobald-Dietrich A, Giegé R, Rudinger-Thirion J. Evidence for the existence in mRNAs of a hairpin element responsible for ribosome dependent pyrrolysine insertion into proteins. *Biochimie*. 2005; 87:813–817. [PubMed: 16164991]
15. Longstaff DG, Blight SK, Zhang L, Green-Church KB, Krzycki JA. *In vivo* contextual requirements for UAG translation as pyrrolysine. *Mol Microbiol*. 2007; 63 :229–241. [PubMed: 17140411]
16. Namy O, Zhou Y, Gundllapalli S, Polycarpo CR, Denise A, Rousset JP, Söll D, Ambrogelly A. Adding pyrrolysine to the *Escherichia coli* genetic code. *FEBS Lett*. 2007; 581:5282–5288. [PubMed: 17967457]
- 17. Nozawa K, O'Donoghue P, Gundllapalli S, Arais Y, Ishitani R, Umehara T, Söll D, Nureki O. Pyrrolysyl-tRNA synthetase – tRNA^{Pyl} structure reveals the molecular basis of orthogonality. *Nature*. 2009; 457:1163–1168. By studying the crystal structure of the *D. hafniense* PylRS–tRNA^{Pyl} complex, the authors identified intricate and unique interactions between the two components, thus providing an explanation for the observed orthogonality of the pair. [PubMed: 19118381]
18. Polycarpo CR, Herring S, Bérubé A, Wood JL, Söll D, Ambrogelly A. Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. *FEBS Lett*. 2006; 580:6695–6700. [PubMed: 17126325]

19. Li WT, Mahapatra A, Longstaff DG, Bechtel J, Zhao G, Kang PT, Chan MK, Krzycki JA. Specificity of pyrrolysyl-tRNA synthetase for pyrrolysine and pyrrolysine analogs. *J Mol Biol.* 2009; 385:1156–1164. [PubMed: 19063902]
20. Fekner T, Li X, Chan MK. Pyrrolysine analogs for translational incorporation into proteins. *Eur J Org Chem.* 2010:4171–4179.
21. Liu WR, Wang YS, Wan W. Synthesis of proteins with defined posttranslational modifications using the genetic noncanonical amino acid incorporation approach. *Mol BioSyst.* 2011; 7:38–47. [PubMed: 21088799]
22. Kobayashi T, Yanagisawa T, Sakamoto K, Yokoyama S. Recognition of non- α -amino substrates by pyrrolysyl-tRNA synthetase. *J Mol Biol.* 2009; 385:1352–1360. [PubMed: 19100747]
23. Huang Y, Wan W, Russell WK, Pai PJ, Wang Z, Russell DH, Liu W. Genetic incorporation of an aliphatic keto-containing amino acid into proteins for their site-specific modifications. *Bioorg Med Chem Lett.* 2010; 20:878–880. [PubMed: 20074948]
24. Fekner T, Li X, Lee MM, Chan MK. A pyrrolysine analogue for protein click chemistry. *Angew Chem Int Ed.* 2009; 48:1633–1635. Using a pyrrolysine analog containing a terminal alkyne, the authors prepared a recombinant protein amenable to labeling via CuAAC with organic azides. This is the first study to use the PylRS-tRNA^{Pyl} system to generate proteins for 'click' chemistry.
25. Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin JW. Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA synthetase/tRNA^{CUA} pair and click chemistry. *J Am Chem Soc.* 2009; 131:8720–8721. [PubMed: 19514718]
26. Li X, Fekner T, Chan MK. N^6 -(2-(*R*)-Propargylglycyl)lysine as a clickable pyrrolysine mimic. *Chem Asian J.* 2010; 5:1765–1769. [PubMed: 20544791]
27. Kaya E, Gutmiedl K, Vrabel M, Müller M, Thumbs P, Carell T. Synthesis of threefold glycosylated proteins using click chemistry and genetically encoded unnatural amino acids. *ChemBioChem.* 2009; 10:2858–2861. [PubMed: 19899090]
28. Li X, Fekner T, Ottesen JJ, Chan MK. A pyrrolysine analogue for site-specific protein ubiquitination. *Angew Chem Int Ed.* 2009; 48:9184–9187. Using a pyrrolysine analog containing a cysteine residue, the authors generated a recombinant protein suitable for nonenzymatic ubiquitylation via expressed protein ligation. This is the first study to install this important post-translational modification in a single step by purely chemical means via coupling between two recombinant fragments.
29. Fekner T, Li X, Chan MK. Nonenzymatic ubiquitylation. *ChemBioChem.* 2011; 12 :21–33. [PubMed: 21181847]
30. Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N^{ϵ} -(*o*-azidobenzoyloxycarbonyl)lysine for site-specific protein modification. *Chem Biol.* 2008; 15:1187–1197. [PubMed: 19022179]
31. Neumann H, Peak-Chew SY, Chin JW. Genetically encoding N^{ϵ} -acetyllysine in recombinant proteins. *Nat Chem Biol.* 2008; 4:232–234. [PubMed: 18278036]
32. Neumann H, Hancock SM, Buning R, Routh A, Chapman L, Somers J, Owen-Hughes T, van Noort J, Rhodes D, Chin JW. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol Cell.* 2009; 36:153–163. [PubMed: 19818718]
33. Lammers M, Neumann H, Chin JW, James LC. Acetylation regulates Cyclophilin A catalysis, immunosuppression and HIV isomerization. *Nat Chem Biol.* 2010; 6 :331–337. [PubMed: 20364129]
- 34. Nguyen DP, Alai MMG, Kapadnis PB, Neumann H, Chin JW. Genetically encoding N^{ϵ} -methyl-L-lysine in recombinant histones. *J Am Chem Soc.* 2009; 131:14194–14195. N^{ϵ} -Boc- N^{ϵ} -methyllysine was translationally incorporated into a recombinant protein using the PylRS-tRNA^{Pyl} system. Following posttranslational removal of the Boc group, the N^{ϵ} -methyllysine residue was revealed. This is the first study to use temporal protection to render an amino acid a competent PylRS substrate, thus expanding a range of recombinant proteins that can be prepared. [PubMed: 19772323]

35. Wang YS, Wu B, Wang Z, Huang Y, Wan W, Russell WK, Pai PJ, Moe YN, Russell DH, Liu WR. A genetically encoded photocaged *N*^ε-methyl-L-lysine. *Mol Biosyst.* 2010; 6:1557–1560. [PubMed: 20711534]
36. Ai HW, Lee JW, Schultz PG. A method to site-specifically introduce methyllysine into proteins in *E. coli*. *Chem Commun.* 2010; 46:5506–5508.
37. Virdee S, Ye Y, Nguyen DP, Komander D, Chin JW. Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol.* 2010; 6:750–757. [PubMed: 20802491]
- 38. Nguyen DP, Alai MMG, Virdee S, Chin JW. Genetically directing ε-N, N-dimethyl-L-Lysine in recombinant histones. *Chem Biol.* 2010; 17:1072–1076. Following translational incorporation of *N*^ε-Boc-lysine and subsequent protecting group manipulations, the authors were able to further modify the side chain of a selected lysine residue. The method provides a means to prepare recombinant proteins containing amino acids that are not substrates for PylRS. [PubMed: 21035729]
- 39. Wang YS, Russell WK, Wang Z, Wan W, Dodd LE, Pai PJ, Russell DH, Liu WR. The *de novo* engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of L-phenylalanine and its derivatives. *Mol Biosyst.* 2011; 7:714–717. The authors evolved PylRSs that accommodate phenylalanine and its derivatives as competent substrates. This study demonstrates that the PylRS–tRNA^{Pyl} system is remarkably flexible and can be modified to accept amino acids that not only bear little resemblance to pyrrolysine but also belong to the substrate set of *M. jannaschii* TyrRS–tRNA^{Tyr}. [PubMed: 21234492]
- 40. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature.* 2010; 464:441–444. Using an evolved orthogonal ribosome that decodes both a range of quadruplet codons and an amber codon, the authors demonstrated that it is possible to simultaneously and site-specifically incorporate two NAAs into a recombinant protein. This is the first study to provide a general method for the synthesis of such proteins. [PubMed: 20154731]
- 41. Wan W, Huang Y, Wang Z, Russell WK, Pai PJ, Russell DH, Liu WR. A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angew Chem Int Ed.* 2010; 49:3211–3214. Following the reassignment of tRNA^{Pyl} into a different canonical stop codon, the authors used the PylRS–tRNA^{Pyl} pair in combination with an evolved *M. jannaschii* TyrRS–tRNA^{Tyr} pair to prepare a recombinant protein containing two different NAAs. This is the first study to report a practical method for the synthesis of this class of proteins.
42. Ambrogelly A, Gundllapalli S, Herring S, Polycarpo C, Frauer C, Söll D. Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. *Proc Natl Acad Sci U S A.* 2007; 104:3141–3146. [PubMed: 17360621]
43. Mukai T, Kobayashi T, Hino N, Yanagisawa T, Sakamoto K, Yokoyama S. Adding L-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem Biophys Res Commun.* 2008; 371:818–822. [PubMed: 18471995]
44. 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.* 2009; 48:4052–4055.
45. Schwarzer D. Hacking the genetic code of mammalian cells. *ChemBioChem.* 2009; 10:1602–1604. [PubMed: 19533721]
46. Groff D, Chen PR, Peters FB, Schultz PG. A genetically encoded ε-N-methyl lysine in mammalian cells. *ChemBioChem.* 2010; 11:1066–1068. [PubMed: 20422671]
47. Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW. Genetically encoded photocontrol of protein localization in mammalian cells. *J Am Chem Soc.* 2010; 132:4086–4088. [PubMed: 20218600]
48. Hancock SM, Uprety R, Deiters A, Chin JW. Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J Am Chem Soc.* 2010; 132:14819–14824. [PubMed: 20925334]

Figure 1.
Pyrrolysine (1) and its original surrogates 2–4.

Figure 2.

Representative examples of compounds incorporated into recombinant proteins using the pyrrolysine incorporation system (*Mm*PyIRS–tRNA^{PyI} for **5**, **7**, **9–14**, **19**, **21–24**; *Mb*PyIRS–tRNA^{PyI} for **6**, **8**, **15**, **17–18**, **25–26**; *Mb*PyIRS–*Mm*tRNA^{PyI} for **19**). Note that **16** and **20** were not incorporated directly and that two systems were used for **19**). Boc = *tert*-butoxycarbonyl, Alloc = allyloxycarbonyl, Cbz = benzyloxycarbonyl, 2-N₃-Cbz = 2-azidobenzyloxycarbonyl, Ac = acetyl, NBOC = 2-nitrobenzyloxycarbonyl.

R¹, R² = functional groups

X = a heteroatom

Figure 3.
Amino acids of unknown acceptability by the PyIRS-tRNA^{Pyl} system.