ACS chemical biology

Letters

pubs.acs.org/acschemicalbiology

Terms of Use

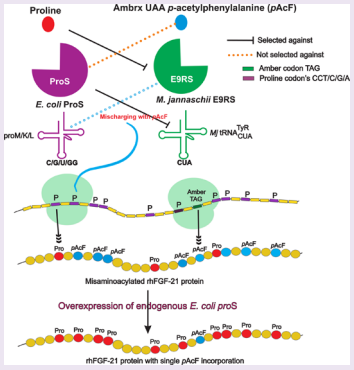
Role of tRNA Orthogonality in an Expanded Genetic Code

Tsotne Javahishvili,*[,†] Anthony Manibusan,[†] Shaila Srinagesh,[†] Darin Lee,[†,§] Semsi Ensari,[†,⊥] Mark Shimazu,[†] and Peter G. Schultz[‡]

[†]Ambrx, Inc., La Jolla, California 92037, United States

[‡]The Scripps Research Institute, La Jolla, California 92037, United States

ABSTRACT: We found that *Methanocaldococcus jannaschii* DSM2661 tyrosyl-tRNA synthetase (*Mj* E9RS), specifically evolved to charge its cognate tRNA with the unnatural amino acid *p*-acetylphenylalanine (*p*AcF) in *E. coli*, misaminoacylates the endogenous *E. coli* prolyl-tRNAs with *p*AcF at a low level (0.5% per proline frequency) in both the absence or presence of its co-evolved amber suppressor tRNA (*M. jannaschii* tyrosyl-tRNA, $tRNA^{MjTyr}_{CUA}$). In contrast to other *E. coli* tRNAs, the identity elements for recognition of the proly tRNAs by the *E. coli* prolyl-tRNA synthetase (C1, G72, and A73) are similar to those in $tRNA^{MjTyr}_{CUA}$. Although the unique acceptor stem identity elements of the prolyl-tRNAs likely lower their recognition by the other endogenous aaRSs in *E. coli*, resulting in enhanced fidelity in the wild type strain, they lead to misaminoacylation by the archae-derived E9RS. Misincorporation of pAcF for proline was resolved to below detectable levels by overexpression of the endogenous *E. coli* prolyl-tRNA synthetase (*pro*S) gene in combination with additional genomic manipulations to further increase the intracellular ratio of the ProS over its cognate proline tRNAs. These experiments suggest another mechanism by which the cell maintains the high fidelity of protein biosynthesis.



The ability to expand the universal genetic code depends on a high level of fidelity for unnatural amino acid (UAA) incorporation into proteins that rivals the natural level of translational precision. There are several levels of genomic and biochemical requirements for such an expanded code to work efficiently, based on relatively well understood mechanisms.[1] Some of these requirements are (a) low physiological levels of cross-talk among the endogenous and exogenous components of the translation machinery while supplying the cell with sufficient amounts of active synthetase and correctly processed tRNA, (b) chemically stable UAAs that are readily up-taken from the media, (c) tolerable read-thru levels of the genomic nonsense or frameshift codons, and (d) correct charging of tRNAs by their cognate aminoacyl-tRNA synthetases (aaRSs). Multiple structural elements have evolved in both the tRNAs and aaRSs that result in this high degree of accuracy in the tRNA recognition, acylation, and editing processes. We have exploited these specificity determinants to develop new tRNA/aaRS pairs that allow us to genetically encode amino acids with novel chemical and biological properties in both prokaryotic and eukaryotic organisms.[2,3] These unnatural amino acids are specified by unique stop or frameshift codons in the gene of interest.[4,5]

To ensure high translational fidelity, any tRNA/aaRS pair encoding a new amino acid must be orthogonal to the host's endogenous tRNAs and aminoacyl-tRNA synthetases, i.e., it cannot cross react to misincorporate the unnatural amino acids in place of any canonical amino acid, or the converse. The generation of such orthogonal tRNA/aaRS pairs is typically based on the distinct identity elements used by bacterial, archaeal, and eukaryotic tRNAs.[6] For example, the first base pair in the archaeal *M. jannashii* (*Mj*) tyrosine tRNA is C1:G72, which in addition to the discriminator base A73 is required for recognition of the cognate tRNA by the *Mj* tyrosyl-tRNA synthetase (Figure 1).[7−10] In contrast, the identity element in the acceptor stem of bacterial tRNAs is a G1:C72 base pair. Thus by transferring the archaeal $tRNA^{Tyr}$-TyrRS pair to *E. coli* and engineering additional negative recognition elements into *Mj* $tRNA^{Tyr}$ for the bacterial aminoacyl-tRNA synthetases, one can generate a highly orthogonal pair that can be evolved to encode noncanonical amino acids with high translational fidelity. Indeed, amber codon suppressor tRNA/aaRS pairs evolved from the *Mj* $tRNA^{Tyr}$/TyrRs pair have been used to site-specifically incorporate a large number of unnatural amino acids (UAAs) into proteins in bacteria, including photo-cross-linking, fluorescent, chemically reactive, and metal binding amino acids, with yields up to 10 g per liter.[13−17] However, recently we found that when the unnatural amino acid *p*AcF is site-specifically incorporated to produce a mutant recombinant fibroblast growth factor 21 (rhFGF-21) in *E. coli*, it is misincorporated for proline with roughly 0.5% frequency/proline. This observation has led us to develop new methods to further enhance the fidelity of unnatural amino acid incorporation and has provided new insights into those factors that influence the fidelity of the translation process.

The investigation of the misincorporation phenomenon of *p*AcF was triggered by an observation of multiple conjugation products in a reaction between rhFGF-21-*p*AcF and a

Received: July 10, 2013
Accepted: January 20, 2014
Published: January 20, 2014

 © 2014 American Chemical Society dx.doi.org/10.1021/cb4005172 | *ACS Chem. Biol.* 2014, 9, 874−879

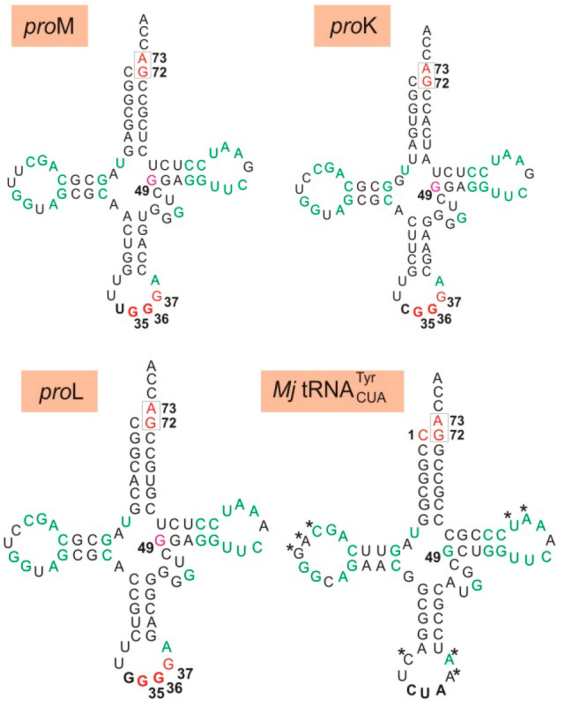


Figure 1. Secondary structures of *E. coli* prolyl-tRNAs *proM*, K, and L, and *Mj* $tRNA^{MjTyr}_{CUA}$.[11,12] Anticodon loops are shown in bold. Nucleotides required for recognition by the respective canonical aaRSs, i.e., identity elements, are in red and numbered. Asterisks (*) denote the positions that were randomized to generate an amber suppressor tRNA library.[11,12] The homology between the suppressor tRNA used in our work and *E. coli* prolyl-tRNAs are shown in green. Critical identity elements, G72 and A73 for *E. coli* prolyl-tRNA synthetase and C1:G72 and A73 for *Mj* tyrosyl-tRNA synthetase, are shown in boxes.

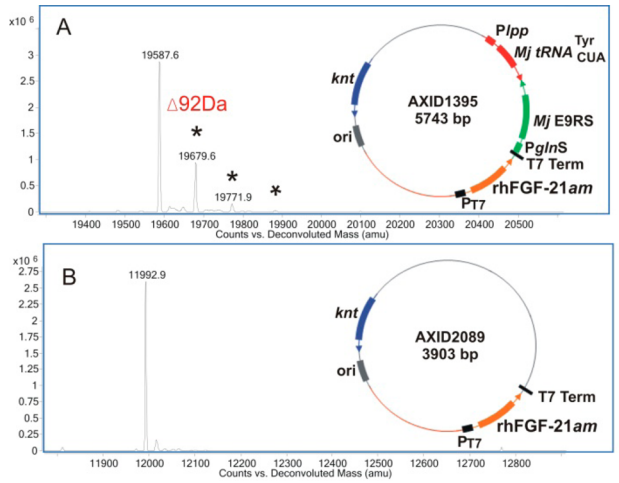


Figure 2. LC−MS evaluation of solubilized rhFGF-21-pAcF inclusion bodies from plasmid constructs with, AXID1395, and without the Ambrx suppresion cassette, AXID2089. (A) Mass spectra of rhFGF-21-pAcF from AXID1395 grown in defined media exhibited multiple +92 Da species, suggesting misincorporation of pAcF for proline at proline codons. The expected intact mass is 19587.1 Da; the observed mass is 19587.6 Da. (B) Removal of the suppresion cassette enriches for truncated rhFGF-21; however, no masses were observed that correlated to misincorporation. The expected mass of truncated rhFGF-21 is 11992.5 Da; the observed mass is 11992.9 Da. Peaks labeled as "*" are Δ92 Da species. Plasmid map symbols: P_{lpp}, *E. coli* murein lipoprotein gene promoter; *Mj* $tRNA_{CUA}$, *M jannaschii* tyrosine tRNA evolved recognizing amber stop codon; *Mj* E9RS, tyrosyl-tRNA synthetase, specifically evolved to charge its cognate tRNA with *p*AcF; P_{glnS}, promoter of *E. coli* glutamyl-tRNA synthetase gene; T7term, transcription terminator sequence for bacteriophage T7 RNA Polymerase; rhFGF-21, recombinant gene encoding the mature human fibroblast growth factor 21 containing amber stop codon in the 108th positon (Q108*am*); P_{T7}, promoter for bacteriophage T7 RNA Polymerase; *ori*, pBR322 origin of replication; *knt*, kanamycin nucleotidyltransferase.

poly(ethylene glycol) (PEG) polymer. rhFGF-21-*p*AcF produced in high density bacterial fermentation (2.5 g per liter) using a *p*AcF specific $tRNA_{CUA}$/aaRS pair afforded other higher molecular weight PEGylated species in addition to the expected mono-PEGylated protein. The earliest processing stage in which the high molecular weight PEGylated species were observed was also investigated, and it was determined that unpurified protein from solubilized inclusion bodies, when subjected to PEGylation, also yielded multiple PEGylated products. Possible causes of the higher molecular weight protein species in the PEGylation reaction were systematically addressed and ruled out, including protein−protein aggregation, protein oligomerization via disulfide bond formation, and high molecular weight contaminants in the PEG raw material reagent. Subsequent reduced intact mass analysis of the mutant protein revealed a ∼17% containment of a +92 Da species (Figure 2).

It was observed by peptide mapping of analyte enriched for the +92 Da species by RP -HPLC that the modification was present across multiple peptides and not limited to a specific site (Figure 3). Mass spectral analysis of the tryptic digest revealed that this contaminant corresponded to substitution of proline with *p*AcF at various sites (Table 1). Because rhFGF-21 has 22 endogenous prolines, misincorporation of *p*AcF for Pro at a 0.5% frequency afforded a clearly detectable impurity. To confirm this observation, wild-type (wt) rhFGF-21 was produced in the presence of the orthogonal $tRNA_{CUA}$/E9RS without any amber mutant in the rhFGF-21 gene, and the +92

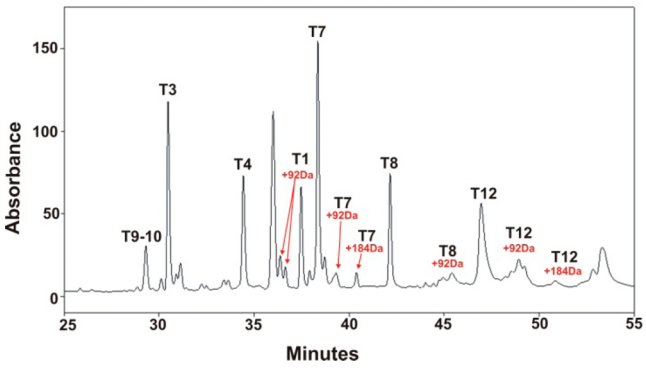


Figure 3. Peptide mapping analysis of rhFGF-21-*p*AcF enriched for the +92 Da species. Multiple new peaks with increased masses of +92 or +184 Da were observed to contain this modification when compared to the native tryptic peptides. The +184 Da modification applied to peptides that contained two sites of misincorporation. These mass shifts are consistent with the amino acid substitution of a *p*AcF and was observed across multiple site of rhFGF-21-*p*AcF.

species was still observed at roughly the same abundance. There was, however, a significant difference in relative misincorporation depending on the type of media that was used and the temperature of induction for rhFGF-21 expression, which could be due to previously reported effects

Table 1. Theoretical Tryptic Digest of rhFGF-21-*p*AcF[a]

peak no.	from−to	tryptic sequence
T1	1−18	MHPIPDSSPLLQFGGQVR
T2	19−20	QR
T3	21−37	YLYTDDAQQTEAHLEIR
T4	38−57	EDGTVGGAADQSPESLLQLK
T5	58−70	ALKPGVIQILGVK
T6	71−73	TSR
T7	74−97	FLCQRPDGALYGSLHFDPEACSFR
T8	98−123	ELLLEDGYNVYQSEAHGLPLHLPGNK
T9	124−127	SPHR
T10	128−132	DPAPR
T11	133−136	GPAR
T12	137−176	FLPLPGLPPAPPEPPGILAPQPPDVGSSDPLSMVGPSQGR
T13	177−182	SPSYAS

[a]Prolines are highlighted where the +92 Da species was observed by peptide mapping. The peak numbers T1−T13 correspond to the peptide mapping analysis in Figure 3.

of bacterial growth rates, relative amino acid and tRNA abundances and codon usage.

Misincorporation due to misrecognition at the level of codon−anticodon interactions is unlikely as there is no real bias of *p*AcF at different proline sites encoded by the canonical codons (Figure 4), and base pairing, including the wobble rule,

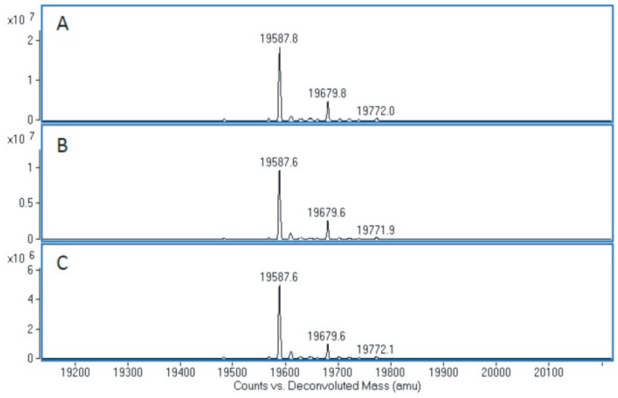


Figure 4. Intact mass analysis of rhFGF-21-*p*AcF with all proline codons changed to either CCT (A), CCA (B), or CCG (C). No differences in the abundance of +92 Da species were observed between the three proline codon site changes, suggesting an unbiased incorporation of *p*AcF. The expected mass of intact rh-FGF-21-*p*AcF is 19587.1 Da; the observed masses are 19587.8, 19587.6, and 19587.6 Da.

does not allow the *amber* suppressing anticodon CUA to recognize any of the 4 proline codons (CCU, CCC, CCA, CCG). It is possible that the *E. coli* prolyl-tRNA synthetase aminoacylates $tRNA^{Pro}$ with *p*AcF to a limited extent. However, these amino acids are quite distinct structurally. Moreover, attempts to increase endogenous proline levels to more effectively compete with *p*AcF for binding to prolyl-RS, either by adding exogenous proline to the media or by knocking out the *putA* gene encoding proline dehydrogenase (the first step in proline catabolism to glutamate), had little effect on the observed +92 species relative abundance. Increased osmotic stress by addition of 0.8 M NaCl resulted in an increase, rather than a decrease, in misincorporation of *p*AcF. These results suggested that it is unlikely that *p*AcF is a substrate for the endogenous prolyl-RS.

An alternative explanation for misincorporation of *p*AcF at proline sites is that the *M. jannaschii* *p*AcF specific aaRS to some degree recognizes the *E. coli* proline tRNA family and aminoacylates these tRNAs with *p*AcF. Indeed, deletion of the *Mj* aaRS gene from a plasmid encoding wt rhFGF-21 and $tRNA_{CUA}$ (plasmid AXID2103) eliminated misincorporation, whereas deletion of the suppressor $tRNA_{CUA}$ (plasmid AXID2083) had no detectable effect on misincorporation (Figure 5). The bacterial $tRNA^{Pro}$ family is the only set of

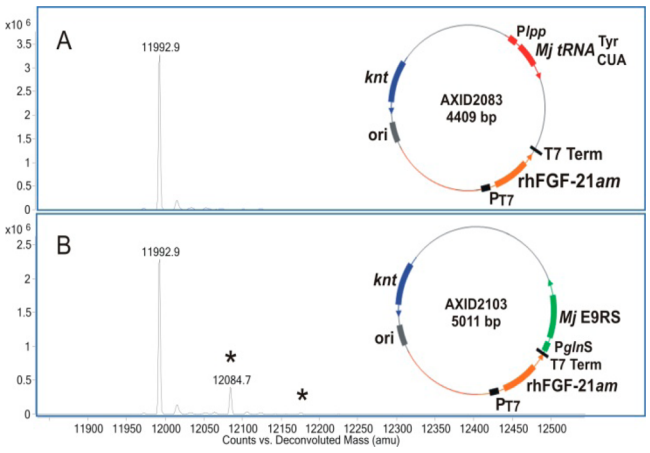


Figure 5. Deconvoluted mass spectra of rhFGF-21-*p*AcF from plasmid constructs containing individual components of the suppression cassette. (A) In AXID2083 there was no evidence of misincorporation observed in the plasmid without the *M. jannaschii* E9RS synthetase . (B) In AXID 2103, the +92 Da species from truncated rhFGF21 was observed with the *M. jannaschii* E9RS synthetase and the absence of the modified *M. jannaschii* Tyr-tRNA in the plasmid construct, indicative of misincorporation. The expected mass of truncated rhFGF-21 is 11992.5 Da; the observed mass is 11992.9 Da for both constructs.

tRNAs to have a C1:G72 base pair in their acceptor stem, which is the same identity element used by archaea (and distinct from the other *E. coli* tRNAs). In fact a single substitution at G72 in $tRNA^{Pro}$ leads to a large decrease in aminoacylation by the endogenous prolyl-tRNA synthetase.[21−23] In addition the nucleotides between A14 and U18 are conserved between the *E. coli* $tRNA^{Pro}$ family and the *Mj* $tRNA_{CUA}$ as is the short variable loop.

If misrecognition of $tRNA^{Pro}$ by the *p*AcF aaRS is the cause of misincorporation, overexpression of the *E. coli* prolyl-tRNA synthetase gene (*pro*S) should reduce or eliminate this phenomenon by competing with the *Mj p*AcF RS for the acylation of endogenous $tRNA^{Pro}$ with proline rather than *p*AcF. To this end we subcloned the *E. coli* K-12 W3110 *pro*S gene, which has a temperature-sensitive (≥34 °C) tryptophan to arginine mutation at residue 375, behind its own promoter in the plasmid AXID2191, and transformed the cell line W3110B57 with this expression plasmid.[19] At 30 °C there was no *p*AcF misincorporation detected due to increased expression of the mutant *pro*S gene (Figure 6), nor was there

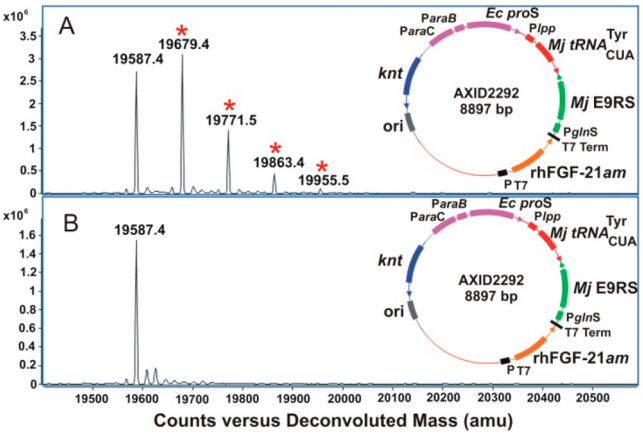


Figure 6. LC−MS analyses of rhFGF-21-*p*AcF from the optimized expression plasmid, AXID2292 at (A) 30 °C and (B) 37 °C. The expected mass of intact rhFGF-21 is 19587.1 Da; the observed mass is 19587.4 Da for both constructs. Peaks labeled as "*" are Δ92 Da species. Map symbols: P_{araC}, *E. coli* arabinose operon dual regulator gene; P_{araB} ; arabinose operon promoter; *Ec pro*S, *E. coli* wild type prolyl-tRNA synthetase gene.

any significant effect on mutant rhFGF21 expression levels. At 37 °C *p*AcF misincorporation was restored at proline residues in rhFGF-21*am* as the mutant ProS is not functional at this lethal temperature, i.e., only the wild-type protein is functional and translated from the transcript of the genomic copy of *pro*S. Thus it appears that misincorporation of *p*AcF for proline in these experiments is due to a low level of misrecognition of $tRNA^{Pro}$ by the *p*AcF aaRS.

Why is misincorporation of the UAA observed only with proline? Proline is the only amino acid with a secondary amine, and as a consequence its incorporation into the growing polypeptide chain is 3−4 orders of magnitude slower than the incorporation of the other canonical amino acids.[24−26] Thus if the endogenous *E. coli* ProS misaminoacylates $tRNA^{Pro}$ with any other canonical amino acid, that amino acid will be incorporated into peptide at the A site (assuming little EF-Tu bias) more rapidly than proline.[25] Therefore since binding to the A site by the tRNA-EF-Tu complex is reversible,[22,27] there will be an inherent bias for misincorporation by misacylated $tRNA^{Pro}$ relative to other misacylated tRNAs. ProS has an active site that can broadly accept a range of small hydrophobic amino acids but contains an editing domain to ensure genetic code integrity. However, misincorporation can also occur by binding and aminoacylation of any of the $tRNA^{Pro}$ by non-proline specific aaRSs. By mutation of the tRNA identity element from G1:C72 to the archaeal C1:G72 base pair, the *E. coli* $tRNA^{Pro}$ increases its orthogonality to the other bacterial tRNA/aaRS pairs and further decreases the potential for misincorporation of another canonical amino acid at proline sites. This strategy, in addition to editing, EFTu binding, and post-aminoacylation enzymatic processing, likely represents yet another mechanism by which *E. coli* can ensure high fidelity in mRNA translation. However, it is also clear that this mechanism is not used universally, so other mechanisms likely exist to ensure high fidelity. Unfortunately, the C1:G72 base pair makes the bacterial $tRNA^{Pro}$ similar to our archaeal *Mj* orthogonal $tRNA_{CUA}$/aaRS pair and leads to misincorporation. This misincorporation is overcome by overexpressing of the *pro*S gene in the host strain.

The cross-talk among the cellular components of the translation machinery (the last step in the Central Dogma) is widely accepted to have been evolutionary balanced so that it is not detrimental to the proteome or the cell itself. One of the results of this balancing act is the intrinsic intracellular ratio of the matured, fully processed tRNAs to their cognate synthetases (which almost certainly changes as cells go through various physiological states). In our synthetic strains with a 21 amino acide code, due to the overlapping recognition elements in both *E. coli* proline and *M. jannaschii* tyrosine tRNAs, *M. jannaschii* tyrosyl-tRNA synthetase aminoacylates available uncharged endogenous proline tRNAs. When the *E. coli* endogenous *pro*S is overexpressed in the cell by increasing its copy number, the phenotype becomes undetectable by LC−MS/RP since the balance shifts toward the endogenous ProS and the remaining excess, if any, of the free proline tRNAs becomes insufficient for the misaminoacylation. Thus, in this case one consequence of adding additional translational components, i.e., an orthogonal tRNA/aaRS pair, is that the levels of the endogenous prolyl-tRNAs must be altered to ensure high fidelity in protein biosynthesis, underscoring the degree to which the protein translational machinery has been evolutionary optimized.

■ METHODS

***E. coli* Cell Line Construction.** The wild-type *E. coli* K-12 W3110 cell line with GenBank accession no. AP009048.1 was purchased from ATCC, Virginia (catalog no. 27325). Through homologous recombination the T7 RNA polymerase gene cassette was inserted into the W3110 genome.[18] In this newly created cell line the *fhu*A gene was replaced with *dhfr* (Trimethoprim resistance marker) to create the W3110B55 cell line. In W3110B55 the *omp* gene was replaced with the chloramphenicol resistance marker (*cat*) to create the W3110B57 cell line. In a similar manner the *pro*S W375R::*cat* (point mutation for substitution of tryptophan with arginine) was generated to create a temperature-sensitive (*ts*) cell line W3110B60.[19] The W375R point mutation in *pro*S conferred a lethal host phenotype at temperatures ≥34 °C.

Construction of Expression Plasmids. The expression plasmids were constructed using pET-20b (+) and pET-24 (+) plasmids (EMD4Biosciences, California). The amber codon suppression cassette consisting of the modified *M. jannaschii* tyrosyl tRNA synthetase was inserted at the *Bam*HI site. The rhFGF-21*am* gene of interest (also known as AXP-000-028*am*) with the amber codon (TAG) within the ORF at glutamine position 108 was inserted downstream of the bacteriophage T7 promoter at the *Nde*I and *Kpn*I restriction sites in this plasmid. The wild-type *E. coli* K-12 W3110 *pro*S gene was inserted into the vector at the *Bgl*II restriction site. The temperature sensitive phenotype of the W3110B60 was complemented by the wild-type copy of the *pro*S gene in the expression plasmid. AXID1395 corresponds to the amber expression plasmid with the Ambrx suppression elements (*Mj* tyrosine $tRNA^{MjTyr}_{CUA}$ and *Mj* tyrosine tRNA synthetase, E9RS) and rhFGF-21-Q108*am* (AXP-000−028*am*). The AXID2089 plasmid has the rhFGF-21-Q108*am* gene but lacks the

amber suppression elements. The AXID2083 plasmid contains the rhFGF-21-Q108*am* gene and the *Mj* tyrosine $tRNA^{MjTyr}_{CUA}$ but lacks the *Mj* E9RS synthetase. The AXID2103 plasmid includes only the rhFGF-21-Q108*am* gene and the *Mj* E9RS synthetase. The AXID2292 plasmid encodes the rhFGF-21-Q108*am*, *Mj* tyrosine $tRNA^{MjTyr}_{CUA}$, *Mj* E9RS synthetase, and the *E. coli* K-12 W3110 proline synthetase (*pro*S) under the arabinose promoter.

Production of *p*AcF Containing rhFGF-21. Chemically competent DH5α-T1 cells (Life Technologies, California) were transformed with the expression plasmid containing rhFGF-21*am*, the system components for *p*AcF incorporation, and the *pro*S gene. Plasmid DNA from a single colony was extracted and used for transformation of chemically competent *E. coli* K-12 W3110B60 cells.[20] rhFGF-21-*p*AcF protein production was confirmed through SDS-PAGE and LC−MS analyses of a 500 mL shake flask experiment using defined medium (minimal medium supplemented with leucine, isoleucine, trace metals, and vitamins) with 50 μg/mL of kanamycin sulfate (Sigma, Missouri) at 37 °C. This result was further validated and confirmed through similar methods in high cell density fermentations consisting of batch and fed-batch phases.

High Cell Density Fermentations. The fermentation process for production of rhFGF-21-*p*AcF consists of two stages: (1) inoculum preparation and (2) fermentor production of rhFGF-21-*p*AcF. The inoculum is started from a single vial, thawed, diluted 1:1000 (v/v) into 50 mL of defined seed medium in a 250 mL baffled Erlenmeyer flask, and incubated at 37 °C and 250 rpm. The fermentor is batched with 4 L of chemically defined medium that utilizes glycerol as a carbon source. The seed culture is added to fermentor to an initial OD_{600nm} of 0.05. Dissolved oxygen is maintained at 30% air saturation using agitation from 480 to 1200 rpm and oxygen enrichment with a head pressure of 6 psig and air flow of 5 slpm. The temperature and pH are controlled at 37 °C and 7.0, respectively. When the culture reaches an OD_{600nm} of 35 ± 5, feeding commences at a rate of 0.25 mL/L/min. Consequently, L-Ala-*p*AcF dipeptide is added at 0.4 g/L. Fifteen minutes after the addition of dipeptide, the culture is induced with L-arabinose at a final concentration of 2 g/L. The culture is harvested at 6 h post induction.

Inclusion Body Preparation. Cell paste (150−250 μg) was resuspended with 1 mL of Bugbuster (EMD4Biosciences, California) and 2 μL of Lysonase (EMD4Biosciences, California) and incubated at 25 °C with shaking (Thermomixer) at 1400 rpm for 15 min. The sample was centrifuged at 16000*g* for 15 min, and the supernatant was removed by aspiration. The Bugbuster, Lysonase treatment and supernatant removal was repeated twice. The inclusion body was solubilized with 500 μL of 20 mM Tris, pH 8.0, 8 M Guanidine-HCl, and 0.1 M dithiothreitol (DTT). The sample was incubated at 45 °C with shaking at 1400 rpm for 30 min. Approximately 100 μL of the solubilized inclusion body was buffer exchanged into 1x PBS pH 7.4 using a Zeba spin column (Pierce, Illinois). The sample was diluted 1:1 with solubilization buffer and then applied to a 0.22 μm spin filter (Millipore, Massachusetts).

HPLC and Mass Spectrometry. Samples were applied onto a POROS R2 10 μm column, 2.1 mm × 30 mm (Applied Biosystems, California) using an Agilent 1200 HPLC in tandem with an Agilent 6510 Q-TOF. The column was equilibrated in 80% mobile phase A (0.05% TFA in 98% HPLC H_2O/2% acetonitrile) and 20% mobile phase B (0.04% TFA in 90% acetonitrile and 10% HPLC H_2O) with an increase to 65% mobile phase B over 15 min with a flow rate of 0.15 mL/min. The acquired spectra were deconvoluted using Agilent's Bioconfirm software (Agilent Technologies, California). The TIC was also integrated with Chemstation settings for quantitation.

Peptide Mapping. The rhFGF-21-*p*AcF protein was reduced in 6 M guanidine-HCl, 0.1 M Tris, pH 8,0, 0.05 M DTT for 1 h at 37 °C. The reduced samples were alkylated with 0.1 M iodoacetamide at RT in the dark for 40 min followed by quenching with 0.1 M DTT. Samples were buffer exchanged into 50 mM Tris, 5 mM calcium chloride, pH 7.5 followed by trypsin addition at 1:20 (trypsin:protein) and incubated for 18 h at 37 °C. Peptide mapping samples were loaded onto an Agilent SB-C18 2.1 mm × 150 mm column with 100% mobile phase A (0.05% TFA in 98% HPLC H_2O/2% acetonitrile) and eluted with a gradient of 0.5%/min with mobile phase B (0.04% TFA in acetonitrile) over 68 min. The flow rate was 0.2 mL/min, and column temperature was set to 50 °C.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tsotne@ambrx.com.

Present Addresses

[§]aTyr Pharma, Inc., San Diego, California 92121.
[⊥]Genentech, Inc., Oceanside, California 92056.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge A. Putnam-Hays and S. Dao for their contributions to the mass spectrometry analyses. This work was supported by National Institutes of Health grant R01GM062159 (PGS).

■ REFERENCES

(1) Ruan, B., A., I., Ambrogelly, A., Becker, H. D., Bunjun, S., Feng, L., Tumbula-Hansen, D., Ibba, M., Korencic, D., Kobayashi, H., Jacquin-Becker, C., Mejlhede, N., Min, B., Raczniak, G., Rinehart, J., Stathopoulos, C., Li, T., and Söll, D. (2001) Genomics and the evolution of aminoacyl-tRNA synthesis. *Acta Biochim. Pol. 48*, 313−321.

(2) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli. *Science 292*, 498−500.

(3) Zhang, Z., Smith, B. A. C., Wang, L., Brock, A., Cho, C., and Schultz, P. G. (2003) A new strategy for the site-specific modification of proteins in vivo. *Biochemistry 42*, 6735−6746.

(4) Goodman, H. M., Abelson, J., Landy, A., Brenner, S., and Smith, J. D. (1968) Amber suppression: a nucleotide change in the anticodon of a tyrosine transfer RNA. *Nature 217*, 1019−1024.

(5) Guo, J., Melançon, C. E., Lee, H. S., Groff, D., and Schultz, P. G. (2009) Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew. Chem., Int. Ed. 48*, 9148−9151.

(6) Liu, H., and Musier-Forsyth, K. (1994) Escherichia coli proline tRNA synthetase is sensitive to changes in the core region of tRNAPro. *Biochemistry 33*, 12708−12714.

(7) Wang, L., and Schultz, P. G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol. 8*, 883−890.

(8) Steer, B. A., and Schimmel, P. (1999) Major anticodon-binding region missing from an archaebacterial tRNA synthetase. *J. Biol. Chem. 274*, 35601−35606.

(9) Zhang, Y., Wang, L., Schultz, P. G., and Wilson, I. A. (2005) Crystal structures of apo wild-type M. jannaschii tyrosyl-tRNA synthetase (TyrRS) and an engineered TyrRS specific for O-methyl-L-tyrosine. *Protein Sci. 14*, 1340−1349.

(10) Tsunoda, M., Kusakabe, Y., Tanaka, N., Ohno, S., Nakamura, M., Senda, T., Moriguchi, T., Asai, N., Sekine, M., Yokogawa, T., Nishikawa, K., and Nakamura, K. T. (2007) Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms. *Nucleic Acids Res. 35*, 4289−4300.

(11) Hasegawa, T., and Yokogawa, T. (2000) Escherichia coli proline tRNA: Structure and recognition sites for prolyl-tRNA synthetase. *Nucleic Acids Symp. Ser. 44*, 7−8.

(12) Wang, L., and Schultz, P. G. (2005) Expanding the genetic code. *Angew. Chem., Int. Ed. 44*, 34−66.

(13) Kobayashi, T., Sakamoto, K., Takimura, T., Sekine, R., Vincent, K., Kamata, K., Nishimura, S., and Yokoyama, S. (2005) Structural basis of nonnatural amino acid recognition by an engineered aminoacyl-tRNA synthetase for genetic code expansion. *Proc. Natl. Acad. Sci. U.S.A. 102*, 1366−1371.

(14) Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N[epsi]-acetyllysine in recombinant proteins. *Nat. Chem. Biol. 4*, 232−234.
(15) Iraha, F., Oki, K., Kobayashi, T., Ohno, S., Yokogawa, T., Nishikawa, K., Yokoyama, S., and Sakamoto, K. (2010) Functional replacement of the endogenous tyrosyl-tRNA synthetase−tRNATyr pair by the archaeal tyrosine pair in Escherichia coli for genetic code expansion. *Nucleic Acids Res. 38*, 3682−3691.
(16) Young, T. S., and Schultz, P. G. (2010) Beyond the canonical 20 amino acids: Expanding the genetic lexicon. *J. Biol. Chem. 285*, 11039−11044.
(17) Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 99*, 11020−11024.
(18) Kang, Y., Durfee, T., Glasner, J. D., Qiu, Y., Frisch, D., Winterberg, K. M., and Blattner, F. R. (2004) Systematic mutagenesis of the Escherichia coli genome. *J. Bacteriol. 186*, 4921−4930.
(19) Bohman, K., and Isaksson, L. (1980) A temperature-sensitive mutant in prolinyl-tRNA ligase of Escherichia coli K-12. *Mol. Gen. Genet. 177*, 603−605.
(20) Chung, C. T., Niemela, S. L., and Miller, R. H. (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A. 86*, 2172−2175.
(21) An, S., Barany, G., and Musier-Forsyth, K. (2008) Evolution of acceptor stem tRNA recognition by class II prolyl-tRNA synthetase. *Nucleic Acids Res. 36*, 2514−2521.
(22) Liu, H., Peterson, R., Kessler, J., and Musier-Forsyth, K. (1995) Molecular recognition of tRNAPro by Escherichia coli proline tRNA synthetase in vitro. *Nucleic Acids Res. 23*, 165−169.
(23) McClain, W. H., Schneider, J., and Gabriel, K. (1994) Distinctive acceptor-end structure and other determinants of Escherichia coli tRNAPro identity. *Nucleic Acids Res. 22*, 522−529.
(24) Pavlov, M. Y., Watts, R. E., Tan, Z., Cornish, V. W., Ehrenberg, M., and Forster, A. C. (2009) Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc. Natl. Acad. Sci. U.S.A. 106*, 50−54.
(25) Vargas-Rodriguez, O., and Musier-Forsyth, K. (2013) Exclusive use of trans-editing domains prevents proline mistranslation. *J. Biol. Chem. 288*, 14391−14399.
(26) Louie, A., Ribeiro, N. S., Reid, B. R., and Jurnak, F. (1984) Relative affinities of all Escherichia coli aminoacyl-tRNAs for elongation factor Tu-GTP. *J. Biol. Chem. 259*, 5010−5016.
(27) Beuning, P. J., and Musier-Forsyth, K. (2000) Hydrolytic editing by a class II aminoacyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A. 97*, 8916−8920.