Published in final edited form as:

Angew Chem Int Ed Engl. 2009 ; 48(48): 9148–9151. doi:10.1002/anie.200904035.

Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Unnatural Amino Acid-Containing Proteins

Jiantao Guo⁺, Charles E. Melançon III⁺, Hyun Soo Lee, Dan Groff, and Peter G. Schultz^{*}

Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)

Abstract

Regions of the *M. jannaschii* tyrosyl tRNA_{CUA} thought to interact with elongation factor Tu were randomized, and the resulting tRNA libraries were subjected to *in vitro* evolution. The tRNAs identified resulted in significantly improved unnatural amino acid-containing protein yields. In some cases, the degree of improvement varied in an amino acid-dependent manner.

Keywords

tRNA; EF-Tu; evolution; amber suppression; unnatural amino acids

tRNAs have evolved to act as highly efficient amino acid carriers and activators during each stage of protein synthesis. Each tRNA must be selectively charged by its cognate aminoacyl-tRNA synthetase (aaRS); the resulting aminoacyl-tRNA must be efficiently bound by

^{*}Fax: (+1) 858-784-9440, schultz@scripps.edu.

⁺These authors contributed equally to this work.

elongation factor Tu (EF-Tu) for transport to the ribosome; after binding to the ribosomal A site, the aminoacyl-tRNA must function efficiently in translation as a substrate for the peptidyl transferase; and finally the tRNA bearing the growing peptide chain must be translocated to the P site, undergo another acyl transfer reaction, and be released from the ribosome. We and others have used orthogonal tRNA/aaRS pairs for the site-specific incorporation of nearly 50 unnatural amino acids in *E. coli*, *S. cerevisiae*, and mammalian cells in response to unique nonsense and frameshift codons.¹ An engineered *M. jannaschii* amber suppressor tyrosyl-tRNA/tRNA-synthetase (MjtRNA^{Tyr}_{CUA}/MjYRS) pair has been the most extensively used system for the evolution of aaRS variants that incorporate unnatural amino acids in *E. coli*. Although unnatural amino acids are typically incorporated into proteins with good efficiency and excellent fidelity, further system optimization resulting in increased protein yields is highly desirable.

Because MjtRNA^{Tyr}_{CUA} is derived from an archaeal tRNA and therefore significantly differs in sequence from *E. coli* tRNAs, it may not function optimally with the *E. coli* translational machinery. Furthermore, *in vitro* binding studies have shown that while correctly acylated tRNAs bind EF-Tu with near uniform affinity, tRNAs bearing non-cognate amino acids show a broad range of affinities for EF-Tu², indicating that the tRNA body and the esterified amino acid make compensatory contributions to EF-Tu binding. A number of genetic, biochemical, and structural studies have implicated specific residues within the tRNA acceptor stem and T stem as being important for EF-Tu binding.³ tRNA misacylation has also recently been shown to perturb binding to the ribosomal A site.⁴ Therefore it is likely that tRNAs acylated with noncognate unnatural amino acids adversely impact the efficiency of protein synthesis due to non-optimal interactions with EF-Tu and/or the ribosome.

We have used *in vitro* evolution to optimize the sequence of MjtRNA^{Tyr}_{CUA} with MjYRS and a panel of six evolved unnatural amino acid incorporating MjYRS variants, and have identified several unique tRNA sequences that confer significantly improved mutant protein expression in *E. coli*. While most tRNAs identified show some improvement in protein yields for all aaRSs tested, the degree of improvement for each tRNA often varies depending on the identity of the esterified amino acid.

Examination of the X-ray crystal structure of the *T. aquaticus* EF-Tu•*E. coli* cysteinyl-tRNA^{Cys}•GDPNP ternary complex⁶ reveals that residues 1–3, 50–54, 63–67, and 73–76, which reside in the acceptor stem and T stem of the tRNA, are in close proximity to EF-Tu (Figure 1). Interactions with EF-Tu almost exclusively involve the tRNA backbone, although contacts also exist between EF-Tu and the base of residue 63. Based on the notion

that mutation of the corresponding positions in MjtRNA^{Tyr}_{CUA} might modulate its binding affinity for EF-Tu, we created a tRNA library (Library I, theoretical diversity 1.05×10^6 , >99% coverage) in which the five base pairs of the T-stem (49–53, 61–65) were randomized. The tRNA library was subjected to a negative selection to remove tRNAs that are substrates for endogenous aaRSs using an amber mutant of the barnase gene as previously described⁵; positive selection of surviving clones to identify functional tRNA sequences was carried out in the presence of aaRS specific for tyrosine (1) (MjYRS), L-3-(2)-naphthylalanine (2) (NapRS7a), *p*-azidophenylalanine (3) (pAzPheRS7b), *p*-iodophenylalanine (4) (pIPheRS7c), *p*-acetylphenylalanine (5) (pAcPheRS7d), and *p*-benzoylphenylalanine (6) (pBpaRS7e) (Scheme 1), chloramphenicol, and an amber mutant of chloramphenicol acetyltransferase (CAT) as previously described. The surviving tRNAs were then screened based on cell fluorescence using an amber suppressible GFPuv construct to identify the most active tRNAs (see supplementary material for experimental details).

Nine unique tRNA sequences were identified with significantly improved protein yields compared to MjtRNA^{Tyr}_{CUA} (supplementary Table S3). The G52-C62 and G53-C61 pairs, which are highly conserved among *E. coli* tRNAs, were invariant in all hits. The consensus sequence for the most active hits is W⁴⁹NRGG⁵³-C⁶¹CYNW⁶⁵ (where W is A or T; N is A, T, G, or C; R is A or G; and Y is C or T).

Next, we constructed a second tRNA library (Library II, theoretical diversity 1.7×10^7 , >99% coverage) in which the consensus sequences at positions 49–53 and 61–65 from the best Lib I hits were retained and the remaining positions in contact with EF-Tu (2, 3, 6, 7, 66, 67, 70, and 71) were randomized. Positions 1 and 72–76 were not mutated because of their high degree of conservation and their importance in recognition by MjYRS. Library II was subjected to the same selection/screening procedure as was used for Library I. In addition to the 6 previously used aaRSs, the aaRS specific for *p*-hydroxy-L-phenyllactic acid (**7**) (PlaRS)1c, was also used for positive selection. Fourteen unique tRNA sequences were identified that lead to significant improvements in protein yields compared to MjtRNA^{Tyr}_{CUA} (supplementary Table S4). Of these, ten unique sequences were found to rank in the top 3 in terms of fluorescence intensity for at least one aaRS, and resulted in mutant protein yields 200–2200% of those obtained with MjtRNA^{Tyr}_{CUA} (Table 1). Yield improvements were far greater with **6** (18 to 21-fold better than MjtRNA^{Tyr}_{CUA}) than with the other amino acids (100–350% better than MjtRNA^{Tyr}_{CUA}). Several of the best hits contained base deletions in the D-loop which arose as artifacts of library construction.

These 10 tRNAs were isolated, crossed against 13 aaRSs (the 7 used in this study and 6 additional aaRSs previously evolved for incorporation of bipyridylalanine (**8**) (Bpy)8a, hydroxyquinolinyllalanine (**9**) (HQ)8b, sulfotyrosine (**10**) (SfY)8c, *p*-azobenzylphenylalanine (**11**) (pABPhe)8d, *o*-nitrobenzyltyrosine (**12**) (ONBY)8e, and 7-hydroxycoumarinylethylglycine (**13**) (Cou)8f) (Scheme 1) and assayed for GFP production in order to identify the tRNA sequence that confers the largest improvement in expressed protein yield for each amino acid. Improvements in yield of 175%–320% were seen with the best tRNAs charged with compounds **1–5**, **8**, **10**, and **11**. However, significantly larger improvements were observed with **9** (405%), **7** (420%), **13** (790%), **12** (2070%), and **6** (2520%) (Figure 2). Interestingly, the compounds with the most dramatic improvements in yield have bulky side chains. Notable exceptions include **7**, which differs from tyrosine only by the replacement of the α -amino group with an α -hydroxyl group; and **11**, which gave only a 255% improvement, yet has a bulky side chain. This data suggests that the structure of the unnatural amino acid can affect the ability of the corresponding aminoacyl-tRNA to function productively in translation. This may be a result of interactions between the amino acid and the EF-Tu binding pocket. Consistent with this hypothesis, it was recently shown that mutation of the EF-Tu amino acid binding pocket resulted in more efficient binding of tRNAs bearing bulky unnatural amino acids.⁹ The 195% yield increase observed with the natural amino acid **1** represents the baseline improvement of MjtRNA^{Tyr}_{CUA} due to optimization of this archaeal tRNA for the *E. coli* translational machinery. Five tRNAs (Tyr1, pAzPhe1, Nap1, Nap3, and pIPhe1) consistently give the largest increases in yield. Of these, Nap1 is the best “general” hit (the tRNA that gives the best overall yield improvements with all amino acids). Yields of purified GFP expressed using Nap1 and a GFPN149TAG construct with **1–7** range from 2.7–16.3 mg/L (380–1175% of yields obtained with MjtRNA^{Tyr}_{CUA}) in rich media (supplementary Figure S9). The fidelity of unnatural amino acid incorporation with Nap1 and **1–7** was determined by LC-ESI-MS of purified GFP proteins produced using

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

evolved tRNAs and is comparable to the high fidelity obtained with MjtRNA^{Tyr}_{CUA} (supplementary Table S2, Figures S11–S17).

Growth rates were determined for *E. coli* TOP10 cells expressing the evolved tRNAs (with GFP N149TAG reporter, aaRS, IPTG, and 1–7) (supplementary Figure S10, Table S1). In all cell lines except those expressing PlaRS, the exponential growth rates for cells expressing the evolved tRNAs were significantly faster ($\mu = 0.47 \pm 0.05 \text{ h}^{-1}$) than for those expressing MjtRNA^{Tyr}_{CUA} ($\mu = 0.30 \pm 0.05 \text{ h}^{-1}$), indicating that the evolved tRNAs are in general less toxic than MjtRNA^{Tyr}_{CUA}. Because all tRNAs were subjected to negative selection to remove those that are substrates for endogenous aaRSs, it is unlikely that altered toxicity results from differences in orthogonality of the tRNAs. The decreased toxicity of the evolved tRNAs could result from decreased readthrough of natural stop codons, or altered interactions of the tRNA with EF-Tu and/or the ribosome which result in enhanced translation of heterologously expressed amber mutant genes. Several examples of suppressor tRNA-dependent cellular toxicity, in which toxicity was suggested to be the result of readthrough of endogenous stop codons, have been reported.¹⁰ The fact that cells expressing all evolved tRNAs charged with 1–6 have nearly the same growth rates yet display a broad range of protein yields, suggests that other factors in addition to ameliorated toxicity contribute to the improvements in protein yield.

A recent study on the EF-Tu binding affinities of a collection of acceptor stem and T stem mutations allows prediction of EF-Tu interaction strength based on sequence.^{3a} This model predicts that our evolved tRNAs bind EF-Tu 0.3–0.9 kcal/mol more tightly than MjtRNA^{Tyr}_{CUA}. The 51–63 base pair, which was changed from a U-G pair in MjtRNA^{Tyr}_{CUA} to a G-C pair in all evolved tRNAs, is predicted to strengthen the EF-Tu binding affinity by 1.0 kcal/mol. The gain in function observed for the evolved tRNAs may be due in large part to this mutation, although convergence to non-wild-type sequences at other positions suggests that these mutations may act to fine-tune tRNA affinity for EF-Tu. It is possible that the predicted increase in binding energy between EF-Tu and the evolved tRNAs compensates for weaker binding of unnatural amino acids in the EF-Tu binding pocket, and that this compensation is the reason for the large yield improvements observed with some amino acids. Detailed analysis of the interaction of the acylated tRNAs with EF-Tu will provide additional insights into the mechanistic basis for the improved activities of these tRNAs.

In conclusion, we have utilized an *in vitro* evolution approach to identify MjtRNA^{Tyr}_{CUA} variants with significantly enhanced activity for the incorporation of unnatural amino acids into proteins, especially the photocrosslinking amino acid 6, which is widely used to map biomolecular interactions. We also found that the degree of yield improvement is in some cases unnatural amino acid-dependent. These tRNAs will facilitate the creation of an optimized and standardized system for the genetic incorporation of unnatural amino acids into proteins in *E. coli*.¹¹

Experimental Section

Materials and methods can be found in the supporting information

Acknowledgments

This work is supported by NIH grants R01 GM062159 (P.G.S.) and F32 GM080067 (C.E.M.), and the Skaggs Institute for Chemical Biology.

References

1. Select references include: a) Xie J, Schultz PG. *Nat. Rev. Mol. Cell. Biol.* 2006; 7:775. [PubMed: 16926858] b) Lemke EA, Summerer D, Geierstanger BH, Brittain SM, Schultz PG. *Nat. Chem. Biol.* 2007; 3:769. [PubMed: 17965709] c) Guo J, Wang J, Anderson JC, Schultz PG. *Ang. Chem.* 2008; 47:722. d) Brustad E, Bushey ML, Lee JW, Groff D, Liu W, Schultz PG. *Ang. Chem.* 2008; 47:8220. e) Neumann H, Hazen JL, Weinstein J, Mehl RA, Chin JW. *J. Am. Chem. Soc.* 2008; 130:4028. [PubMed: 18321101] f) Neumann H, Peak-Chew SY, Chin JW. *Nat. Chem. Biol.* 2008; 4:232. [PubMed: 18278036] g) Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. *Chem. Biol.* 2008; 15:1187. [PubMed: 19022179] h) Peters FB, Brock A, Wang J, Schultz PG. *Chem. Biol.* 2009; 16:148. [PubMed: 19246005] i) Li W-T, Mahapatra A, Longstaff DG, Bechtel J, Zhao G, Kang PT, Chan MK, Krzycki JA. *J. Mol. Biol.* 2009; 385:1156. [PubMed: 19063902] j) Kobayashi T, Yanagisawa T, Sakamoto K, Yokoyama S. *J. Mol. Biol.* 2009; 385:1352. [PubMed: 19100747] .
2. a) LaRiviere FJ, Wolfson AD, Uhlenbeck OC. *Science.* 2001; 294:165. [PubMed: 11588263] b) Asahara H, Uhlenbeck OC. *Proc. Nat. Acad. Sci.* 2002; 99:3499. [PubMed: 11891293] c) Dale T, Sanderson LE, Uhlenbeck OC. *Biochem.* 2004; 43:6159. [PubMed: 15147200]
3. a) Schrader JM, Chapman SJ, Uhlenbeck OC. *J. Mol. Biol.* 2009; 386:1255. [PubMed: 19452597] b) Eargle J, Black AA, Sethi A, Trabuco LG, Luthey-Schulten Z. *J. Mol. Biol.* 2008; 377:1382. [PubMed: 18336835] c) Forster C, Chakraborty K, Sprinzl M. *Nucl. Acids Res.* 1993; 21:5679–5683. [PubMed: 8284215] d) Rudinger J, Hillenbrandt R, Sprinzl M, Giege R. *EMBO J.* 1996; 15:650–657. [PubMed: 8599948]
4. Dale T, Fahlman RP, Olejniczak M, Uhlenbeck OC. *Nucl. Acids Res.* 2009; 37:1202. [PubMed: 19129224]
5. Wang L, Schultz PG. *Chem. Biol.* 2001; 8:883. [PubMed: 11564556]
6. Nissen P, Thirup S, Kjeldgaard M, Nyborg J. *Structure.* 1999; 7:143. [PubMed: 10368282]
7. a) Wang L, Brock A, Schultz PG. *J. Am. Chem. Soc.* 2002; 124:1836. [PubMed: 11866580] b) Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. *J. Am. Chem. Soc.* 2002; 124:9026. [PubMed: 12148987] c) Xie J, Wang L, Wu N, Brock A, Spraggon G, Schultz PG. *Nat. Biotech.* 2004; 22:1297. d) Wang L, Zhang Z, Brock A, Schultz PG. *Proc. Nat. Acad. Sci.* 2003; 100:56. [PubMed: 12518054] e) Chin JW, Martin AB, King DS, Wang L, Schultz PG. *Proc. Nat. Acad. Sci.* 2002; 99:11020. [PubMed: 12154230]
8. a) Xie J, Liu W, Schultz PG. *Ang. Chem.* 2007; 46:9239. b) Lee HS, Spraggon G, Schultz PG, Wang F. *J. Am. Chem. Soc.* 2009; 131:2481. [PubMed: 19193005] c) Liu C, Schultz PG. *Nat. Biotechnol.* 2006; 24:1436. [PubMed: 17072302] d) Bose M, Groff D, Xie J, Brustad E, Schultz PG. *J. Am. Chem. Soc.* 2006; 128:388. [PubMed: 16402807] e) Deiters A, Groff D, Ryu Y, Xie J, Schultz PG. *Ang. Chem.* 2006; 45:27281. f) Wang J, Xie J, Schultz PG. *J. Am. Chem. Soc.* 2006; 128:8738. [PubMed: 16819861]
9. Doi Y, Ohtsuki T, Shimizu Y, Ueda T, Sisido M. *J. Am. Chem. Soc.* 2007; 129:14458–14462. [PubMed: 17958427]
10. a) Doerig RE, Suter B, Gray M, Kubli E. *EMBO J.* 1988; 7:2579–2584. [PubMed: 3142765] b) Leuker CE, Ernst JF. *Mol. Gen. Genet.* 1994; 245:212–217. [PubMed: 7816029] c) Carneiro VTC, Pelletier G, Small I. *Plant Mol. Biol.* 1993; 22:681–690. [PubMed: 8343603]
11. Young TS, Ahmad I, Schultz PG. (submitted).

Figure 1.

The EF-Tu/tRNA interface. a) Putative *M. jannaschii* tRNA^{Tyr}_{CUA} nucleotides that interact with EF-Tu are shown in red. Previously mutated positions are shown in green. b) A diagram derived from the *T. aquaticus* EF-Tu/*E. coli* cysteinyl-tRNA^{Cys} X-ray structure⁶ showing EF-Tu/tRNA interactions. The tRNA residues that interact with EF-Tu are numbered and shown in red. Interacting residues of EF-Tu are shown in green. Note that all interactions between tRNA and EF-Tu involve the tRNA backbone except in the case of residue 63.

Figure 2.

Activity of the 10 best evolved tRNAs with each aaRS shown as fold improvement over MjtRNA^{Tyr}_{CUA}. tRNAs and aaRSs are arranged from right to left in order of increasing yield improvement. Inset shows the region of the main graph between -1- and 5-fold improvement.

Scheme 1.
Structures of amino acids used in this study.

Table 1

Evolved tRNA variants with improved amber suppression activity. The sequence of each tRNA at randomized positions and the expressed protein yield compared to MjtRNA^{Tyr}_{CUA} are listed. The three most active tRNAs obtained from each selection are shown. S = G or C; W = A or T; B = G, T, or C; V = A, G, or C.

aaRS	tRNA	position					% of
		2, 3	6, 7	16-18	49-51	63-67	70, 71
	MjtRNA ^{Tyr} _{CUA}	CG	GG	CAG	GCT	GGCCC	MjtRNA ^{Tyr} _{CUA} activity
Nap	Nap1	CG	CG	CAG	ATG	CATCG	CG
	Nap2	CC	CC	-A-	TCG	CGAGG	GG
	Nap3	CC	CT	CAG	ACG	CGTAG	GG
pAzPhe	pAzPhe1	CC	CT	-AG	ACG	CGTAG	GG
	pAzPhe2	same as Nap1					281 ± 17
	pAzPhe3	same as Nap2					219 ± 4
Tyr	Tyr1	same as Nap2					281 ± 11
	Tyr2	CG	CT	CAG	ATG	CATAG	CG
	Tyr3	same as Nap3					221 ± 8
pIPhe	pIPhe1	CG	CG	CAG	TCG	CGACG	CG
	pIPhe2	same as Nap1					345 ± 17
	pIPhe3	same as pAzPhe1					332 ± 30
pAcPhe	pAcPhe1	same as Nap2					290 ± 16
	pAcPhe2	same as Nap1					274 ± 9
	pAcPhe3	same as Nap3					254 ± 5
pBpa	pBpa1	same as pAzPhe1					211 ± 12
	pBpa2	same as Nap1					2266 ± 12
	pBpa3	CG	CT	CAG	ACG	CGTAG	CG
Pla	Pla1	CC	CT	CAG	ATG	CATAG	GG
	Pla2	CC	CA	CAG	TGG	CCATG	GG
	Pla3	CC	CA	CAG	ATG	CATTG	GG

aaRS	tRNA	position						% of
		2, 3	6, 7	16-18	49-51	63-67	70, 71	
								Mj:tRNA ^{Tyr} _{CUA}
	Mj:tRNA ^{Tyr} _{CUA}	CG	GG	CAG	GCT	GGCCC	CG	activity
	consensus	CS	CN	CAG	WBG	CVWNG	SG	