Published in final edited form as:

Chembiochem. 2014 August 18; 15(12): 1744–1749. doi:10.1002/cbic.201402235.

Response and Adaptation of *Escherichia coli* to Suppression of the Amber Stop Codon

Qian Wang^{a,b}, Tingting Sun^{a,c}, Jianfeng Xu^a, Zhouxin Shen^d, Steven P. Briggs^d [Prof.], Demin Zhou^c, and Lei Wang^a

Lei Wang: lwang@salk.edu

^aJack H. Skirball Center for Chemical Biology & Proteomics, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037 (USA)

^bCollege of Veterinary Medicine, China Agricultural University, Beijing, 100193 (China)

^cState Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University Beijing, 100191 (China)

^dSection of Cell and Development Biology, University of California at San Diego, La Jolla, CA 92093 (USA)

Abstract

Some extant organisms reassign the amber stop codon to a sense codon through evolution, and suppression of the amber codon with engineered tRNAs has been exploited to expand the genetic code for incorporating non-canonical amino acids (ncAAs) in live systems. However, it is unclear how the host cell would respond and adapt to the amber suppression. Here we suppressed the amber codon in *Escherichia coli* with an orthogonal tRNA/synthetase pair and cultured the cells under such a pressure for about 500 generations. We discovered that *E. coli* quickly counteracted with transposon insertion to inactivate the orthogonal synthetase. Persistent amber suppression evading transposon inactivation led to global proteomic changes with a marked up-regulation of an uncharacterized protein YdiI, for which we identified an unexpected function of expelling plasmids. These results should be valuable for understanding codon reassignment in code evolution and for improving the efficiency of ncAA incorporation.

Keywords

expansion of the genetic code; unnatural amino acid; evolution of the genetic code; amber suppression; codon reassignment

The canonical genetic code contains 61 sense codons specifying 20 amino acids and 3 nonsense codons (UAA, UAG and UGA) specifying the stop signal for protein translation. Sense codons are decoded by tRNAs, while the nonsense codons are recognized by proteins called class I release factors (RFs), which promote peptide release from the tRNA in the ribosome. The genetic code was once thought to be a “frozen accident” because it was universally preserved in all then known organisms and any changes would affect all proteins

simultaneously and be deleterious to the host.[¹] Small deviations from the canonical code were later discovered in the mitochondrial and nuclear codes of an increasing number of organisms.[²] These include the reassignment of sense codons from one amino acid to another and, more frequently, the reassignment of a nonsense codon to an amino acid. For instance, the eukaryotic release factor 1 (eRF1) of *Tetrahymena* restricts its recognition to UGA, and UAA/UAG are reassigned to Gln; the eRF1 of *Euplotes* recognizes UAA/UAG only as stop codons, and UGA is used to encode Cys.[³] These deviations in codon assignment suggest that the code is flexible and might be still evolving in extant lineages. However, extant organisms harboring such changes are at the end-point or status quo of the code evolution. Knowledge of initial organism response, concurrent cellular adaptation, and eventual fixation of codon reassignments are lacking.[⁴]

Suppression of nonsense codons have been exploited for the incorporation of both natural and non-canonical amino acids (ncAAs) into proteins.[⁵] Occurring only once per gene, the relative scarcity of nonsense codons may mitigate any damage caused by codon reassignment. Natural suppressor tRNAs decoding stop codons as common amino acids have been identified in *E. coli* and other organisms.[⁶] In addition, orthogonal tRNA/synthetase pairs have been generated to incorporate ncAAs into proteins in response to a stop codon directly in live cells.[^{5c}, ⁷] Such an orthogonal tRNA/synthetase pair does not crosstalk with endogenous tRNA/synthetase pairs of the host cell, and functionally couples with the cell's protein translational machinery. The anticodon of the orthogonal tRNA is changed to pair with a stop codon for specific recognition, and the orthogonal synthetase is engineered to use a desired ncAA as its substrate. This approach has enabled ncAAs with a variety of functional groups to be genetically incorporated into proteins in bacteria, eukaryotic cells and even mammals.[^{5c}, ⁸] Genetically encoding ncAAs provides the attractive potential of investigating proteins and biological processes *in vivo*. [⁹] However, as the stop codon chosen to encode the ncAA is also used by multiple endogenous genes for translation termination, it is important to understand how the host cell would be affected when this stop codon is suppressed by the orthogonal tRNA genome-wide.

To begin to address these questions, we introduced an orthogonal tRNA/synthetase pair into *E. coli* to suppress the amber stop codon UAG, and maintained the suppression in *E. coli* for over 500 generations. We discovered that *E. coli* initially uses transposon to counteract the pressure, and that long term adaptation of *E. coli* to amber suppression involves an uncharacterized protein YdiI, which performs an unexpected function of expelling plasmids. These results should facilitate the synthetic reassignment of the amber codon in *E. coli* and the improvement of ncAA incorporation efficiency via amber suppression.

To introduce strong amber suppression in the *E. coli* genome, an amber suppressor tyrosyl tRNA derived from archaeobacteria *Methanococcus jannaschii* (J17–tRNA^{Tyr}_{CUA}) and the cognate tyrosyl-tRNA synthetase (*Mj*TyrRS) were expressed in *E. coli*. The

J17–tRNA^{Tyr}_{CUA}/*Mj*TyrRS pair is orthogonal to endogenous *E. coli* tRNA/synthetase pairs and incorporates tyrosine in response to the amber stop codon, TAG, in high efficiency[¹⁰]. To monitor the amber suppression, the gene for the enhanced green fluorescent protein (EGFP) containing a TAG codon at a permissive site (Tyr182) was coexpressed as the

fluorescence reporter in one single plasmid pBK-tYGT (Figure 1A). *E. coli* DH10 β cells transformed with the plasmid were fluorescent when excited at 470 nm. Cells harboring pBK-tYGT (doubling time = 43 min) grew significantly slower than cells harboring pBK-GT (doubling time = 27 min), a control plasmid without the tRNA and synthetase genes (Figure 1B), suggesting that strong amber suppression negatively impacts cell propagation.

To investigate how *E. coli* would react upon strong amber suppression, cells containing pBK-tYGT were continually passaged either in liquid media or on plates using kanamycin to maintain the plasmid. For liquid culture, cells were diluted 2.5×10^8 folds for subculture, and they could reach OD₆₀₀ 2 in 24 hours. We found that non-fluorescent cells emerged and the percentage of green fluorescent cells decreased with growth time (Figure 1C). After 3 passages with one passage per day, the liquid culture quickly lost its fluorescence. Plasmid extracted from the non-fluorescent cells shifted on agarose electrophoresis in comparison to those from the fluorescent cells (Figure 2). Separate PCR amplification of the tRNA, GFP-TAG, and *MjTyrRS* genes showed that the plasmid shift resulted from size change in the synthetase gene cassette. DNA sequencing of the isolated plasmids verified there were insertions in the synthetase gene, and the insertions were revealed all to be *E. coli* endogenous transposons IS1 or IS10. The 5 representative insertion sites and direction are listed in Table 1. The high frequency transposition occurred in the pBK-tYGT plasmid only, where as the same plasmid backbone harboring tRNA only or synthetase only was stable in the same growth condition, indicating that the transposition is in response to the amber suppression.

Transposition as a mechanism for stress adaptation has been reported in *E. coli* cells under UV light or continuous stationary phase growth.^[11] It is believed that insertion sequences function as mutation generator to acquire growth fitness in evolution. They can dynamically regulate gene expression by abolishing, duplicating or translocating target genes. We sequenced 9 insertions in total, and they were all located within the synthetase gene cassette, abolishing the gene integrity.

To determine how *E. coli* cells would respond to long term strong amber suppression, more controllable passaging was performed on plates. In each round, 10 fluorescent colonies were randomly picked and suspended in media, and $1/10^5$ of the suspension was spread on the subsequent plate containing kanamycin. The continuous use of fluorescent cells ensured that the amber suppression was not crippled by transposons. The passaging was carried out for 22 rounds, which added up to approximate 450~500 generations. The initial and final round of cells were collected and subjected to quantitative whole-proteome profiling using mass spectrometry. Proteins whose amount differed more than 2 folds are listed in Table 2.

Among proteins with large changes, the elongation factor EF-Tu is the only protein directly related with translation, which was decreased to 0.39. Since EF-Tu facilitates acylated tRNA entering into the ribosome A site, reduced EF-Tu expression would slow down translation and cell growth. A decrease in EF-Tu would also decrease the delivery of the amber suppressor tRNA to the TAG site, whereas the competitor release factor 1 is not affected, leading to more efficient termination than suppression at the TAG sites, which would mitigate the amber suppression pressure. The genes of four identified proteins end with a

TAG stop codon, and their protein products were all decreased. A possible mechanism for the decrease of these proteins may relate with tmRNA mediated protein degradation of no-stop mRNA products.^[12] The family of outer membrane proteins have 9 members with the most decreased amount, suggesting an impaired membrane translocation system.^[13] Cold shock proteins (Csp) are another family identified but showing up-regulated expression. Csp families boost after cold shock, and usually are considered as RNA or protein chaperons to maintain necessary transcription and translation efficiency.^[14] The most up-regulated protein is a hypothetic protein YdiI, which showed a remarkable 16 fold increase (Table 2). We thus investigated its role in adapting *E. coli* cells to amber suppression.

YdiI contains a Hot Dog folding domain and was predicted as a putative thioesterase. Enzyme genomics confirmed its esterase activity and structural genomics showed high similarity to other CoA conjugate thioesterase.^[15] We cloned the *ydiI* gene into the pLEI plasmid and could overexpress it with IPTG induction. When pLEI-ydiI was cotransformed with pBK-tYGT into *E. coli* cells, we found that the overexpression of *ydiI* changed colony morphology (Figure 3A). Colonies expressing *ydiI* were no longer uniformly fluorescent; less fluorescent “white” spots appeared inside the colonies. When DH10 β cells freshly transformed with pLEI-ydiI and pBK-tYGT grew for 48 hours in liquid media in the presence of 0.1 mM IPTG, their green fluorescence intensity was dramatically reduced compared to that of cells transformed with pBK-tYGT and an empty vector pLEI. Decrease of EGFP fluorescence could result from less efficient suppression of the TAG codon introduced at the Tyr182 site (which would increase truncated EGFP), or from overall lower expression of the plasmid pBK-tYGT. To compare the observed reduction and distinguish the two possibilities, we purified the full-length and truncated EGFP proteins by Ni²⁺ affinity chromatography and separated them by SDS-PAGE (Figure 3B). The YdiI-overexpressing cells produced 20 fold less full-length EGFP, but didn't accumulate much more truncated EGFP. These results indicated that EGFP expression was reduced in the presence of YdiI, but not because of increased premature termination at the introduced TAG codon.

To check if the EGFP gene or the orthogonal tRNA/TyrRS were deactivated, we then sequenced the two plasmids isolated from the cells. However, no mutation, insertion, or deletion in the tRNA, synthetase, or EGFP gene expression cassettes was found. Surprisingly, plasmids extracted from YdiI-overexpressing cells showed significant decrease in the amount of both the pBK-tYGT and the pLEI-ydiI plasmid, suggesting that the decrease of EGFP expression was mainly due to a reduced plasmid copy number (Figure 3C).

To verify if YdiI overexpression expels plasmids from *E. coli*, we grew *E. coli* cells with and without YdiI expression and quantified the number of colonies capable of surviving in the antibiotics resistant by each plasmid (Figure 4A). The percentage of antibiotic resistant cells relative to the total number of cells growing in the antibiotic-free plate was determined as a measure for plasmid retainment. In the absence of YdiI expression, no significant change in plasmid retaining was observed for either plasmid pBK-tYGT containing a ColEI origin of replication (ori) or plasmid pLEI containing a p15A ori over time (Figure 4B). In contrast, when YdiI expression was induced by IPTG, it reduced kanamycin resistant cells

(conferred by plasmid pBK-tYGT) from 46% to 20% in 12 hr. The decrease of the chloramphenicol resistant cells (conferred by plasmid pLEI-YdiI) was from 40% to 6.5% in 6 hr and further to 0.5% in 12 hr. When both antibiotics were present, only 6.5% cells could survive after 6 hr and 0.1% at 12 hr. These results indicate that YdiI has the unexpected ability to expel plasmids from *E. coli* efficiently, the mechanism of which warrants further investigation.

In summary, upon strong amber suppression in the genome, *E. coli* DH10 β cells immediately try to inactivate the responsible gene using transposon insertion. If transposons miss the target and amber suppression persists, the expression of YdiI is up-regulated, which reduces the copy number of plasmids to attenuate the suppression effect. Amber suppression is currently the prevalent method to genetically incorporate ncAAs into proteins in live cells. The incorporation efficiency of ncAAs has not matched that of natural amino acids.[⁴, ¹⁶] Results in this report suggest that responses from host cells counteract the suppression effect and may set a ceiling. This limit may be overcome by using a transposon-free strain[¹⁷] as well as manipulating the proteins identified here. In addition, codon reassignment has been hypothesized to occur during the evolution of the genetic code.[¹⁸] A detailed understanding of how cells respond to amber suppression will help to evaluate whether and how a stop codon can be reassigned to a sense codon, which will provide direct experimental evidence for the challenging evolutionary question.

Experimental Section

Plasmid construction

EGFP with a TAG codon at site Tyr182 was first cloned into plasmid pLEIG[¹⁹] to replace the α GFP gene. The (J17-tRNA^{Tyr}_{CUA}) and EGFP-TAG expression cassettes from the resultant plasmid were then subcloned into plasmid pBK-JYRS[^{5c}] to afford pBK-tYGT. Plasmid pBK-GT was derived from pBK-tYGT by deleting the genes for the tRNA and the *Mj*TyrRS. pLEI was made from pLEIG by deleting the α GFP gene and the tRNA expression cassette. The *ydiI* gene was amplified from *E. coli* genomic DNA using primers *SpeI_ydiI_5*: ccA CTA GTa tga tat gga aac gaa aaa tcc ccc t and *BglII_ydiI_3*: ccA GAT CTc aaa atg gcg gtc gtc aat cg. The PCR product was cloned into pLEI using *Spe* I and *Bgl* II sites to afford pLEI-ydiI, which expresses YdiI upon IPTG induction.

Liquid culture and fluorescent cell counting

At the end of each liquid culture after 24 hr, cells were diluted 10⁶ folds and 100 μ L of diluted culture was spread on the plate with 50 μ g/mL kanamycin. Fluorescence of colonies were checked by a macro fluorescence imaging system with excitation of 470 nm (bandwidth 40 nm). The percentage of fluorescent colonies was calculated and plotted in Figure 1C.

Mass spectrometric profiling of proteomic change

P0 and P22 *E. coli* cells were harvested and washed 3 times by PBS buffer. The cells were lysed in 2% Rapi Gest (Waters Corp.) and Hepes buffer (10 mM Hepes, 150 mM NaCl, pH 7.2) by a Branson Sonifier 450 Ultrasonic Homogenizer. The proteins were reduced and

alkylated using 2 mM Tris(2-carboxyethyl)phosphine at 95°C for 5 min and 5 mM iodoacetamide at 37°C in dark for 30 min, respectively. The proteins were then digested with 1:50 trypsin overnight. 50 µg of digested peptides of P0 and P22 sample was labeled by iTRAQ 115 and 117 reagents, respectively. The labeled peptides were mixed together subject to Nano-LC-MS/MS analysis. Automated 2D nanoflow LC-MS/MS analysis was performed using LTQ tandem mass spectrometer (Thermo Electron Corporation) employing automated data-dependent acquisition. An Agilent 1100 HPLC system was used to deliver a flow rate of 300 nL min⁻¹ to the mass spectrometer through a splitter. Chromatographic separation was accomplished using a 3 phase capillary column. Using an in-house constructed pressure cell, 5µm Zorbax SB-C18 packing material was packed into a fused silica capillary tubing (200µm ID, 360 µm OD, 20 cm long) to form the first dimension RP column (RP1). A similar column (200µm ID, 5 cm long) packed with 5 µm PolySulfoethyl (PolyLC) packing material was used as the SCX column. A zero dead volume 1µm filter (Upchurch, M548) was attached to the exit of each column for column packing and connecting. A fused silica capillary (100µm ID, 360 µm OD, 20 cm long) packed with 5 µm Zorbax SB-C18 packing material was used as the analytical column (RP2). One end of the fused silica tubing was pulled to a sharp tip with the ID smaller than 1µm using a laser puller as the electrospray tip. The peptide mixtures were loaded onto the RP1 column using the same in-house pressure cell. Peptides were first eluted from RP1 column to SCX column using a 0 to 80% acetonitrile gradient for 150 min. Then the peptides were fractionated by the SCX column using a series of salt gradients (20, 30, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 120, 150, 180, 200, 1000 mM ammonium acetate for 20 min), followed by high resolution reverse phase separation using an acetonitrile gradient of 0 to 80% for 120 min. The full MS scan range of 400 – 2000 m/z was divided into 3 smaller scan ranges (400 – 800, 800 – 1050, 1050 – 2000) to improve the dynamic range. Both CID (Collision Induced Dissociation) and PQD (Pulsed-Q Dissociation) scans of the same parent ion were collected for protein identification and quantitation. Each MS scan was followed by 4 pairs of CID-PQD MS/MS scans of the most intense ions from the parent MS scan. A dynamic exclusion of 1 min was used to improve the duty cycle of MS/MS scans. About 20,000 MS/MS spectra were collected for each sample. The raw data was extracted and searched using Spectrum Mill (Agilent, ver. A.03.02). The CID and PQD scans from the same parent ion were merged together. MS/MS spectra with a sequence tag length of 1 or less were considered as poor spectra and discarded. The rest of the MS/MS spectra were searched against the NCBI Ref Seq protein database limited to E. coli (16,324 sequences). The enzyme parameter was limited to full tryptic peptides with a maximum miscleavage of 1. All other search parameters were set to Spectrum Mill's default settings (carbamidomethylation of cysteines, iTRAQ modification, +/- 2.5 Da for precursor ions, +/- 0.7 Da for fragment ions, and a minimum matched peak intensity of 50%). A concatenated forward-reverse database was constructed to calculate the in-situ false discovery rate (FDR). Proteins with shared peptide(s) were grouped together into protein groups. A protein identification cut-off of 1% FDR at the protein group level was used. A total of 1,132 protein groups corresponding to 3,130 Ref Seq proteins were identified, among them 11 protein groups and 29 proteins were from the reverse database. Relative protein quantitation was performed by calculating the 117/115 (P22/P0) iTRAQ reported ion intensity ratios. Protein iTRAQ intensities were calculated by summing the peptide iTRAQ intensities from each protein group. Peptides

shared among different protein groups were removed before quantitation. Among the 1,132 identified protein groups, 1,097 had strong iTRAQ reporter intensities to obtain relative quantitation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

L.W. acknowledges support from US National Institutes of Health (1DP2OD004744-01 and P30CA014195).

References

1. Crick FH. J. Mol. Biol. 1968; 38:367–379. [PubMed: 4887876]
2. a) Osawa S, Jukes T, Watanabe K, Muto A. Microbiol. Rev. 1992; 56:229–264. [PubMed: 1579111] b) Knight RD, Freeland SJ, Landweber LF. Nat. Rev. Genet. 2001; 2:49–58. [PubMed: 11253070]
3. a) Lozupone C, Knight R, Landweber L. Curr. Biol. 2001; 11:65–74. [PubMed: 11231122] b) Inagaki Y, Doolittle WF. Nucleic Acids Res. 2001; 29:921–927. [PubMed: 11160924]
4. Johnson DB, Wang C, Xu J, Schultz MD, Schmitz RJ, Ecker JR, Wang L. ACS Chem. Biol. 2012; 7:1337–1344. [PubMed: 22662873]
5. a) Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. Science. 1989; 244:182–188. [PubMed: 2649980] b) Nowak MW, Kearney PC, Sampson JR, Saks ME, Labarca CG, Silverman SK, Zhong WG, Thorson J, Abelson JN, Davidson N, Schultz PG, Dougherty DA, Lester HA. Science. 1995; 268:439–442. [PubMed: 7716551] c) Wang L, Brock A, Herberich B, Schultz PG. Science. 2001; 292:498–500. [PubMed: 11313494]
6. a) Benzer S, Champe SP. Proc. Natl. Acad. Sci. U. S. A. 1962; 48:1114–1121. [PubMed: 13867417] b) Beier H, Grimm M. Nucleic Acids Res. 2001; 29:4767–4782. [PubMed: 11726686]
7. Wang L, Schultz PG. Angew. Chem. Int. Ed. Engl. 2005; 44:34–66. [PubMed: 15599909]
8. a) Sakamoto K, Hayashi A, Sakamoto A, Kiga D, Nakayama H, Soma A, Kobayashi T, Kitabatake M, Takio K, Saito K, Shirouzu M, Hirao I, Yokoyama S. Nucleic Acids Res. 2002; 30:4692–4699. [PubMed: 12409460] b) Liu W, Brock A, Chen S, Chen S, Schultz PG. Nat. Methods. 2007; 4:239–244. [PubMed: 17322890] c) Wang W, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee KF, Slesinger PA, Wang L. Nat. Neurosci. 2007; 10:1063–1072. [PubMed: 17603477] d) Shen B, Xiang Z, Miller B, Louie G, Wang W, Noel JP, Gage FH, Wang L. Stem Cells. 2011; 29:1231–1240. [PubMed: 21681861] e) Kang JY, Kawaguchi D, Coin I, Xiang Z, O'Leary DD, Slesinger PA, Wang L. Neuron. 2013; 80:358–370. [PubMed: 24139041]
9. Reviews: Wang L, Xie J, Schultz PG. Annu. Rev. Biophys. Biomol. Struct. 2006; 35:225–249. [PubMed: 16689635] Wang Q, Parrish AR, Wang L. Chem. Biol. 2009; 16:323–336. [PubMed: 19318213] Liu CC, Schultz PG. Annu. Rev. Biochem. 2010; 79:413–444. [PubMed: 20307192] Recent examples: Nikic I, Plass T, Schraidt O, Szymanski J, Briggs JA, Schultz C, Lemke EA. Angew. Chem. Int. Ed. Engl. 2014; 53:2245–2249. [PubMed: 24474648] Schmidt MJ, Borbas J, Drescher M, Summerer D. J. Am. Chem. Soc. 2014; 136:1238–1241. [PubMed: 24428347] Lee YJ, Wu B, Raymond JE, Zeng Y, Fang X, Wooley KL, Liu WR. ACS Chem. Biol. 2013; 8:1664–1670. [PubMed: 23735044] Gattner MJ, Vrabel M, Carell T. Chem Commun (Camb). 2013; 49:379–381. [PubMed: 23192406] Li F, Zhang H, Sun Y, Pan Y, Zhou J, Wang J. Angew. Chem. Int. Ed. Engl. 2013; 52:9700–9704. [PubMed: 23873613] Yang Y, Lin S, Lin W, Chen PR. Chem Bio Chem. 2014 Coin I, Katritch V, Sun T, Xiang Z, Siu FY, Beyermann M, Stevens RC, Wang L. Cell. 2013; 155:1258–1269. [PubMed: 24290358] Zhu S, Riou M, Yao CA, Carvalho S, Rodriguez PC, Bensaude O, Paoletti P, Ye S. Proc. Natl. Acad. Sci. U. S. A. 2014; 111:6081–6086. [PubMed: 24715733]
10. Wang L, Schultz PG. Chem. Biol. 2001; 8:883–890. [PubMed: 11564556]

11. a) Skaliter R, Eichenbaum Z, Shwartz H, Ascarelli-Goell R, Livneh Z. *Mutat Res.* 1992; 267:139–151. [PubMed: 1373849] b) Eichenbaum Z, Livneh Z. *Genetics.* 1998; 149:1173–1181. [PubMed: 9649512]
12. Moore SD, Sauer RT. *Annu Rev Biochem.* 2007; 76:101–124. [PubMed: 17291191]
13. Ruiz N, Kahne D, Silhavy TJ. *Nat Rev Microbiol.* 2006; 4:57–66. [PubMed: 16357861]
14. Phadtare S. *Curr Issues Mol Biol.* 2004; 6:125–136. [PubMed: 15119823]
15. Badger J, et al. *Proteins.* 2005; 60:787–796. [PubMed: 16021622]
16. a) Mukai T, Hayashi A, Iraha F, Sato A, Ohtake K, Yokoyama S, Sakamoto K. *Nucleic Acids Res.* 2010; 38:8188–8195. [PubMed: 20702426] b) Johnson DB, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, Xiang Z, Ecker JR, Briggs SP, Wang L. *Nat. Chem. Biol.* 2011; 7:779–786. [PubMed: 21926996] c) Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman DB, Reppas NB, Emig CJ, Bang D, Hwang SJ, Jewett MC, Jacobson JM, Church GM. *Science.* 2011; 333:348–353. [PubMed: 21764749]
17. Posfai G, Plunkett G 3rd, Feher T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M, Burland V, Harcum SW, Blattner FR. *Science.* 2006; 312:1044–1046. [PubMed: 16645050]
18. Schultz DW, Yarus M. *J Mol Biol.* 1994; 235:1377–1380. [PubMed: 8107079]
19. Wang L, Xie J, Deniz AA, Schultz PG. *J. Org. Chem.* 2003; 68:174–176. [PubMed: 12515477]

Figure 1.

Strong amber suppression slowed down bacteria growth and led to decrease in amber suppression over time. A) The orthogonal amber suppressor tRNA (J17-tRNA^{Tyr}_{CUA}), its cognate orthogonal synthetase (*MjTyrRS*), and the TAG-containing EGFP gene were assembled in plasmid pBK-tYGT to afford strong amber suppression and green fluorescence readout. Plasmid pBK-GT expresses EGFP-TAG only and was used as the control. P: promoter; T: terminator. B) Growth curves of DH10β cells harboring plasmid pBK-tYGT or

pBK-GT. Error bars represent s. d., n = 3. C) Percentage of fluorescent cells decreased with growth time. Error bars represent s. d., n = 3.

Figure 2.

Non-fluorescent cells resulted from amber suppression had transposon insertions in the synthetase gene cassette of plasmid pBK-tYGT. Plasmids extracted from non-fluorescent colonies N1-N8 were amplified with primers specific for each gene cassette. Fluorescent colonies F1-F2 and untransformed plasmid pBK-tYGT were used as control. The shift of plasmids was coordinate with the shift of the synthetase PCR products. M, molecular marker.

Figure 3.

YdiI overexpression reduces EGFP reporter gene expression by decreasing the plasmid copy number. A) Plasmid pBK-tYGT was co-transformed with pLEI-ydiI or an empty vector pLEI into *E. coli* DH10 β cells. Fluorescence images were taken for colonies grown on plates. Colony morphology altered in cells overexpressing YdiI. B) Full-length and truncated EGFP with N-terminal His6 tag were purified and separated by 12% SDS-PAGE. Note 40-fold loading for ydiI overexpressing cells. F: full-length EGFP, T: truncated EGFP. C) Plasmids extracted from same amount of cells were separated in agarose gel and stained

with ethidium bromide. Plasmid yield in *ydiI* overexpressing cells was dramatically lower than the control. Note plasmids were not linearized.

Figure 4.

Expression of YdiI expels plasmids from *E. coli*. A) Experimental procedures to measure what percentage of cells could survive on plates containing different antibiotics. In cells transformed with pBK-tYGT and pLEI-ydiI, YdiI was expressed by pLEI-ydiI upon IPTG induction. Control cells were transformed with pBK-tYGT and pLEI, a control plasmid without the ydiI gene. pBK-tYGT is kanamycin (Kan) resistant, and pLEI plasmids are chloramphenicol (Cm) resistant. B) Percentage of surviving cells transformed with pBK-tYGT and pLEI. C) Percentage of surviving cells transformed with pBK-tYGT and pLEI-ydiI. Error bars represent s.e.m., n = 3.

Table 1

Transposon insertion in the synthetase gene cassette.

Colony	Transposon type	Insertion site ^[a]	Insertion direction
N1	IS10	427	Reverse
N3	IS10	1040	Reverse
N5	IS1	404	Reverse
N6	IS10	1040	Forward
N7	IS1	748	Forward

^[a] Insertion sites were numbered by the start site of the MjTyrRS gene.

Table 2

Proteins up- or down-regulated > 2 fold after persistent, long-term amber suppression in *E. coli*.

Accession ^[a]	Protein	P22/P0 ^[b]	Accession	Protein	P22/P0
NP_417479	ExbB	0.15	NP_415278	GalK	0.46
NP_416719	OmpC	0.19	NP_415280	GalE	0.47
NP_415772	OmpW	0.20	NP_418392	ArgE	0.47
NP_418232	WzzE ^[c]	0.21	NP_418392	ArgE	0.48
NP_415895	OmpN	0.27	NP_417616	YraM	0.49
NP_415085	NmpC	0.27	NP_418258	UvrD	2.00
NP_415449	OmpF	0.27	NP_417573	YqjG	2.01
NP_417950	PitA	0.27	NP_417079	YfiQ	2.03
NP_418593	Hfq	0.32	NP_417348	YgeY	2.05
NP_418476	DnaB	0.32	NP_417432	AnsB	2.08
NP_417887	GlgA ^[c]	0.34	NP_414978	QueC	2.09
NP_415335	OmpX	0.34	NP_416476	HchA	2.11
NP_417875	MalQ ^[c]	0.34	NP_418222	IivC	2.14
NP_416693	YejM	0.37	NP_414995	AcrB	2.36
NP_415710	LdcA	0.37	NP_416567	YegH	2.38
NP_414724	LpxB	0.38	NP_415581	YceB	2.43
NP_418407	EF-Tu	0.39	NP_415820	PspA	2.52
NP_416175	PurR	0.41	NP_414702	Dgt	2.68
NP_417244	CysJ	0.42	NP_415430	Cmk	2.76
NP_418142	IbpA	0.42	NP_416200	YdiH	3.12
NP_415348	MoeA	0.43	NP_417354	YgfK	3.93
NP_416930	AmiA	0.43	NP_415510	CspG	4.23
NP_417999	DppC	0.44	NP_415212	YbfF	5.15
NP_418141	IbpB	0.44	NP_418012	CspA	6.78
NP_416678	YeiR	0.46	NP_416075	CspB	8.34
NP_418193	AtpE ^[c]	0.46	NP_416201	YdiI	16.19

^[a] NCBI Accession number.

b) Ratios of proteins from final round (P22) and initial round (P0) cells are listed as P22/P0.

c) Proteins whose encoding genes are ended with TAG.