This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

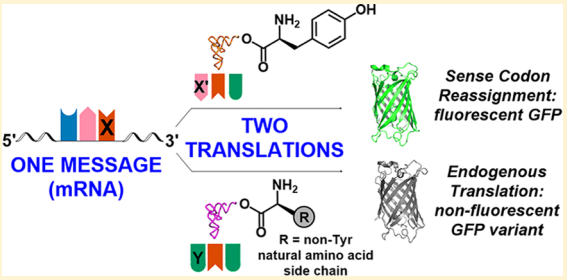
Evaluating Sense Codon Reassignment with a Simple Fluorescence Screen

Wil Biddle,[‡] Margaret A. Schmitt,[†] and John D. Fisk*[,‡,†,§]

[†]Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, Colorado 80523, United States
[‡]Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, United States
[§]School of Biomedical Engineering, Colorado State University, Fort Collins, Colorado 80523, United States

Supporting Information

ABSTRACT: Understanding the interactions that drive the fidelity of the genetic code and the limits to which modifications can be made without breaking the translational system has practical implications for understanding the molecular mechanisms of evolution as well as expanding the set of encodable amino acids, particularly those with chemistries not provided by Nature. Because 61 sense codons encode 20 amino acids, reassigning the meaning of sense codons provides an avenue for biosynthetic modification of proteins, furthering both fundamental and applied biochemical research. We developed a simple screen that exploits the absolute requirement for fluorescence of an active site tyrosine in green fluorescent protein (GFP) to probe the pliability of the degeneracy of the genetic code. Our screen monitors the restoration of the fluorophore of GFP by incorporation of a tyrosine in response to a sense codon typically assigned another meaning in the genetic code. We evaluated sense codon reassignment at four of the 21 sense codons read through wobble interactions in *Escherichia coli* using the *Methanocaldococcus jannaschii* orthogonal tRNA/aminoacyl tRNA synthetase pair originally developed and commonly used for amber stop codon suppression. By changing only the anticodon of the orthogonal tRNA, we achieved sense codon reassignment efficiencies between 1% (Phe UUU) and 6% (Lys AAG). Each of the orthogonal tRNAs preferentially decoded the codon traditionally read via a wobble interaction in *E. coli* with the exception of the orthogonal tRNA with an AUG anticodon, which incorporated tyrosine in response to both the His CAU and His CAC codons with approximately equal frequencies. We applied our screen in a high-throughput manner to evaluate a 10^9-member combined tRNA/aminoacyl tRNA synthetase library to identify improved sense codon reassigning variants for the Lys AAG codon. A single rapid screen with the ability to broadly evaluate reassignable codons will facilitate identification and improvement of the combinations of sense codons and orthogonal pairs that display efficient reassignment.



One of the driving motivations of both chemical and synthetic biology is the expansion of the set of building blocks that can be employed in the templated synthesis of biopolymers. The expansion of the set of amino acids that can be utilized in translation is particularly challenging. Nature expends a great deal of energy to maintain the fidelity of translation: 75% of a bacterial cell's energy budget is spent in the production of proteins, and much of that energy is devoted to proofreading at multiple steps in the translation process.[1−3] Genetic code expansion is further hampered by the fact that all 64 codon triplets have an assigned function. However, the genetic code is degenerate: 61 sense codons specify 20 canonical amino acids. While the amino acid specified by each triplet of nucleotides has not changed during evolution (with some rare exceptions), codon usage frequencies and the complements of adapter tRNA molecules used to translate the code have diverged considerably.[4,5] The fact that codon usage varies widely across organisms and that different species employ different sets of tRNAs to decode their genomes implies that a fair degree of plasticity exists in the machinery that specifies the genetic code.[6,7]

Breaking the degeneracy of the genetic code and reassigning the meaning of sense codons has the potential to expand the genetic code to 22 (or more) amino acids, greatly increasing the encodable properties of proteins. Genetic code expansion via nonsense [i.e., amber stop (UAG) or four-base codon] suppression has been limited to incorporation of non-canonical amino acids (ncAAs) at a single site because competition with termination signals or frameshift mutations curtail the amount of protein produced when multiple suppressions are attempted in a single protein.[8−11] Recently, the laboratories of Sakamoto, Wang, and Church each engineered genomic changes in *Escherichia coli* that mitigate the usual cytotoxic effect of deletion of the release factor that competes for decoding the amber stop signal.[12−14] These efforts have produced cells in which the meaning of the amber codon is "free" and can be utilized to encode multiple copies of a non-canonical amino acid, generating a 21 amino acid

Received: August 4, 2015
Revised: October 6, 2015
Published: November 4, 2015

 © 2015 American Chemical Society

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

genetic code. However, none of the methods can be easily extended to additional codons.

An alternative approach to genetic code expansion, residue specific reassignment, allows multisite incorporation of ncAAs but requires replacement of every occurrence of one natural amino acid with the ncAA. Reassignment is controlled by switching the growth medium such that the targeted natural amino acid is removed and replaced by a ncAA that is a close structural analogue.[15−17] The ncAA is utilized in translation in place of the removed natural amino acid, resulting in proteins with high levels of non-canonical amino acid incorporation at multiple sites without appreciable reductions in yield. The resulting genetic code is an altered 20 amino acid genetic code because one natural amino acid must be removed entirely.

Using a strategy that combines aspects of the amber suppression and residue specific methods, Kwon, Kirshenbaum, and Tirrell broke the degeneracy of the genetic code and demonstrated that one of two codons specifying phenylalanine (Phe) in *E. coli* could be reassigned to allow multisite incorporation of ncAAs.[18,19] *E. coli* utilize one tRNA species with a GAA anticodon to decode two Phe codons: UUU and UUC. Naphthylalanine was incorporated in response to the UUU codon (usually read by a wobble interaction) by supplying an orthogonal yeast tRNA modified with an AAA anticodon and an orthogonal yeast aminoacyl tRNA synthetase. The UUC codon continued to direct incorporation of Phe, resulting in a 21 amino acid genetic code.

Expanding the genetic code by reassigning the meaning of sense codons should be broadly generalizable. Unfortunately, predicting which sense codons are amenable to reassignment and which orthogonal machinery is best suited for the task is made challenging by the largely unknown and idiosyncratic recognition and discrimination features of each organism's complement of tRNAs and aminoacyl tRNA synthetases (aaRSs). The interactions between the tRNA and aaRS molecules that drive the fidelity of the genetic code have been only partially mapped, and the space between these interactions, the extent to which additional orthogonal tRNA/aaRS pairs can be added, is largely unknown. Understanding the limits to which modifications can be made without breaking the translational system has practical implications beyond the expansion of the set of encodable amino acids. Measurements of the global effects of translational system modifications in which the reading of specific sense codons is altered, including the extent to which a given sense codon is naturally subject to errors, contribute to understanding the molecular mechanisms of translation-related diseases and certain cancers.[20−22]

Five key conditions must be met in *E. coli* to break the degeneracy of the genetic code. (1) A natural amino acid must be specified by multiple codons such that one of the codons can be reassigned to the ncAA. In the standard genetic code, 18 of the 20 natural amino acids are encoded by more than one codon. (2) The codon targeted for sense codon reassignment should be read by the set of natural *E. coli* tRNAs via a wobble interaction in which the tRNA anticodon is able to decode a codon triplet through recognition of two positions by Watson−Crick base pairing.[23] The third codon position is recognized via an energetically less favorable wobble pairing of nucleotide bases other than G with C or A with U. In *E. coli*, 40 tRNA species decode 61 sense codons.[24] The subset of 21 codons that are read through wobble interactions are potential targets for sense codon reassignment.[18] (3) The orthogonal tRNA must be able to outcompete the natural translational system for the codon to be reassigned. A combination of increased concentration, preferential energetic interactions of the Watson−Crick versus wobble base pairing, and more subtle effects of the precise geometry of the decoding hairpin in the codon−anticodon interaction contribute to the ability of tRNAs to compete for decoding the same sense codon.[4,25] (4) The nucleotide changes in the anticodon of the orthogonal tRNA required to read the selected sense codon must not abrogate the interaction with the cognate orthogonal aminoacyl tRNA synthetase. The anticodon is often an important identity element that allows a specific aaRS to recognize its appropriate tRNA. Changing the anticodon can affect the efficiency with which the ncAA is attached to the sense codon reassigning tRNA.[26] (5) The orthogonality of the tRNA/aaRS pair must be maintained as the tRNA anticodon is modified and the aaRS is evolved to recognize the new tRNA. The complete set of tRNA/aaRS interactions is not known even for *E. coli*, and how each modified heterologous pair will interact with the natural system is extremely difficult to predict.

In this report, we describe a simple screen to evaluate the reassignment of sense codons in *E. coli* through the introduction of modified forms of an orthogonal tRNA/aminoacyl tRNA synthetase pair. Our system employs the orthogonal tyrosine-incorporating tRNA and aaRS pair from *Methanocaldococcus jannaschii* and a GFP reporter protein with modifications to the codon specifying the fluorophore tyrosine residue. We used the fluorescence-based screen to evaluate the reassignment of four *E. coli* sense codons and improve the extent of reassignment of one codon through the directed evolution of tRNA and aaRS variants. Because our intent was to improve the interactions between an orthogonal tRNA with an altered anticodon and the orthogonal aminoacyl tRNA synthetase, rather than change the identity of the amino acid charged to the tRNA, libraries were designed to explore the space of the anticodon loop of the tRNA and the anticodon binding domain of the aaRS. The expansion of the genetic code via sense codon reassignment can be viewed as a two-part problem. The simple screen we developed addresses the first facet, identification of the particular sense codons that can be most productively targeted by a particular orthogonal pair. The second facet of the problem involves modulating the levels of aminoacyl tRNA synthetase function (through adjustments in expression levels or activity) and controlling protein expression conditions to effectively compete with the natural translational apparatus. Our screen allows improvement of the aminoacylation efficiency for a specific tRNA by providing a high-throughput measurable for incorporation.

The selection of initial codons to evaluate was guided by the eventual goal of using variants of the *M. jannaschii* tRNA/aaRS pair to incorporate non-canonical amino acids in response to sense codons. The *M. jannaschii* pair has been subjected to repeated directed evolution experiments, leading to identification of variants capable of incorporating nearly 100 ncAAs into proteins in response to the amber stop codon.[8] The choice of codons for this initial evaluation also considered maintenance of orthogonal pair function with anticodon modification, cellular codon usage, tRNA levels, and the cellular tolerance toward substitution of the amino acid targeted for reassignment. Beyond the utility of the fluorescence screen for evaluating sense codon reassignment, the screen is sufficiently sensitive to measure natural background levels of missense incorporation of tyrosine.

EXPERIMENTAL PROCEDURES

General Methods. Reagent details and sources, DNA manipulation and mutagenesis details, and vector and

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

oligonucleotide primer sequences are described in the Supporting Information.

Antibiotics for Vector Maintenance. Spectinomycin was used at 50 μg/mL to maintain the vectors harboring the tRNA and aaRS genes. Carbenicillin was used at 50 μg/mL to maintain the vectors harboring the GFP reporter gene. Unless otherwise noted, these antibiotics were present in all liquid and solid media.

GFP Fluorescence-Based Sense Codon Reassignment Efficiency Assays: Reassignment Efficiency *In Vivo*. Superfolder green fluorescent protein (GFP) reporter plasmids (pGFP66xxx, where xxx specifies the codon at position 66) were cotransformed with vectors expressing the modified orthogonal translational components (pWB_Ultra-Tyr-yyy, where yyy indicates the anticodon on the tRNA) into SB3930 *E. coli* cells (λ^-, Δ*hisB463*). After overnight growth, colonies were picked into 200 μL of LB medium in a 96-well plate. Cells were grown to saturation (usually 12 h) while being shaken at 37 °C. Cells were diluted 10-fold into LB medium with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for induction of aaRS and GFP reporter. Assays were performed in a Fluorotrac 200 clear bottom 96-well plate (Greiner 655096) and monitored in a BioTek Synergy H1 plate reader at 37 °C with continuous double orbital shaking. The optical density (OD_{600}) and fluorescence of each well were measured every 20 min; the optical density was measured at 600 nm, and the fluorescence was measured with an excitation at 485 nm and detection at 515 nm with an 8 nm band-pass. The relative fluorescence of each 200 μL culture was calculated by dividing the LB medium-blanked fluorescence by the OD_{600}. The 100% relative fluorescence unit (reported as fluorescence units per unit of optical density at 600 nm, RFU) value for sense codon reassignment efficiency was defined by taking an average of three cultures expressing wild-type GFP (pGFP66tat) in cells harboring a negative control translational machinery plasmid (pWB_Ultra-Tyr-no_tRNA) to maintain a similar metabolic burden on the cell. The sense codon reassignment efficiency for each tRNA variant was calculated by dividing the individual RFU values from each of three colonies by the average 100% reference RFU value and then averaging the results.

Protein Purification and Quantification for Reassignment Efficiency. Colonies of SB3930 cells harboring vectors expressing both the orthogonal translational components and a GFP reporter gene were picked into 200 μL of LB and grown to saturation while being shaken at 37 °C. An 80 μL aliquot from each culture was diluted with 3 mL of LB medium containing 1 mM IPTG. The aaRS and reporter protein were expressed while being shaken at 37 °C for 12 h; 2.6 mL of each expression culture was pelleted by centrifugation at 17000*g* for 2 min, and the supernatant was discarded. Cell pellets were frozen overnight at −80 °C. The pellets were thawed at room temperature, and 200 μL of B-PER lysis reagent (Thermo Scientific) was added to each cell pellet and vortexed to mix. Lysis reaction mixtures were incubated at room temperature for 30 min, and whole cell lysates were clarified by pelleting at 17000*g* for 5 min. Ni-NTA spin columns (Qiagen) were equilibrated with NPI-10 buffer [50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole (pH 8.0)]; 200 μL of cleared lysate was loaded onto the columns and eluted following the recommendations from Qiagen. Proteins were washed twice with NPI-20 buffer [50 mM NaH_2PO_4, 300 mM NaCl, and 20 mM imidazole (pH 8.0)] and then eluted twice with 350 μL of NPI-500 [50 mM NaH_2PO_4, 300 mM NaCl, and 500 mM imidazole (pH 8.0)] for a total elution volume of 700 μL. Imidazole and other salts were removed from the eluted protein samples using Amicon 3000 molecular weight cutoff spin columns (Millipore, 0.5 mL volume). Briefly, 150 μL of each protein elution was combined with 350 μL of phosphate-buffered saline (PBS) and spun through the filter at 17000*g* and 4 °C for 30 min. After the first spin, approximately 50 μL of condensate remained, to which 450 μL of PBS was added. After the second spin that was identical to the first, 150 μL of PBS was added to the 50 μL of condensate and the filter was inverted. The dilute condensate was spun out of the filter for 2 min at 1000*g*, resulting in a final imidazole concentration of 3.8 mM. Two samples of 200 μg/mL bovine serum albumin (BSA) were diluted in NPI-500 elution buffer and desalted using the filtration process described above to prepare a standard curve for protein quantification using the Pierce Micro BCA Protein Assay kit.

Fourfold and 8-fold dilutions of each protein sample (150 μL each) were placed into a Fluorotrac 200 clear bottom 96-well plate, and the fluorescence of each sample was measured (excitation at 485 nm, emission read at 515 nm). Following quantification of fluorescence, 150 μL of BCA working reagent (25:24:1) was added to each well to determine the protein concentration. The plate was shaken for 30 s to mix and then incubated at 37 °C for 2 h before the absorbance of each well was read at 562 nm. Protein concentrations were determined on the basis of a BSA standard curve (see the Supporting Information).

Construction of the 256-Member tRNA Anticodon Loop Library. Primer GX (see the Supporting Information) was used to construct a small (256-member) library of tRNA anticodon loop variants via Kunkel mutagenesis.[27,28] The mutagenic primer was designed to simultaneously randomize tRNA positions 32, 33, 37, and 38 flanking the anticodon (Figure 2a) and change the anticodon from AUG to CUU. We created the library using vector pWB_Ultra-Tyr-AUG as a template to reduce background fluorescence from non-library members; the AUG anticodon in the starting material tRNA would not recognize the AAG codon in the GFP reporter fluorophore position, and therefore, Lys would be incorporated into a non-fluorescent protein. The library was transformed into NEB 5-alpha cells harboring the pGFP66aag reporter plasmid and plated onto LB agar with 0.5 mM IPTG to induce expression of the aaRS and reporter proteins. Of 295 colonies on a 10000-fold diluted plate (2.95×10^6 total unique transformants), seven were visibly fluorescent when viewed on a dark reader (DR46B Transilluminator, Clare Chemical Research). Colonies from this plate, including the seven visibly fluorescent colonies, were evaluated using the *in vivo* assay described above. Clones from the library with fluorescence on the same order of magnitude as that of pWB_Ultra-Tyr-CUU with the same pGFP66aag reporter were colony purified on LB agar plates. Colonies from these plates were then analyzed to evaluate colony-to-colony variation.

Construction of the Combined tRNA Anticodon Loop and aaRS tRNA Binding Domain Library. The starting material for the large combined tRNA/aaRS library was pWB_Ultra-XhoI-Inactive, a variant of pWB_Ultra-Tyr-AUG that contains two XhoI restriction sites within the aaRS gene and an additional XhoI site in the anticodon position of the tRNA. Restriction sites were introduced using primers GZ, HA, and HB (see the Supporting Information). The primers used to construct the library remove the XhoI restriction sites, providing a facile method for determining the efficiency of mutagenesis. Most library members will not be green because most mutations will be deleterious to tRNA/aaRS or tRNA/mRNA recognition. The library was created using primer GX (anticodon loop random-

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

ization) and primers HE, HF, and HG, which target nine amino acid residues in three regions in the aaRS. The amino acids chosen for randomization (Y230, C231, P232, A233, F261, H283, P284, M285, and D286) are within 5 Å of the tRNA anticodon in the crystal structure of the *M. jannaschii* TyrRS complexed with the *M. jannaschii* tyrosyl-tRNA [Protein Data Bank (PDB) entry 1J1U] (Figure 2b).[29] In designing the library, we used PyMol software to approximate the tRNA anticodon as CUU (instead of the Tyr GUA).[30] Degenerate codons for each position of the library were chosen to include the original wild-type residue for both the aaRS and the tRNA anticodon loop. The theoretical diversity of the combined library is 1.77×10^{10}. The library was generated using Kunkel mutagenesis and then transformed into SS320 cells to remove the single-stranded dU DNA template. The transformation was allowed to recover in 50 mL of SOC without antibiotics for 60 min at 37 °C. A small aliquot was plated on LB agar (50 μg/mL spectinomycin) to determine the mutation efficiency (1×10^9 unique transformants). The remainder of the transformation recovery was diluted up to 150 mL of 2xYT medium with 0.2% glucose (to mimic SOC recovery medium) and 50 μg/mL spectinomycin and allowed to grow for an additional 60 min at 37 °C; 18.4 μg of DNA was isolated using a Qiagen maxi prep kit (product 12162). Sixteen micrograms of this DNA was digested with XhoI to remove unmutated template DNA. The resulting DNA was transformed into NEB 5-alpha cells harboring the pGFP66aag reporter plasmid and recovered in 50 mL of SOC medium at 37 °C for 60 min. A small aliquot was plated on LB agar with 1 mM IPTG to determine the transformation efficiency (2.5×10^8 unique transformants). The remainder of the transformation recovery was transferred into 600 mL of LB medium with 1 mM IPTG and grown overnight to a final OD_{600} of 1.7. The LB agar plates from each transformation were analyzed, and polymerase chain reaction fragments amplified from individual colonies were restricted with XhoI to determine mutation efficiency. After transformation into SS320 cells (to remove uridine-containing template DNA), 3 of 11 colonies analyzed were library members (three of three XhoI restriction sites in the starting material had been removed). The maxiprep−restrict−retransform step eliminated the remaining starting material: 10 of 10 colonies analyzed had all three XhoI restriction sites removed.

Flow Cytometry and Cell Sorting. Approximately 1.6×10^9 cells from a 600 mL library culture (final OD_{600} of 1.7) were pelleted in a benchtop centrifuge at 8000*g* for 3 min and then resuspended in a sterile 0.9% NaCl solution. These cells were sorted via fluorescence-activated cell sorting (FACS) on a Dako-Cytomation MoFlo Legacy instrument using the 488 nm laser line. To remove non-cell particles and cell clumps, cells were initially gated on forward and side scatter. Accepted cells were then gated on fluorescence using a 530/40 nm band-pass filter in front of the GFP emission detector. Cells were sorted using the Purify 1-2 mode and either a 70 μm or a 100 μm flow cell tip. As a control, we used NEB 5-alpha cells harboring the most efficient clone from the 256-member anticodon loop library and the unmodified Tyr aaRS, pWB_Ultra-Tyr-CUU^{AU}, and the pGFP66aag reporter expressed in a fashion identical to that of the combined library. Approximately 6×10^7 library cells were analyzed in the first sort, and 3×10^5 cells with fluorescence greater than 100 arbitrary fluorescence units were collected in PBS. The collected cells were diluted directly into LB medium and incubated while being shaken overnight at 37 °C. The saturated culture of FACS-enriched cells was plated on LB agar with 1 mM IPTG to yield isolated colonies. The remainder of the library culture was frozen in 35% glycerol at −80 °C. Isolated colonies from the first round of FACS screening were evaluated using the *in vivo* fluorescence assay described above.

For the second round of FACS, 1 mL of the cells frozen after the first sort (representing approximately 10% of the enriched clones) was diluted into 20 mL of LB with 1 mM IPTG and grown while being shaken at 37 °C for 10 h. Then 200 μL of the cell culture (approximately 1.4×10^9 cells) was pelleted at 8000*g* for 3 min and resuspended in 0.9% NaCl; 3.7×10^7 of these cells were sorted using a narrowed collection gate, and 1.0×10^6 cells with fluorescence greater than the top 5% of our control cells were collected. These cells were diluted into LB, amplified, and plated on LB agar with 1 mM IPTG. Individual clones were screened using the *in vivo* fluorescence assay as described above.

RESULTS AND DISCUSSION

Principle of the GFP-Based Screen. Our screen evaluates the ability of anticodon-modified tRNAs to incorporate tyrosine (Tyr) in response to a sense codon that is assigned another identity in the standard *E. coli* genetic code. Residues 65−67 of superfolder GFP specify the Thr-Tyr-Gly sequence that autocatalytically folds into the tripeptide fluorophore. Replacement of Tyr at position 66 with any other natural amino acid effectively abolishes the fluorescence of the protein. (Phenylalanine or histidine substitution leads to proteins with greatly reduced and shifted fluorescence.[31] Tryptophan incorporation without additional mutations leads to unfolded protein and no fluorophore formation.) We generated a series of reporter plasmids in which the fluorophore tyrosine position in GFP was mutated to each sense codon under evaluation for reassignment. Amino acids incorporated into the fluorophore tyrosine position by the *E. coli* translational machinery lead to proteins that are not fluorescent; when the orthogonal translational machinery incorporates tyrosine, fluorescent proteins are produced (see the graphic in the abstract). The amount of fluorescence observed for each reporter is a measure of the extent to which the orthogonal aaRS is able to recognize and aminoacylate its cognate tRNA with a modified anticodon and the ability of the orthogonal tRNA to compete with an *E. coli* tRNA to decode the sense codon specifying the fluorophore Tyr. Similar gain of function via missense mutation enzyme reporter systems have been employed to measure natural levels of missense incorporation.[32−35] The levels of sense codon reassignment measured under the conditions of this screen (e.g., rich medium) represent the lower limit of reassignment that should be achievable under the controlled conditions utilized for typical protein expression employing the residue specific reassignment strategy. Even low-level sense codon reassignment is useful because missed incorporations do not lead to termination, and multiple non-canonical amino acids can be incorporated into a single protein, producing a statistical mixture, but in high yield.

Selection of Sense Codons for Initial Evaluation. We selected four of the 21 sense codons that are read through wobble interactions by endogenous *E. coli* tRNAs: lysine (Lys) AAG, histidine (His) CAU, asparagine (Asn) AAU, and phenylalanine (Phe) UUU for our initial evaluation of sense codon reassignment. Our choice of codons considered cellular codon usage, tRNA levels, and the cellular tolerance toward substitution of the amino acid targeted for reassignment as well as the predicted efficiency of aminoacylation of anticodon-modified *M. jannaschii* tRNAs. The recognition elements that define the *M. jannaschii* tRNA have been mapped through kinetic measurements of tRNA variants.[36] On the basis of an assumption of additivity in

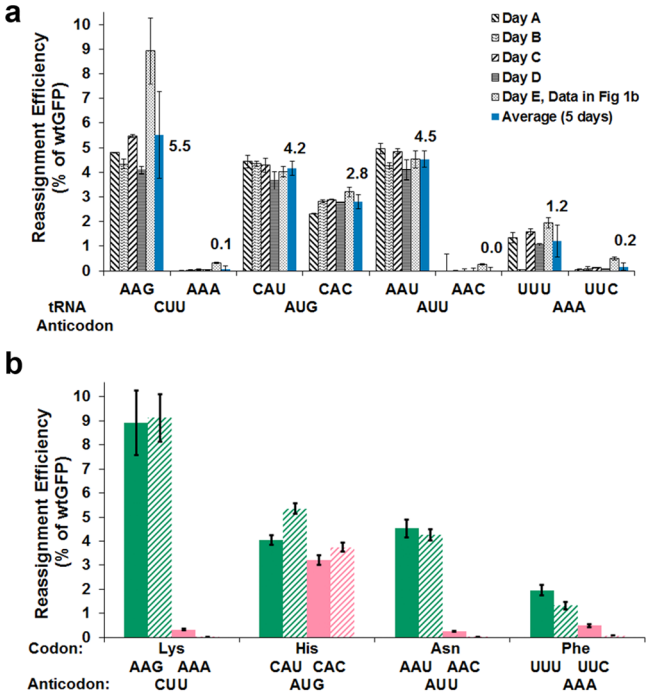
DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

the effects of measured changes in the anticodon sequence on aminoacylation, we calculated expected aminoacylation efficiency reductions for tRNA species with altered anticodons.[36] Predicted losses in aminoacylation efficiency across the anticodon sequences directed at the 61 sense codons range between 1- and 3600-fold relative to the natural tyrosine anticodon. The nucleotide changes required to convert a Tyr (GUA anticodon) to an amber suppressor (CUA anticodon) tRNA result in a 97-fold decrease in the recognition of the tRNA species by the aaRS. The sense codons we selected for our initial evaluation have predicted losses in aminoacylation efficiency of 116-fold (Lys AAG), 108-fold (Asn AAU), 360-fold (His CAU), and 864-fold (Phe UUU). These reductions in predicted aminoacylation efficiency reflect at most a 9-fold decrease in expected tRNA charging efficiency relative to that of the amber suppressor (predicted 97-fold reduction), suggesting that the altered tRNA species should be charged by the aaRS.[36]

Evaluation of Sense Codon Reassignment. We prepared variants of a form of the *M. jannaschii* tRNA gene (optimized for amber suppression)[37,38] in which the anticodon was modified to either CUU, AUG, AUU, or AAA to base pair via Watson−Crick interactions with the codons for Lys (AAG), His (CAU), Asn (AAU), or Phe (UUU) that are read through wobble interactions by endogenous *E. coli* tRNAs. In all cases, both the GFP reporter and orthogonal aaRS genes were under the control of an inducible promoter.[38,39] A single plasmid containing the *M. jannaschii* tRNA and aaRS genes was cotransformed into SB3930 *E. coli* cells along with a GFP reporter plasmid.

The observed relative fluorescence (reported as fluorescence units per unit of optical density at 600 nm, RFU) for the four sense codons ranges from 1 to 6% of the relative fluorescence of wild-type GFP (Figure 1a). The trend in sense codon reassignment efficiencies, reported as the relative fluorescence observed in each of the test cases divided by the relative fluorescence of wild-type GFP, loosely correlates with the predicted reductions in aminoacylation of anticodon-modified *M. jannaschii* tRNAs by the *M. jannaschii* aaRS. Reassignment of the Lys AAG codon has an average efficiency of 6%; the Asn and His codons AAU and CAU were reassigned with 4% efficiency, and the Phe UUU codon was reassigned with 1% efficiency. The 1−6% levels of sense codon reassignment we observe are similar to the levels seen in other systems designed to force missense incorporation.[32−34,40] The variation in sense codon reassignment levels between colonies on a given day is small, typically <5% of the measured value. Variation in observed reassignment efficiency for a given codon between tests performed on different days is up to 16% of the measured values, with the exception of a high outlier for Lys AAG and a low outlier for Phe UUC. The day-to-day variability is primarily due to variations in the amount of wild-type GFP produced.

Comparison of *In Vivo* and Purified Protein Fluorescence. In our screen, the fluorescence observed in a bacterial culture is an absolute measurement of the production of GFP with a functional tyrosine-containing fluorophore. The reassignment efficiency calculation assumes that the normalization of fluorescence per cell approximates fluorescence per protein (i.e., the amount of all forms of GFP produced in all cases are comparable). To confirm that cell density relates to protein production level, we isolated and quantified the GFP produced in each reassignment system. Sense codon reassignment efficiencies determined from protein concentration-corrected fluorescence measurements parallel those determined from the in cell assay (Figure 1b). The reassignment efficiencies determined from purified protein samples are based not on the amount of protein obtained but on the relative fluorescence of approximately equal amounts of protein obtained in all cases. The sensitivity of the in cell assay is hampered by the background fluorescence of cells



c

tRNA Anticodon	Sense Codon	Reassignment Percentage (In Cell Screen)	Reassignment Percentage (Purified Protein)	Codon Discrimination[a]
CUU	AAG	8.9 ± 1.3	9.1 ± 1.0	2275
	AAA	0.3 ± 0.03	0.004 ± 0.002	
AUG	CAU	4.0 ± 0.2	5.4 ± 0.2	1.5
	CAC	3.2 ± 0.2	3.7 ± 0.2	
AUU	AAU	4.5 ± 0.4	4.3 ± 0.2	614
	AAC	0.3 ± 0.02	0.007 ± 0.004	
AAA	UUU	2.0 ± 0.2	1.3 ± 0.2	16.3
	UUC	0.5 ± 0.06	0.08 ± 0.01	

[a] Codon discrimination is the ratio of reassignment percentage for the sense codon read by an *E. coli* tRNA through wobble interactions relative to the reassignment percentage for the sense codon read by an *E. coli* tRNA through Watson-Crick base pairing.

Figure 1. Reassignment efficiencies reported as the observed optical density-corrected fluorescence divided by the optical density-corrected fluorescence of wild-type GFP for each combination of orthogonal pair and GFP reporter. (a) Representation of day-to-day variation in observed reassignment efficiencies. Data are grouped by the anticodon of the tRNA utilized in each evaluation. Labeled blue bars represent the average reassignment efficiency observed for each codon. Data for each of the eight codons from five different experiments are shown. Data for day E are those for which protein samples were also isolated and quantified (see panel b). (b) Data are grouped by the anticodon of the tRNA utilized in each evaluation. Green bars represent data for the reassigned codon (i.e., Watson−Crick base pairing to the orthogonal tRNA). Rose bars represent data for the alternative sense codon specifying the replaced amino acid (i.e., wobble base pairing to the orthogonal tRNA). Solid bars represent reassignment efficiencies determined from the in cell screen. Hashed bars represent reasignment efficiencies determined from purified protein samples. (c) Values for reassignment efficiencies of each codon evaluated. Codon discrimination for each orthogonal tRNA is reported on the basis of an analysis of purified protein. Both in cell and purified protein data were collected for the same set of cell cultures. Error bars represent variation across three individual colonies for each of the eight sense codons and wild-type GFP.

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

and growth medium. The limit of detection in cells allows the facile identification of sense codon reassignments of ~1 part in 1000. The isolated protein assay has an expanded detection limit because of a decreased background. Sense codon substitutions can be detected at 1 part in 25000, corresponding to a substitution frequency of 4×10^{-5}. The background missense incorporation rate is generally estimated to be between 10^{-3} and 10^{-4} per codon.[35,41] Using isolated proteins, this fluorescence assay should be sensitive enough to measure background incorporation of tyrosine at sense codons in *E. coli*. The sensitivity of the purified protein assay is based on the absolute detection of fluorescence in a sample and could be further improved by evaluating larger protein samples.

Evaluation of Codon Discrimination. The guiding principle of breaking the degeneracy of the genetic code is that wobble versus Watson–Crick base pairing can be used to specifically reassign one of two (or more) codons decoded by a single tRNA species. To examine the ability of the anticodon-modified orthogonal tRNAs to discriminate between decoding the reassigned codon and the alternative sense codon specifying the replaced amino acid, we evaluated the same set of anticodon-modified tRNAs in GFP gene reporter systems in which the fluorophore Tyr is specified by the alternative Lys (AAA), His (CAC), Asn (AAC), and Phe (UUC) codons that are read via Watson–Crick interactions by the endogenous *E. coli* tRNAs. For the sense codons in our initial evaluation, codon discrimination measured using purified proteins ranges across 3 orders of magnitude from 1.5-fold for the CAU versus CAC histidine codons to 2275-fold between the lysine AAG and AAA codons (Figure 1c). Codon discrimination between Lys AAG and AAA requires discrimination between a G and an A in the codon by a C at position 34 (see tRNA numbering in Figure 2a) of the anticodon-modified *M. jannaschii* tRNA. For His, Asn, and Phe codons, discrimination between the Watson–Crick and wobble pairings involves an A at position 34 of the anticodon recognizing a U or C in the codon. A-C pairs are expected to be energtically unfavorable.[18,25] Despite the A-C pairing, and in contrast to the levels of discrimination observed for Asn and Phe, the AUG anticodon-modified *M. jannaschii* tRNA does not discriminate well between the two His codons. The lack of discrimination may be due to modification of A34 in the tRNA to inosine. An IUG anticodon would recogize the CAU (His), CAC (His), and CAA (Glu) codons with similar efficiencies.[23] The only *E. coli* tRNA with an A to I modification is the Arg2 tRNA that contains an ACG anticodon.[42] The exact substrate determinants of the A to I deaminase tadA responsible for tRNA Arg2 modification have not been mapped, but the anticodon and anticodon loop sequences appear to be important recognition elements.[42] With the exception of the central anticodon position (35), six of seven nucleobases in the *E. coli* Arg2 and AUG anticodon-modifed *M. jannaschii* tRNA anticodon and anticodon loop sequences are identical.

High-Throughput in Cell Fluorescence Evaluations. In the *M. jannaschii* system, the tRNA anticodon changes required for sense codon reassignment are expected to adversely affect aminoacylation and may also change the anticodon stem loop geometry, impacting decoding on the ribosome. To achieve useful levels of sense codon reassignment for orthogonal aaRSs that incorporate non-canonical amino acids, we expect that the aminoacylation efficiencies of the anticodon-modified tRNAs will need to be improved. We applied our fluorescence-based screen in a high-throughput manner for the directed evolution of improved sense codon reassigning tRNA/aaRS pairs. We first

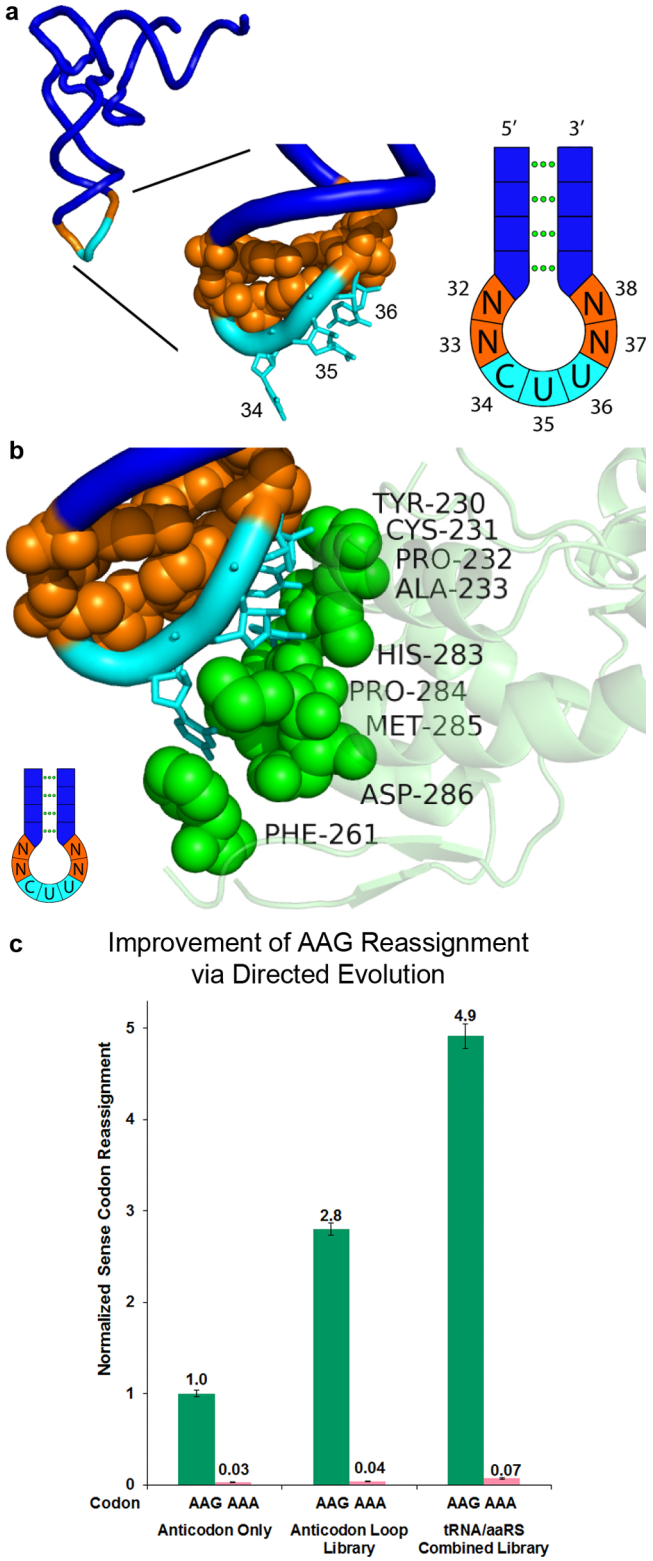


Figure 2. Library design and improvement of sense codon reassignment at the Lys AAG codon. (a) Enlargement of the *M. jannaschii* tyrosyl-tRNA anticodon stem loop with anticodon nucleotides shown as sticks and the remaining nucleotides in the anticodon loop shown in space filling representation (PDB entry 1J1U). The two-dimensional cartoon representation is color-coded to match the backbone trace of the tRNA, and nucleotides in the anticodon loop are numbered. The four anticodon loop positions that flank the anticodon nucleotides were randomized in a 256-member tRNA anticodon loop library. (b) Crystal structure of *M. jannaschii* tyrosyl-aaRS complexed with the *M. jannaschii*

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355–7364

Figure 2. continued

tyrosyl-tRNA (PDB entry 1J1U) with positions targeted for diversity in the combined tRNA/aaRS library shown in space filling representation. We have altered the anticodon of the tRNA to CUU. Our combined tRNA/aaRS library for Lys reassignment includes diversity at the four positions in the tRNA anticodon loop as well as the amino acids on the aaRS that are within 5 Å of the anticodon. (c) Increase in the efficiency of sense codon reassignment for the best tRNA identified from our initial 256-member tRNA library (5′ CU CUU AU 3′) as well as an improved orthogonal pair variant identified via a FACS sort of the combined tRNA/aaRS library (clone F7 in Figure 3b, tRNA 5′ CU CUU AC 3′ and aaRS H283L, P284C, M285C, D286P). Fluorescence data for the alternative sense codon specifying the replaced amino acid (i.e., wobble base pairing to the orthogonal tRNA) are also shown to demonstrate retention of codon discrimination by improved variants.

prepared a small library of tRNA anticodon loop variants of the lysine AAG codon reassigning *M. jannaschii* tRNA (CUU anticodon). All of the nucleotides in the anticodon loop flanking the anticodon (nucleotides 32, 33, 37, and 38) were varied to produce a library of 256 variants (Figure 2a). Similar tRNA libraries have been employed to improve the efficiency and orthogonality of amber codon reassigning variants with different levels of success.[43−45] The tRNA library was transformed into cells containing the reporter plasmid with a lysine AAG codon specifying the fluorophore Tyr position. Fluorescent clones from the transformed mini-library were grown in 96-well plates, and their fluorescence was measured. The most fluorescent clone analyzed from this library showed a surprisingly large improvement in sense codon reassignment efficiency of 2.8-fold over the starting tRNA (Figure 2c). The improved tRNA maintained the high level of discrimination between the AAG and AAA codons. The selected variant had a single modification at position 38 from A to U and maintained the G 37 A modification previously identified upon optimization of the *M. jannaschii* tRNA for amber suppression.[37,44]

To evaluate co-evolution of the orthogonal tRNA and aminoacyl tRNA synthetase, we generated a combined *M. jannaschii* tyrosyl-aaRS tRNA binding domain/tRNA library (Figure 2b). The combined library included the 256-member tRNA anticodon loop library and diversified nine amino acid residues that are within 5 Å of the tRNA anticodon in the cocrystal structure of the *M. jannaschii* tRNA/aaRS complex (PDB entry 1J1U).[29] The library employed restricted codon sets at the nine varied positions to keep the total library theoretical diversity to 10^{10}. The library was prepared via Kunkel mutagenesis; the unmutated template was removed by restriction enzyme digestion, and the library was transformed into cells harboring the GFP reporter plasmid with AAG specifying the fluorophore tyrosine position. Restriction analysis of randomly picked clones after digestion and transformation revealed minimal carry-through of unmutated starting template DNA: 10 of 10 clones were full library members. The library was sorted via FACS; the distribution of fluorescence observed for the initial sort of 7×10^7 cells included a long tail of many highly fluorescent variants (Figure 3a). The top 0.5% of cells were collected, amplified, and resorted. The cell population for the second sort showed a large increase in average fluorescence. The top 3% of these cells were collected. Randomly picked isolated clones were grown in 96-well plates, and their fluorescence was quantified (Figure 3b). The majority of the analyzed clones (17 of 33, 52%) were more fluorescent than the positive control cells from the mini-library selected modified tRNA with unmodified aaRS. The most fluorescent clone from the analyzed variants (clone F7) appeared to be nearly twice as fluorescent (Figure 3b). The liquid culture of clone F7 was replated for further analysis. Seven replicate colonies of clone F7 were analyzed, and the reassignment efficiency improved 4.9-fold relative to that of the starting tRNA/aaRS pair, which included only an anticodon modification (Figure 2c). In both the tRNA mini-library hit and in clone F7, position 38 in the tRNA changed from the original A, but the two variants did not share the same mutation; position 38 of the tRNA from the mini-library is U and from clone F7 is C. The aaRS from the clone F7 pair included changes to four of nine amino acids. Amino acids at positions 230−233 in the aaRS are closest to position 36 of the tRNA. This region of the library was represented by minimal diversity and contained only 256 protein sequence variants. The fact that none of the amino acids at these positions changed from those of wild-type *M. jannaschii* aaRS is not surprising. Greater diversity was allowed at position 261 (15 amino acids + 1 stop), but a mutation from the wild-type Phe was not present in clone F7. The majority of diversity in the library was available for amino acids 283−286, and clone F7 does include mutations at these positions. The His-Pro-Met-Asp in

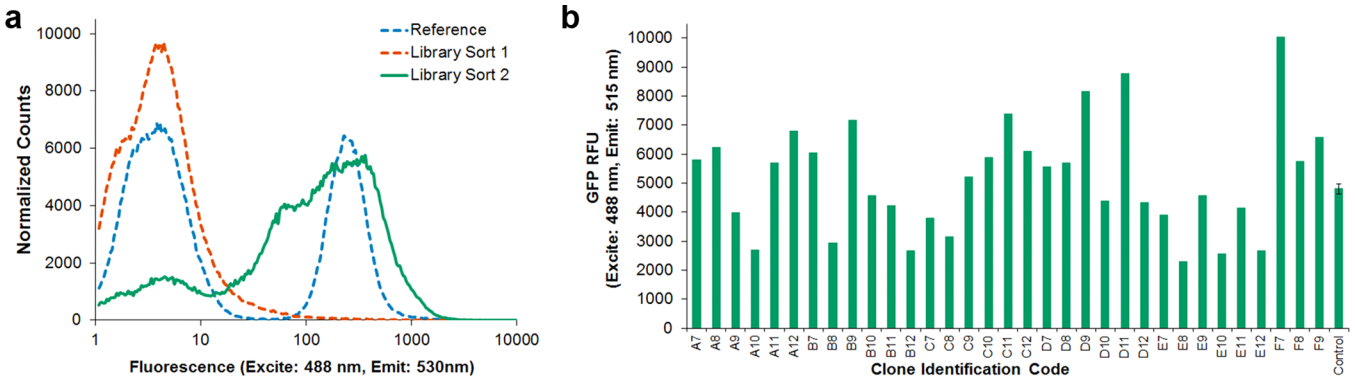


Figure 3. (a) Distribution of the fluorescence of sorted cells with counts normalized to 500000 to show multiple sorts on the same axes. Reference cells express the tRNA variant that was identified in the 256-member library. In our first sort of 6.5×10^7 cells from the combined tRNA/aaRS library, we collected all cells that showed fluorescence greater than 100 fluorescence units (library sort 1 trace). We amplified the fraction collected in sort 1 and resorted (library sort 2 trace). Cells in the second sort with fluorescence greater than that displayed by our reference cells were collected for further analysis. (b) Relative fluorescence (RFU) for 33 randomly picked clones after plating FACS sort 2. Data for the tRNA variant identified from the mini-library (control) are shown for comparison. Clone F7 was selected for further characterization (see also Figure 2c).

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

wild-type aaRS changed to a Leu-Cys-Cys-Pro. These mutations suggest backbone rearrangements in the structure. Movement of the proline from position 284 to 286 suggests a shortening of an α-helix.

The fact that the tRNA anticodon loop and aaRS evolve together to improve sense codon reassignment highlights the fact that the effects of changing the aaRS and the tRNA are not independent. Using the GFP-based screen for directed evolution of the orthogonal tRNA and aminoacyl tRNA synthetase together allows identification of improved pairs that balance many factors, including discrimination, efficiency, and orthogonality. The extent to which tRNA and aaRS modifications that affect the efficiency of sense codon reassignment can be transferred between aaRSs evolved for different non-canonical amino acids is difficult to assess. Attempts to identify tRNA modifications that improve amber suppression efficiency by modulating interactions with Ef-Tu identified tRNAs that produced generally improved suppression for multiple aaRSs evolved to incorporate different ncAAs. However, some tRNA modifications appeared to be beneficial to only one aaRS.[44] Additional studies looking at the transferability of modifications to the aaRS anticodon binding domain suggest that transferability is related to the similarity between the ncAA sizes.[45] Regardless of the transferability, measurement and improvement of tyrosine-incorporating orthogonal aaRS variants provides a probable high-end estimate for the efficiency of ncAA incorporation that can be attained at a particular sense codon and will be useful as starting points for future studies employing non-canonical amino acid-incorporating aaRSs. Our expectation is that each orthogonal tRNA/aaRS pair will have a different set of idiosyncratic cross reactions with the translational system into which it is transplanted.

Comparison to Reported Levels of Missense Incorporation. The extent of sense codon reassignment obtained with the evolved lysine AAG-decoding tRNA/aaRS pair (clone F7) is among the highest reported to date in rich medium. Döring and Marlière described an editing-deficient mutant of the valine aaRS that allows incorporation of aminobutyric acid in response to valine codons at a level of 24%.[32] Min and Söll reported the introduction of a non-discriminating archeal aspartyl-tRNA synthetase into *E. coli*. This change resulted in the aminoacylation of the aspartyl-tRNA with both Asp and Asn.[33] The Asp for Asn substitutions of 16−38% were detected using a screen that monitored restoration of enzyme function as a result of missense incorporation. Ruan and Söll expanded on the work of Min and evaluated four different missense substitutions (Cys for Pro, Glu for Gln, Ser for Thr, and Asp for Asn).[34] For these systems, levels of replacement are between 1.8 and 28%. Together, these reports describe the highest levels of missense incorporation observed in cells without controlling the medium composition or employing auxotrophic strains.

Effect of Sense Codon Reassignment on Cell Health. A possible complicating factor in our screening and selection strategy is the effect of sense codon reassignment on cell health. Of the four codons for which data are presented in this manuscript, cell growth is significantly negatively impacted only when either the directed evolution-improved machinery for reassignment of lysine AAG codons (clone F7) or the machinery to reassign His codons is expressed (Figures S4 and S5). Only minor effects on the growth of cells are observed for all other reassignment systems. In general, cell growth rates are slowed slightly relative to control systems in which no reassignment takes place. Cell growth rate decreases are within the ranges of the wild-type GFP-producing control systems for all reassignment systems, except His and clone F7 for Lys AAG. The standard deviation in the calculated exponential growth rates is approximately 5−7.5% of the measured values. The final optical densities observed in culture are slightly reduced for the majority of systems in which reassignment takes place. Final optical densities for cells expressing either the directed evolution-improved machinery for Lys AAG codons (clone F7) or histidine reassignment decreased further. This observation may suggest cumulative effects of incorporation of tyrosine at histidine or lysine codons throughout the *E. coli* genome.

Our results are consistent with several reports that measure the tolerance of bacterial cells toward missense incorporation. In the series of reports describing directed attempts to induce high levels of missense incorporation in *E. coli*, only minor effects on growth were noted. Ruan and Söll specifically evaluated growth effects on *E. coli* cells with normal protein quality control machinery and missense incorporation systems that led to as much as 28% missense incorporation. No effects on cell growth were observed. However, upon evaluation of the missense incorporation systems in *E. coli* cells with compromised protein quality control mechanisms, Ruan and Söll report pronounced growth inhibition effects when high levels of missense mutations are induced.

The lysine AAG codon encodes 22% of the lysine residues in the *E. coli* genome. Our observed levels of sense codon reassignment, and even nearly complete reassignment of the AAG codon, would be within the range of missense levels generally tolerated by bacteria. Approximately half of the codons read via wobble interactions are used in fewer than a third of the instances in which a given amino acid is specified in the *E. coli* genome. This subset of wobble codons represents target codons for ready reassignment. We expect that the cell growth effects will correlate with the relative use of the particular codon targeted for reassignment and the catalytic importance of the substituted amino acid (e.g., histidine). The exact extent to which either of these factors plays a role in observed effects on cellular health has not been established. Our screen could be readily applied to evaluate the relative importance of particular amino acids in living systems.

CONCLUSIONS

We developed a simple, high-throughput screen to measure and improve sense codon reassignment by the most commonly employed orthogonal tRNA/aaRS pair for non-canonical amino acid incorporation at amber stop codons. By changing only the anticodon of the orthogonal tRNA, we achieved sense codon reassignment efficiencies between 1% (Phe UUU) and 6% (Lys AAG). Each of the orthogonal tRNAs preferentially decoded the codon traditionally read via a wobble interaction in *E. coli* with the exception of the orthogonal tRNA with an AUG anticodon, which incorporated tyrosine in response to both the His CAU and His CAC codons with approximately equal frequency. We applied our screen in a high-throughput manner to evaluate a 10^9-member combined tRNA/aminoacyl tRNA synthetase library to identify improved sense codon reassigning variants for the Lys AAG codon. We believe that a single screen with the ability to broadly evaluate sense codons and the available set of orthogonal tRNA/aaRS pairs will facilitate quick identification of the combinations that can be employed to efficiently reassign the meaning of sense codons.

Producing proteins using expanded genetic codes combines exquisite positional control over sites of modification with a large

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

toolbox of encodable functionalities. Unlike the amber stop codon suppression methodology, sense codon reassignment has the potential to expand the genetic code further by allowing the reassignment of more than one codon in the same system. With the present suite of nearly 100 aaRSs evolved to incorporate ncAAs, there are 100 possible 21 amino acid genetic codes. We are interested in employing the set of previously evolved aminoacyl tRNA synthetase variants to incorporate multiple copies of non-canonical amino acids in response to sense codons in *E. coli*. We began our evaluation of sense codon reassignment in *E. coli* using the *M. jannaschii* tyrosyl-tRNA/aaRS pair, but the fluorescence-based screen will be applicable to other organisms and to other orthogonal pairs that can be evolved to incorporate tyrosine. Differences in the complements of tRNAs, codon usage, and vagrancies in determinants of aminoacylation between organisms may require targeting alternative sets of sense codons across different organisms. At present, three well-studied, mutually orthogonal sets of tRNA/aaRSs can be employed for stop and four-base codon-directed ncAA incorporation (two orthogonal pairs derived from the *M. jannaschii* pair and the *M. barkeri* pyrrolysyl-tRNA/aaRS pair).[46−48] The *M. barkeri* pyrrolysyl-tRNA/aaRS pair is evolutionarily closely related to phenylalanine aaRSs and has been evolved to introduce para-substituted phenyl ring-containing amino acids. Several other pairs have been described but not widely employed.[11,49−52] Extending the sense codon reassignment strategy to exploit these ncAA-incorporating orthogonal pairs is a step toward generating expanded genetic codes where multiple copies of multiple ncAAs can be incorporated into proteins simultaneously. The ability to incorporate more than one type of ncAA in the same protein expands the genetic code exponentially to 10000 possible 22 amino acid genetic codes. Beyond the utility toward expanded genetic codes, the screening system is sensitive enough to measure natural rates of tyrosine missense incorporation, and evolved tyrosine-incorporating aaRSs will allow the evaluation of the biological effects of high-level directed amino acid substitutions at various codons of interest.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00870.

General methods, reagent details and sources, DNA manipulation and mutagenesis details, and vector and oligonucleotide primer sequences (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO 80523. E-mail: nickfisk@colostate.edu. Phone: (970) 491-4115.

Funding

This work was supported by funds from Colorado State University and a Boettcher Foundation Webb-Waring Biomedical Research Award (to J.D.F.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are especially grateful to Leslie Armstrong-Lea and the Colorado State University Proteomics and Metabolomics Facility (PMF) for access to and expertise with FACS. We thank Erin E. Carlson, Rebecca E. Connor, and Alan J. Kennan for helpful comments on the manuscript. We thank Andrew Bradbury for a generous gift of plasmids. We thank Jay Keasling for pBbB1a-GFP (Addgene plasmid no. 35340), Nina Papavasiliou for pCDFDuet-1::ADAT3-ADAT2 (Addgene plasmid no. 15917), and Peter Schultz for pEVOL-pAzF (Addgene plasmid no. 31186) and pEVOL-pBpF (Addgene plasmid no. 31190). DNA sequencing was performed in the PMF at Colorado State University.

■ ABBREVIATIONS

aaRS, aminoacyl tRNA synthetase; BSA, bovine serum albumin; FACS, fluorescence-activated cell sorting; GFP, green fluorescent protein; IPTG, isopropyl β-D-1-thiogalactopyranoside; ncAA, non-canonical amino acid.

■ REFERENCES

(1) Harold, F. M. (1986) *The vital force a study of bioenergetics*, W. H. Freeman, New York.

(2) Russell, J. B., and Cook, G. M. (1995) Energetics of bacterial-growth - balance of anabolic and catabolic reactions. *Microbiol. Rev. 59*, 48−62.

(3) Stouthamer, A. H. (1973) A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van Leeuwenhoek 39*, 545−565.

(4) Agris, P. F. (2004) Decoding the genome: a modified view. *Nucleic Acids Res. 32*, 223−238.

(5) Koonin, E. V., and Novozhilov, A. S. (2009) Origin and evolution of the genetic code: the universal enigma. *IUBMB Life 61*, 99−111.

(6) Plotkin, J. B., and Kudla, G. (2011) Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet. 12*, 32−42.

(7) Knight, R. D., Freeland, S. J., and Landweber, L. F. (2001) Rewiring the keyboard evolvability of the genetic code. *Nat. Rev. Genet. 2*, 49−58.

(8) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. In *Annual Review of Biochemistry* (Kornberg, R. D., Raetz, C. R. H., Rothman, J. E., and Thorner, J. W., Eds.) Vol. 79, pp 413−444, Annual Reviews, Palo Alto, CA.

(9) Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins. *Science 244*, 182−188.

(10) Bain, J. D., Glabe, C. G., Dix, T. A., Chamberlin, A. R., and Diala, E. S. (1989) Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *J. Am. Chem. Soc. 111*, 8013−8014.

(11) Furter, R. (1998) Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in *Escherichia coli*. *Protein Sci. 7*, 419−426.

(12) Mukai, T., Hayashi, A., Iraha, F., Sato, A., Ohtake, K., Yokoyama, S., and Sakamoto, K. (2010) Codon reassignment in the *Escherichia coli* genetic code. *Nucleic Acids Res. 38*, 8188−8195.

(13) Johnson, D. B. F., Xu, J. F., Shen, Z. X., Takimoto, J. K., Schultz, M. D., Schmitz, R. J., Xiang, Z., Ecker, J. R., Briggs, S. P., and Wang, L. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol. 7*, 779−786.

(14) Isaacs, F. J., Carr, P. A., Wang, H. H., Lajoie, M. J., Sterling, B., Kraal, L., Tolonen, A. C., Gianoulis, T. A., Goodman, D. B., Reppas, N. B., Emig, C. J., Bang, D., Hwang, S. J., Jewett, M. C., Jacobson, J. M., and Church, G. M. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science 333*, 348−353.

(15) Dougherty, M. J., Kothakota, S., Mason, T. L., Tirrell, D. A., and Fournier, M. J. (1993) Synthesis of a genetically engineered repetitive polypeptide containing periodic selenomethionine residues. *Macromolecules 26*, 1779−1781.

(16) Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid,

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364

selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur. J. Biochem. 230*, 788−796.
(17) Budisa, N. (2006) *Engineering the genetic code expanding the amino acid repertoire for the design of novel proteins*, Wiley-VCH, Weinheim, Germany.
(18) Kwon, I., Kirshenbaum, K., and Tirrell, D. A. (2003) Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc. 125*, 7512−7513.
(19) Kwon, I., and Lim, S. I. (2015) Tailoring the substrate specificity of yeast phenylalanyl-tRNA synthetase toward a phenylalanine analog using multiple-site-specific incorporation. *ACS Synth. Biol. 4*, 634−643.
(20) Drummond, D. A., and Wilke, C. O. (2008) Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell 134*, 341−352.
(21) Schimmel, P., and Guo, M. (2009) A tipping point for mistranslation and disease. *Nat. Struct. Mol. Biol. 16*, 348−349.
(22) Kirchner, S., and Ignatova, Z. (2014) Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat. Rev. Genet. 16*, 98−112.
(23) Crick, F. H. C. (1966) Codon-Anticodon Pairing - Wobble Hypothesis. *J. Mol. Biol. 19*, 548−555.
(24) Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res. 26*, 148−153.
(25) Meroueh, M., and Chow, C. S. (1999) Thermodynamics of RNA hairpins containing single internal mismatches. *Nucleic Acids Res. 27*, 1118−1125.
(26) Giege, R., Sissler, M., and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res. 26*, 5017−5035.
(27) Sidhu, S. S., and Weiss, G. A. (2004) Phage display a practical approach. In *Practical approach series* (Clackson, T., and Lowman, H. B., Eds.) Oxford University Press, Oxford, U.K.
(28) Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U. S. A. 82*, 488−492.
(29) Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol. 10*, 425−432.
(30) *The PyMOL Molecular Graphics System*, version 1.3r1 (2010) Schrödinger, LLC, Portland, OR.
(31) Heim, R., and Tsien, R. Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol. 6*, 178−182.
(32) Döring, V., Mootz, H. D., Nangle, L. A., Hendrickson, T. L., de Crécy-Lagard, V., Schimmel, P., and Marlière, P. (2001) Enlarging the amino acid set of *Escherichia coli* by infiltration of the valine coding pathway. *Science 292*, 501−504.
(33) Min, B., Kitabatake, M., Polycarpo, C., Pelaschier, J., Raczniak, G., Ruan, B., Kobayashi, H., Namgoong, S., and Söll, D. (2003) Protein synthesis in *Escherichia coli* with mischarged tRNA. *J. Bacteriol. 185*, 3524−3526.
(34) Ruan, B. F., Palioura, S., Sabina, J., Marvin-Guy, L., Kochhar, S., LaRossa, R. A., and Söll, D. (2008) Quality control despite mistranslation caused by an ambiguous genetic code. *Proc. Natl. Acad. Sci. U. S. A. 105*, 16502−16507.
(35) Kramer, E. B., and Farabaugh, P. J. (2006) The frequency of translational misreading errors in *E. coli* is largely determined by tRNA competition. *RNA 13*, 87−96.
(36) Fechter, P., Rudinger-Thirion, J., Tukalo, M., and Giege, R. (2001) Major tyrosine identity determinants in *Methanococcus jannaschii* and *Saccharomyces cerevisiae* tRNA(Tyr) conserved but expressed differently. *Eur. J. Biochem. 268*, 761−767.
(37) Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol. 395*, 361−374.
(38) Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in *Escherichia coli*. *Biochemistry 52*, 1828−1837.
(39) Ryu, Y. H., and Schultz, P. G. (2006) Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods 3*, 263−265.
(40) Doring, V., and Marliere, P. (1998) Reassigning cysteine in the genetic code of *Escherichia coli*. *Genetics 150*, 543−551.
(41) Parker, J. (1989) Errors and alternatives in reading the universal genetic-code. *Microbiol. Rev. 53*, 273−298.
(42) Wolf, J., Gerber, A. P., and Keller, W. (2002) tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J. 21*, 3841−3851.
(43) Wang, L., and Schultz, P. G. (2001) A general approach for the generation of orthogonal tRNAs. *Chem. Biol. 8*, 883−890.
(44) Guo, J. T., Melancon, 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.
(45) Wang, N. X., Ju, T., Niu, W., and Guo, J. T. (2015) Fine-tuning interaction between aminoacyl-tRNA synthetase and tRNA for efficient synthesis of proteins containing unnatural amino acids. *ACS Synth. Biol. 4*, 207−212.
(46) Neumann, H., Slusarczyk, A. L., and Chin, J. W. (2010) *De novo* generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J. Am. Chem. Soc. 132*, 2142−2144.
(47) Blight, S. K., Larue, R. C., Mahapatra, A., Longstaff, D. G., Chang, E., Zhao, G., Kang, P. T., Green-Church, K. B., Chan, M. K., and Krzycki, J. A. (2004) Direct charging of tRNA(CUA) with pyrrolysine *in vitro* and *in vivo*. *Nature 431*, 333−335.
(48) Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*. *Science 292*, 498−500.
(49) Anderson, J. C., and Schultz, P. G. (2003) Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry 42*, 9598−9608.
(50) Santoro, S. W., Anderson, J. C., Lakshman, V., and Schultz, P. G. (2003) An archaebacteria-derived glutamyl-tRNA synthetase and tRNA pair for unnatural amino acid mutagenesis of proteins in *Escherichia coli*. *Nucleic Acids Res. 31*, 6700−6709.
(51) Hughes, R. A., and Ellington, A. D. (2010) Rational design of an orthogonal tryptophanyl nonsense suppressor tRNA. *Nucleic Acids Res. 38*, 6813−6830.
(52) Chatterjee, A., Xiao, H., and Schultz, P. G. (2012) Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A. 109*, 14841−14846.

DOI: 10.1021/acs.biochem.5b00870
Biochemistry 2015, 54, 7355−7364