Incorporation of Modified Amino Acids by Engineered Elongation Factors with Expanded Substrate Capabilities

Vanessa E. DeLey Cox,[†] Megan F. Cole,[‡] and Eric A. Gaucher*,[§,∥]

[†]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States
[‡]Department of Biology, Emory University, Atlanta, Georgia 30322, United States
[§]School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332, United States
[∥]Department of Biology, Georgia State University, Atlanta, Georgia 30303, United States

S Supporting Information

ABSTRACT: Noncanonical amino acid (ncAA) incorporation has led to significant advances in protein science and engineering. Traditionally, *in vivo* incorporation of ncAAs is achieved *via* amber codon suppression using an engineered orthogonal aminoacyl-tRNA synthetase:tRNA pair. However, as more complex protein products are targeted, researchers are identifying additional barriers limiting the scope of currently available ncAA systems. One barrier is elongation factor Tu (EF-Tu), a protein responsible for proofreading aa-tRNAs, which substantially restricts ncAA scope by limiting ncaa-tRNA delivery to the ribosome. Researchers have responded by engineering ncAA-compatible EF-Tus

for key ncAAs. However, this approach fails to address the extent to which EF-Tu inhibits efficient ncAA incorporation. Here, we demonstrate an alternative strategy leveraging computational analysis to broaden EF-Tu's substrate specificity. Evolutionary analysis of EF-Tu and a naturally evolved specialized elongation factor, SelB, provide the opportunity to engineer EF-Tu by targeting amino acid residues that are associated with functional divergence between the two ancient paralogues. Employing amber codon suppression, in combination with mass spectrometry, we identified two EF-Tu variants with non-native substrate compatibility. Additionally, we present data showing these EF-Tu variants contribute to host organismal fitness, working cooperatively with components of native and engineered translation machinery. These results demonstrate the viability of our computational method and lend support to corresponding assumptions about molecular evolution. This work promotes enhanced polyspecific EF-Tu behavior as a viable strategy to expand ncAA scope and complements ongoing research emphasizing the importance of a comprehensive approach to further expand the genetic code.

KEYWORDS: *EF-Tu, noncanonical amino acid, orthogonal translation system, genetic code expansion, polyspecificity*
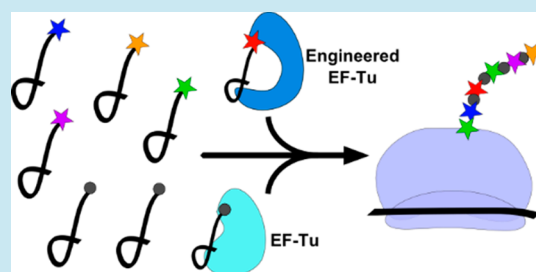
G enetic code expansion is a central goal of protein research and engineering with a broad range of applications. The ability to reliably incorporate noncanonical amino acids (ncAAs) in a site-specific manner has expanded the protein engineering toolbox to enable the functionalization of proteins with affinity, spectroscopic, and chemical tags.[1] Consequently, bio-orthogonal modification of proteins with ncAAs is a powerful and emerging tool critical to the development of both fundamental protein science and applied biotechnologies. The most common technique for the translation of proteins containing site-specific ncAA mutations is amber codon suppression.[2] This technique leverages an orthogonal translation system (OTS), consisting of a dedicated aminoacyl-tRNA synthetase (aaRS):tRNA pair, which mediates incorporation of a specific ncAA in the target protein at a repurposed amber codon.[3] In bacteria, ncAA incorporation is typically accomplished *via* OTSs developed from bio-orthogonal aaRS:tRNA pairs derived from *Methanocaldococcus jannaschii* or *Methanosarcina* species. Production yields of proteins containing ncAAs have been improved through

development of an engineered *E. coli* strain in which release factor 1 has been deleted and genomic amber codons have been replaced with the ochre stop codon, allowing reassignment of the amber codon to an ncAA.[4] Additional advances promoting ncAA incorporation include cell-free translation systems, optimized translation component concentrations, and genomic incorporation of OTSs.[5−7] However, while these advances have led to incorporation of some ncAAs at high yield, the routine application of the OTS strategy is consistently hindered by considerable and recurring barriers.[8]

Persistent challenges include cross-reactive OTSs, incompatibility with endogenous elongation factors, and discrimination by additional translation components. These factors affect both the yield and purity of ncAA-containing proteins as well as the fitness and viability of the host microorganism.[9−13] Furthermore, the diversity of these modifications is reduced to a specific set of ncAAs compatible with existing and engineered

translation machinery, thereby significantly reducing the readily available scope of potential chemistries and applications. These challenges highlight an immediate need to develop improved engineering strategies beyond OTS development that will enable translation of increasingly complex peptide products, with multisite incorporation of multiple ncAAs.[7,14]

One obstacle limiting expansion of the genetic code is elongation factor Tu (EF-Tu), a guanosine triphosphatase (GTPase).[10,15−18] EF-Tu serves two functions in translation. While most commonly recognized for translocation of aa-tRNA complexes to the ribosome, it also plays a critical role in quality control by proofreading aa-tRNAs.[19] All 20 aa-tRNAs associate with EF-Tu having carefully tuned interactions that prevent misacylated tRNAs from being efficiently delivered to the ribosome for translation.[20] Similar to misacylated tRNAs, ncaa-tRNAs are non-native substrates and can be discriminated by EF-Tu, thus preventing their incorporation into a translated protein.[21] Past efforts have typically circumvented EF-Tu's editing mechanism by targeting ncAAs that are tolerated as substrates. For particularly intractable ncAAs, often with bulky or highly charged side chains, orthogonal EF-Tus have been developed.[15,16,22,23] However, these efforts fail to recognize EF-Tu's comprehensive effect on translation. Even OTSs that can mediate ncAA incorporation *via* wild-type EF-Tu benefit from an engineered EF-Tu.[15,24] As a result, engineering EF-Tu to accept an expanded set of ncaa-tRNA substrates represents a unique opportunity for expanding ncAA incorporation.

Within the framework of this strategy, there are two approaches to broaden EF-Tu's substrate acceptance. One is to knockout EF-Tu's proofreading capabilities and develop a variant that can accommodate additional ncAAs as well as the canonical 20. However, this method requires a trade-off between the degree of polyspecificity desired to translate ncAAs and the specificity required for host organism survival. Here, we present an alternative strategy: engineering a novel EF-Tu with broader ncAA compatibilities to be used in complement with native EF-Tu. This strategy parallels an evolved mechanism for cellular cotranslational incorporation of selenocysteine (Sec), the 21st proteinogenic amino acid, which uses a dedicated elongation factor, SelB, in concert with EF-Tu.[25]

Computational methods that exploit models of molecular evolution have been previously leveraged to develop enzymes with expanded substrate scope. These strategies are based on the concept that enzymes evolved specialized activity from generic activities, a theory that is supported by research demonstrating ancestral proteins exhibit broader substrate compatibility than their modern counterparts.[26,27] In order to apply these methods to engineering an EF-Tu with enhanced polyspecific substrate compatibility, we assume, on the basis of sequence similarity, that SelB and EF-Tu are paralogues. This, in turn, suggests EF-Tu and SelB share a common ancestor that exhibited greater substrate promiscuity than the modern proteins. Motivated by this theory, EF-Tu and SelB protein families were selected for computational analysis to identify sites involved in functional divergence between EF-Tu and SelB. This information was then utilized to engineer substrate-promiscuous EF-Tus.

Herein, we describe our efforts to transform the manner in which EF-Tu is utilized to incorporate ncAAs. Leveraging an evolutionary-based method, reconstructing evolutionary adaptive paths (REAP), we engineered EF-Tu variants to better accommodate three non-native substrates. By mass spectrometry, we demonstrate two variants, from a collection of eight, have expanded substrate capabilities. By monitoring cell culture density, we also show these EF-Tu variants support host organism fitness. These results lend credence to our choice of evolutionary-based method and also suggest that EF-Tu and SelB had a common ancestor with expanded substrate polyspecificity. We discuss how this approach complements current research highlighting the advantages of improved OTSs and promotes a more comprehensive approach critical to achieving future goals that expand the genetic code.

RESULTS AND DISCUSSION

1. Computational Approach to Protein Engineering.

REAP has been previously employed to guide development of enzyme libraries with expanded substrate acceptance.[28,29] In brief, this method employs inferred evolutionary mutation rates of amino acid positions to predict which amino acid replacements are most likely to impart novel protein activity (Figure 1).[30,31] REAP analysis is based on the assumption that
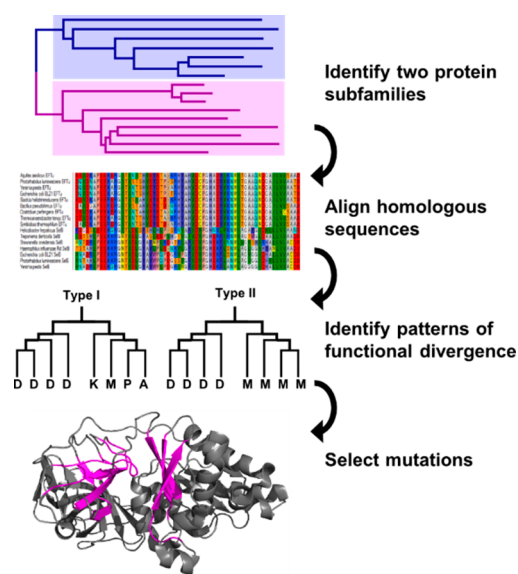

Figure 1. General schematic illustrating REAP methodology. This scheme shows the comparison of two clades highlighted in blue and pink. Homologous sequences from each clade are aligned and analyzed computationally to identify Type I and Type II functional divergence. Results can be used to estimate the probability that a mutation will affect protein activity, leading to development of a functionally diverse protein library.

amino acids that impact function are conserved during the evolution of a protein family and the corresponding assumption that residues lacking conservation are likely not correlated to activity or stability. REAP functions by ranking residues according to their degree of conservation in one lineage compared to the degree of conservation in another lineage. Amino acid sites with low inferred replacement rates are predicted to have a high correlation to function and are thus targeted during library design. Correspondingly, sites with high replacements rates are predicted to have minimal influence on protein behaviors and are excluded from library design. A central tenet of this method is that a REAP-developed library can enrich the functional diversity of a library while reducing the number of variants required for testing.

amino acid position	63	64	65	66	67	68	78	214	215	216	230	259	261	262	263	272	273	275	276
EF-Tu																			
Escherichia coli BL21	N	T	S	H	V	E	H	I	E	D	R	E	F	R	K	E	N	G	V
Aquifex aeolicus	N	I	T	H	V	E	H	I	E	D	R	E	F	R	K	D	N	G	V
Bacillus pseudofirmus	S	T	A	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	A
Burkholderia cenocepacia	N	T	A	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	I
Clostridium perfringens	N	T	A	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	A
Dechloromonas aromatica	N	T	A	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	A
Geobacter sulfurreducens	A	T	S	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	A
Haemophilus influenza Rd	N	T	S	H	V	E	H	I	E	D	R	E	F	R	K	E	N	G	A
Helicobacter hepaticus	A	T	S	H	I	E	H	V	E	D	R	E	F	R	K	D	N	G	I
Photorhabdus luminescens	S	T	S	H	V	E	H	I	E	D	R	E	F	R	K	E	N	G	V
Pseudomonas aeruginosa	N	T	S	H	V	E	H	I	E	D	R	E	F	R	K	E	N	G	I
Ralstonia eutropha	N	T	A	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	L
Shewanella oneidensis	N	T	S	H	I	E	H	I	E	D	R	E	F	R	K	E	N	G	I
Sinorhizobium meliloti	S	T	A	H	V	E	H	I	E	D	R	E	F	R	K	D	N	G	A
Slackia heliotrinireducens	S	I	A	H	I	E	H	V	E	D	R	E	F	R	K	D	N	G	C
Symbiobac thermophilum	N	T	S	H	V	E	H	I	E	D	R	E	F	R	K	D	N	G	A
Thermoanaerobacter tengc	N	T	T	H	V	E	H	V	E	D	R	E	F	R	K	D	N	G	V
Treponema denticola	N	T	R	H	L	E	H	I	E	D	R	E	F	N	K	D	N	G	L
Yersinia pestis	N	T	S	H	V	E	H	I	E	D	R	E	F	R	K	E	N	G	V
SelB																			
Escherichia coli BL21	D	L	G	Y	A	Y	F	I	D	R	T	H	Q	N	Q	Q	R	A	L
Aquifex aeolicus	D	I	G	F	A	Y	I	V	D	S	S	Q	H	G	V	E	R	A	L
Bacillus pseudofirmus	E	L	G	Y	A	P	I	I	D	Q	T	Q	H	H	Q	Q	R	A	I
Burkholderia cenocepacia	E	L	G	Y	A	Y	L	V	D	R	T	H	Q	N	R	E	R	A	L
Clostridium perfringens	N	L	G	F	T	F	I	V	D	R	T	Q	H	D	E	Q	R	A	L
Dechloromonas aromatica	D	L	G	Y	A	Y	F	V	D	R	T	R	Q	D	A	Q	R	A	L
Geobacter sulfurreducens	E	L	G	F	A	H	I	V	D	R	T	Q	H	G	R	Q	R	A	V
Haemophilus influenza Rd	D	L	G	Y	A	Y	F	I	D	R	T	H	Q	N	T	Q	R	A	L
Helicobacter hepaticus	D	L	S	F	S	H	F	I	D	R	S	Q	H	D	K	H	R	A	L
Photorhabdus luminescens	D	L	G	Y	A	Y	F	I	D	R	T	H	Q	N	Q	Q	R	A	L
Pseudomonas aeruginosa	D	L	G	Y	L	Y	F	V	D	R	T	H	Q	N	R	Q	R	A	L
Ralstonia eutropha	E	L	G	Y	A	Y	L	V	D	R	T	H	Q	N	R	Q	R	A	L
Shewanella oneidensis	D	L	G	Y	A	F	F	L	D	R	T	H	Q	G	M	T	R	A	L
Sinorhizobium meliloti	D	L	G	F	A	Y	F	V	D	R	T	H	Q	N	Q	Q	R	A	L
Slackia heliotrinireducens	E	L	G	F	A	R	V	V	D	R	T	Q	H	D	K	N	R	A	L
Symbiobac thermophilum	D	I	G	F	A	R	V	I	D	R	T	Q	H	G	E	Q	R	A	V
Thermoanaerobacter tengc	D	L	G	F	A	Y	I	V	D	R	T	Q	H	E	R	Q	R	A	I
Treponema denticola	E	L	G	F	A	S	I	V	D	R	T	Q	H	H	K	T	R	A	L
Yersinia pestis	D	L	G	Y	A	Y	F	I	D	R	T	H	Q	N	Q	Q	R	A	L

Figure 2. Multiple sequence alignment of EF-Tu and SelB sequences. Sites selected *via* REAP analysis are shown. Type I (blue) and Type II (red) are color-indicated.
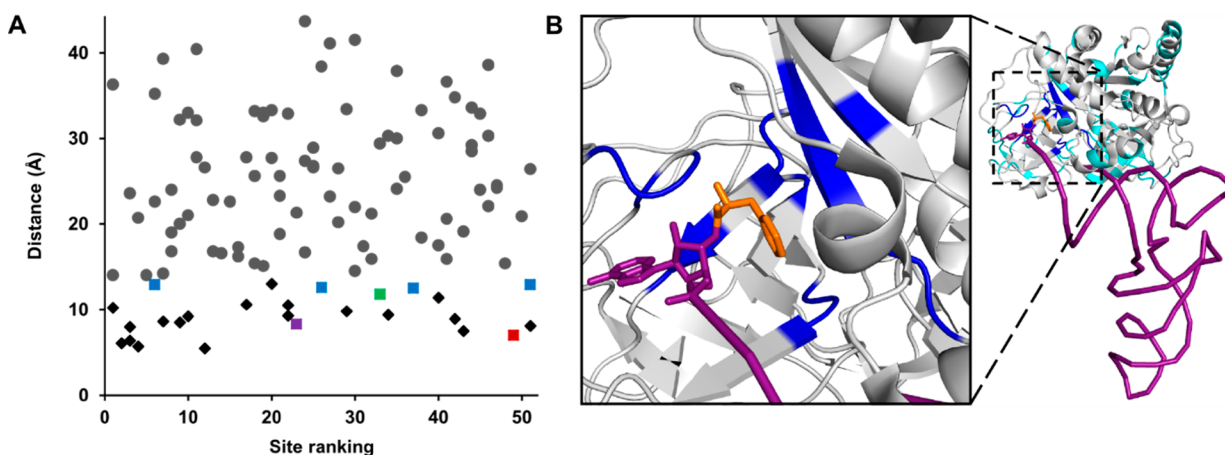

Figure 3. Selection of mutations in the REAP-designed EF-Tu library. (A) Plot shows residues identified by REAP. Ranking of position *versus* distance from target. Black diamonds denote positions selected for replacement. Gray circles indicate amino acid sites outside the 13 Å distance cutoff. Colored squares represent residues within the distance cutoff that were excluded from the library for various reasons: aliphatic residues (blue), alanine (purple), methionine (red), tyrosine (green). (B) Amino acids identified by REAP analysis highlighted on crystal structure of EF-Tu (gray) complexed with tRNA[Phe] (purple). Inset highlights sites mutated to generate EF-Tu library (blue). Residues not selected for library are also identified (cyan). Phenylalanine (orange) is situated in the amino acid binding pocket. Based on Protein Data Bank structure 1OB2.

Conserved amino acid sites are classified during REAP analysis as exhibiting either Type I or Type II functional divergence. Type I indicates an amino acid is conserved in only one lineage of a protein phylogeny.[32,33] This indicates the residue is critical for function in one protein family (where it is conserved), but not the other (in which the site is variable). Alternatively, amino acid sites exhibiting Type II functional divergence show conservation in both branches of the
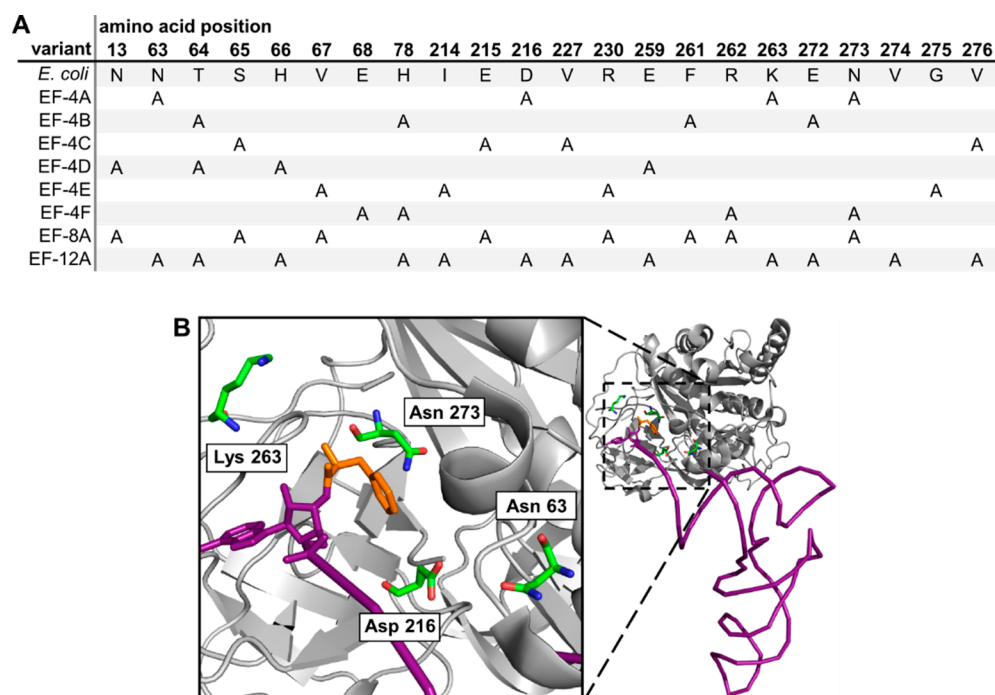
A

variant	13	63	64	65	66	67	68	78	214	215	216	227	230	259	261	262	263	272	273	274	275	276
E. coli	N	N	T	S	H	V	E	H	I	E	D	V	R	E	F	R	K	E	N	V	G	V
EF-4A		A									A						A		A			
EF-4B			A					A							A			A				
EF-4C					A				A			A										A
EF-4D	A	A		A									A									
EF-4E				A				A					A								A	
EF-4F						A	A									A			A			
EF-8A	A		A		A				A				A	A	A				A			
EF-12A		A	A		A				A		A	A	A		A	A		A		A		A

B

Figure 4. REAP-derived library variants. (A) Chart of mutations made to each EF-Tu variant. Sequence of wild-type *E. coli* EF-Tu is shown for reference. (B) EF-Tu (gray) with amino acid residues mutated in variant EF-4A (inset). Protein is complexed with phenylalanine (orange) and tRNA[Phe] (purple).

phylogeny, although the amino acid identity at the conserved position differs between families.[34] This type of divergence suggests that while the amino acid position is important to protein activity in both families, its role in protein function may differ.

2. Selection of Relevant Amino Acid Residues. To design a small EF-Tu library, REAP analysis compared EF-Tu and SelB sequences. Examination of sequence similarity suggests that EF-Tu and SelB can be classified as functionally divergent homologues, making them appropriate protein families for a REAP application. EF-Tu and SelB sequences from 19 prokaryotic families were aligned and evaluated to identify amino acid positions predicted to influence substrate compatibility (Figure 2). The aligned sequences were analyzed *via* three computational models using DIVERGE software.[35] Two models, which employ different parameters for analysis, were used to identify Type I functional divergence.[32,33] Sites associated with Type II functional divergence were identified using a third model.[34] Residues were ranked according to their posterior probability (Type I) or posterior ratio (Type II) producing a rank-ordered list of amino acid positions, with the top-ranked sites being predicted to have a greater influence on activity (Table S1).

A preliminary list of targeted amino acid sites was produced by parsing the top-ranked residues according to their distance from the target substrate (Figure 3A).[36] Because REAP identifies residues based on conservation rates, a metric influenced by many factors, the list of REAP-inferred sites was refined *via* distance discrimination, which has been previously used to engineer EF-Tu variants. Distances were calculated using the Cγ of the binding target amino acid and the Cα of the EF-Tu residue. Residues exceeding 13 Å were removed from the list leaving 26 predicted positions in close proximity to the target. Of the 26 sites, 7 residues were excluded, thereby culling the final list to 19 residues (Figure

3B). Residues omitted from the library included aliphatic residues between 12 and 13 Å since they were not expected to have a significant effect on substrate acceptance, and an alanine residue that was not eligible for mutation in an alanine-scanning library. Methionine, having the highest entropy rotamer of the amino acid side chains, was also excluded. Lastly, although the Cα of Y76 was within 13 Å, the Cγ of the side chain fell outside the distance cutoff, suggesting that substitution of alanine would not impact target specificity.

To gauge REAP's ability to identify relevant residues, we also selected three decoy positions. These positions were chosen on the basis of visual analysis of either a multiple sequence alignment or EF-Tu's crystal structure. Site N13 was chosen on the basis of its conservation in EF-Tu proteins. This position was excluded from the library because protein lengths were normalized for analysis (see Materials and Methods). In addition, two positions, V227 and V274, were selected on the basis of EF-Tu's crystal structure and their proximity to the target at 9.1 and 5.8 Å, respectively. Since distance discrimination is the prevailing strategy used to select mutation sites in EF-Tu, these positions were deemed likely candidates for mutation and were incorporated into the library.

Alanine scanning was employed to assess the functional implications of each position selected *via* REAP. Definitive evaluation of EF-Tu mediated incorporation of non-native substrates would require target protein purification *via* affinity chromatography and confirmation *via* mass spectrometry, a low-throughput, high-content workflow. The total library size was reduced by grouping alanine replacements in combinations of 4, 8, or 12 mutations since the REAP-derived library contained 22 targeted positions, a larger number of amino acid positions than previous efforts (Figure 4A). By generating a small, targeted library from the computational analysis, this comprehensive strategy for EF-Tu variant analysis was feasible

for all variants that merited further investigation, even the entire library.

3. ncAA-Compatible EF-Tu Variants. To characterize the EF-Tu library, we used an amber codon suppression assay requiring the cotranslational insertion of an ncAA at an in-frame amber codon.[16] The target gene, chloramphenicol acetyltransferase (CAT), contained an amber mutation at the permissive D112 position. CAT confers antibiotic resistance to *E. coli* resulting in an assay that directly correlates ncAA incorporation with cellular survival reported as half the maximal inhibitory concentration (IC_{50}). Rates of survival above wild-type EF-Tu (EF-coli) indicate the REAP-engineered EF-Tu variant can facilitate incorporation of the ncAA with greater efficiency than EF-coli.

O-Phospho-L-serine (Sep) was a strong ncAA candidate for our system, because it had been previously identified as an ncAA that benefits from an engineered EF-Tu.[16] Different strategies were employed to overcome this barrier to Sep incorporation, but even OTSs that were somewhat compatible with wild-type EF-Tu showed improved yields when paired with an engineered EF-Tu.[24,37] One effort to incorporate Sep developed an orthogonal triplet consisting of tRNA[Sep], SepRS, and EF-Sep to enable cotranslational insertion of Sep.[16] This engineered triplet provided a platform for assessing the substrate compatibility of our modified EF-Tus. Our EF-Tu variants were assayed in combination with the Sep-OTS, specifically tRNA[Sep] and SepRS.

Of the REAP-designed EF-Tu variants, variant EF-4A (N63A/D216A/K263A/N273A) resulted in the highest IC_{50} values as determined by the CAT translation assay. Variant EF-4C conferred survivability similar to EF-coli with other variants presenting substantially lower IC_{50} values (Table S2). To deconvolute the contribution of the four point mutations comprising EF-4A, single-mutation variants were assayed (EF-N63A, EF-D216A, EF-K263A, and EF-N273A) (Figure 4B). Of these variants, EF-D216A showed improved survivability relative to both EF-coli and the quadruple mutant EF-4A (Figure 5). IC_{50} values associated with variants EF-N63A and EF-N273A were not statistically distinguishable from EF-coli. Variant EF-K263A presented IC_{50} values below wild type (Table S2).

However, host organism survival conferred by CAT expression does not exclusively require incorporation of Sep at the amber mutation. Rather, bacteria survival could be a result of EF-Tu mediated incorporation of any available aa-tRNA that can pair with the amber codon. To identify this mechanism of survival, the CAT expression assay was performed withholding either tRNA[Sep] or SepRS. In the event that a misacylated tRNA was incorporated at the amber mutation, IC_{50} values would remain unchanged when SepRS was withheld. Conversely, an endogenous tRNA mispairing with the amber codon would be indicated by unchanged IC_{50} values when tRNA[Sep] was withheld. Analysis of experiments lacking tRNA[Sep] showed no host survival, thereby confirming that endogenous tRNAs are not capable of mispairing with the amber codon. Complementary experiments withholding SepRS showed largely unchanged IC_{50} values, indicating tRNA[Sep] is cross-reactive with endogenous aaRSs (Figure 5). EF-4A and EF-N273A show improved host survival when SepRS was withheld, suggesting these variants might have greater aptitude with a misacylated tRNA. In contrast, EF-D216A shows somewhat decreased host survival, suggesting perhaps it may have greater compatibility with the ncaa-tRNA. The compat-
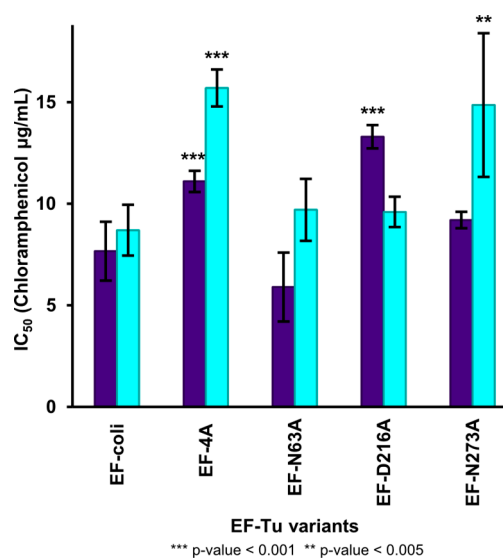
Figure 5. Characterization of EF-4A and single-mutation EF-Tu variants. *In vivo* suppression *via* EF-Tu variants with Sep-OTS (dark purple) or without SepRS (cyan) as measured by synthesis of CAT (quantified by IC_{50} value). Data shown represent triplicate averages except for EF-coli (Sep-OTS), EF-N63A (Sep-OTS), and EF-N273A (no SepRS), which show data from five replicates. EF-4A assayed without SepRS shows data from four replicates. All error bars represent standard deviation. *P*-values are relative to EF-coli.

ibility of EF-4A and EF-D216A with non-native substrates was further evaluated *via* electrospray ionization mass spectrometry (ESI-MS).

4. Mass Spectrometry Confirms Substrate-Promiscuous EF Variants. In order to determine the specific substrate compatibility of variants EF-4A and EF-D216A, CAT proteins expressed *via* these EF-Tu variants were purified, and ESI-MS was employed to investigate the breadth of amino acids incorporated at the permissive position. Analysis of CAT proteins translated *via* EF-4A and EF-D216A showed peaks consistent with incorporation of both Sep and Ser at position 112, suggesting enhanced EF-Tu compatibility with non-native substrates, specifically ncaa-tRNAs and misacylated tRNAs (Figure 6). EF-D216A was additionally capable of mediating Gln incorporation at the amber codon, making it compatible with three non-native substrates. Since the aim of this effort

— S E Y H D X F R Q F L —

	amino acids incoporated at amber mutation		
EF-Tu variants	**Sep**	**Ser**	**Gln**
EF-4A	X	X	
EF-D216A	X	X	X
EF-Sep	X		

Figure 6. Mass spectrometry confirmed amino acids incorporated at amber mutation in CAT protein. Protein translation was mediated by EF-Tu variants listed. The relevant region of the CAT amino acid sequence is shown for reference with an "X" indicating the permissive position. A representative group of protein spectra matched is shown in Table S5.

was the expansion of EF-Tu's substrate scope, not the incorporation of a specific ncAA, it is not necessary to distinguish between types of non-native substrates. To eliminate post-translational dephosphorylation of Sep as a possible route to Ser incorporation, ESI-MS was used to analyze CAT protein expressed *via* EF-Sep mediated translation. If a post-translational modification were responsible, Ser112 incorporation would be evident in this CAT protein as well; however, only peaks consistent with Sep incorporation were evident. These data indicate that Ser incorporation was not the result of a post-translational dephosphorylated Sep. Combined with results from the CAT expression assay, these data are indicative of EF-4A and EF-D216A having expanded substrate compatibility with non-native substrates, specifically Sep-tRNA^{Sep}, Ser-tRNA^{Sep}, and Gln-tRNA^{Sep}.

This approach to EF-Tu engineering requires balancing the expanded polyspecificity desired for non-native substrate acceptance with the risk of inaccurate translation of the target gene. In this case, EF-Tu variants mediated expression of a mixed protein product, CAT proteins with Sep, Ser, Gln at position 112. While this may initially seem problematic, we argue this challenge is readily overcome by improvements to synthetase and tRNA engineering. Our data align with prior work that shows this particular orthogonal tRNA (o-tRNA), tRNA^{Sep}, is cross-compatible with endogenous aaRSs, indicating that misincorporation, while permitted by an engineered EF-Tu, is actually caused by a cross-reactive OTS.[16] While a substrate-specific EF-Tu variant can prevent misincorporation, this strategy merely shifts the burden of accurate translation from the aaRS:tRNA pair to the EF-Tu and fails to address the cross-compatible OTS as the underlying cause.

Since cross-reactive OTSs are common obstacles to genetic expansion, recent articles strongly advocate for more rigorous o-tRNA and orthogonal aaRS (o-RS) engineering.[9,13,15,38] Improving the precision of OTS engineering transfers the responsibility of accurate translation from EF-Tu back to the OTS, a distribution of labor that mimics native translation in which the primary responsibility for accuracy falls to the canonical aaRS:tRNA pairs, not downstream translation components.[39] By mirroring the native distribution of responsibilities, dedicated OTSs that ensure accurate tRNA acylation pave the way for researchers to use translation components with expanded capabilities, including substrate-promiscuous EF-Tus. This application of downstream polyspecificity is reflected in the role of the ribosome, which is known to exhibit broad substrate acceptance and still produce accurately translated proteins due to the fact that translation components further upstream ensure accurate tRNA acylation.[40−42] If a rigorously engineered OTS were used, there is no evidence a promiscuous EF-Tu would undermine accurate translation. Similarly, we would anticipate that synthetic acylation methods (*e.g.*, flexizymes) would be compatible with an EF-Tu exhibiting alternative substrate compatibility.[43]

5. EF-Tu Variant Supports Organismal Fitness. While broader substrate acceptance is a highly desirable feature of an EF-Tu, there may be a limit to the degree of infidelity that is possible for the cell to tolerate. An EF-Tu with enhanced polyspecificity risks the potential of being so indiscriminate that it is detrimental to host organism fitness. To examine the impact of engineered EF-Tus on organismal fitness, we compared substrate-promiscuous variants (EF-4A and EF-D216A) against an ncAA-specific variant (EF-Sep) and the

wild type (EF-coli). Each EF-Tu variant was expressed in bacteria grown in 2xYT media to detect leaky expression of EF-Tu, 2xYT media with 2% glucose added for catabolic repression, and 2xYT media with 0.5 mM IPTG for induction of EF-Tu expression. Each growth curve represents triplicate averages that were subsequently fitted to modified growth models that estimate the maximum specific growth rate and lag time.[44] Table S8 presents these two parameters and their respective errors for each assay. Importantly, the cell line used, BL21Δ*serB*, lacks the gene encoding Sep phosphatase; as such, growth curves were not reproduced using an alternative engineered cell line.[16]

While growth curves for all EF-Tus were similar, cultures in which EF-4A and EF-D216A expression had been induced suggested these variants marginally improved host organism fitness. On average, these variants showed somewhat elevated maximal growth rates and a shorter lag time before entering an exponential growth phase relative to EF-coli and EF-Sep (Figure 7). They also demonstrated more reliable reproducibility with consistently small standard deviations. While any benefit to the host organism is minimal, it is significant that these substrate-promiscuous variants do not impair host organism fitness. Rather, these data support the application of engineered polyspecific EF-Tu variants for use in concert with native translation machinery, further recommending our strategy as a route to ncAA incorporation. These growth curves suggest that expanding EF-Tu's substrate scope is compatible with the endogenous translation machinery and does not negatively impact native translation. Hence, an EF-Tu variant with non-native polyspecific behavior appears to be an asset to genetic code expansion.

■ CONCLUSION

As expansion of the genetic code targets increasingly complex protein products, EF-Tu discrimination of ncAAs is emerging as a critical factor limiting ncAA scope. In this approach, we identified multiple EF-Tu variants that facilitated incorporation of non-native substrates, both ncaa-tRNAs and misacylated tRNAs. Computational methods rooted in theories of molecular evolution guided development of a targeted EF-Tu library. Compatibility of EF-Tu variants with non-native substrates was assessed *via* OTS-mediated amber codon suppression and confirmed *via* ESI-MS analysis of purified CAT protein expressed *via* EF-Tu mediated translation. Using these techniques, two EF-Tu variants with expanded substrate compatibility were identified. Growth assays demonstrated that a cooperative EF-Tu with expanded substrate scope is a viable addition to cellular translation without sabotaging cell growth. These data suggest that expanding EF-Tu's substrate compatibility may be compatible with the natural limits imposed by endogenous gene expression. This research supports future goals to expand the genetic code including multisite ncAA incorporation, multiple ncAA incorporation, and proteome-wide incorporation, which can be impacted by EF-Tu's proofreading capabilities.[8]

The strategy of employing multiple elongation factors with different substrate compatibilities in parallel is an evolved mechanism for cellular cotranslational incorporation of Sec. Our results support the application of this naturally occurring strategy to engineer the genetic code and expand ncAA scope. Specifically, this research demonstrated EF-Tu variants with expanded substrate compatibility can work effectively in concert with endogenous translation machinery. The success
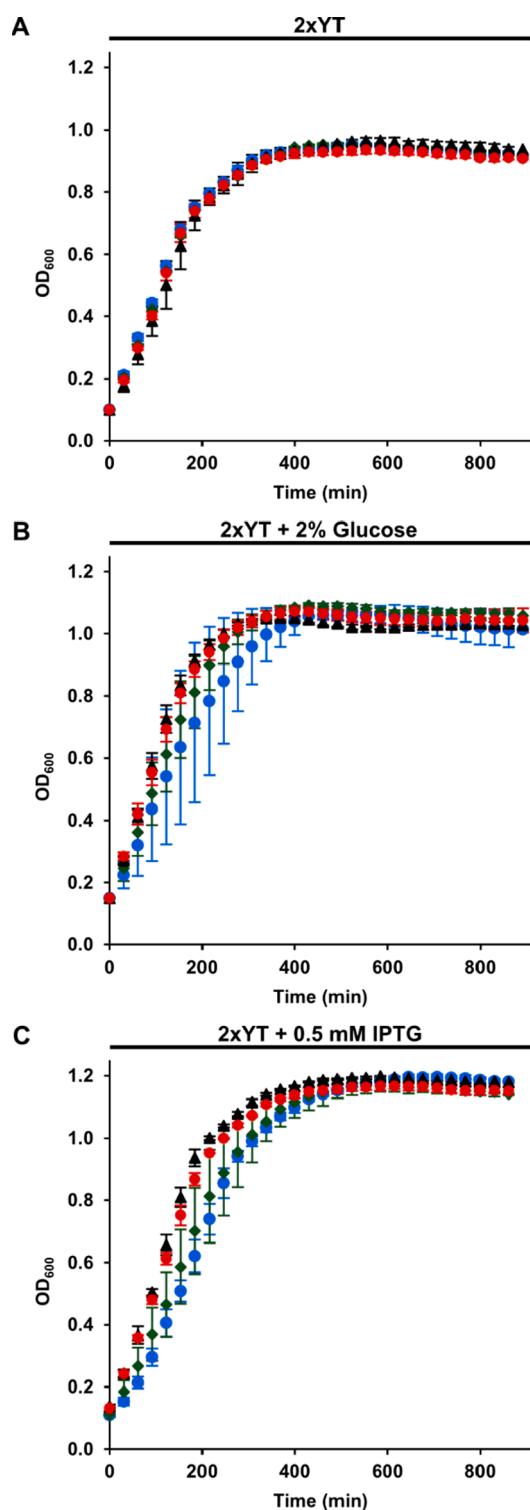
Figure 7. Growth assays for EF-Tu variants expressed in BL21Δ*serB* cell line. Triplicate averages are shown for EF-4A (black triangles), EF-D216A (red circles), EF-Sep (blue circles), and EF-coli (green diamonds). Cultures were grown in 2xYT media (A), 2xYT media with 2% glucose added (B), and 2xYT media with 0.5 mM IPTG added (C). Error bars show standard deviation.

of the REAP-derived library also offers support to our underlying assumptions, specifically that EF-Tu and SelB may be paralogues and may have once shared a common ancestor that exhibited broader polyspecific activity. Further analysis of the EF-Tu library emphasizes the impact of single

mutations to engineered EF-Tu variants. To generate the library, each EF-Tu variant contained multiple mutations, a commonly used approach; however, a single-mutation variant showed enhanced substrate compatibility relative to the quadruple variant. Additionally, the single-mutation variants offered improved insight into the contribution of individual residues to EF-Tu behavior. These data suggest the possibility that epistatic interactions among amino acid residues may limit researchers' ability to identify sites influencing substrate acceptance, thus recommending single-mutation variants for future engineering of EF-Tus compatible with non-native substrates.

Data generated by the EF-Tu library also presented an opportunity to evaluate how effectively REAP identified residues that expand substrate acceptance. The most impactful mutation, D216, was ranked within the top ten sites associated with Type II functional divergence and within the top 15% (out of 279 total residues) overall. Only six sites selected for library development ranked higher. Although the impact of D216 has been debated, our data support evidence that this position strongly affects EF-Tu's substrate specificity.[17,22] Of the residues that ranked higher than D216, site N273 was also mutated in variant EF-4A. Although position N273 was not as impactful as site D216, follow-up experiments suggest that N273 can also influence substrate binding, contrary to previous findings. Since both residues associated with expanded EF-Tu activity were ranked within the top positions identified by REAP, these data lend validation to our computational method. They also suggest that REAP can identify relevant positions whose importance may be otherwise overlooked. Since this research highlights the influence of individual mutations on EF-Tu, the contribution of the individual decoy positions cannot be fully characterized, as they were not evaluated singly in the context of the wild-type sequence. However, because positions relevant to ncAA compatibility were identified *via* assessment of mutations in combination, we can conclude that in combination, the decoy positions were not as effective at expanding EF-Tu's non-native substrate compatibility as those identified by REAP, providing support for our methodology.

This work directly complements current research seeking to further expand the breadth of non-natural protein translation. Prior work targeting multisite ncAA incorporation has demonstrated the vital importance of both improved OTS and EF-Tu engineering.[13,15,38] Additionally, more precisely engineered OTSs could allow the translation machinery to accommodate even a highly polyspecific EF-Tu with limited risk of inaccurate translation. Expanded EF-Tu substrate acceptance also has the potential to reduce, if not completely eliminate, EF-Tu:ncAA compatibility as a challenge inhibiting ncAA incorporation. The substrate-promiscuous EF-Tus described herein, for example, could be promising in combination with other ncAAs. Additionally, they may be tractable platforms for development of additional EF-Tus with novel function in the form of either further expanded substrate compatibility or alternate substrate specificity. Continued expansion of the genetic code to incorporate alternative polymer chemistries, non-natural peptide backbone structures, and increasingly exotic ncAAs is anticipated to demand increasingly extensive and creative bioengineering solutions.[7,14] Components that have previously been somewhat tolerant of ncAA incorporation, like EF-Tu, are beginning to come to the forefront as obstacles that must be addressed to achieve these

challenging goals.[11,12] These concurrent efforts illustrating the urgent need for comprehensive and creative strategies to expand the genetic code support the argument for novel approaches to engineer EF-Tu.

■ MATERIALS AND METHODS

REAP Library. The REAP alignment was generated using 38 sequences from 19 species of bacteria that express both EF-Tu and SelB. Due to a large discrepancy in average sequence length between EF-Tu and SelB sequences, we normalized the length of the 38 selected sequences to create a more accurate phylogeny. Generally, 25 residues were eliminated from the N-terminus of each SelB sequence and 343 residues were deleted from the C-terminus. For EF-Tu sequences, 47 residues were removed from the N-terminus; the C-terminus was not adjusted. REAP DNA sequences are found in Table S4.

REAP analysis was completed using DIVERGE2.0 software.[35] The multiple sequence alignment was generated in Clustal Omega; the phylogeny was generated within DIVERGE2.0 using a Poisson distribution. Output was calculated for Gu99, Gu01, and Type II.[32−34]

***In Vivo* Assay.** Variants were assayed using a system for the cotranslational insertion of Sep *in vivo*. This system included an orthogonal triplet, a tRNA (tRNA^{Sep}), an aminoacyl-tRNA^{Sep} synthetase (SepRS), and EF-Tu variant (EF-Sep) specifically engineered for Sep. These genes were located on two plasmids: pCAT112TAG-SepT (Addgene, plasmid number 34624), and pKD-SepRS-EFSep (Addgene, plasmid number 34623). The EF-Tu variant EF-Sep was used as a positive control and the standard to which the REAP variants were compared. Wild-type *E. coli* EF-Tu, which is not compatible with Sep, was used as a negative control. It is relevant to note that all experiments contained endogenous wild-type EF-Tu. The BL21Δ*serB* cell line (Addgene, bacterial strain number 34929), which critically lacks Sep phosphatase, was used. All plasmids and cell lines described here were gifts from Jesse Rinehart and Dieter Söll.

Calculate IC$_{50}$. Plasmids of choice (pKD and pCAT) were transformed into BL21Δ*serB* competent cells (Addgene, catalog number 34929). A single colony was selected from each transformation, grown overnight, and made into a glycerol freezer stock (25% sterile glycerol, 25% sterile water, and 50% bacteria culture). For each assay, glycerol freezer stocks were streaked out and a single colony was picked and grown for ∼24 h. The culture was then diluted to OD$_{600}$ 0.15 in media supplemented with 2 mM Sep, grown to OD$_{600}$ 0.6–0.8 and induced (0.5 mM IPTG, Sigma-Aldrich). Cultures were allowed to express for 20 h and then diluted in saline and plated, in duplicate, on agar plates with a range of chloramphenicol (Thermo Fisher Scientific) concentrations. Colonies were counted daily.

All liquid and solid cultures were grown at 30 °C. All liquid cultures were grown in LB media supplemented with 0.08% glucose. Kanamycin (25 μg/mL, kanamycin sulfate, VWR) and tetracycline (10 μg/mL, tetracycline hydrochloride 98%, Alfa Aesar) were present in all liquid cultures and agar plates. Sep (2 mM, *O*-phospho-L-serine, Sigma-Aldrich) was present in agar plates used for the CAT assay.

Protein Purification. In order to purify the CAT protein, a hexahistidine tag was added to the carboxyl-terminus of the CAT112TAG gene (*via* Gibson assembly). The His-tag was added to the carboxyl-terminus to prevent truncated peptides from being purified. Appropriate glycerol freezer stocks were made as described above. Glycerol freezer stocks were streaked out and a single colony was picked and grown overnight. Then, 1–1.5 mL starter culture was added to 0.5−3 L media supplemented with 2 mM Sep, grown to OD 0.6−0.8 and induced (0.5 mM IPTG). Protein was expressed for 20 h then spun down and frozen at −80 °C. Cultures were resuspended in 5 mL protein extraction reagent (BugBuster, EMD Millipore) and 2.5 μL Benzonase nuclease (250 U/μL purity >90% EMD Millipore) per 1 g cell pellet. Resuspended pellets were incubated at room temperature for 60 min on a rocking platform and then spun down (11 419*g*). The supernatant for each sample was collected and applied to 1.5 mL Ni-NTA resin (Superflow prepacked columns, Qiagen) using a vacuum manifold (QIAvac 24 Plus, Qiagen). All filter sterilized buffers contained 50 mM NaH$_2$PO$_4$ and 300 mM NaCl pH 8.0 with either 10, 20, or 500 mM imidazole added. Columns were prepped by decanting the storage buffer, and then applying 10 mL of 10 mM imidazole buffer. Next, 30 mL of supernatant were applied to column, followed by 10 mL of 20 mM imidazole buffer. This step was repeated, applying 30 mL supernatant followed by 10 mL of 20 mM imidazole buffer, until all supernatant had been applied to the column, ending with 10 mL of 20 mM imidazole buffer. Finally, protein was eluted in 0.5 mL aliquots with 500 mM imidazole buffer (4.5 mL total). Eluate aliquots were run on an SDS-PAGE gel to estimate protein concentration in aliquots. When deemed necessary, aliquots were combined and concentrated using centrifugal concentrators with a 10 000 MWCO membrane (Spin-X UF, Corning).

In-Gel Digestion and Mass Spectrometry. In-gel digestion, nano-LC−MS/MS, and peptide identification was performed as previously described with the following modifications.[45] Protein digestion was performed using chymotrypsin. Reverse phase chromatography was performed using an in-house packed column (40 cm long × 75 μm ID × 360 OD, Dr. Maisch GmbH ReproSil-Pur 120 C18-AQ 1.9 μm beads) and a 120 min gradient. Raw files were searched using the Mascot algorithm (version 2.5.1) against a protein database constructed of combining the FASTA file for CAT protein (modified to generate 20 versions each with a different natural or modified amino acid at position 112) with a contaminant database (cRAP, downloaded 11−21−16 from http://www.thegpm.org) *via* Proteome Discoverer 2.1. Variable modifications include oxidation of Met, carboxyamidomethylation of Cys, and phosphorylation of Ser, Thr, or Tyr. Only peptide spectral matches with an expectation value of less than 0.01 ("High Confidence") were used (Table S6). As a control, wild-type CAT protein was translated *via* wild-type EF-Tu and as expected, only the wild-type amino acid, aspartic acid, was translated at position 112 (Table S7). CAT protein translation mediated by variants EF-coli, EF-N63A, EF-K263A, and EF-N273A were not analyzed using mass spectrometry because protein expression levels were too low to isolate purified CAT protein.

Growth Curves. For each sample, glycerol freezer stocks were streaked out. Three colonies were selected from each plate and grown overnight in 2xYT media. The following day, 5 μL of the overnight culture was diluted in 195 μL fresh 2xYT media, 2xYT media supplemented with 2% glucose, or 2xYT media supplemented with 0.5 mM IPTG. These three media stocks contained 2 mM Sep. Samples were grown 24 h with shaking in a SpectraMax M2e microplate reader (Molecular Devices) and absorbance was measured (OD$_{600}$) at 10.25 min

intervals. Three wells with only 200 μL 2xYT media with 2 mM Sep (supplemented with nothing, glucose, or IPTG) served as references for absorbance measurements. All liquid cultures were grown at 30 °C in media supplemented with 25 μg/mL kanamycin and 10 μg/mL tetracycline. During data analysis, OD_{600} values were averaged for the three blank reference cells and the value was subtracted from the corresponding growth curves. Data sets were normalized on the basis of the cultures' starting density with growth curves beginning at OD_{600} 0.1 for t = 0 min.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00305.

Tables S1−S8 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel: (404) 413-5432. E-mail: egaucher@gsu.edu.

ORCID Ⓞ

Eric A. Gaucher: 0000-0001-6836-0828

Author Contributions

V.E.D.C. performed experiments, analyzed results, and wrote the manuscript. M.F.C. and V.E.D.C. performed and evaluated the computational analysis. M.F.C. designed the EF-Tu library. E.A.G. supervised the project.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

OTS, orthogonal translation system; Sep, phosphoserine; EF-Tu, elongation factor Tu; aaRS, aminoacyl-tRNA synthetase; REAP, reconstructing evolutionary adaptive paths; ncAA, noncanonical amino acid.

■ REFERENCES

(1) Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015) Designing logical codon reassignment - Expanding the chemistry in biology. *Chemical Science 6*, 50−69.

(2) Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem. 79*, 413−444.

(3) Mendel, D., Cornish, V. W., and Schultz, P. G. (1995) Site-directed mutagenesis with an expanded genetic code. *Annu. Rev. Biophys. Biomol. Struct. 24*, 435−462.

(4) Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H. R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., Mosberg, J. A., Rohland, N., Schultz, P. G., Jacobson, J. M.,

(5) Martin, R. W., Des Soye, B. J., Kwon, Y. C., Kay, J., Davis, R. G., Thomas, P. M., Majewska, N. I., Chen, C. X., Marcum, R. D., Weiss, M. G., Stoddart, A. E., Amiram, M., Charna, A. K. R., Patel, J. R., Isaacs, F. J., Kelleher, N. L., Hong, S. H., and Jewett, M. C. (2018) Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun. 9*, 1−9.

(6) Li, J., Gu, L. C., Aach, J., and Church, G. M. (2014) Improved cell-free RNA and protein synthesis system. *PLoS One 9*, 1−11.

(7) Amiram, M., Haimovich, A. D., Fan, C. G., Wang, Y. S., Aerni, H. R., Ntai, I., Moonan, D. W., Ma, N. J., Rovner, A. J., Hong, S. H., Kelleher, N. L., Goodman, A. L., Jewett, M. C., Söll, D., Rinehart, J., and Isaacs, F. J. (2015) Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol. 33*, 1272−1279.

(8) Des Soye, B. J., Patel, J. R., Isaacs, F. J., and Jewett, M. C. (2015) Repurposing the translation apparatus for synthetic biology. *Curr. Opin. Chem. Biol. 28*, 83−90.

(9) Fan, C. G., Ho, J. M. L., Chirathivat, N., Söll, D., and Wang, Y. S. (2014) Exploring the substrate range of wild-type aminoacyl-tRNA synthetases. *ChemBioChem 15*, 1805−1809.

(10) Ieong, K. W., Pavlov, M. Y., Kwiatkowski, M., Forster, A. C., and Ehrenberg, M. (2012) Inefficient delivery but fast peptide bond formation of unnatural L-aminoacyl-tRNAs in translation. *J. Am. Chem. Soc. 134*, 17955−17962.

(11) Hong, S. H., Ntai, I., Haimovich, A. D., Kelleher, N. L., Isaacs, F. J., and Jewett, M. C. (2014) Cell-free protein synthesis from a release factor 1 efficient Escherichia coli activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol. 3*, 398−409.

(12) Liu, Y., Kim, D. S., and Jewett, M. C. (2017) Repurposing ribosomes for synthetic biology. *Curr. Opin. Chem. Biol. 40*, 87−94.

(13) Maranhao, A. C., and Ellington, A. D. (2017) Evolving orthogonal suppressor tRNAs to incorporate modified amino acids. *ACS Synth. Biol. 6*, 108−119.

(14) Liu, C. C., Jewett, M. C., Chin, J. W., and Voigt, C. A. (2018) Toward an orthogonal central dogma. *Nat. Chem. Biol. 14*, 103−106.

(15) Gan, R., Perez, J. G., Carlson, E. D., Ntai, I., Isaacs, F. J., Kelleher, N. L., and Jewett, M. C. (2017) Translation system engineering in Escherichia coli enhances non-canonical amino acid incorporation into proteins. *Biotechnol. Bioeng. 114*, 1074−1086.

(16) Park, H. S., Hohn, M. J., Umehara, T., Guo, L. T., Osborne, E. M., Benner, J., Noren, C. J., Rinehart, J., and Söll, D. (2011) Expanding the genetic code of Escherichia coli with phosphoserine. *Science 333*, 1151−1154.

(17) Doi, Y., Ohtsuki, T., Shimizu, Y., Ueda, T., and Sisido, M. (2007) Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system. *J. Am. Chem. Soc. 129*, 14458−14462.

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

(19) LaRiviere, F. J., Wolfson, A. D., and Uhlenbeck, O. C. (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science 294*, 165−168.

(20) Schrader, J. M., Chapman, S. J., and Uhlenbeck, O. C. (2011) Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc. Natl. Acad. Sci. U. S. A. 108*, 5215−5220.

(21) Dale, T., Sanderson, L. E., and Uhlenbeck, O. C. (2004) The affinity of elongation factor Tu for an aminoacyl-tRNA is modulated by the esterified amino acid. *Biochemistry 43*, 6159−6166.

(22) Fan, C. G., Ip, K., and Söll, D. (2016) Expanding the genetic code of Escherichia coli with phosphotyrosine. *FEBS Lett. 590*, 3040−3047.

(23) Haruna, K., Alkazemi, M. H., Liu, Y. C., Söll, D., and Englert, M. (2014) Engineering the elongation factor Tu for efficient selenoprotein synthesis. *Nucleic Acids Res. 42*, 9976−9983.

(4) Rinehart, J., Church, G. M., and Isaacs, F. J. (2013) Genomically recoded organisms expand biological functions. *Science 342*, 357−360.

(24) Rogerson, D. T., Sachdeva, A., Wang, K. H., Haq, T., Kazlauskaite, A., Hancock, S. M., Huguenin-Dezot, N., Muqit, M. M. K., Fry, A. M., Bayliss, R., and Chin, J. W. (2015) Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol. 11*, 496−503.

(25) Yoshizawa, S., and Böck, A. (2009) The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta, Gen. Subj. 1790*, 1404−1414.

(26) Khersonsky, O., and Tawfik, D. S. (2010) Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu. Rev. Biochem. 79*, 471−505.

(27) Risso, V. A., Gavira, J. A., Mejia-Carmona, D. F., Gaucher, E. A., and Sanchez-Ruiz, J. M. (2013) Hyperstability and substrate promiscuity in laboratory resurrections of precambrian beta-lactamases. *J. Am. Chem. Soc. 135*, 2899−2902.

(28) Laos, R., Shaw, R. W., Leal, N. A., Gaucher, E. A., and Benner, S. A. (2013) Directed evolution of polymerases to accept nucleotides with nonstandard hydrogen bond patterns. *Biochemistry 52*, 5288−5294.

(29) Cole, M. F., and Gaucher, E. A. (2011) Exploiting models of molecular evolution to efficiently direct protein engineering. *J. Mol. Evol. 72*, 193−203.

(30) Gaucher, E. A. (2007) Ancestral sequence reconstruction as a tool to understand natural history and guide synthetic biology: realizing and extending the vision of Zuckerkandl and Pauling, In *Ancestral Sequence Reconstruction* (Liberles, D. A., Ed.), pp 20−33, Oxford University Press, New York.

(31) Gaucher, E. A., Das, U. K., Miyamoto, M. M., and Benner, S. A. (2002) The crystal structure of eEF1A refines the functional predictions of an evolutionary analysis of rate changes among elongation factors. *Mol. Biol. Evol. 19*, 569−573.

(32) Gu, X. (1999) Statistical methods for testing functional divergence after gene duplication. *Mol. Biol. Evol. 16*, 1664−1674.

(33) Gu, X. (2001) Maximum-likelihood approach for gene family evolution under functional divergence. *Mol. Biol. Evol. 18*, 453−464.

(34) Gu, X. (2006) A simple statistical method for estimating Type-II (Cluster-Specific) functional divergence of protein sequences. *Mol. Biol. Evol. 23*, 1937−1945.

(35) Gu, X., and Vander Velden, K. (2002) DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. *Bioinformatics 18*, 500−501.

(36) Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F. C., and Nyborg, J. (1995) Crystal-structure of the ternary complex of Phe-tRNA(Phe), EF-Tu, and a GTP analog. *Science 270*, 1464−1472.

(37) Pirman, N. L., Barber, K. W., Aerni, H. R., Ma, N. J., Haimovich, A. D., Rogulina, S., Isaacs, F. J., and Rinehart, J. (2015) A flexible codon in genomically recoded Escherichia coli permits programmable protein phosphorylation. *Nat. Commun. 6*, 1−6.

(38) Kunjapur, A. M., Stork, D. A., Kuru, E., Vargas-Rodriguez, O., Landon, M., Söll, D., and Church, G. M. (2018) Engineering posttranslational proofreading to discriminate nonstandard amino acids. *Proc. Natl. Acad. Sci. U. S. A. 115*, 619−624.

(39) Ibba, M., and Söll, D. (2004) Aminoacyl-tRNAs: setting the limits of the genetic code. *Genes Dev. 18*, 731−738.

(40) Josephson, K., Hartman, M. C. T., and Szostak, J. W. (2005) Ribosomal synthesis of unnatural peptides. *J. Am. Chem. Soc. 127*, 11727−11735.

(41) Fahnestock, S., Neumann, H., Shashoua, V., and Rich, A. (1970) Ribosome-catalyzed ester formation. *Biochemistry 9*, 2477−2483.

(42) Fahnestock, S., and Rich, A. (1971) Ribosome-catalyzed polyester formation. *Science 173*, 340−343.

(43) Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006) A highly flexible tRNA acylation method for non-natural potypeptide synthesis. *Nat. Methods 3*, 357−359.

(44) Zwietering, M. H., Jongenburger, I., Rombouts, F. M., and van 't Riet, K. (1990) Modeling of the bacterial-growth curve. *Appl. Environ. Microbiol. 56*, 1875−1881.

(45) Rinker, T. E., Philbrick, B. D., Hettiaratchi, M. H., Smalley, D. M., McDevitt, T. C., and Temenoff, J. S. (2018) Microparticle-mediated sequestration of cell-secreted proteins to modulate chondrocytic differentiation. *Acta Biomater. 68*, 125−136.