ACS central science

http://pubs.acs.org/journal/acscii

Article

A Translation-Independent Directed Evolution Strategy to Engineer Aminoacyl-tRNA Synthetases

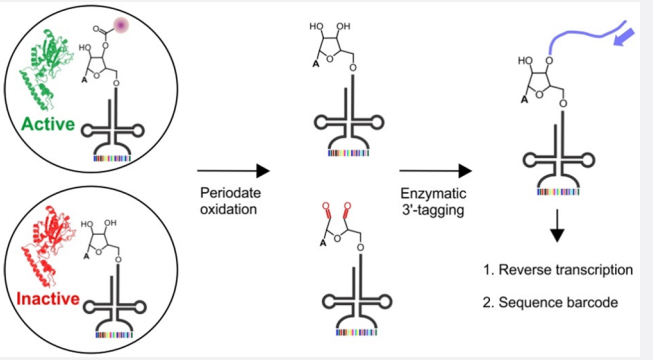
Chintan Soni, Noam Prywes, Matthew Hall, Malavika A. Nair, David F. Savage, Alanna Schepartz, and Abhishek Chatterjee*

Cite This: https://doi.org/10.1021/acscentsci.3c01557

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Using directed evolution, aminoacyl-tRNA synthetases (aaRSs) have been engineered to incorporate numerous noncanonical amino acids (ncAAs). Until now, the selection of such novel aaRS mutants has relied on the expression of a selectable reporter protein. However, such translation-dependent selections are incompatible with exotic monomers that are suboptimal substrates for the ribosome. A two-step solution is needed to overcome this limitation: (A) engineering an aaRS to charge the exotic monomer, without ribosomal translation; (B) subsequent engineering of the ribosome to accept the resulting acyl-tRNA for translation. Here, we report a platform for aaRS engineering that directly selects tRNA-acylation without ribosomal translation (START). In START, each distinct aaRS mutant is correlated to a cognate tRNA containing a unique sequence barcode. Acylation by an active aaRS mutant protects the corresponding barcode-containing tRNAs from oxidative treatment designed to damage the 3′-terminus of the uncharged tRNAs. Sequencing of these surviving barcode-containing tRNAs is then used to reveal the identity of the aaRS mutants that acylated the correlated tRNA sequences. The efficacy of START was demonstrated by identifying novel mutants of the *Methanomethylophilus alvus* pyrrolysyl-tRNA synthetase from a naïve library that enables incorporation of ncAAs into proteins in living cells.



INTRODUCTION

Over the last two decades, the genetic code of cells in various domains of life has been expanded to include hundreds of noncanonical amino acids (ncAAs) with diverse chemical structures.[1−4] This expansion has been accomplished by developing engineered aminoacyl-tRNA synthetase/tRNA pairs that selectively incorporate various ncAAs in response to a unique codon (e.g., a repurposed nonsense codon).[1−5] The ability to site-specifically incorporate enabling ncAAs into proteins using this technology has unlocked powerful new ways to probe and manipulate protein function for both basic science and biotechnology applications.[1−3,6,7]

Despite such exciting progress, this technology has been largely restricted to the incorporation of simple L-α-amino acids and, to a lesser extent, some α-hydroxy acids.[8−12] Pioneering early work using *in vitro* translation systems has demonstrated that the ribosome has the potential to accept a much wider variety of monomers for translation such as D-α-amino acids,[13] β- and γ-amino acids,[14−17] long-chain amino acids,[18] aminobenzoic acid,[19] α-aminoxy and α-hydrazino acids,[20] α-thio acids,[21] and others.[22−24] The remarkable tolerance of the ribosome for such noncanonical monomers and further expansion of its noncanonical substrate scope through engineering[25−28] offer a possible path for repurposing mRNA-templated ribosomal translation to synthesize non-peptide polymers.

Although incorporation of structurally divergent monomers into peptides has been broadly explored *in vitro*, examples of achieving the same in living cells remain scarce. The key bottleneck limiting progress on this front is the lack of engineered aminoacyl-tRNA synthetases (aaRSs) capable of efficiently charging such monomers. In the handful of examples, where such exotic monomers were incorporated into proteins in *E. coli*, such as β^3-*p*-bromo-homophenylalanine[26] and a β^2-hydroxy acid,[29] tRNA acylation relied on promiscuous recognition that certain wild-type aaRSs fortuitously exhibited for these substrates. The ability to systematically engineer aaRSs to charge a wider variety of noncanonical monomers with high fidelity and efficiency is needed to overcome this limitation. However, established directed evolution strategies for altering the substrate specificity of

Received: December 14, 2023
Revised: May 7, 2024
Accepted: May 8, 2024

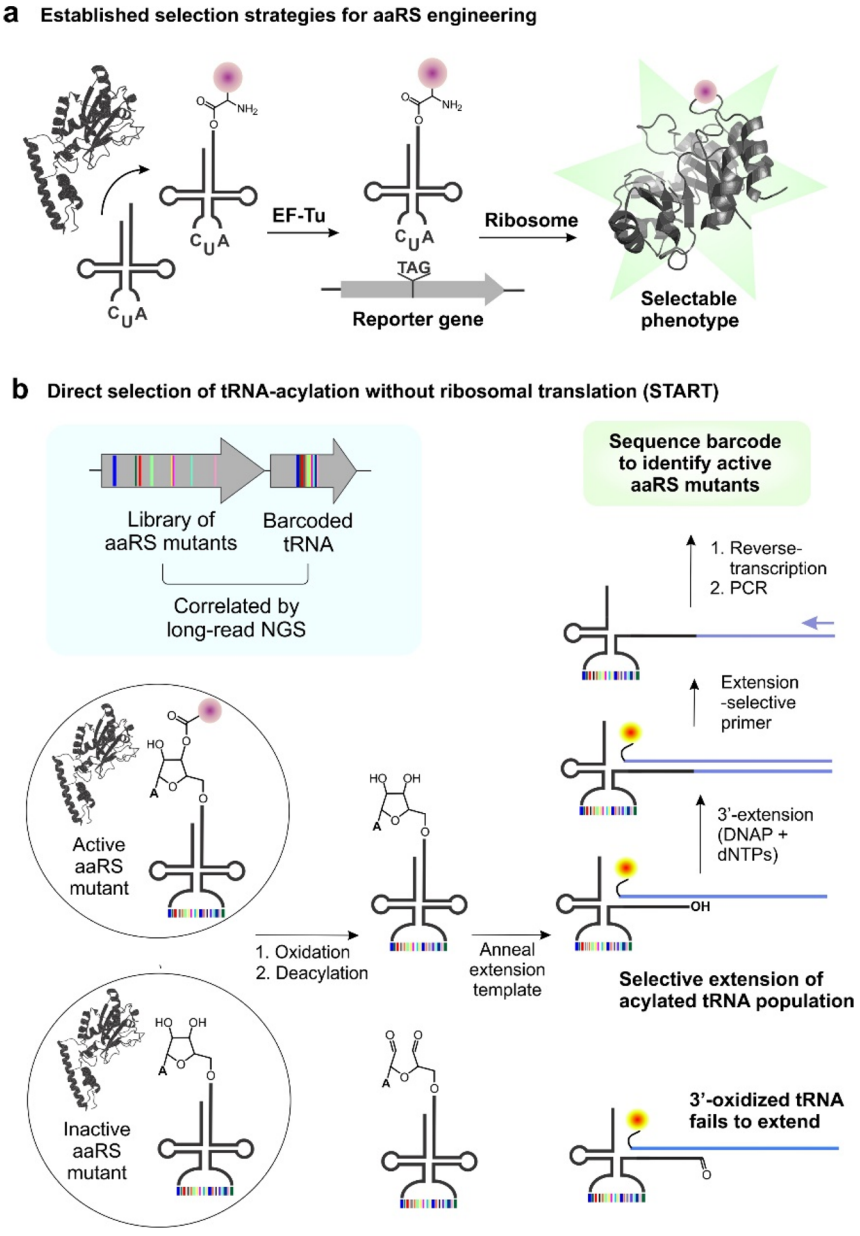


Figure 1. Directed evolution strategies for engineering the substrate specificity of aaRSs. (a) Established selection strategies depend on ribosomal translation to connect the activity of aaRS to the expression of a reporter protein with a selectable phenotype. This strategy is incompatible with noncanonical monomers that are poor substrates for the ribosome. (b) In START, members in an aaRS mutant library are individually correlated to a sequence barcode located within the cognate tRNA. Successful charging by a mutant aaRS protects its cognate barcode-containing tRNAs from a periodate-mediated oxidation, enabling their subsequent enrichment via templated 3′-extension and RT-PCR.

aaRSs rely on the ribosomal translation of reporter proteins with a selectable phenotype (such as antibiotic resistance or fluorescence; Figure 1a).[30−39] Unfortunately, such translation-dependent selection schemes are incompatible with non-canonical monomers that are poor substrates for the ribosome. Although the ribosome could be engineered to accept novel monomers, such efforts would require access to engineered aaRSs capable of generating monomer-charged tRNAs. Engineering of both the aaRS and the ribosome at the same time is challenging, as it would require co-localization of two rare mutants from each library in the same cell, which is associated with low statistical odds. Consequently, a two-step solution is needed, where the aaRS is first engineered to charge the noncanonical monomer without relying on translational readout, which can then be used for ribosome engineering.

Here we report a general approach for engineering aaRSs that involves direct selection of tRNA acylation without ribosomal translation (START; Figure 1b). Our approach was inspired by an established strategy to enrich acylated tRNA pools from living cells, where the 2′,3′-dihydroxy functionality at the 3′-terminus of uncharged tRNAs is selectively oxidized using periodate.[40−47] The acylated tRNAs are protected from this damage and can be subsequently tagged with a unique oligonucleotide sequence at its intact 3′-terminus, allowing their subsequent identification through selective reverse-transcription and PCR (RT-PCR).[42−47] Although this strategy is effective for enriching and characterizing acylated tRNA sequences, it does not reveal the identity of the aaRS responsible for their acylation. To do so, it is essential to connect the identity of each aaRS variant to its cognate tRNA

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX

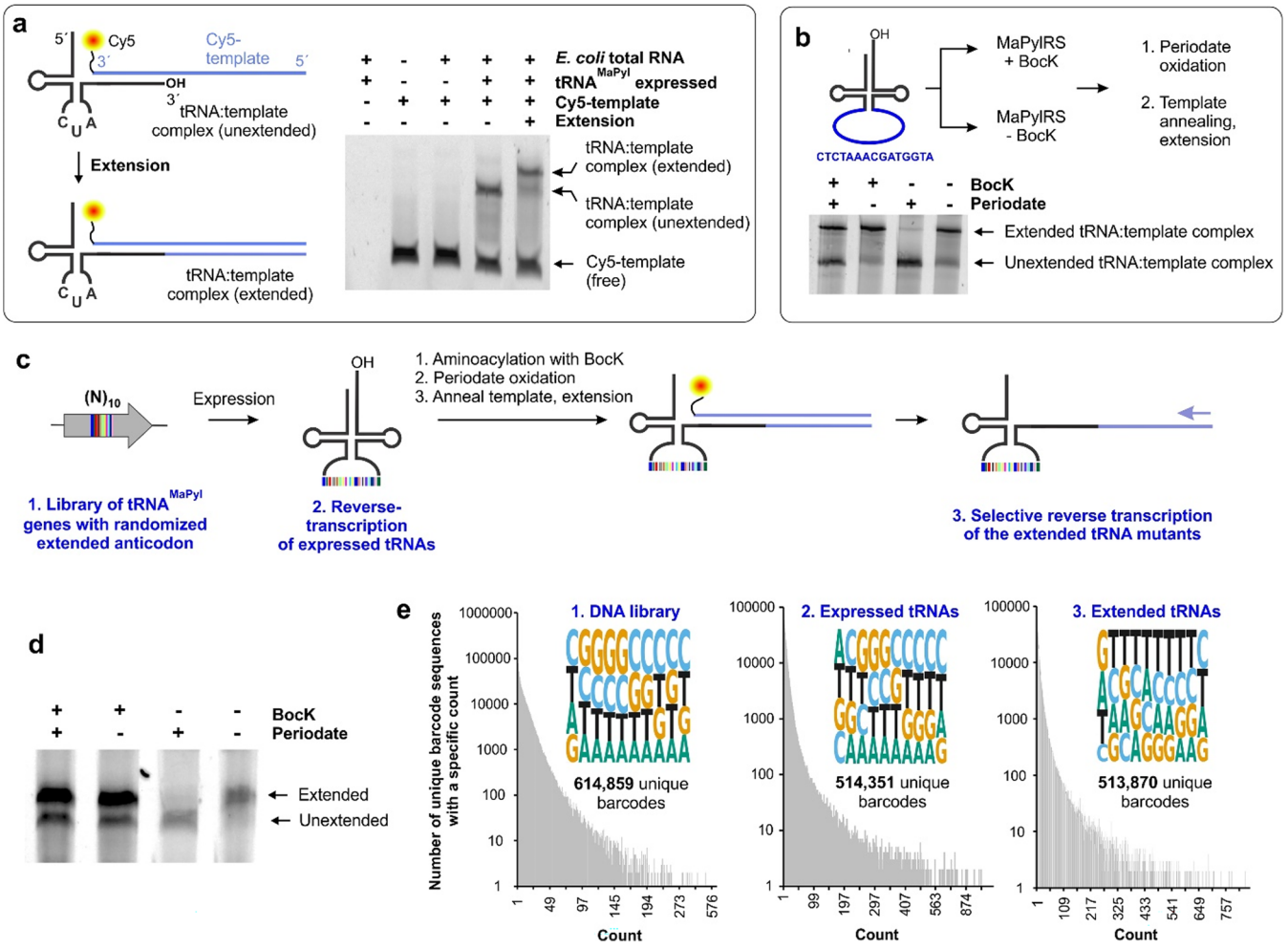


Figure 2. The anticodon loop of tRNA^{MaPyl} can be expanded to include a sequence barcode. (a) 3′-extension method used to selectively tag tRNA^{MaPyl} with an intact 3′-end, using a 3′-Cy5-labeled DNA template. Native PAGE followed by fluorescence imaging shows that the designed Cy5-DNA template selectively hybridizes with tRNA^{MaPyl} and enables the formation of the extension product. (b) A tRNA^{MaPyl} mutant with a severely expanded anticodon loop can be effectively charged by MaPylRS, as revealed by the tRNA-extension assay. In the absence of tRNA acylation, periodate treatment prevents the formation of the extension product, but when coexpressed with MaPylRS and a cognate ncAA substrate (BocK), robust formation of extension product was observed, indicating successful acylation of the extended tRNA^{MaPyl}. (c) Scheme for characterizing the expression and potential aminoacylation of the members of the tRNA^{MaPyl} barcode library. (d) A large majority of the barcode-containing tRNA^{MaPyl} library members are successfully acylated in the presence of MaPylRS and BocK, as revealed by protection from periodate oxidation in the tRNA-extension assay. (e) Illumina sequencing of the barcode-containing tRNA^{MaPyl} library DNA (left), after it is expressed in *E. coli* (middle), and members that survive periodate treatment when coexpressed with MaPylRS in the presence of BocK (right) reveal similar distribution and composition. These analyses indicate that a majority of the barcode-containing tRNA^{MaPyl} sequences are successfully expressed and acylated by MaPylRS.

sequence. In START, we achieve this connection by introducing a sequence barcode into a permissive site within the tRNA. Using *Methanomethylophilus alvus* pyrrolysyl-tRNA (tRNA^{MaPyl}) as the model system, we demonstrate the feasibility of inserting a sequence barcode in the anticodon loop with a limited impact on tRNA expression and charging. Next, the START selection scheme was optimized to allow robust enrichment of an acylated barcode-containing tRNA^{MaPyl} from a mixed population that also contained an uncharged tRNA^{MaPyl} with a distinct barcode. Then, we created an active site library of *M. alvus* pyrrolysyl-tRNA synthetase (MaPylRS), where each unique mutant is encoded by associated tRNA^{MaPyl} barcodes, and subjected this library to the START selection scheme to identify MaPylRS mutants capable of charging different ncAAs. Finally, we show that the START selection scheme is compatible with non-α-amino acid monomers. START represents the first translation-independent directed evolution platform for aaRS engineering, which will be invaluable for developing incorporation systems for unique noncanonical monomers.

RESULTS AND DISCUSSION

A Chemical Strategy to Select for Charged tRNAs. A translation-independent directed evolution strategy would ideally allow the selection of active aaRS mutants simply based on their ability to acylate a cognate tRNA with a desired noncanonical monomer. Upon aaRS-catalyzed acylation of a tRNA on its 3′-terminal ribose, it becomes protected from oxidation by sodium periodate, which selectively oxidizes vicinal diol groups. This method has long been leveraged to characterize the charging status of tRNAs in cells.[40−47] Recently, this strategy has been combined with modern enrichment and sequencing methods to perform such analyses

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX

in a more high-throughput manner.[42,43,45−47] After the selective periodate-mediated oxidation of the uncharged tRNAs and the deacylation of the charged tRNAs using mildly alkaline conditions, a unique oligonucleotide sequence can be introduced onto the intact 3′-terminus of the acylated population either through ligation[43,45,46] or by polymerase-mediated extension of the 3′-terminus using a DNA template that is hybridized onto the tRNA sequence.[44] The installed oligonucleotide sequence can be subsequently used to selectively reverse-transcribe and PCR amplify the charged tRNA sequences.

The methods described above for the selective enrichment of acylated tRNAs have the potential to serve as the foundation for a translation-independent selection strategy for tRNA acylation. To explore this possibility, we focused on the *M. alvus*-derived pyrrolysyl-tRNA synthetase (MaPylRS)/tRNA^{MaPyl} pair as a model system.[48] The pyrrolysyl pair represents the leading platform for expanding the genetic code with both ncAAs and non-α-amino acid monomers.[1−3,9,24,29,49] We designed a 3′-fluorophore-labeled DNA template (Figure S1) that will hybridize with tRNA^{MaPyl} and allow the extension of its 3′-terminus by DNA polymerase (Figure 2a). This template selectively hybridized with tRNA^{MaPyl} expressed in *E. coli*, as observed by nondenaturing PAGE followed by fluorescence imaging; no such complex was found when total RNA from *E. coli* not expressing the tRNA^{MaPyl} was used instead (Figure 2a). Incubation of this DNA:tRNA complex with the Klenow fragment of *E. coli* DNA polymerase (lacks 3′ → 5′ exonuclease activity) and dNTPs led to successful extension of the 3′-terminus, as shown by an upward shift in the PAGE analysis. Treatment of tRNA^{MaPyl} with sodium periodate prevented the appearance of this extension product, consistent with the oxidation of its free 3′-terminus (Figure S2). However, when tRNA^{MaPyl} was coexpressed with MaPylRS in the presence of a known ncAA substrate (N^{ϵ}-Boc-L-lysine; BocK), it was protected from periodate-mediate damage and successfully yielded the extension product (Figure S2). These results confirm that the tRNA-extension strategy can be used to selectively introduce an oligonucleotide tag at the 3′-terminus of the acylated tRNA^{MaPyl} population, allowing for their subsequent enrichment by RT-PCR.

Expanding the Anticodon Loop of tRNA^{MaPyl} to Insert a Sequence Barcode. To use this strategy for evolving MaPylRS, its sequence information must be encoded within tRNA^{MaPyl} such that the identity of variants responsible for tRNA acylation can be retrieved by sequencing the acylated tRNA^{MaPyl} pool following their enrichment. We envisioned achieving this encoding by introducing a sequence barcode within a permissive site of tRNA^{MaPyl}. The anticodon loop of tRNA^{MaPyl} is a promising location to introduce the barcode, given pyrrolysyl synthetases do not interact with the anticodon region.[50−55] Indeed, the anticodons of pyrrolysyl-tRNAs have been altered to enable suppression of other nonsense codons, as well as four-base codons, indicating significant plasticity.[53−56]

To explore whether the anticodon loop of tRNA^{MaPyl} can tolerate more dramatic expansions, we generated a variant where the native sequence in the anticodon loop was replaced with a random 15-nucleotide sequence (Figure 2b). The resulting expanded tRNA^{MaPyl} mutant was coexpressed in *E. coli* with MaPylRS in either the presence or absence of a cognate ncAA substrate (BocK). Total RNA isolated from these cells was subjected to a tRNA-extension assay with or without periodate treatment (Figure 2b). The successful expression of the expanded tRNA^{MaPyl} mutant was confirmed by the appearance of the tRNA:DNA hybrid band. When incubated with DNA polymerase and dNTPs, expanded tRNA^{MaPyl} successfully yielded the extension product in the absence of periodate treatment. MaPylRS was found to efficiently acylate the expanded tRNA^{MaPyl} variant with BocK, as the formation of the extension product was largely insensitive to periodate when the ncAA was included in the growth medium, but almost fully disappeared in its absence (Figure 2b).

Although this result was encouraging, we sought to further confirm that the anticodon loop of tRNA^{MaPyl} would tolerate a broad variety of sequences without significantly compromising its expression or aminoacylation by MaPylRS, which is essential to accessing sufficient barcode diversity for encoding a MaPylRS mutant library. To this end, we replaced the anticodon loop of tRNA^{MaPyl} with a sequence of 10 fully randomized nucleotides to create a library of roughly 1×10^6 mutants (Figure 2c). Illumina sequencing (approximately 5×10^6 usable reads) of the resulting plasmid DNA encoding this tRNA library revealed the presence of >600,000 unique barcodes (Figure 2e). Each nucleotide was well-represented at each position of the barcode library (Figure 2e). Additionally, >90% of the barcodes in the library were found to have an abundance that is within a single standard deviation from the average (~8.5 counts/barcode), indicating reasonably good diversity and distribution of the barcode sequences. To evaluate the expression patterns of the barcode-containing tRNA^{MaPyl} genes, this plasmid library was transformed into *E. coli*, and the expressed tRNA^{MaPyl} sequences were amplified by RT-PCR and analyzed by Illumina sequencing. Over 500,000 unique sequences were identified in the expressed tRNA^{MaPyl} barcode library, and the representation of each nucleotide in the barcode was comparable to the DNA library (Figure 2e; Figure S3). These observations confirm that a majority of the barcode-containing tRNA^{MaPyl} sequences can be expressed in *E. coli*.

Finally, to explore if MaPylRS can charge barcode-containing tRNA^{MaPyl} sequences, the library was coexpressed with MaPylRS in *E. coli* in the presence of BocK. The resulting barcode-containing tRNA^{MaPyl} pool was oxidized by periodate to damage the uncharged population, and the acylated sequences were subjected to tRNA extension using a 3′-Cy5-labeled DNA template (Figure 2c). Analysis of this extension reaction by PAGE followed by fluorescence imaging (Figure 2d) showed that most of the tRNA^{MaPyl} pool was protected from periodate oxidation in the presence of BocK (but not when the ncAA was absent), indicating that a large majority of barcode-containing tRNA^{MaPyl} sequences are efficiently charged by MaPylRS. This finding was further confirmed by amplifying the acylated barcode-containing tRNA^{MaPyl} population by RT-PCR and analyzing them by Illumina sequencing. Greater than 500,000 unique barcode sequences were identified in the acylated fraction, and their composition and distribution were similar to the library at the DNA and expressed tRNA level (Figure 2e). Together, these experiments show that expanding the tRNA^{MaPyl} anticodon loop is a viable approach for incorporating a diverse sequence barcode that does not drastically disrupt the expression or charging by MaPylRS.

Optimization of START to Enrich Charged, Barcode-Containing tRNA^{MaPyl} Sequences from a Mixed Pop-

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX

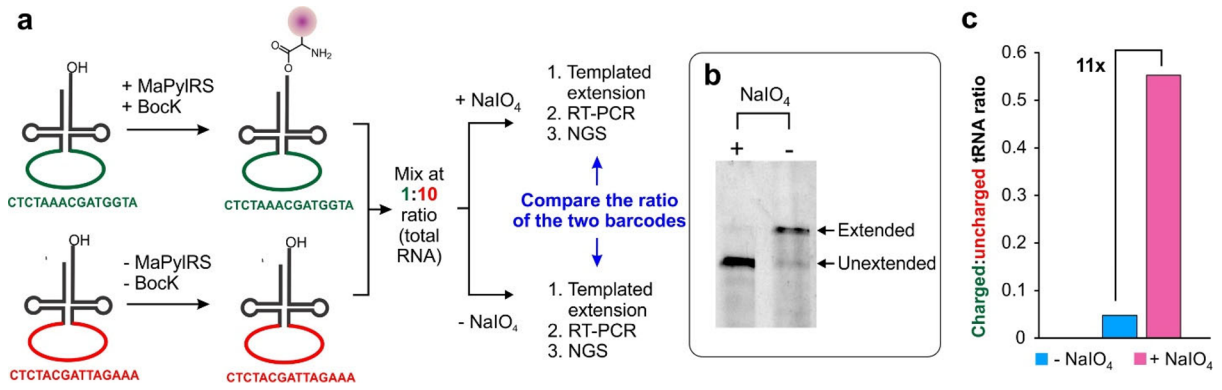


Figure 3. START enables the enrichment of an acylated barcode-containing $tRNA^{MaPyl}$ from a defined mixed population. (a) Scheme of the experiment showing two distinct barcode-containing $tRNA^{MaPyl}$ are separately expressed, where one is charged with BocK and the other is not. A defined mixture of these two populations is subjected to the START scheme, and their relative abundance is quantitatively measured in the presence and absence of the periodate-mediated selection using amplicon sequencing. (b) PAGE followed by fluorescence imaging of the tRNA extension reaction of this mixed population in the presence or absence of periodate oxidation. (c) Amplicon sequencing reveals that the acylated $tRNA^{MaPyl}$ barcodes are enriched approximately 11-fold upon periodate selection.

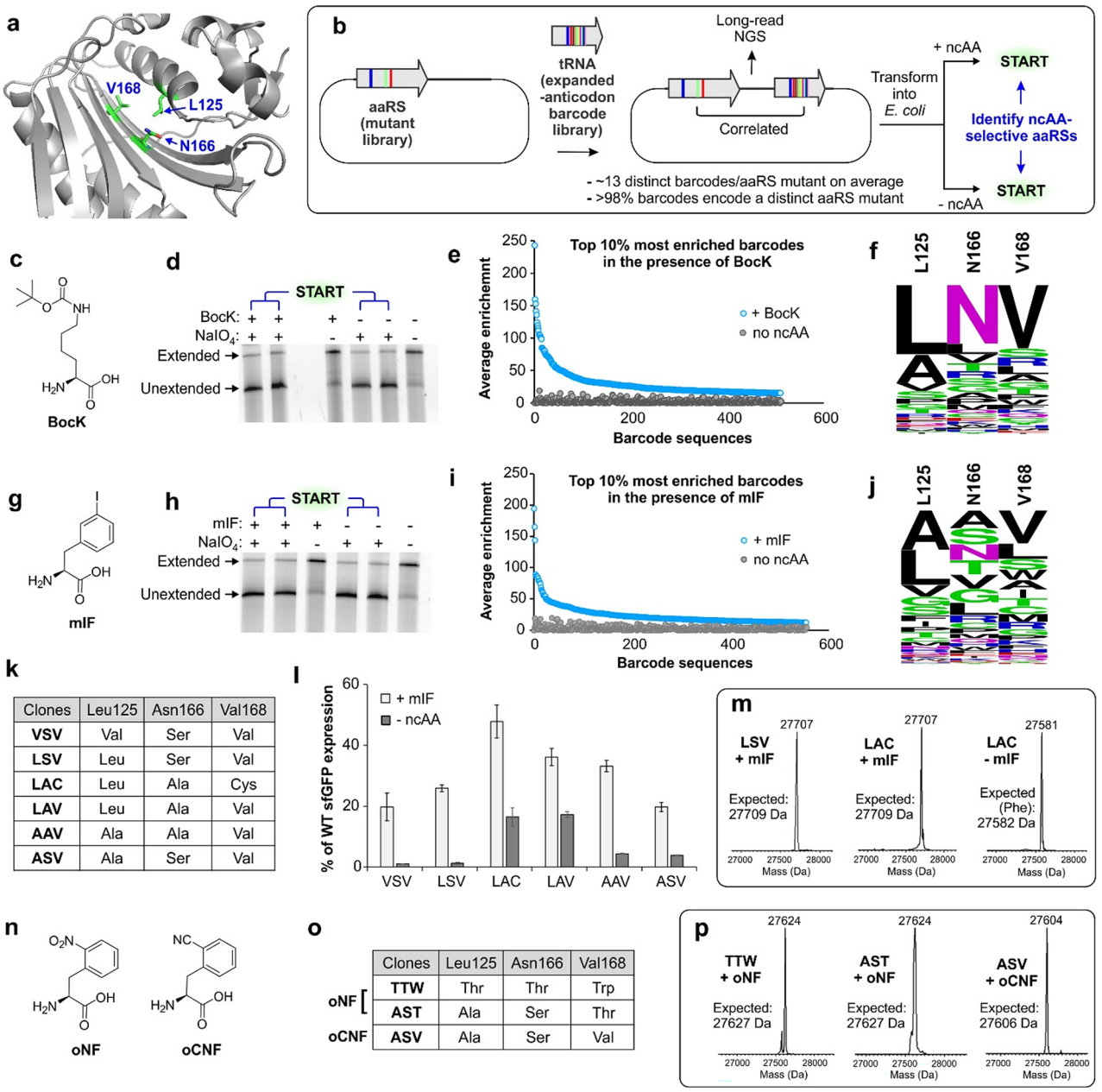
ulation. The basic steps of the START scheme involve periodate oxidation to damage the uncharged tRNAs, selective DNA-templated extension of the charged tRNAs to introduce a unique sequence at the 3′-terminus, followed by its use to amplify these sequences by RT-PCR. To improve the enrichment of the desired population, we employed a gel purification step to isolate the extended DNA:RNA hybrid from the unextended counterparts based on its different mobility on nondenaturing PAGE. We explored different lengths of the DNA template to alter the extended product size and optimize its separation from other species (Figure S4). Next, for evaluating the performance of this selection scheme, we created two standard plasmids to represent an "active" or "inactive" member in the library (Figure 3a). Each plasmid encoded a distinct barcode-containing $tRNA^{MaPyl}$ sequence, but one of these also encoded wild-type MaPylRS, which will acylate the corresponding $tRNA^{MaPyl}$ in the presence of BocK and represent an active library member. The other plasmid lacked MaPylRS and would represent an inactive library member. These plasmids were separately transformed into *E. coli*, and using the aforementioned tRNA-extension assay, we confirmed that the $tRNA^{MaPyl}$ variant coexpressed with MaPylRS was protected from oxidative damage in the presence of BocK, whereas its counterpart lacking MaPylRS was not (Figure S5). In order to quantify the degree of enrichment afforded by our strategy, the mock inactive and active $tRNA^{MaPyl}$ populations were mixed in 10:1 ratio, and the composition of these mixed population was characterized by amplicon sequencing (~50,000 reads) before and after subjecting it to the START scheme (Figure 3a,b). The $tRNA^{MaPyl}$ variant coexpressed with MaPylRS was found to undergo roughly an 11-fold enrichment upon selection (Figure 3c), suggesting that this strategy should allow differentiation between barcode-containing $tRNA^{MaPyl}$ variants associated with active or inactive MaPylRS mutants.

Identification of ncAA-Selective MaPylRS Mutants Using START. Next, we explored if START can be used to identify MaPylRS mutants with altered substrate specificities from a naïve library. To this end, we simultaneously randomized three key residues (Leu125, Asn166, and Val168) around the substrate binding pocket of MaPylRS to NNK codons (Figure 4a), creating a library with a theoretical diversity of 32,768. Into the plasmid encoding the MaPylRS library, we further introduced our barcode-containing $tRNA^{MaPyl}$ library containing an $(N)_{10}$ randomized sequence at the anticodon loop (Figure 4b) using standard restriction cloning. The resulting library was covered using approximately 1.5×10^5 transformants to ensure that (A) each barcode corresponds to a distinct MaPylRS mutant and (B) each MaPylRS is associated with multiple distinct barcodes. Since this library utilizes only ~15% of all possible barcode sequences (~10^6), the chances of two different MaPylRS mutants receiving the same barcode are low. Additionally, as barcode diversity in the resulting library far exceeds the diversity of the MaPylRS library (roughly 50-fold), each distinct MaPylRS mutant should be associated with multiple unique barcodes, which would be beneficial to reduce the risk of false positives.

To characterize this library and correlate each unique MaPylRS mutant to the associated sequence barcodes, we used PacBio HiFi long-read DNA sequencing. From ~1.2×10^5 usable reads retrieved from the sequencing analysis, we confirmed the presence of >93% of all possible MaPylRS mutants. We also found that >91% of unique MaPylRS mutants were represented by more than one barcode, with the average being ~13 barcodes/mutant (Figure S6a). Furthermore, approximately 93% of the barcodes were found to encode a single MaPylRS mutant (Figure S6b), which minimizes potential confusion when establishing genotype−phenotype connections.

Following characterization, the barcoded MaPylRS library was transformed to *E. coli* for subsequent selection. It is important to note that many MaPylRS active site mutants may charge one of the 20 canonical amino acids, which will protect the associated barcode-containing tRNAs from the periodate oxidation, generating false positives. Traditional aaRS engineering strategies typically employ a negative selection step to deplete such cross-reactive mutants.[1,30,31,33,34] However, using the deep sequence coverage from NGS, it should be possible to characterize the enrichment pattern of nearly all mutants present in the MaPylRS library in response to the START scheme, and doing so in parallel in the presence and absence of a desired ncAA should help identify mutants that were selectively enriched in the presence of ncAA (Figure 4b). A similar approach was recently used for orthogonal tRNA evolution in mammalian cells.[57]

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX



Clones	Leu125	Asn166	Val168
VSV	Val	Ser	Val
LSV	Leu	Ser	Val
LAC	Leu	Ala	Cys
LAV	Leu	Ala	Val
AAV	Ala	Ala	Val
ASV	Ala	Ser	Val

	Clones	Leu125	Asn166	Val168
oNF	**TTW**	Thr	Thr	Trp
	AST	Ala	Ser	Thr
oCNF	**ASV**	Ala	Ser	Val

Figure 4. Selection of ncAA-selective MaPylRS mutants using START. (a) The active site of MaPylRS, where the three key residues randomized to generate the library are highlighted. (b) Scheme for generating and selecting the MaPylRS library, where each mutant is correlated to distinct barcode-containing $tRNA^{MaPyl}$ sequences. (c) Structure of BocK. (d) PAGE followed by fluorescence imaging of the tRNA-extension reaction on the library in the presence and absence of BocK. (e) Top 10% most enriched sequence barcodes that show at least 2-fold higher enrichment in the presence of BocK (relative to no BocK). Enrichment of each barcode in the presence (blue) or the absence (gray) of BocK is shown. (f) WebLogo analyses of the residues observed in the active sites of the MaPylRS mutants correlated to these top 10% barcodes reveal the wild-type MaPylRS sequence as the most prevalent. (g) The structure of mIF. (h) PAGE followed by fluorescence imaging of the tRNA-extension reaction on the library in the presence and absence of mIF. (i) Top 10% most enriched sequence barcodes that show at least 2-fold higher enrichment in the presence of mIF (relative to no mIF). Enrichment of each barcode in the presence (blue) or the absence (gray) of mIF is shown. (j) WebLogo analyses of the residues observed in the active sites of the MaPylRS mutants correlated to these top 10% barcodes reveal that the prevalent signature is distinct from that of the wild-type MaPylRS. (k) Selected clones for further characterizations, based on their average enrichment levels, the number of associated barcodes that show enrichment, and similarity to the consensus sequence of the most enriched sequences (panel j). (l) The mIF-charging activity of each MaPylRS mutant, upon coexpression with $tRNA_{CUA}^{MaPyl}$, measured as the expression of a sfGFP-151-TAG reporter in the presence/absence of mIF and normalized to the expression of a wild-type sfGFP reporter. The characteristic fluorescence of the reporter was measured in cells resuspended in PBS. (m) ESI-MS analysis of purified sfGFP-151-TAG reporter for LSV and LAC mutants in the presence of mIF shows masses consistent with the incorporation of mIF at the TAG codon. In the absence of mIF, the LAC mutant likely charges phenylalanine, as indicated by the mass of the isolated reporter protein. (n) The structures of oNF and oCNF. (o) MaPylRS clones selected using START (see Figures S8 and S9 in the SI for details) to be characterized for charging oNF and oCNF. (p) ESI-MS analysis of purified sfGFP-151-TAG reporters expressed using these mutants in the presence of the appropriate ncAA shows masses consistent with the incorporation of oNF or oCNF at the TAG codon.

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX

Using this strategy, the barcoded MaPylRS library was selected for its ability to charge two distinct ncAAs, BocK (Figure 4c–f) and *m*-iodo-L-phenylalanine (mIF) (Figure 4g–j), performed in duplicates. After selection, the surviving barcodes were amplified and characterized by Illumina sequencing. To identify barcodes associated with ncAA-selective MaPylRS mutants, we filtered the NGS data based on three key criteria: (A) it must be represented in the collection of barcodes obtained by long-read sequencing characterization of the original library, (B) it must show at least a 2-fold enrichment upon selection in the presence of ncAA, and (C) its relative enrichment in the presence of the ncAA must be at least 2-fold higher than in the absence. The top 10% (~500) of the surviving barcodes were arranged in a decreasing order of average enrichment in the presence of the ncAA (Figure 4e and i) and were used to retrieve the sequences of the corresponding MaPylRS mutants. These sequences were used to generate a WebLogo sequence map that represents the relative abundance of the observed amino acid residues at each of the three randomized positions within this selected mutant pool. We were encouraged to find that the most enriched sequences for the BocK-selection matched with the wild-type MaPylRS (Figure 4f). This is expected since BocK is an excellent substrate for wild-type MaPylRS. In contrast, the most enriched mutants for the mIF selection had a distinct sequence signature (Figure 4j). We used the following criteria to select six distinct MaPylRS mutants (Figure 4k, Figure S7) from this pool of enriched sequences for individually assessing the ability to charge mIF: (A) having multiple associated barcodes that show enrichment, (B) high average enrichment observed across the correlated barcodes, and (C) similarity to the consensus sequence, i.e., the amino acid residues most frequently observed in each of the randomized positions in the most enriched MaPylRS mutants. These six MaPylRS mutants were coexpressed in *E. coli* with $tRNA_{CUA}^{MaPyl}$ and assessed for their ability to express sfGFP-151-TAG in the presence and absence of mIF (Figure 4l). Gratifyingly, all six mutants exhibited robust reporter expression in the presence of mIF (between 20% and 50% wild-type sfGFP). Although four mutants (those containing VSV, LSV, AAV, and ASV at positions L125, N166, and V168, respectively) exhibited low activity in the absence of mIF, LAC and LAV mutants maintained significant (albeit significantly less than +mIF) reporter expression. ESI-MS analysis of the purified reporter protein expressed using the LSV MaPylRS mutant confirmed incorporation of only mIF, when it was supplied in the growth medium (Figure 4m). Although the MaPylRS LAC mutant showed significant reporter expression in the absence of any ncAA, it was found to selectively charge mIF when it was present during expression (Figure 4m). This behavior has been previously documented for other engineered aaRSs as well and may result from its superior affinity for the ncAA relative to the canonical counterpart.[58−60] We also isolated the reporter protein expressed using the MaPylRS LAC variant in the absence of mIF, and we observed a mass consistent with the incorporation of phenylalanine (Figure 4m).

Using the same MaPylRS library, we further applied the START protocol to identify MaPylRS mutants for charging ncAAs with two different ortho-substituted phenylalanine derivatives, *o*-nitro-L-phenylalanine (oNF) and *o*-cyano-L-phenylalanine (2-CNF). Following the same selection and analysis methods as described above, we identified TTW and AST as potential top candidates for charging oNF (Figure S8), and ASV was a potential top candidate for charging oCNF (Figure S9). All three MaPylRS mutants enabled robust expression of the sfGFP-151-TAG reporter in the presence of the appropriate ncAAs, and the ESI-MS analysis of the purified reporter protein confirmed selective incorporation of oNF (for TTW and AST) and oCNF (for ASV). These experiments further confirm that START can be used to identify novel aaRS mutants that accept noncanonical monomers as substrates from naïve libraries.

START Is Compatible with Non-α-Amino Acid Monomers. Since the selection system reported here relies solely on the ability of an aaRS to acylate its cognate tRNA, it should be applicable to develop mutant aaRSs that charge noncanonical monomers whose structures diverge from those of the α-amino acids. It has been shown that the pyrrolysyl synthetases such as MaPylRS do not strongly recognize the α-amino group of the native substrate.[9] This feature has enabled the use of this family of aaRSs to charge $tRNA^{Pyl}$ with multiple substrate analogs in which the α-amino group is replaced with other functionalities, including α-hydroxyacids, desamino-acids, β^2-hydroxyacids, etc., without further synthetase engineering.[10−12,24,26,29] Using this demonstrated polyspecificity of native MaPylRS, we sought to explore whether the core selection system described herein is also compatible with non-α-amino acid monomers. *E. coli* cells coexpressing MaPylRS and $tRNA^{MaPyl}$ were treated with BocK, OH-BocK, or H-BocK (Figure 5a) or in the absence of a potential substrate, and the resulting tRNA population was subjected to the tRNA-extension assay with or without a periodate treatment (Figure 5b). As expected, in the absence of any potential substrate, $tRNA^{MaPyl}$ was uncharged and was susceptible to periodate oxidation, which prevented tRNA extension. The presence of BocK, OH-BocK, and H-BocK each facilitated the formation of the tRNA-extension product even with periodate treatment, confirming that the charging of this noncanonical monomer

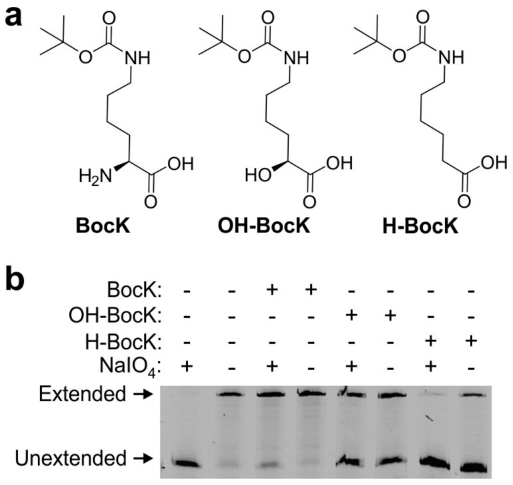


Figure 5. START scheme is compatible with non-α-amino acid monomers. (a) The structures of BocK, OH-BocK, and H-BocK. (b) Native PAGE followed by fluorescence imaging of tRNA-extension reaction performed on $tRNA^{MaPyl}$ coexpressed with MaPylRS in the presence of BocK, OH-BocK, or H-BocK, or in the absence of any cognate substrate, and with or without periodate treatment. Successful formation of the extension product in the presence of each substrate, which is a known substrate for MaPylRS, confirms the compatibility of this selection with non-α-amino acid monomers.

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX

protects the tRNA from periodate oxidation and that this strategy may be potentially used to identify aaRS mutants selective for structurally diverse noncanonical monomers.

CONCLUSION

There has been considerable interest in repurposing the mRNA-templated polypeptide synthesis by the ribosome to create novel sequence-defined polymers with noncanonical molecular architectures. However, translational incorporation of novel noncanonical monomers necessary for this purpose would require engineered variants of both aaRSs and the ribosome. Established aaRS engineering strategies are reliant on ribosomal translation and *vice versa*, thereby creating an interdependence that prevents systematic introduction of structurally novel noncanonical monomers into the genetic code. We offer a solution to this challenge by developing the first directed evolution strategy for engineering aaRSs, START, which does not rely on ribosomal translation. We show that tRNAs can be equipped with sequence barcodes, which can be correlated to individual mutants in large aaRS libraries through long-read, next-generation sequencing. Acylation of the correlated partner tRNA by a novel aaRS mutant protects it from periodate-mediated oxidation, allowing their subsequent enrichment. The identity of the corresponding aaRS mutant was then retrieved from the barcode sequence. The efficacy of this strategy was demonstrated by identifying novel mutants of MaPylRS selective for two ncAAs from a naïve library. The compatibility of this selection system with non-α-amino acids was further demonstrated by using the polyspecificity of MaPylRS for such substrates. It should be possible to extend the strategy described here to other aaRS/tRNA pairs. Although many aaRSs recognize the anticodon of its cognate tRNA as an identity element, which may complicate the introduction of a barcode in this region, it has been possible to engineer aaRS anticodon binding domains to be permissive for non-natural expanded anticodon sequences.[61,62] Furthermore, such sequence barcodes can also be inserted into alternative locations within the tRNA to avoid aaRS identity elements. Indeed, it has been shown that aaRSs can effectively acylate significantly altered versions of their cognate tRNAs, including severely miniaturized versions.[63−66]

While this manuscript was under review, a similar strategy named tRNA display was reported for translation-independent engineering of *M. barkeri* PylRS (MbPylRS), wherein the cognate $tRNA^{Pyl}$ was split at the anticodon region, and the MbPylRS mRNA was physically connected to the 3′-half of this split tRNA.[67] While physically connecting the aaRS mRNA to a split tRNA offers certain advantages—such as the ability to perform multiple rounds of selection—splitting the tRNA while maintaining aaRS recognition is a complex undertaking that might be challenging to extend beyond the pyrrolysyl pair. By contrast, inserting a sequence barcode into a tRNA is considerably simpler, as discussed above, and could be more readily generalized. Nonetheless, the success of these two related translation-independent selection methods further corroborates the robustness of the underlying selection strategy, which offers an exciting opportunity to develop aaRS mutants to charge structurally divergent noncanonical monomers for the ribosomal synthesis of evolvable, sequence-defined polymers with unprecedented structure and function.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.3c01557.

Experimental methods, nucleotide sequences, supplementary figures and tables (PDF)

AUTHOR INFORMATION

Corresponding Author

Abhishek Chatterjee − *Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, United States;* orcid.org/0000-0002-6231-5302; Email: abhishek.chatterjee@bc.edu

Authors

Chintan Soni − *Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, United States*

Noam Prywes − *Innovative Genomics Institute, University of California, Berkeley, California 94720, United States; Howard Hughes Medical Institute, University of California, Berkeley, California 94720, United States*

Matthew Hall − *Department of Biology, Boston College, Chestnut Hill, Massachusetts 02467, United States;* orcid.org/0000-0003-1260-9203

Malavika A. Nair − *Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, United States*

David F. Savage − *Innovative Genomics Institute, University of California, Berkeley, California 94720, United States; Howard Hughes Medical Institute and Department of Molecular and Cellular Biology, University of California, Berkeley, California 94720, United States;* orcid.org/0000-0003-0042-2257

Alanna Schepartz − *Department of Molecular and Cellular Biology, Department of Chemistry, and California Institute for Quantitative Biosciences, University of California, Berkeley, California 94720, United States; Chan Zuckerberg Biohub, San Francisco, California 94158, United States; ARC Institute, Palo Alto, California 94304, United States;* orcid.org/0000-0003-2127-3932

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.3c01557

Notes

The authors declare the following competing financial interest(s): AC and CS are co-inventors on a patent application that incorporates methods outlined in this manuscript. AC is a co-founder and senior advisor at BrickBio, Inc.

ACKNOWLEDGMENTS

This work was supported by the NSF Center for Genetically Encoded Materials (C-GEM; CHE 2002182). We thank Prof. Huiqing Zhou and Mitch Tepe (Boston College) for helpful discussions and assistance with the RNA workflow.

REFERENCES

(1) Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **2017**, *550* (7674), 53−60.

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

(3) Young, D. D.; Schultz, P. G. Playing with the molecules of life. *ACS Chem. Biol.* **2018**, *13* (4), 854−870.
(4) Icking, L.-S.; Riedlberger, A. M.; Krause, F.; Widder, J.; Frederiksen, A. S.; Stockert, F.; Spädt, M.; Edel, N.; Armbruster, D.; et al. iNClusive: a database collecting useful information on non-canonical amino acids and their incorporation into proteins for easier genetic code expansion implementation. *Nucleic Acids Res.* **2024**, *52*, D476−D482.
(5) Vargas-Rodriguez, O.; Sevostyanova, A.; Söll, D.; Crnković, A. Upgrading aminoacyl-tRNA synthetases for genetic code expansion. *Curr. Opin. Chem. Biol.* **2018**, *46*, 115−122.
(6) Manandhar, M.; Chun, E.; Romesberg, F. E. Genetic Code Expansion: Inception, Development, Commercialization. *J. Am. Chem. Soc.* **2021**, *143* (13), 4859−4878.
(7) Sun, S. B.; Schultz, P. G.; Kim, C. H. Therapeutic applications of an expanded genetic code. *Chembiochem: a European journal of chemical biology* **2014**, *15* (12), 1721−9.
(8) Guo, J.; Wang, J.; Anderson, J. C.; Schultz, P. G. Addition of an α-Hydroxy Acid to the Genetic Code of Bacteria. *Angew. Chem., Int. Ed.* **2008**, *47* (4), 722−725.
(9) Kobayashi, T.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. Recognition of non-alpha-amino substrates by pyrrolysyl-tRNA synthetase. *Journal of molecular biology* **2009**, *385* (5), 1352−60.
(10) Li, Y. M.; Yang, M. Y.; Huang, Y. C.; Li, Y. T.; Chen, P. R.; Liu, L. Ligation of expressed protein α-hydrazides via genetic incorporation of an α-hydroxy acid. *ACS Chem. Biol.* **2012**, *7* (6), 1015−22.
(11) Bindman, N. A.; Bobeica, S. C.; Liu, W. R.; van der Donk, W. A. Facile Removal of Leader Peptides from Lanthipeptides by Incorporation of a Hydroxy Acid. *J. Am. Chem. Soc.* **2015**, *137* (22), 6975−8.
(12) Spinck, M.; Piedrafita, C.; Robertson, W. E.; Elliott, T. S.; Cervettini, D.; de la Torre, D.; Chin, J. W. Genetically programmed cell-based synthesis of non-natural peptide and depsipeptide macrocycles. *Nature Chem.* **2023**, *15* (1), 61−69.
(13) Katoh, T.; Iwane, Y.; Suga, H. Logical engineering of D-arm and T-stem of tRNA that enhances d-amino acid incorporation. *Nucleic Acids Res.* **2017**, *45* (22), 12601−12610.
(14) Katoh, T.; Suga, H. Ribosomal Elongation of Cyclic γ-Amino Acids using a Reprogrammed Genetic Code. *J. Am. Chem. Soc.* **2020**, *142* (11), 4965−4969.
(15) Adaligil, E.; Song, A.; Cunningham, C. N.; Fairbrother, W. J. Ribosomal Synthesis of Macrocyclic Peptides with Linear γ(4)- and β-Hydroxy-γ(4)-amino Acids. *ACS Chem. Biol.* **2021**, *16* (8), 1325−1331.
(16) Fujino, T.; Goto, Y.; Suga, H.; Murakami, H. Ribosomal Synthesis of Peptides with Multiple β-Amino Acids. *J. Am. Chem. Soc.* **2016**, *138* (6), 1962−9.
(17) Sando, S.; Abe, K.; Sato, N.; Shibata, T.; Mizusawa, K.; Aoyama, Y. Unexpected preference of the E. coli translation system for the ester bond during incorporation of backbone-elongated substrates. *J. Am. Chem. Soc.* **2007**, *129* (19), 6180−6.
(18) Lee, J.; Schwarz, K. J.; Kim, D. S.; Moore, J. S.; Jewett, M. C. Ribosome-mediated polymerization of long chain carbon and cyclic amino acids into peptides in vitro. *Nat. Commun.* **2020**, *11* (1), 4304.
(19) Katoh, T.; Suga, H. Ribosomal Elongation of Aminobenzoic Acid Derivatives. *J. Am. Chem. Soc.* **2020**, *142* (39), 16518−16522.
(20) Katoh, T.; Suga, H. Consecutive Ribosomal Incorporation of α-Aminoxy/α-Hydrazino Acids with l/d-Configurations into Nascent Peptide Chains. *J. Am. Chem. Soc.* **2021**, *143* (45), 18844−18848.
(21) Takatsuji, R.; Shinbara, K.; Katoh, T.; Goto, Y.; Passioura, T.; Yajima, R.; Komatsu, Y.; Suga, H. Ribosomal Synthesis of Backbone-Cyclic Peptides Compatible with In Vitro Display. *J. Am. Chem. Soc.* **2019**, *141* (6), 2279−2287.
(22) Ad, O.; Hoffman, K. S.; Cairns, A. G.; Featherston, A. L.; Miller, S. J.; Söll, D.; Schepartz, A. Translation of Diverse Aramid- and 1,3-Dicarbonyl-peptides by Wild Type Ribosomes in Vitro. *ACS central science* **2019**, *5* (7), 1289−1294.
(23) Lee, J.; Coronado, J. N.; Cho, N.; Lim, J.; Hosford, B. M.; Seo, S.; Kim, D. S.; Kofman, C.; Moore, J. S.; Ellington, A. D.; Anslyn, E. V.; Jewett, M. C. Ribosome-mediated biosynthesis of pyridazinone oligomers in vitro. *Nat. Commun.* **2022**, *13* (1), 6322.
(24) Fricke, R.; Swenson, C. V.; Roe, L. T.; Hamlish, N. X.; Shah, B.; Zhang, Z.; Ficaretta, E.; Ad, O.; Smaga, S.; Gee, C. L.; Chatterjee, A.; Schepartz, A. Expanding the substrate scope of pyrrolysyl-transfer RNA synthetase enzymes to include non-α-amino acids in vitro and in vivo. *Nature Chem.* **2023**, *15* (7), 960−971.
(25) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. Enhanced D-amino acid incorporation into protein by modified ribosomes. *J. Am. Chem. Soc.* **2003**, *125* (22), 6616−7.
(26) Melo Czekster, C.; Robertson, W. E.; Walker, A. S.; Söll, D.; Schepartz, A. In Vivo Biosynthesis of a β-Amino Acid-Containing Protein. *J. Am. Chem. Soc.* **2016**, *138* (16), 5194−7.
(27) Maini, R.; Dedkova, L. M.; Paul, R.; Madathil, M. M.; Chowdhury, S. R.; Chen, S.; Hecht, S. M. Ribosome-Mediated Incorporation of Dipeptides and Dipeptide Analogues into Proteins in Vitro. *J. Am. Chem. Soc.* **2015**, *137* (35), 11206−11209.
(28) Chen, S.; Ji, X.; Gao, M.; Dedkova, L. M.; Hecht, S. M. In Cellulo Synthesis of Proteins Containing a Fluorescent Oxazole Amino Acid. *J. Am. Chem. Soc.* **2019**, *141* (14), 5597−5601.
(29) Hamlish, N. X.; Abramyan, A. M.; Schepartz, A. Incorporation of multiple β2-backbones into a protein in vivo using an orthogonal aminoacyl-tRNA synthetase. *bioRxiv* 2023, 2023.11.07.565973.
(30) Liu, D. R.; Schultz, P. G. Progress toward the evolution of an organism with an expanded genetic code. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (9), 4780−5.
(31) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the genetic code of Escherichia coli. *Science (New York, N.Y.)* **2001**, *292* (5516), 498−500.
(32) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. An expanded eukaryotic genetic code. *Science (New York, N.Y.)* **2003**, *301* (5635), 964−7.
(33) Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol.* **2002**, *20* (10), 1044−1048.
(34) Ellefson, J. W.; Meyer, A. J.; Hughes, R. A.; Cannon, J. R.; Brodbelt, J. S.; Ellington, A. D. Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol.* **2014**, *32* (1), 97−101.
(35) Bryson, D. I.; Fan, C.; Guo, L. T.; Miller, C.; Söll, D.; Liu, D. R. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat. Chem. Biol.* **2017**, *13* (12), 1253−1260.
(36) Suzuki, T.; Miller, C.; Guo, L. T.; Ho, J. M. L.; Bryson, D. I.; Wang, Y. S.; Liu, D. R.; Söll, D. Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat. Chem. Biol.* **2017**, *13* (12), 1261−1266.
(37) Fischer, J. T.; Söll, D.; Tharp, J. M. Directed Evolution of Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase Generates a Hyperactive and Highly Selective Variant. *Frontiers in molecular biosciences* **2022**, *9*, 850613.
(38) Stieglitz, J. T.; Van Deventer, J. A. High-Throughput Aminoacyl-tRNA Synthetase Engineering for Genetic Code Expansion in Yeast. *ACS Synth. Biol.* **2022**, *11* (7), 2284−2299.
(39) Hohl, A.; Karan, R.; Akal, A.; Renn, D.; Liu, X.; Ghorpade, S.; Groll, M.; Rueping, M.; Eppinger, J. Engineering a Polyspecific Pyrrolysyl-tRNA Synthetase by a High Throughput FACS Screen. *Sci. Rep.* **2019**, *9* (1), 11971.
(40) Dyer, J. R. Use of periodate oxidations in biochemical analysis. *Methods of biochemical analysis* **1956**, *3*, 111−52.
(41) Dittmar, K. A.; Sørensen, M. A.; Elf, J.; Ehrenberg, M.; Pan, T. Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO reports* **2005**, *6* (2), 151−7.
(42) Evans, M. E.; Clark, W. C.; Zheng, G.; Pan, T. Determination of tRNA aminoacylation levels by high-throughput sequencing. *Nucleic Acids Res.* **2017**, *45* (14), e133−e133.
(43) Erber, L.; Hoffmann, A.; Fallmann, J.; Betat, H.; Stadler, P. F.; Mörl, M. LOTTE-seq (Long hairpin oligonucleotide based tRNA high-throughput sequencing): specific selection of tRNAs with 3′-

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX

CCA end for high-throughput sequencing. *RNA biology* **2020**, *17* (1), 23−32.
(44) Cervettini, D.; Tang, S.; Fried, S. D.; Willis, J. C. W.; Funke, L. F. H.; Colwell, L. J.; Chin, J. W. Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase-tRNA pairs. *Nat. Biotechnol.* **2020**, *38* (8), 989−999.
(45) Passarelli, M. C.; Pinzaru, A. M.; Asgharian, H.; Liberti, M. V.; Heissel, S.; Molina, H.; Goodarzi, H.; Tavazoie, S. F. Leucyl-tRNA synthetase is a tumour suppressor in breast cancer and regulates codon-dependent translation dynamics. *Nat. Cell Biol.* **2022**, *24* (3), 307−315.
(46) Davidsen, K.; Sullivan, L. B. A robust method for measuring aminoacylation through tRNA-Seq. eLife Sciences Publications, Ltd, 2023.
(47) Tsukamoto, Y.; Nakamura, Y.; Hirata, M.; Sakate, R.; Kimura, T. i-tRAP (individual tRNA acylation PCR): A convenient method for selective quantification of tRNA charging. *RNA (New York, N.Y.)* **2023**, *29* (1), 111−22.
(48) Willis, J. C. W.; Chin, J. W. Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nature Chem.* **2018**, *10* (8), 831−837.
(49) Wan, W.; Tharp, J. M.; Liu, W. R. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochimica et biophysica acta* **2014**, *1844* (6), 1059−70.
(50) Nozawa, K.; O'Donoghue, P.; Gundllapalli, S.; Araiso, Y.; Ishitani, R.; Umehara, T.; Söll, D.; Nureki, O. Pyrrolysyl-tRNA synthetase-tRNA(Pyl) structure reveals the molecular basis of orthogonality. *Nature* **2009**, *457* (7233), 1163−7.
(51) Seki, E.; Yanagisawa, T.; Kuratani, M.; Sakamoto, K.; Yokoyama, S. Fully Productive Cell-Free Genetic Code Expansion by Structure-Based Engineering of Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase. *ACS Synth. Biol.* **2020**, *9* (4), 718−732.
(52) Ambrogelly, A.; Gundllapalli, S.; Herring, S.; Polycarpo, C.; Frauer, C.; Söll, D. Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (9), 3141−6.
(53) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli. *Biochemistry* **2013**, *52* (10), 1828−1837.
(54) Zheng, Y.; Addy, P. S.; Mukherjee, R.; Chatterjee, A. Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cells. *Chemical science* **2017**, *8* (10), 7211−7217.
(55) Niu, W.; Schultz, P. G.; Guo, J. An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS Chem. Biol.* **2013**, *8* (7), 1640−5.
(56) Dunkelmann, D. L.; Willis, J. C. W.; Beattie, A. T.; Chin, J. W. Engineered triply orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids. *Nature Chem.* **2020**, *12* (6), 535−544.
(57) Jewel, D.; Kelemen, R. E.; Huang, R. L.; Zhu, Z.; Sundaresh, B.; Cao, X.; Malley, K.; Huang, Z.; Pasha, M.; Anthony, J.; van Opijnen, T.; Chatterjee, A. Virus-assisted directed evolution of enhanced suppressor tRNAs in mammalian cells. *Nat. Methods* **2023**, *20* (1), 95−103.
(58) Young, D. D.; Young, T. S.; Jahnz, M.; Ahmad, I.; Spraggon, G.; Schultz, P. G. An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry* **2011**, *50* (11), 1894−900.
(59) Italia, J. S.; Addy, P. S.; Wrobel, C. J. J.; Crawford, L. A.; Lajoie, M. J.; Zheng, Y.; Chatterjee, A. An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. *Nat. Chem. Biol.* **2017**, *13* (4), 446−450.
(60) Italia, J. S.; Latour, C.; Wrobel, C. J. J.; Chatterjee, A. Resurrecting the Bacterial Tyrosyl-tRNA Synthetase/tRNA Pair for Expanding the Genetic Code of Both E. coli and Eukaryotes. *Cell chemical biology* **2018**, *25* (10), 1304−1312.
(61) Chatterjee, A.; Xiao, H.; Schultz, P. G. 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.* **2012**, *109* (37), 14841−6.
(62) Neumann, H.; Slusarczyk, A. L.; Chin, J. W. De Novo Generation of Mutually Orthogonal Aminoacyl-tRNA Synthetase/tRNA Pairs. *J. Am. Chem. Soc.* **2010**, *132* (7), 2142−2144.
(63) Buechter, D. D.; Schimmel, P.; de Duve, C. Aminoacylation of RNA Minihelices: Implications for tRNA Synthetase Structural Design and Evolution. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28* (4), 309−322.
(64) Frugier, M.; Florentz, C.; Giegé, R. Efficient aminoacylation of resected RNA helices by class II aspartyl-tRNA synthetase dependent on a single nucleotide. *EMBO journal* **1994**, *13* (9), 2218−26.
(65) Shiba, K.; Ripmaster, T.; Suzuki, N.; Nichols, R.; Plotz, P.; Noda, T.; Schimmel, P. Human alanyl-tRNA synthetase: conservation in evolution of catalytic core and microhelix recognition. *Biochemistry* **1995**, *34* (33), 10340−9.
(66) Martinis, S. A.; Schimmel, P. Small RNA Oligonucleotide Substrates for Specific Aminoacylations. In *tRNA: Structure, Biosynthese, and Function*; Wiley, 1994; pp 349−370.
(67) Dunkelmann, D. L.; Piedrafita, C.; Dickson, A.; Liu, K. C.; Elliott, T. S.; Fiedler, M.; Bellini, D.; Zhou, A.; Cervettini, D.; Chin, J. W. Adding α,α-disubstituted and β-linked monomers to the genetic code of an organism. *Nature* **2024**, *625* (7995), 603−610.

https://doi.org/10.1021/acscentsci.3c01557
ACS Cent. Sci. XXXX, XXX, XXX−XXX