nature portfolio

https://doi.org/10.1038/s41586-026-10949-y

Supplementary information

Automated prototyping of genetic codes

In the format provided by the authors and unedited

Supplementary Information — Table of Contents

Supplementary Discussion 1 | Development of a multiplexed and automated method to measure tRNA aminoacylation

Traditionally, purified aaRS proteins and in vitro transcribed tRNAs were incubated in vitro to assess the degree of aminoacylation by a specific amino acid[65-67]. More recently, periodate oxidation and next generation sequencing (NGS)[44,68] was applied to globally quantify the percentage of each isoacceptor tRNA that is aminoacylated from an in vivo pool of tRNAs. Given the anticipated challenges of engineering otRNAs, we required a multiplexed method to measure tRNA aminoacylation with on-demand control of tRNA modifications (**Main text**).

We reasoned that instead of cloning libraries of tRNA genes into cells and subsequently lysing the cellular populations to sequence their tRNAs[44,68], we could instead use the cell lysate of *E. coli* and in cell-free conditions recapitulate the charging of in vitro produced tRNAs by the cell's repertoire of aaRSs. Especially as the production of tRNAs is highly conserved to maintain a CCA end, we predicted that avoiding challenges of in vivo tRNA production could endow us with an unbiased methodology to study tRNA aminoacylation for non-natural tRNAs along with their modification context. We hypothesized that in vitro transcription of tRNAs and introduction into cell-free lysates from *E. coli* would allow for recapitulation of the native tRNA modifications and interactions with aaRSs as in vivo.

We began by adapting previous in vivo tRNAseq methods to cell-free translation systems and in vitro transcribed (IVT) tRNAs. Furthermore, existing tRNAseq protocols involve many manual steps that have so far precluded their ready adoption to synthetic tRNA libraries. We thus chose to automate the tRNA sequencing protocol for the first time. Open-source robotic liquid handling systems such as Opentrons[69] are enabling advances in synthetic biology by allowing rapid automation of procedures coupled to computational analysis. This permits scaling of throughput for multiplexed procedures and increased reproducibility[70]. In this method, which we call tRNA-Sequencing of Charging by Automated NGS (tSCAN), we sought to develop an improved and automated method to read aminoacylation status of tRNAs using in vitro transcribed tRNAs in cell-free translation systems (**Main text, Fig. 2a-b**).

We constructed a universal IVT tRNA design from chip-synthesized oligo pools employing a T7 promoter. In order to produce tRNAs without non-templated nucleotides[71,72] at their 3' ends we inserted an HDV ribozyme[61] after the end of the tRNA sequence (**Supplementary Fig. 1**). Previously magnetic beads have been used for RNA purification and size selection[73-76]. We adapted SPRIselect[77] beads and performed size selection on IVT purified tRNAs. We obtained robust purifications of tRNA at ~4µg/µL after bead purification (**Supplementary Fig. 2**). We automated both IVT reaction setup and the purification procedure on the Opentrons T2 (OT-2) using the temperature control and magnetic bead purification modules and obtained similar size distributions (**Supplementary Fig. 3**) and yields of ~4µg/µL(**Supplementary Fig. 2**).

Once we validated automated production and purification of tRNAs we next proceeded to test whether we could apply tSCAN to measure the aminoacylation of IVT tRNAs in *E. coli* cell lysate. We chose to examine the degree of orthogonality of exogenous tRNAs to the *E. coli* translation system, as use of native *E. coli* tRNAs would not be easily distinguishable from the WT tRNAs in a cell-free lysate. We synthesized all tRNAs from *Methanomethylophilus alvus* (*M. alvus*) and *Methanosarcina mazei* (*M. mazei*), two species that are frequently used for genetic code expansion[45-47,78], as chip-synthesized oligo pools[48] and in vitro transcribed and

purified the tRNAs (**Methods**). We incubated ~80µg of the IVT tRNAs into *E. coli* cell-free translation system (**Methods**) at 37ºC for 2 hrs. This was lower than the concentration of native tRNAs in the *E. coli* lysate[79] due to volume constraints in the composition of the reaction, but in high excess of *E. coli* aaRSs, allowing for similar reaction kinetics as would occur *in vivo* [51,80]. We then extracted the tRNA (**Methods**) and performed validation of tSCAN automated on the OT-2 fluidic system compared to manual execution of all steps.

In the design of the automated procedure we wrote custom Python code to control all steps on the OT-2 instrument and split up the main unit operations of the protocol into reusable steps (**Fig. 2a).** We utilized separation between data representation and machine control for all reaction steps. All unit operations were divided into reaction setup steps or purification steps, from which we inherited the individual instances for each step of the procedure. We performed periodate oxidation, and β-elimination. We then performed 3' adapter ligation and reverse transcription (**Fig. 2a-b**, **Supplementary Fig. 4a, Supplementary Fig. 5**). SPRIselect magnetic bead purification was used at all steps. The 5' end of the cDNA was ligated with DNA adapters. Sequencing adapters were added by PCR and sequencing was performed by Illumina MiSeq. As we performed parallel reaction steps on the OT-2 alongside manual procedures we compared all steps using Agilent TapeStation, and finally an agarose gel for the dsDNA barcoded library. The manual and automated procedures were highly similar (**Supplementary Fig. 5a-d**).

The tRNAs that are aminoacylated with an amino acid are protected from periodate oxidation and will appear with the CCA 3' end intact, whereas tRNAs that are not aminoacylated will lose the terminal adenosine. By quantifying ratios of intact to truncated tRNA reads for each tRNA species, it is possible to quantify the percentage of aminoacylation for that tRNA. We found that as expected, the $tRNA^{Pyl}$ from *M. alvus* and from *M. mazei,* respectively, showed low tRNA aminoacylation efficiency, which is in agreement with them being bio-orthogonal, or minimally cross-reactive, with the *E. coli* translation system (**Fig. 2c-d**). Intriguingly, we found that additional tRNAs showed low aminoacylation, including $tRNA^{Ala}$ in *M. alvus* (**Fig. 2c**), and $tRNA^{Leu}$ in *M. mazei* (**Fig. 2d**), which had not previously been reported as orthogonal, but may be additional orthogonal tRNAs in *E. coli*. The majority of the *M. alvus* and *M. mazei* tRNAs were aminoacylated by the *E. coli* lysate (**Fig. 2c,d**). As a control to determine the frequency of loss of terminal adenosine in the tRNA population before periodate oxidation, we also performed tSCAN without the periodate oxidation and β-elimination steps on an in vitro transcribed pool of *E. coli* tRNAs (**Supplementary Fig. 6a**) and *M. mazei* tRNAs (**Supplementary Fig. 6b**). We found that the terminal adenosine was maintained, in contrast to the full tSCAN procedure (**Fig. 2a-b**).

Supplementary Discussion 2 | tRNA aminoacylation in lysate and PURE

Combining our analysis of all tRNA isoacceptors with variation of their 3' CCA end (**Extended Data Figs. 2-4),** we plotted the aminoacylation of each tRNA species depending on whether the tRNAs were incubated in the PURE system and had no tRNA modifications (**Fig. 3a-b**), or in cell lysate where they could be modified (**Fig. 3c-d**). We further divided the tRNAs by whether their corresponding aaRS is a Class I aaRS (**Fig. 3a,c**) or a Class II aaRS (**Fig. 3b,d**).

We observed that tRNAs aminoacylated by Class II aaRSs in general showed lower aminoacylation levels. This was in particular evident with WT serine and threonine tRNAs (**Fig. 3b, d**). This was consistent with previous reports[49] of aminoacylation levels of serine and threonine in *E. coli* cells. Interestingly, when the WT CCA end was altered to CGA, CUA, or CAA in serine and threonine tRNAs that were incubated in lysate (and had modifications), the aminoacylation levels increased (**Fig. 3d**) but this pattern was not observed in most unmodified tRNAs (**Fig. 3b**). There was an overall reduced aminoacylation frequency in many unmodified tRNAs compared to tRNAs incubated in lysate, and this pattern was more pronounced with Class II aaRSs (**Fig. 3b**).

Within the unmodified tRNAs there were pronounced differences between tRNAs aminoacylation levels depending on the terminal 3' sequence. For example, cysT was active with a CCA end (16.5%), but poorly active with any other 3' sequence (4-7%) (**Fig. 3a**). LeuW showed high aminoacylation with CCA, CUA, CAA, but low aminoacylation with CGA (**Fig. 3a**). GlyT had high activity only with a CCA end but no other sequence (**Fig. 3b**). However cysT and ileY with 3' CCA ends showed lower aminoacylation in lysate compared to cysT and ileY with alternative 3' ends.

We performed modeling with AlphaFold3[81] in order to gain insight into why in the unmodified tRNAs there was such a contrast between CCA and CGA 3' sequences. We modeled *E. coli* LeuRS (a class I aaRS) with tRNA leuW-CCA (**Extended Data Fig. 7a**) or with tRNA leuW-CGA (**Extended Data Fig. 7b**). When tRNA leuW contains a CGA 3' end, there is a steric hinderance between adenosine 75 and guanosine 74 of the tRNA with proline 105 of the aaRS. This hinderance is not present in the WT leuW tRNA where cytosine is at position 84. We also modeled GlyRS, a class II aaRS) with tRNA glyT-CCA (**Extended Data Fig. 7c**) or with tRNA glyT-CGA (**Extended Data Fig. 7d).** π-π stacking is observed between Adenosine 75 of WT tRNA glyT and tyrosine 80 of glyRS. A similar interaction is also present in a number of other aaRS-tRNA pairs[82]. In glyT-CGA, this interaction is not observed, which may destabilize the tRNA-aaRS interaction for proper aminoacylation.

In summary, these data provided a picture of lysate improving the aminoacylation of most otRNAs. As the average concentration of *E. coli* aaRSs in PURE is below 1ng/μL, over a thousand times less than the tRNA concentration[50], and similar to the ratio of tRNAs to aaRSs in lysate[51,80] we were confident that we were measuring meaningful enzymatic differences in aaRS-tRNA interactions in each translation system context (*i.e.* lysate vs. PURE) and in the context of each 3' tRNA sequence.

While these data (**Fig 3**) strengthened our hypothesis of modifications of tRNAs impacting their aminoacylation, we wished to confirm that the otRNAs were indeed aminoacylated with their cognate amino acids in lysate, and rule out artefacts of tRNA extraction from two different in vitro translation systems. For this we needed to determine the identity of amino acids aminoacylated to tRNAs. As the amino acid identity is not captured in an NGS-based analysis of tRNA charging, we turned to high-resolution liquid chromatography mass spectrometry (LC-MS)[66,83]. We purified *E. coli* Leu, Pro, and Val aaRSs and performed in vitro charging with their corresponding IVT tRNAs (leuW, proK, and valT, respectively) containing either CCA or CGA 3' ends. After performing in vitro aminoacylation reactions (**Methods**), we performed RNase A digestion to form a terminal aminoacyl adenosine (**Extended Data Fig. 5a, top**) which can allow for detection of the identity of the amino acid charged to the tRNA by LC-MS. We observed that all the CCA-tRNAs were aminoacylated by their corresponding aaRSs (**Extended Data Fig. 5b**), but for CGA-otRNAs the valT tRNA was aminoacylated, but we did not detect aminoacylation of the leuW or proK tRNAs (**Extended Data Fig. 5b**). This was consistent with our tSCAN measurements in PURE, where valT was 61% charged, while leuW was 17% charged, and proK was 12.8% charged.

Our above in vitro aminoacylation experiments relied on in vitro purified aaRSs which is a condition analogous to the PURE system. But we would need to examine the identity of amino acids charged to tRNAs in the lysate directly to know if under those conditions the charged amino acids indeed corresponded to the cognate amino acid for that tRNA-aaRS pair. As far as we know, there is no method that can direct identify an amino acid esterified to a tRNA whose natural counterpart differs by only one base. LC-MS with RNAse A digestion was used as a method to assess identity of amino acids esterified to tRNAs, as has been used previously[66,83]. However, a mixture of tRNAs is present within the lysate, which makes it impossible to determine which amino acid is being charged onto which tRNA. Using biotin or other probing techniques for tRNAs[84,85] would also not be feasible here as the tRNAs with CCA and CGA ends share the same tRNA sequence. We then proposed to transcribe the tRNAs that are placed into the lysate having an isotopic label at each adenosine by employing N^{15}-ATP during *in vitro* transcription of the tRNA pool. Bulk digestion of tRNAs with the target tRNA being isotopically labeled can enable detection of the aminoacylation status specifically of the tRNA species of interest (**Extended Data Fig. 6a-b**). We called this method tRNA-Species Charging ANalysis by Mass (tSCAN-M). We *in vitro* transcribed CCA-tRNAs and CGA-tRNAs and incubated them within the lysate-based cell-free translation system. We then added RNase A directly to the lysate to digest all tRNAs and precipitated proteins with 1% formic acid followed by LC-MS (**Supplementary Extended Data Fig. 6a Extended Data Fig. 5a, bottom**). We found that indeed aminoacylation is restored in lysate, compared to in vitro aminoacylation reactions (**Extended Data Fig. 5c).** The $tRNA^{Val}$ was aminoacylated in both lysate and in vitro aminoacylation reactions having CCA or CGA end (**Extended Data Fig. 5b-c, right panel**) but the $tRNA^{Leu}$ and $tRNA^{Pro}$ were aminoacylated only with a CCA end in the vitro aminoacylation reactions (**Extended Data Fig. 5b, left panel**). In contrast both CGA and CCA tRNAs were aminoacylated in lysate (**Extended Data Fig. 5c**). These data are consistent with our tSCAN results, where lewW, proK, and valT tRNAs were almost completely aminoacylated in lysate (92%, 75%, and 95%, respectively). As a result of these experiments, we confirmed that otRNAs with 3' CGA ends are indeed aminoacylated with their cognate amino acids in an *E. coli* lysate

system. In contrast to purified translation system components (*i.e.* PURE), the entire translation system is needed to perform this activity.

Supplementary Discussion 3 | Orthogonal translation validation

We synthesized and transcribed a minimal otRNA pool (MP1) with 20 tRNAs for sense codons and one initiator methionine tRNA (**Fig. 5b**, **Extended Data Fig. 8a**) having 3' CGA ends and performed tests of crosstalk with WT ribosomes. First, we tested with tSCAN whether the alanine and serine tRNAs with CGA terminal trinucleotides with swapped anticodons were aminoacylated with the same efficiency as their non-codon-swapped counterparts. We found equivalent aminoacylation for both tRNAs (**Supplementary Fig. 8**). We designed peptide reporters that can be robustly detected by LC-MS/MS[10] to determine whether alanine or serine was incorporated at three positions within the peptide. To validate in vitro production of our reporter peptides, we performed translation in an *E. coli* cell-free translation system and included the dsDNA template for MHHHHHHSKVPGAGVPGAGVPGAGKGGG[10,12] (ELP-AAA). We affinity purified the peptide on Ni-NTA beads and trypsin digested and analyzed by LC-MS/MS (**Methods**). The trypsin-digested product, GPGKVPGAGVPGAGVPGAGK was clearly detected by LC-MS/MS (**Supplementary Fig. 9a**). Similarly, to determine if this peptide would be readily detected by LC-MS/MS if a serine to alanine substitution occurred, we included the dsDNA for the codon-swapped version of peptide ELP-AAA, MHHHHHHAKVPGSGVPGSGVPGSGKGGG (ELP-SSS) in our cell-free translation system and also detected the GPGKVPGSGVPGSGVPGSGK peptide (**Supplementary Fig. 9b**).

We next applied these peptides for tests of crosstalk with WT ribosomes using LC-MS/MS. We introduced our minimal otRNA pool MP1 into cell-free translation reactions containing double-stranded DNA (dsDNA) templates encoding either the ELP-AAA or swapped ELP-SSS peptides and performed automated translation reactions and peptide purifications. We detected the same peptide motif across conditions to compare levels of alanine or serine incorporation at their respective positions. In the event of WT ribosomes translating the peptide with native tRNAs, VPGAGVPGAGVPGAGK should be detected from the ELP-AAA peptide dsDNA template, but if tRNAs with a CGA end are accepted for translation, then VPGSGVPGSGVPGSGK should be detected. Using dsDNA for the ELP-AAA peptide as a template for translation, we detected VPGAGVPGAGVPGAGK (**Fig. 5c-d**) and when we performed translation with the no MP1 pool addition control (**Fig. 5e-f**). With translation carried out using dsDNA for the swapped ELP-SSS reporter template, we detected VPGSGVPGSGVPGSGK, but not VPGAGVPGAGVPGAGK or any alanine substitutions into the peptide (**Fig. 5g-h**). We saw equivalent results in the no MP1 pool addition control (**Fig. 5i-j**). In summary, we could not detect evidence of non-specific accommodation of otRNAs containing a CGA-end into the WT ribosome for translation.

After confirmation that otRNAs do not have crosstalk with the native translation system, we continued to AGENTEX tests of minimal genetic codes. Previously it was shown with a PURE *in vitro*-reconstituted translation system that tRNAs with CGA ends could be charged with flexizymes and accepted by ribosomes containing G2251C and G2553C mutations[33]. We reasoned that otRNAs charged by aaRSs should equally be employed as substrates for translation

by ribosomes having the G2251C and G2553C mutation (CGA ribosome). Production of lysates from a strain containing CGA ribosomes would eliminate the need for separate affinity purification of ribosomes, while enabling translation of alternative genetic codes in a lysate translation system and re-use of tRNAs via continuous enzymatic aminoacylation by native aaRSs.

We cloned a high-copy number plasmid containing the ribosomal 23S rRNA with G2251C and G2553C mutations, and produced an *E. coli* cell lysate from this strain. As a control we also generated an *E. coli* cell lysate with WT ribosomes, where the WT 23S rRNA gene was cloned onto the same plasmid backbone[17] (**Methods**) so as to have an equivalent quantity of WT ribosomes produced. We then tested whether we could use a 22-codon genetic code (MP2) for polypeptide translation with AGENTEX (**Main text**, **Extended Data Fig. 8b**).

We automated all steps of AGENTEX on the OT-2 robotic system. The compressed genetic codes were generated to specify a minimal set of codons for the canonical amino acids and ability to modify anticodons of tRNAs to maximize codon space. Automated designs of tRNAs were then chip-synthesized as an oligonucleotide pool. The in vitro transcription reactions and tRNA purifications were performed as before. The dsDNA templates were produced and purified with Ampure beads, and peptides were purified His-Tag Isolation and Pulldown magnetic beads (**Supplementary Fig. 10**). Reaction conditions for all enzymatic steps were input in reusable spreadsheets to automate all unit operations.

We developed a high-sensitivity luminescence assay to allow for rapid testing of recoded genetic codes with AGENTEX, based on the HiBiT peptide and complementation with LgBiT protein (**Main text, Fig. 6b**). We produced the dsDNA for engineered peptide h-TAG-1 from single stranded DNA (ssDNA) template and used h-TAG-1 dsDNA for translation in our custom cell-free system (**Fig. 6b**). To test the relative activity of lysates we produced with CGA and WT ribosomes (**Methods**), and the sensitivity of the h-TAG-1 reporter, we designed a peptide with a valine (GUA) codon replacing a UAG codon and no swapped codons for alanine and serine (h-GTA-1). We performed translation of h-GTA-1 compared to h-TAG-1 without UAG-suppression in both lysates, and also compared to no peptide control. We found that h-GTA-1 luminescence was similar in both CGA and WT ribosome lysate, indicating a similar translation efficiency in both translation systems (**Supplementary Fig. 11)**. The relative luminescence of h-GTA-1 was ~4x higher compared to h-TAG-1 and ~13x higher than no peptide dsDNA added into the translation reaction, a significant dynamic range that could allow sensitive determination of translation crosstalk. We next proceeded to test h-TAG-1 for tests of AGENTEX translation and crosstalk. All otRNA and template dsDNA production steps, as well as translation assays were automated with AGENTEX. There are two serine positions in HiBiT and h-TAG-1, which we designated with GCG (alanine) codons. The seryl tRNA in the MP2 pool (**Extended Data Fig. 8b**) contained a synthetic CGC anticodon which was predicted to decode the GCG codon but is not naturally found in *E. coli* tRNAs, but is found in human tRNA TRA-CGC4-1[86]. With successful translation using the MP2 code, the full HiBiT peptide should be translated, leading to luminescence upon LgBiT complementation.

We performed translation with or without pazF in lysates either with WT or CGA ribosomes. We observed that translation of the correct message (*i.e.* h-TAG-1 with expanded genetic code) is

more successful with the CGA ribosome and MP2 tRNA pool (**Main text, Fig. 6c**). However, due to the high background in the no-pazF condition, and CGA ribosome lysate compared to WT ribosome lysate we hypothesized that the pazFRS was mischarging the $tRNA^{Tyr}$ with canonical amino acids, as has been reported previously without a quantification of the nonspecific incorporation[87]. When we performed translation in both WT and CGA lysates the absence of pazFRS with or without pazF, the luminesce in all no-pazFRS conditions was not significantly different from background translation in the WT ribosome lysate (**Fig. 6c**), indicating that the high luminescence signal observed in CGA ribosome lysate (+pazF, +pazFRS) compared to WT ribosome lysate (+ pazF, + pazFRS) is indicative of translation of the correct genetic code (~6-fold signal over background, $P < 0.001$).

We then performed further high sensitivity LC-MS/MS experiments to determine if any crosstalk was present with CGA or WT ribosomes in AGENTEX (**Main text, Fig.6d-h**). We used the ZenoTOF7600+ mass spectrometer[53]for femtomolar sensitivity in peptide detection. The MHHHHHHAKGPGKVPGSGVPGAGVPGSGKGGG peptide (ELP-SAS) was used to determine whether two serines and one alanine (expanded genetic code) or two alanines and one serine (standard genetic code) were incorporated into the peptide (**Fig. 6d**). Also as this peptide is 32 amino acids in length, it would be a test of the ability for consecutive decoding of a peptide of this length solely by otRNAs. We used a cell lysate with CGA ribosomes and performed automated translation prototyping using AGENTEX in the presence or absence of the otRNA pool MP1 (**Extended Data Fig. 8a**). We detected the expected peptide corresponding to the expanded genetic code: two serines and one alanine (**Supplementary Fig. 12**) in the presence of pool MP1. To further measure crosstalk between WT and CGA ribosomes in AGENTEX, we compared translation of ELP-SAS using pool MP1, with lysates made with WT ribosomes compared to lysates made with CGA ribosomes. We monitored translation with LC-MS/MS across all experimental conditions and monitored peptide GVPG**J**GVPG**J**GK, where J = A or S, as this peptide was detected with high quality spectra in all experimental conditions, allowing an unbiased and sensitive assay for crosstalk (**Main text, Fig. 6d-f**).

Supplementary Discussion 4 | Alternative Genetic codes

Intriguingly we observed a marked difference between the Class I and Class II aaRSs (**Fig. 3**) in their ability to aminoacylate in vitro-transcribed tRNAs in the PURE system, where tRNA modifications are not present. In general, Class II aaRSs had lower activity for CCA and non-CCA tRNAs. There were a number of aaRSs that were able to efficiently charge CCA tRNAs but showed low activity for tRNAs with non-CCA ends, including LeuRS (Class I) and hisRS and proRS (Class II). These aaRSs do not recognize the anticodon of the tRNA[88-90], as is also the case for many Class II aaRSs, which rely more strongly on acceptor stem recognition for proper aminoacylation[91,92]. We hypothesize that alternative 3' end sequences can disrupt tRNA-aaRS recognition but is stabilized by RNA modifications present on the tRNA, resulting in near-universal charging of tRNAs by both Class I and Class II aaRSs in the native translation system (**Extended Data Fig. 2**). Our modeling of CCA and CGA tRNA recognition by their cognate aaRSs also revealed that disruption of tRNA base contacts in the amino acid binding pocket may reduce aminoacylation efficiency (**Extended Data Fig. 7**). However most crystal structures of tRNAs and aaRS have used in vitro transcribed, unmodified, tRNAs, and the impact of modifications to stabilize tRNA-aaRS contacts for most aaRSs is unclear. Further studies to

uncover the effect of tRNA modifications on aaRS-tRNA interactions may reveal the mechanism of selection for tRNAs with alternative CCA ends in the presence or absence of tRNA modifications.

We also observed that serine and threonine tRNAs with the natural CCA end are aminoacylated at a reduced frequency compared to other tRNAs (**Fig. 3b,d; Fig. 4**). This is in agreement with previous reports that serine and threonine tRNAs are charged at a substantially reduced level (below 10% for serine) in comparison to the tRNA repertoire[49]. This may be an adaptation to reduce cytotoxicity and may be mediated in serine by L-serine-deaminase.[49] Intriguingly, varying the first and second base of the CCA sequence increased serine and threonine aminoacylation levels (**Fig. 3d, Fig. 4**). This may be a consequence of the inability of these non-CCA tRNAs to interact with the 23S rRNA of the ribosome and participate in translation, thus there was not a selective pressure to evolve serine or threonine deacylation mechanisms in these tRNAs. This reinforces that the ribosome may be the control point of the genetic code via tRNA acceptor stem interactions. Those tRNAs whose 3' ends are not accepted by the ribosome do not participate in translation and thus their aminoacylation status does not perturb the cell's proteome.

There is extensive variation within and between organisms in their tRNA sequences, and variation in the RNA and protein sequence of the ribosome while still supporting life[93-95]. The genetic code, while nearly universal, also has a few notable exceptions that have been reported[96-99]. The CCA sequence at the end of the tRNA molecule, however, is universally conserved. This conservation suggests that the acceptor stem plays a crucial role in translation and has been subject to strong evolutionary pressure to maintain its sequence. Previously, it was assumed that tRNA-aaRS interactions were responsible for maintaining the conservation of the CCA sequence because of its placement proximal to the position of amino acid aminoacylation. However, our findings reveal that the ribosome is the primary interaction point for maintenance of this sequence, with secondary contributions from the aaRS aminoacylation pattern that is modulated by RNA modifications (**Extended Data Fig. 9**). There is evidence that the 23S rRNA in the ribosomal large subunit interacts with the CCA-end of tRNAs[7,8], which is required for proper decoding of the tRNA within the peptidyl transferase center and translation of polypeptides. According to our hypothesis, this feature has been conserved since before the last universal common organism (LUCA) and is essential for defining the genetic code.

Comparative-genomic reconstructions of LUCA suggest that the core translation system as found in modern organisms already existed, including at least 18 out of 20 aaRSs, several translation factors, at least 40 ribosomal proteins, and several rRNA and tRNA modification enzymes[100]. However, most likely the original ribosome and translation system were composed entirely of RNA in an RNA world[58,59,101-103] prior to aaRS evolution. It may be possible that ribosomes and tRNAs cooperated to specify a genetic code before the origin of DNA genomes[57], and that a vestige of this is found in the mutual interaction of the CCA end of tRNAs and the 23S rRNA. This reinforces previous models of the origin of the genetic code, where early tRNAs bound their cognate amino acids stereochemically for translation into polypeptides prior to evolution of aaRS enzymes[96]. Whether by this stereochemical model[58,96] or through a frozen accident[58,96,100], the modern genetic code is a multi-layered compromise between the RNA and protein worlds.

A number of recent studies have highlighted how proteomes of organisms can be stably[104] and heritably[105] maintained non-genomically. Our work suggests that the genetic code itself is maintained in a careful balance within the translation apparatus, in a multi-layered system that may predate DNA genomes (**Extended Data Fig. 9b**). As aaRSs are capable of charging tRNAs with variation of the 3' CCA sequence, we propose that the ribosome is the main filter for accommodation of CCA tRNAs, which could allow specialized ribosomes to translate parallel genetic codes within the same cell (**Main text**). However, there is selective pressure to not vary the standard proteome under most conditions. Furthermore, the covalent modifications of tRNA further modulate their recognition by the translation system. We found that multiple aaRSs altered their aminoacylation of tRNAs depending on whether RNA modifications were present, and this was further accentuated by mutations to the CCA sequence. Indeed, the extensive modifications of WT tRNAs may be necessary to maintain tight regulation of their function (**Extended Data Fig. 9b**). This nuanced aaRS-tRNA code may be a second layer that evolved to further protect the standard genetic code and prevent emergence of alternative proteomes.

Supplementary Discussion 5 | AGENTEX analysis and applications

The natural genetic code evolved with redundancies, but so far it has been a challenge to compress it to a minimal set of 20 tRNAs due to the limitations of genome recoding while maintaining cellular fitness. While a proposed evolutionary advantage of a degenerate genetic code minimizes the effect of point mutations during translation[106], with AGENTEX there is no longer the constraint of translation of the entire proteome and this opens new opportunities for the development of an expanded genetic alphabet using a parallel genetic code. At the same time, with AGENTEX the effects of different coding schemes can be studied one protein at a time, allowing fast optimization of genetic code design. Furthermore, a benefit of a compressed genetic code is potential reduction in translation errors due to decoding of each amino acid with fewer tRNAs, reducing the probability of accumulation of tRNA mutations or mis-acylation[107-109]. The use of lysate in AGENTEX allows not only prototyping of new genetic codes, but laboratory-scale insights can be rapidly ported to large-scale cell-free bioproduction, for which there is already industrial precedent[24,28]. In contrast, it is logistically-prohibitive to use PURE for industrial-scale production[110]. Also, PURE-based systems suffer from lack of multiplexability because IVT tRNAs without their natural RNA modifications are not optimally aminoacylated by aaRSs, as we comprehensively determined in this work (**Fig. 3, Extended Data Fig. 5**), and many in vivo codon-anticodon interactions dependent on RNA modifications[27]. As a result, it is challenging to study the effects of tRNA anticodons bearing their natural modifications on translation with PURE, and this would limit the insights from genetic code prototyping that can be directly ported into living systems.

Furthermore, the aminoacylation of otRNAs with 3' CGA ends is on average 1.5-fold more efficient in lysate than tRNAs with 3' CCA ends or 1.8-fold more efficient than otRNAs with 3' CGA ends in PURE **(Fig. 3)**. Compared to prior methods[34], AGENTEX permits codon reassignment without the requirement of removing all endogenous tRNAs, aminoacylating each tRNA species individually, or replacement of tRNAs in a translation system. This has particular promise for the synthesis of biopolymers composed of multiple distinct nsAAs. A procedure that would take a researcher a full week can be performed in a single day[34]. In principle this allows

from 21 (as we tested) to 34 tRNAs to be used in a translation system to encode a unique monomer for each tRNA molecule. This can allow up to 34 x 1.5 performance gain, allowing over a 50-fold efficiency improvement in constructing a single genetic code over previous in vitro approaches, not including the automation we use in this technology. The automation of DNA and RNA purifications and translation reactions increases throughput even further. Procedures that would take ~12 hrs of continuous effort by one researcher are reduced to 1-2 hrs of hands-on time. This further provides a 6-fold efficiency gain, to over 300-fold improvement over previous approaches. Despite these gains, we have provided the details of all procedures in tSCAN and AGENTEX so that researchers without automation capabilities can manually perform these protocols, as even the manual versions of these methods enables a ~50-fold improvement in chemical coding space exploration per researcher. For translation of peptides where multiple distinct open codons are not necessary, AGENTEX has similar translation efficiency to manual approaches and still provides a 6-fold efficiency gain due to the automation of peptide translation.

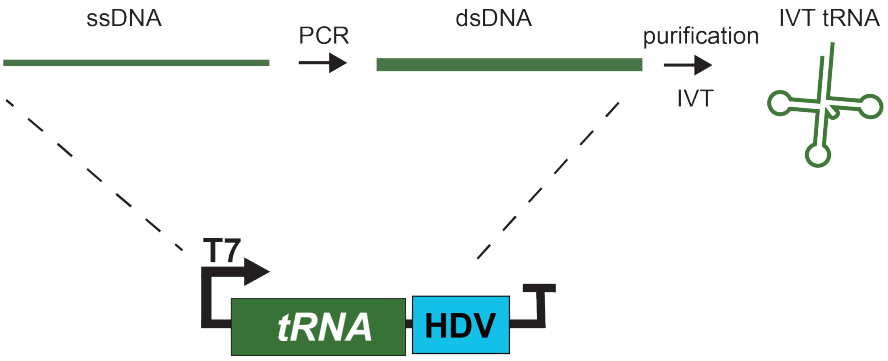
In addition to genetic code expansion, AGENTEX has biocontainment built-in into its design. With the increasing sophistication of engineered translation systems, there is concern for misuses of these technologies, for example construction of mirror-life which can outcompete natural organisms for resources[111]. The makeup of AGENTEX reduces the risk of horizontal gene transfer of its components because while the mutant ribosome is encoded in the host organism for lysate production, all of the tRNAs are provided in RNA form only, minimizing genetic transfer. Furthermore, all known organisms process tRNAs during tRNA maturation to have intact CCA ends in vivo[38,43], which would convert otRNAs into their natural counterparts. The parts of our alternative genetic code will not operate properly even if accidentally leaked into the environment and incorporated into other organisms.

Besides providing insights into the maintenance and origins of natural genetic codes, our findings have implications for current efforts to redesign the translation system and develop in vivo artificial genetic codes. Our discovery that variation of the CCA end of all tRNAs can be accepted by all *E. coli* aaRSs provides an opportunity to free up half of the translation system for redesign, by eliminating the bottleneck of co-evolving every tRNA-aaRS pair for mutual recognition and orthogonality. If all tRNA-aaRS pairs can be parallelized into a parallel genetic code, this has applications for controllable metabolisms switchable by different ribosome-tRNA systems, new genetic firewalls that can resist anticodon mutations[10,108], and proteins composed of non L-α amino acid monomers such as D-amino acids. Previous approaches to genetic code expansion have primarily relied on genome editing[2] or synthesis[1,4] to introduce new open codons and on introduction of orthogonal tRNA and aaRS pairs from diverse organisms[99,112,113] which minimally cross react with the native translation system. Use of tRNAs in orthogonal translation systems focused on the anticodon or on mutations of tRNAs for improved ribosome delivery[114], but engineering of the CCA sequence was not utilized for genetic code expansion in native translation systems. Two aspects of our findings may be applied to translation system redesign (**Supplementary Fig. 14a**). First, the covalent modifications on tRNAs along with their terminal 3’ sequence impact tRNA aminoacylation by *E. coli* aaRSs (**Fig. 3**). Thus by creating a system that can selectively control modification of tRNAs in vivo with tRNA modifying enzymes, it may be possible to prevent aminoacylation of tRNAs with diverging 3’ CCA ends, such as tRNA cysT with CGA, CCA, or CAA end (**Supplementary Fig. 14b**). This then would allow it to be

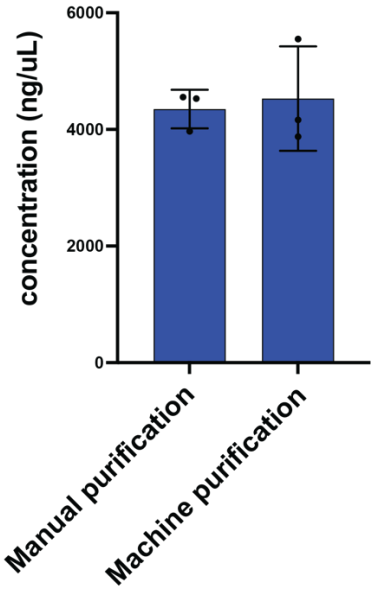
used with an engineered aaRS that can charge nsAAs to cysT while the native cysRS will not aminoacylate it. Previously, transient knockdown of tRNA modifying enzymes has been demonstrated in *E. coli*, though with some impacts on translation fidelity[115]. An alternative approach could employ programmable modification of specified tRNA molecules[116], thus altering the aminoacylation of only specific tRNAs. Second, it was previously shown that ribosomes with G2251C and G2553C mutations could utilize chemically acylated tRNAs to translate peptides using tRNAs with CGA 3' sequences[33]. Previously due to requirements for chemical charging of tRNAs to bypass aaRS evolution, it would be a challenge to implement such a system in vivo. We show that ribosomes with G2251C and G2553C mutations can accommodate otRNAs with a CGA 3' end aminoacylated by *E. coli* aaRSs (**Fig. 6**). As we found that all *E. coli* aaRSs are capable of aminoacylating tRNAs with non-CCA ends, this opens the possibility of in vivo parallel genetic codes where ribosomes with altered large subunit sequences can produce proteins using tRNAs with alternative anticodon sequences and compatible 3' ends (**Supplementary Fig. 14c**). This may allow utilization of up to 14 codons for nsAA incorporation in recoded strains with release factor 1 deletion[2,4], as well as aid with development of additional safeguards to help contain engineered organisms in laboratory environments only[9,117].

We anticipate that further improvements to the technologies presented in this work will improve in vitro translation efficiency and make feasible an in vivo equivalent of our engineered translation system. Our AGENTEX tests of translation with pools MP1 and MP2 used a lysate where the 23S rRNA with G2251C and G2553C mutations was present alongside WT ribosomes. We used a high copy plasmid (>50 copies) to increase signal but the total CGA ribosomes were limited by the availability of small ribosomal subunits in the lysate. As a result, even though we did not detect any evidence of genetic code crosstalk, we observed a greater abundance of the standard code peptide produced from WT ribosomes **(Supplementary Fig. 13)**. While this is sufficient for rapid tests of genetic codes, we anticipate that introduction of engineered ribosomes with both subunits and an orthogonal anti-Shine Delgarno sequence in the 16S rRNA[118,119] to direct orthogonal mRNA for translation could increase translation yields from the orthogonal genetic code both in vitro and in vivo, and improve bioproduction. Increasing the concentration of mRNA and otRNAs could also lead improvements in protein production. Finally, further cell-free reaction yields could be improved by removal of native ribosomes from the lysate to divert all bioproduction to the engineered system.

Supplementary Figures

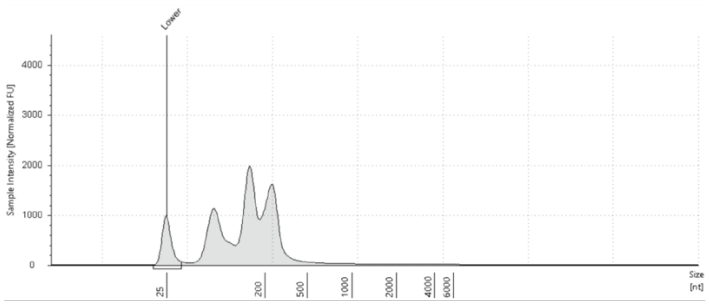


Supplementary Figure 1. Design of tRNAs used in this study. Single-stranded oligonucleotides (ssDNA) containing a T7 promoter, the tRNA sequence followed by the HDV ribozyme and T7 terminator were amplified with PCR into double-stranded DNA (dsDNA). In vitro transcription yielded tRNA molecules.

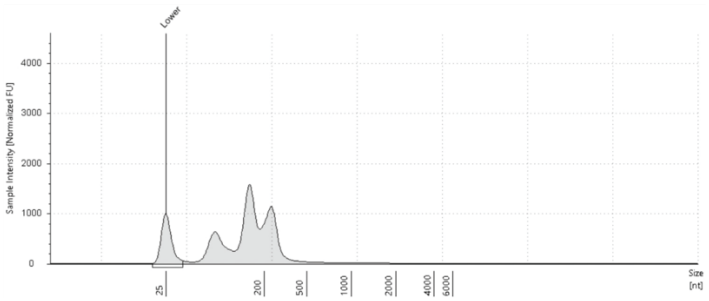


Supplementary Figure 2. Yields from tRNA purifications. In vitro transcribed tRNAs were purified by manual purification using magnetic beads or through the Opentrons robotic procedure. The tRNA concentrations were measured by nanodrop in order to compare manual and robotic tRNA purification efficiencies. Data represent $n = 3$ independent replicates (mean ± SD).

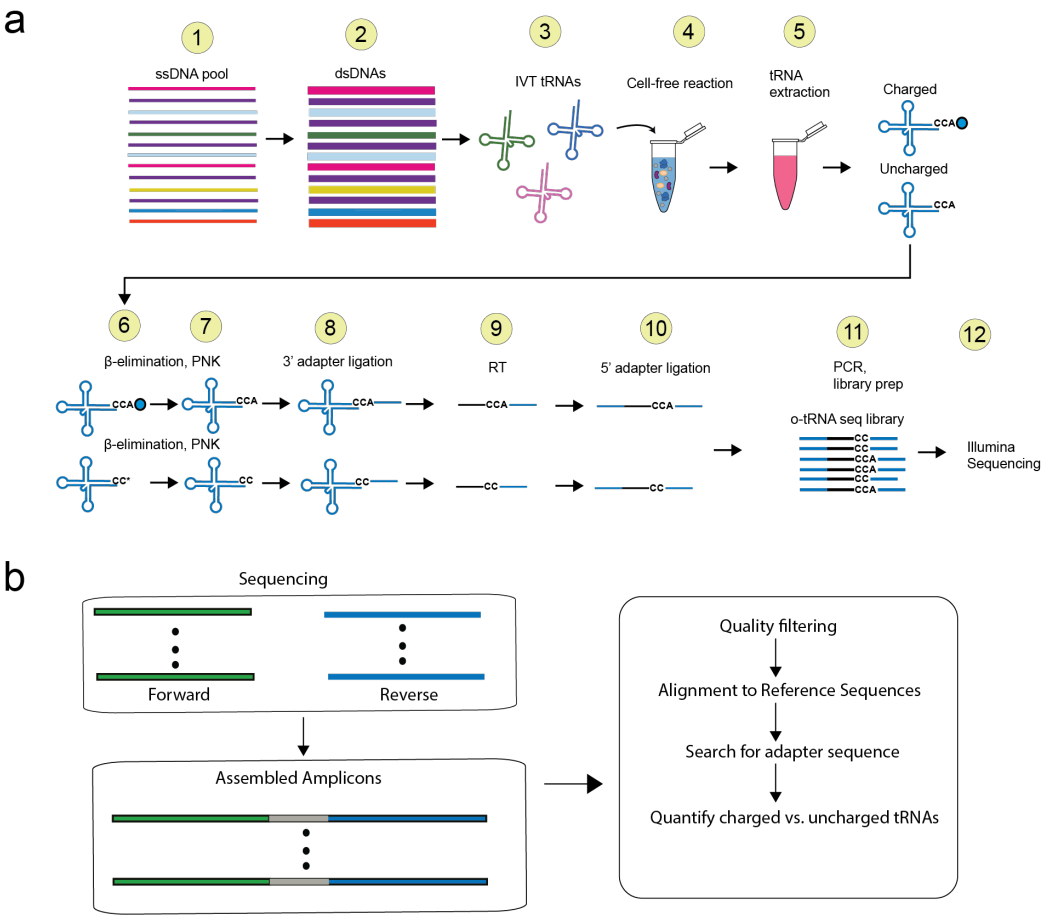
a



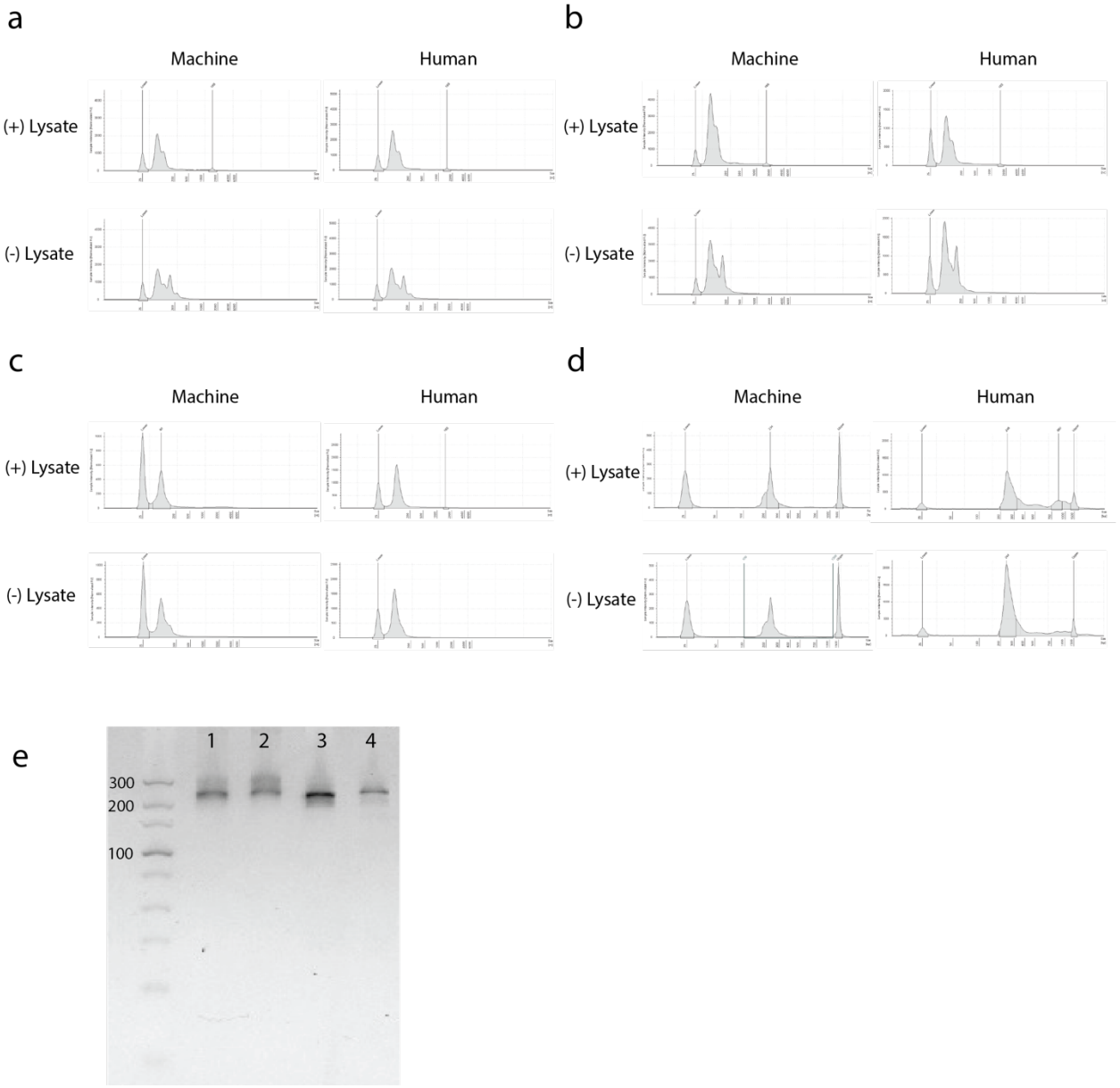
b



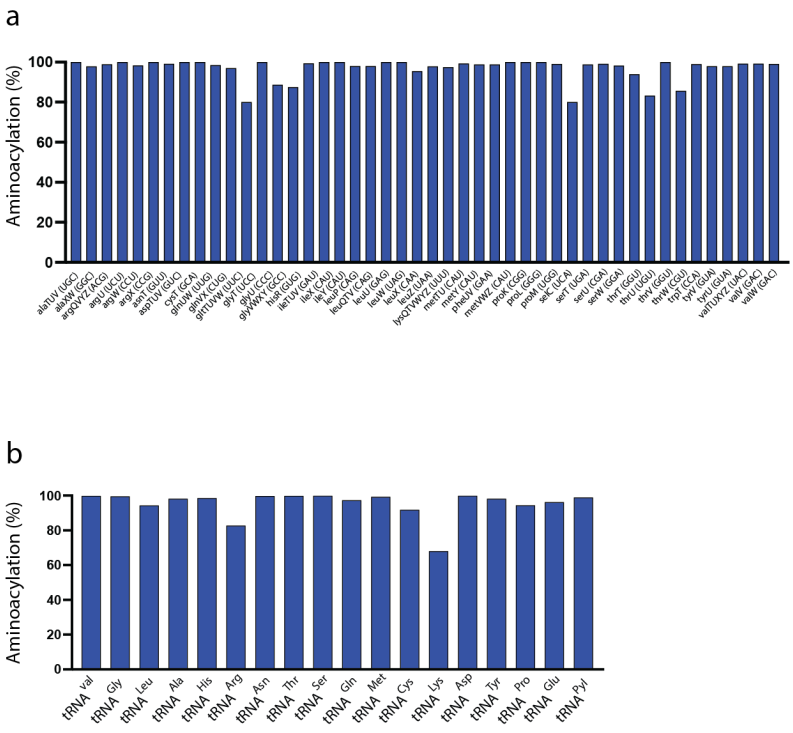
Supplementary Figure 3. Purification of in vitro transcribed tRNA. In vitro transcribed RNAs were **(a)** purified by manual purification using magnetic beads or **(b)** through the Opentrons robotic procedure. The purified tRNAs were analyzed by using Agilent TapeStation in order to assess purity and size distribution of the tRNA molecules.



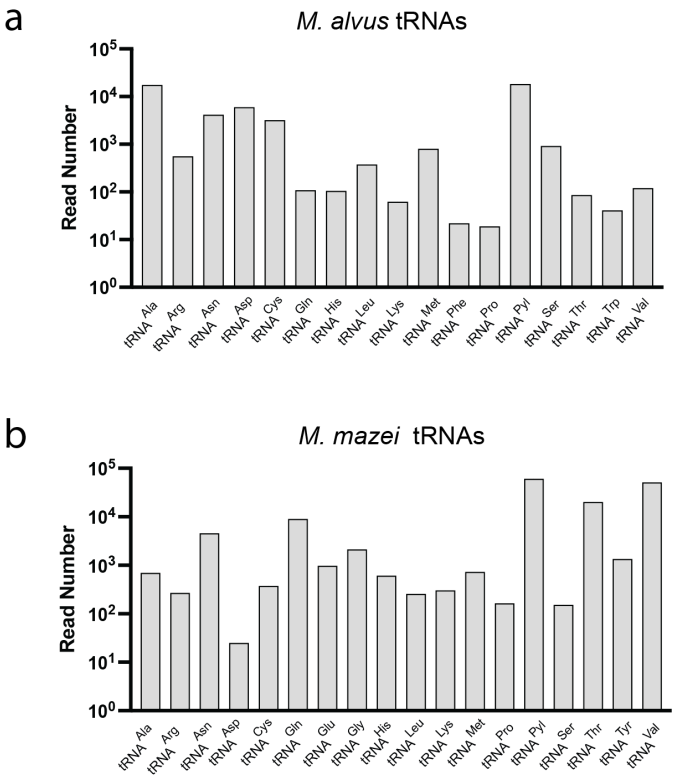
Supplementary Figure 4. Workflow for tSCAN and data analysis pipeline. (a) Schematic of the tSCAN workflow. Chip-synthesized oligonucleotide pools are amplified to dsDNA, transcribed to RNA, and incubated in *E. coli* cell lysates. The tRNA is then extracted, and periodate oxidation, β-elimination and dephosphorylation are performed (1-7) (**Fig. 2b**). The RNA is purified at each step using magnetic beads. After β-elimination and dephosphorylation, purified tRNAs lack a terminal adenosine if they were not aminoacylated prior to the procedure. A 3' adaptor is ligated to the end of the tRNA (8) and primer-dependent reverse transcription is performed (9). Then the cDNA is ligated with a 5' adapter (10) and the cDNA is PCR amplified using primers specific to 5' and 3' adapters and Illumina sequencing barcodes are added (11). The library is then read by Illumina NGS (12). **(b)** Reads from Illumina sequencing are assembled and quality filtering is performed. Reads are aligned to reference sequences for all tRNAs and then searched for the 3' adapter sequence. Depending on the presence of a terminal adenosine immediately prior to the 3' adapter sequence, aminoacylated or non-aminoacylated tRNAs can be identified and quantified.



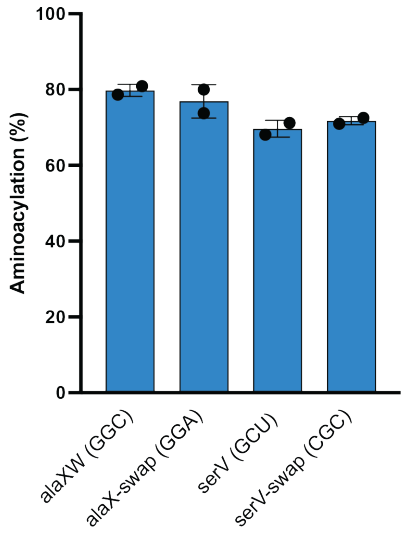
Supplementary Figure 5. tRNA sequencing protocol automated on the Opentrons OT2 pipetting robot or performed manually. (a-d) All steps of tSCAN were performed manually or robotically and the tRNA was analyzed by Agilent TapeStation: **(A)** β-elimination, **(b)** PNK dephosphorylation, **(c)** 3' adapter ligation and **(d)** reverse transcription. **(e)** The cDNA was amplified with PCR to generate dsDNA libraries with Illumina adapters and run on an agarose gel to compare the samples. Each sample is a unique biological replicate, showing DNA from the manual procedure (1-2) and robotic procedure (3-4).



Supplementary Figure 6. Omission of periodate oxidation and β-elimination within the tSCAN procedure. The tSCAN protocol was performed on in vitro transcribed tRNAs without periodate oxidation and β-elimination steps in order to assess the frequency of terminal adenosines in the original tRNA pools. **(a)** The *E. coli* CGA tRNA pool (as in **Extended Data Fig. 2a**) or **(b)** *M. mazei* tRNA pool (as in **Fig. 2d**) was incubated in *E. coli* cell free translation system for 2 hrs at 37°C, extracted, and directly treated with PNK and ligated with 3' adapter without performing periodate oxidation or β-elimination. All subsequent steps of the tSCAN protocol were unchanged.

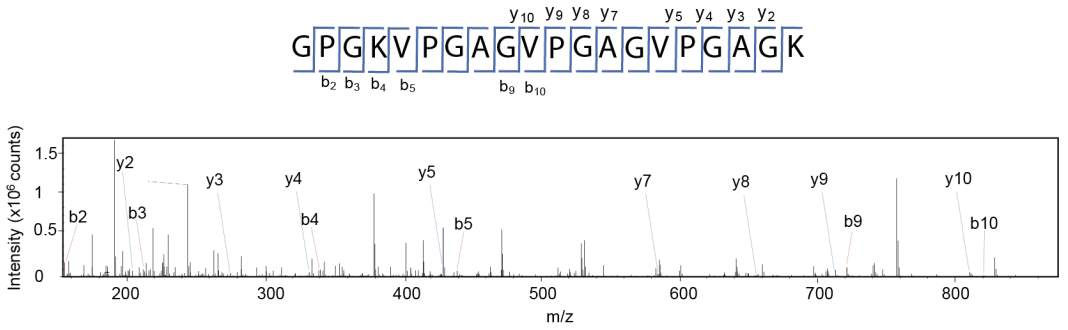


Supplementary Figure 7. Read numbers for M. alvus and M. mazei tRNAs whose aminoacylation was measured with tSCAN in Figure 2. Quality filtering was performed, each tRNAwas aligned to the tRNAs of the respective organism, and reads that had an intact 3' adapter sequence after CC (not aminoacylated tRNA) or CCA (aminoacylated tRNA) terminal trinucleotides were quantified (**Methods**).

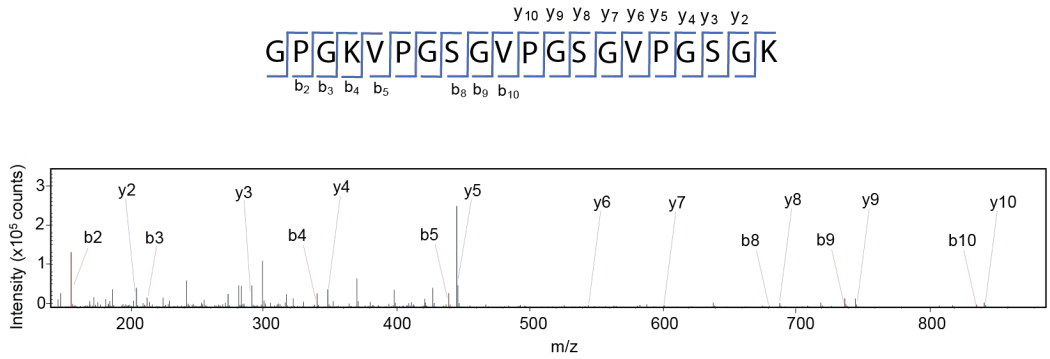


Supplementary Figure 8. Measurement of the aminoacylation levels of ala and ser tRNAs with 3' CGA-ends using tSCAN. A synthetic pool of all *E. coli* tRNA isoacceptors containing alaXW$_{GGC}$ and serV$_{GCU}$, as well as their counterparts with swapped anticodons, was incubated in *E. coli* lysate-based cell free translation system. Aminoacylation was quantified with tSCAN. Data represent $n = 2$ independent replicates (mean ± SD).

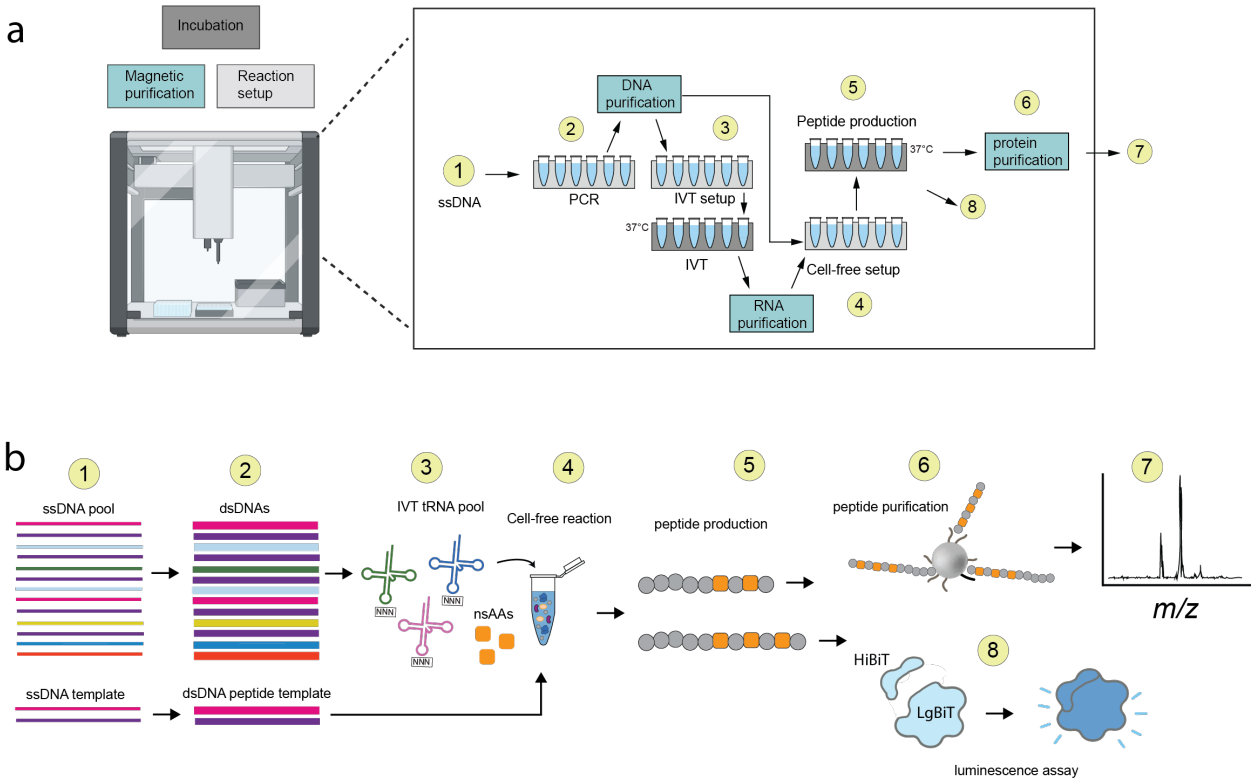
a



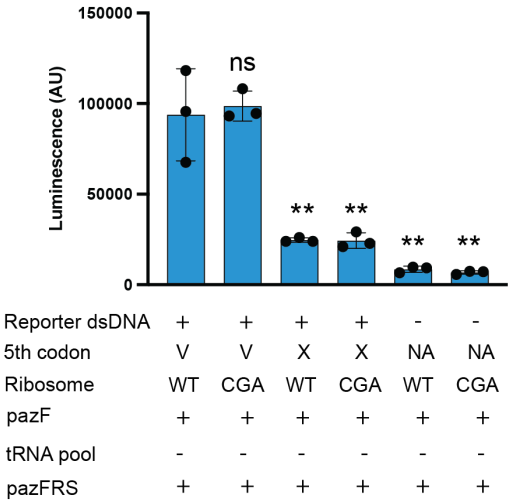
b



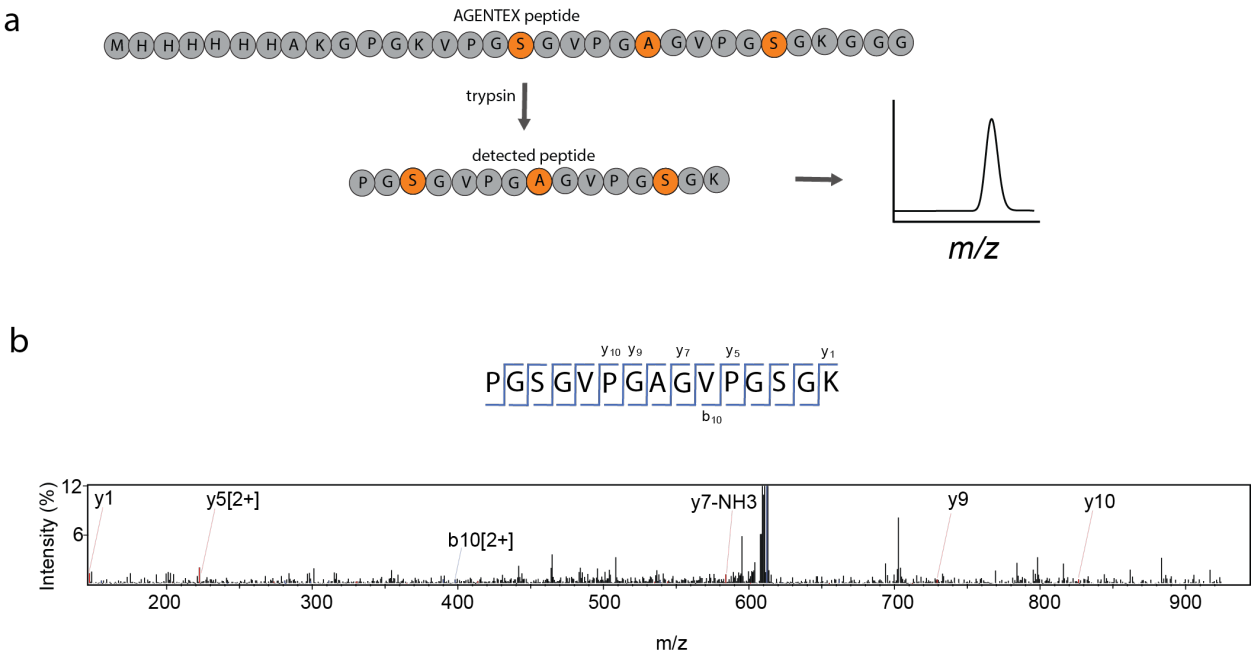
Supplementary Figure 9. Validation of the production of elastin-like polypeptide (ELP) Reporter and Swapped reporter peptides in *E. coli* cell free translation system. The **(a)** Reporter peptide ELP-AAA, or **(b)** swapped reporter peptide ELP-SSS were produced in the *E. coli* cell free translation system, affinity purified, and confirmed by LC-MS/MS.



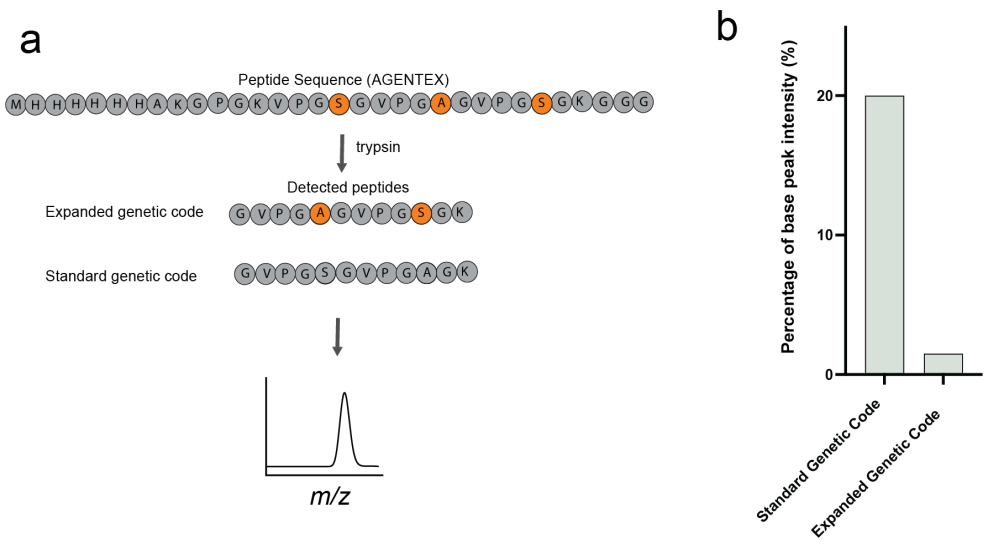
Supplementary Figure 10. AGENTEX workflow. (a) The AGENTEX procedure is divided into unit operations on the Opentrons OT-2 robotic system. All procedures consist of incubations, magnetic purifications, or reaction setups. Each step of the procedure is inherited from one of these base-programs. Translation of polypeptides can be validated by LC-MS/MS (7) or assessed by luminescence assay using a peptide reporter. **(b)** The steps within the AGENTEX protocol. Automated PCR setup is performed on ssDNA templates, to produce dsDNA templates for otRNA pool transcription and polypeptide translation. The dsDNAs are purified and stored for next steps. IVT tRNAs are purified with magnetic beads and introduced into lysate-based translation reactions containing CGA ribosomes along with nsAAs and peptide template dsDNA or orthogonal aaRS dsDNA. Polypeptides production can be assayed by a luminescence assay relying on a modified HiBiT peptide and LgBiT complementation to produce active luciferase. Polypeptides can also be purified using magnetic affinity beads and sequenced with LC-MS/MS.



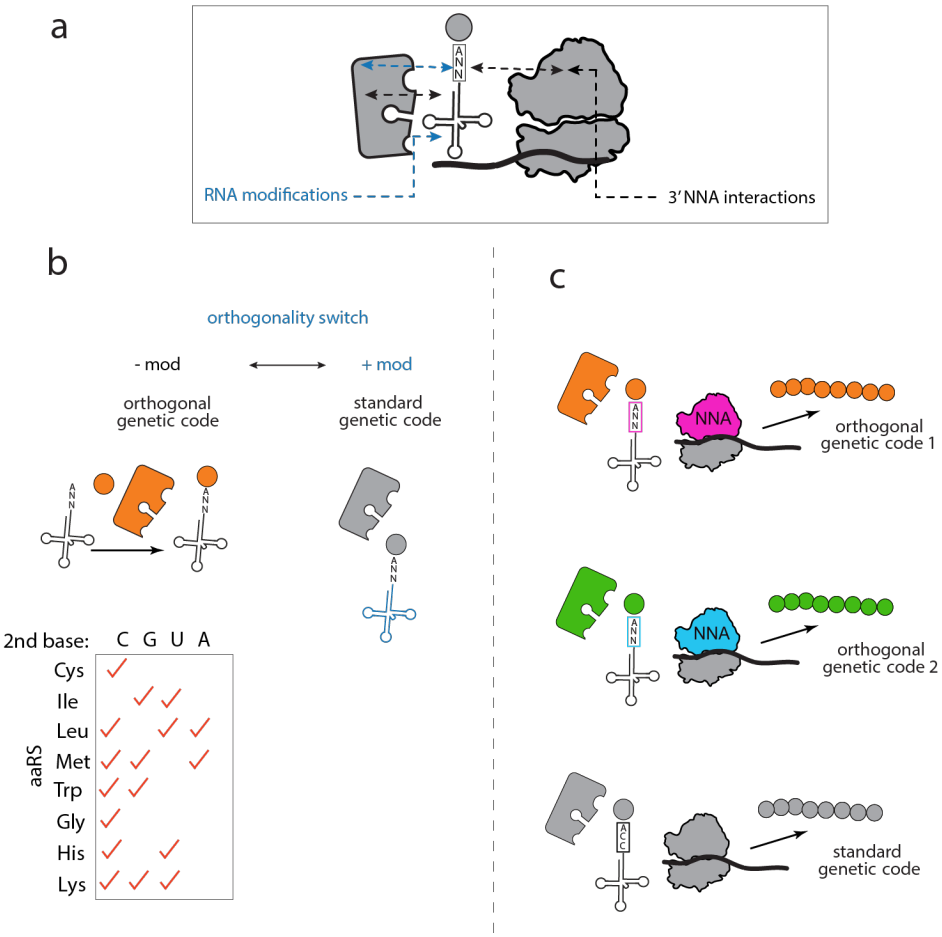
Supplementary Figure 11. Controls for HiBiT peptide translation. The h-GTA-1 peptide was compared to h-TAG-1 peptide in HiBiT luminescence assay using *E. coli* lysate-based cell free translation system prepared with WT or CGA ribosomes. Data represent $n = 3$ independent replicates, with SD shown. A Student's *t* test was used to calculate statistical significance relative to the h-GTA-1 reporter peptide and WT ribosome lysate condition (leftmost bar), where ns (not significant) = $P > 0.05$, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$.



Supplementary Figure 12. Validation of AGENTEX expanded genetic code with LC-MS/MS. We detected a single alanine incorporation into the reporter peptide ELP-SAS along with serines at positions 1 and 3 using LC-MS/MS on affinity purified peptides produced with AGENTEX from lysate with CGA ribosomes, and otRNA pool MP1, where alanine and serine anticodons were swapped.



Supplementary Figure 13. Quantification of relative peptide production from the standard genetic code and expanded genetic code in the same lysate translation reaction. (a) Translation of ELP-SAS was carried out with AGENTEX using lysate containing both WT and CGA ribosomes (G2251C, G2553C), (**Fig. 6d-f**). In ELP-SAS, the Ser and Ala tRNAs from pool MP1 specify whether the standard or expanded genetic code was used to translate the peptide. We utilized a ZenoTOF 7600+ mass spectrometer for MS/MS identification of the peptides. **(b)** The quantitative analysis of the ratio of peptides GVPGAGVPGSGK (**Fig. 6e**) to GVPGSGVPGAGK (**Fig. 6f**) could not be performed at the MS1 level as they are isobaric. To quantify the ratio of these two isobaric peptides we used MS/MS spectra ions that are unique to each sequence. We observed for peptide GVPGSGVPGAGK that ions b5+ (398.20 Da) and y5+ (429.25 Da) are about 20% of total intensity of the spectrum base peak (the y10[2+] ion). The same ions for the spectra GVPGAGVPGSGK have unique masses of 382.21 Da for b5+ and 445.24 Da for y5+ ions at around 1.5% of the same spectrum base peak (the y10[2+] ion). From this we quantified peptide GVPGAGVPGSGK as ~7.5% of the abundance of peptide GVPGSGVPGAGK based on unique MS/MS ion intensities.



Supplementary Figure 14. The potential for the flexibility in the tRNA 3' end and the aaRS-tRNA modification code in development of new-to-nature translation systems. (a) Several aspects of our work may be used in the design of new-to-nature translation systems. **(b)** With a switch that can transiently control tRNA modification status, aaRS-tRNA interactions may be altered depending on the identity of the 3' CCA sequence of the tRNA. This can permit construction of otRNAs that will not be aminoacylated by their native aaRS and could be repurposed for aminoacylation of nonstandard amino acids without importing a tRNA from another organism. **(c)** By altering the large subunit of the ribosome to accept otRNAs, multiple parallel translation systems may be engineered within the cell, containing tRNAs with altered anticodons and non-CCA 3' sequences matching their cognate ribosome. As we found that all of these tRNAs could be aminoacylated by *E. coli* aaRSs (**Fig. 4, Fig. 5**) some tRNAs could be used within this parallel code to introduce cognate amino acids into polypeptides, while others could be repurposed with engineered aaRSs to be aminoacylated with nsAAs.

Supplementary Figure 15. Raw Gel Image corresponding to gel shown in Supplementary Fig. 5e.

Supplementary References

65 Swanson, R. *et al.* Accuracy of in vivo aminoacylation requires proper balance of tRNA and aminoacyl-tRNA synthetase. *Science* **242**, 1548-1551 (1988).

66 Finkler, M. *et al.* Protocol to identify amino acids bound to tRNA by aminoacylation using mass spectrometry. *STAR protocols* **4**, 102504 (2023).

67 Hoben, P. & Söll, D. in *Methods in Enzymology* Vol. 113 55-59 (Elsevier, 1985).

68 Behrens, A., Rodschinka, G. & Nedialkova, D. D. High-resolution quantitative profiling of tRNA abundance and modification status in eukaryotes by mim-tRNAseq. *Molecular cell* **81**, 1802-1815. e1807 (2021).

69 Stephenson, A. *et al.* Physical Laboratory Automation in Synthetic Biology. *ACS Synthetic Biology* **12**, 3156-3169 (2023).

70 Jessop-Fabre, M. M. & Sonnenschein, N. Improving reproducibility in synthetic biology. *Frontiers in bioengineering and biotechnology* **7**, 18 (2019).

71 Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic acids research* **15**, 8783-8798 (1987).

72 Krupp, G. RNA synthesis: strategies for the use of bacteriophage RNA polymerases. *Gene* **72**, 75-89 (1988).

73 Levison, P. R. *et al.* Recent developments of magnetic beads for use in nucleic acid purification. *Journal of Chromatography A* **816**, 107-111 (1998).

74 Cox, J. C., Rudolph, P. & Ellington, A. D. Automated RNA selection. *Biotechnology progress* **14**, 845-850 (1998).

75 Wang, X. *et al.* Development of a simplified and inexpensive RNA depletion method for plasmid DNA purification using size selection magnetic beads (SSMBs). *Genes & Diseases* **8**, 298-306 (2021).

76 Ter-Ovanesyan, D. *et al.* Ultrasensitive measurement of both SARS-CoV-2 RNA and antibodies from saliva. *Analytical Chemistry* **93**, 5365-5370 (2021).

77 Lee, B. N. Automated miRNA Isolation Using SPRI: Changing Binding Conditions Allows for Recovery from Cell Culture, FFPE, and Tissue Samples. *Genetic Engineering & Biotechnology News* **32**, 24-25 (2012).

78 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 Synthetic Biology* **9**, 718-732 (2020).

79 Hong, S. H. *et al.* Cell-free protein synthesis from a release factor 1 deficient Escherichia coli activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS synthetic biology* **3**, 398-409 (2014).

80 Dong, H., Nilsson, L. & Kurland, C. G. Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. *J Mol Biol* **260**, 649-663 (1996).

81 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 1-3 (2024).

82 Torres-Larios, A., Sankaranarayanan, R., Rees, B., Dock-Bregeon, A.-C. & Moras, D. Conformational movements and cooperativity upon amino acid, ATP and tRNA binding in threonyl-tRNA synthetase. *Journal of molecular biology* **331**, 201-211 (2003).

83 Giessen, T. W. *et al.* A Synthetic Adenylation-Domain-Based tRNA-Aminoacylation Catalyst. *Angewandte Chemie* **127**, 2522-2526 (2015).
84 Hassan, T. *et al.* Isolation and identification of cell-specific microRNAs targeting a messenger RNA using a biotinylated anti-sense oligonucleotide capture affinity technique. *Nucleic acids research* **41**, e71-e71 (2013).
85 Mohler, K., Mann, R. & Ibba, M. Isoacceptor specific characterization of tRNA aminoacylation and misacylation in vivo. *Methods* **113**, 127-131 (2017).
86 Chan, P. P. & Lowe, T. M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic acids research* **37**, D93-D97 (2009).
87 Vanderschuren, K. *et al.* Tuning protein half-life in mouse using sequence-defined biopolymers functionalized with lipids. *Proc Natl Acad Sci U S A* **119** (2022).
88 Cvetesic, N., Palencia, A., Halasz, I., Cusack, S. & Gruic-Sovulj, I. The physiological target for Leu RS translational quality control is norvaline. *The EMBO journal* **33**, 1639-1653 (2014).
89 Tian, Q., Wang, C., Liu, Y. & Xie, W. Structural basis for recognition of G-1-containing tRNA by histidyl-tRNA synthetase. *Nucleic acids research* **43**, 2980-2990 (2015).
90 Arnez, J. G., Augustine, J. G., Moras, D. & Francklyn, C. S. The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. *Proceedings of the National Academy of Sciences* **94**, 7144-7149 (1997).
91 Cavarelli, J. & Moras, D. Recognition of tRNAs by aminoacyl-tRNA synthetases. *The FASEB journal* **7**, 79-86 (1993).
92 Giegé, R. & Eriani, G. The tRNA identity landscape for aminoacylation and beyond. *Nucleic Acids Research* **51**, 1528-1570 (2023).
93 Sato, N. S., Hirabayashi, N., Agmon, I., Yonath, A. & Suzuki, T. Comprehensive genetic selection revealed essential bases in the peptidyl-transferase center. *Proceedings of the National Academy of Sciences* **103**, 15386-15391 (2006).
94 Radford, F., Rinehart, J. & Isaacs, F. J. Mapping the in vivo fitness landscape of a tethered ribosome. *Science Advances* **9**, eade8934 (2023).
95 Nissley, A. J. *et al.* Structure of an Archaeal Ribosome with a Divergent Active Site. *bioRxiv*, 2024.2011. 2012.623245 (2024).
96 Koonin, E. V. & Novozhilov, A. S. Origin and evolution of the genetic code: the universal enigma. *IUBMB life* **61**, 99-111 (2009).
97 Sengupta, S. & Higgs, P. G. A unified model of codon reassignment in alternative genetic codes. *Genetics* **170**, 831-840 (2005).
98 Sengupta, S. & Higgs, P. G. Pathways of genetic code evolution in ancient and modern organisms. *Journal of molecular evolution* **80**, 229-243 (2015).
99 Ambrogelly, A., Palioura, S. & Söll, D. Natural expansion of the genetic code. *Nature chemical biology* **3**, 29-35 (2007).
100 Wolf, Y. I. & Koonin, E. V. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biology Direct* **2**, 1-25 (2007).
101 Woese, C. R. On the evolution of the genetic code. *Proceedings of the National Academy of Sciences* **54**, 1546-1552 (1965).
102 Orgel, L. E. Evolution of the genetic apparatus. *Journal of molecular biology* **38**, 381-393 (1968).
103 Gilbert, W. Origin of life: The RNA world. *nature* **319**, 618-618 (1986).

104 Tsour, S. *et al.* Alternate RNA decoding results in stable and abundant proteins in mammals. *bioRxiv* (2024).
105 Eroglu, M. *et al.* Noncanonical inheritance of phenotypic information by protein amyloids. *Nature Cell Biology*, 1-13 (2024).
106 Fimmel, E., Gumbel, M., Starman, M. & Strüngmann, L. Computational analysis of genetic code variations optimized for the robustness against point mutations with wobble-like effects. *Life* **11**, 1338 (2021).
107 Buhrman, H., van der Gulik, P. T., Kelk, S. M., Koolen, W. M. & Stougie, L. Some mathematical refinements concerning error minimization in the genetic code. *IEEE/ACM transactions on computational biology and bioinformatics* **8**, 1358-1372 (2011).
108 Chiappino-Pepe, A. *et al.* Preventing escape and malfunction of recoded cells due to tRNA base changes. *bioRxiv*, 2024.2007. 2018.604179 (2024).
109 Janzen, E., Shen, Y., Liu, Z., Blanco, C. & Chen, I. A. Error minimization and specificity could emerge in a genetic code as by-products of prebiotic evolution. *bioRxiv*, 2021.2005. 2014.444235 (2021).
110 Hillebrecht, J. R. & Chong, S. A comparative study of protein synthesis in in vitro systems: from the prokaryotic reconstituted to the eukaryotic extract-based. *BMC biotechnology* **8**, 1-9 (2008).
111 Adamala, K. P. *et al.* Confronting risks of mirror life. *Science* **386**, 1351-1353 (2024).
112 Liu, C. C., Jewett, M. C., Chin, J. W. & Voigt, C. A. Toward an orthogonal central dogma. *Nature chemical biology* **14**, 103-106 (2018).
113 Cervettini, D. *et al.* Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase–tRNA pairs. *Nature Biotechnology* **38**, 989-999 (2020).
114 Katoh, T., Iwane, Y. & Suga, H. tRNA engineering for manipulating genetic code. *RNA biology* **15**, 453-460 (2018).
115 Masuda, I. *et al.* Loss of N1-methylation of G37 in tRNA induces ribosome stalling and reprograms gene expression. *eLife* **10**, e70619 (2021).
116 Lentini, J. M., Alsaif, H. S., Faqeih, E., Alkuraya, F. S. & Fu, D. DALRD3 encodes a protein mutated in epileptic encephalopathy that targets arginine tRNAs for 3-methylcytosine modification. *Nat Commun* **11**, 2510 (2020).
117 Rovner, A. J. *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518**, 89-93 (2015).
118 Hui, A. & de Boer, H. A. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli. *Proc Natl Acad Sci U S A* **84**, 4762-4766 (1987).
119 Rackham, O. & Chin, J. W. A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol* **1**, 159-166 (2005).