*Published online 17 April 2009*

*Nucleic Acids Research, 2009, Vol. 37, No. 11* ***3747–3755***
*doi:10.1093/nar/gkp240*

# Low modularity of aminoacyl-tRNA substrates in polymerization by the ribosome

**Anthony C. Forster***

Department of Pharmacology and Vanderbilt Institute of Chemical Biology, Vanderbilt University Medical Center, 2222 Pierce Ave., Nashville, TN 37232, USA

Received February 14, 2009; Accepted March 30, 2009

## ABSTRACT

**Aminoacyl-transfer RNAs contain four standardized units: amino acids, an invariant 3′-terminal CCA, trinucleotide anticodons and tRNA bodies. The degree of interchangeability of the three variable modules is poorly understood, despite its role in evolution and the engineering of translation to incorporate unnatural amino acids. Here, a purified translation system is used to investigate effects of various module swaps on the efficiency of multiple ribosomal incorporations of unnatural aminoacyl-tRNA substrates per peptide product. The yields of products containing three to five adjacent L-amino acids with unnatural side chains are low and cannot be improved by optimization or explained simply by any single factor tested. Though combinations of modules that allow quantitative single unnatural incorporations are found readily, finding combinations that enable efficient synthesis of products containing multiple unnatural amino acids is challenging. This implies that assaying multiple, as opposed to single, incorporations per product is a more stringent assay of substrate activity. The unpredictability of most results illustrates the multifactorial nature of substrate recognition and the value of synthetic biology for testing our understanding of translation. Data indicate that the degree of interchangeability of the modules of aminoacyl-tRNAs is low.**

## INTRODUCTION

Protein synthesis is highly modular. Messenger RNAs (mRNAs) are constructed from trinucleotide codon modules. Elongator aminoacyl-transfer RNAs (AA-tRNAs) contain four standardized units: a trinucleotide anticodon, a tRNA body (with a secondary structure usually consisting of four stems), an invariant 3′-terminal CCA, and an amino acid (AA); three of these four modules are highly variable (Supplementary Figure S1). AA-tRNAs are interchangeable on elongation factor Tu/EF1α and on the ribosomal A and P sites. The order of codons on the mRNA and the order of AAs within the protein are also interchangeable; such interchangeability is the basis for the fields of protein mutagenesis and protein engineering. In contrast, the degree of interchangeability of the three variable modules of AA-tRNAs is poorly understood, despite its importance for engineering translation to incorporate unnatural AAs.

AA-tRNA modules were presumably shuffled extensively during evolution by gene duplication, anticodon mutation and charging with different AAs (1,2). But the present degree of modularity of the AA-tRNA domains in protein synthesis cannot be deduced from extensive knowledge of AA-tRNA structures because the four modules act together in *cis*, not *trans*. Though 'domain' boundaries are very precise (Supplementary Figure S1), shuffling may affect translation activity by altering *cis* interactions across domain boudaries or by altering interactions of two domains with another translation macromolecule in *trans*. For example, although anticodon mutants sometimes function well *in vivo* [*e.g.* as suppressor tRNAs (3)], nucleotides adjacent to the anticodon likely affect the efficiency of codon recognition by the anticodon [the extended anticodon hypothesis (4)]. Another example is that EF-Tu and the ribosome may need to bind to multiple domains in the AA-tRNA with compensatory affinities [the thermodynamic compensation hypothesis (5)].

Experimental investigation of the degree of AA-tRNA modularity in translation is challenging. Shuffling AA-tRNA modules often causes pleiotropic effects *in vivo* that extend beyond substrate recognition in translation. For example, tRNA mutations frequently affect tRNA nucleoside modification or the processing of precursor-tRNA (3). Anticodons, in addition to recognizing codons, are frequently major positive determinants for the specificity of AA charging by AA-tRNA synthetases (6). tRNA bodies, in addition to binding to ribosomes, can also contain negative determinants for charging. Most studies of the effects of AA-tRNA domain shuffling in translation have been done by adding *in vitro*-synthesized

*To whom correspondence should be addressed. Tel: +1 615 936 3112; Fax: +1 615 936 5555; Email: a.forster@vanderbilt.edu

AA-tRNA mutants, often loaded with unnatural AAs, to crude *in vitro* translation systems. However, the highly complex and incompletely understood constitutions of crude cell preparations complicate interpretation. For example, mutant AA-tRNAs must compete with natural AA-tRNAs, or release factors in the case of suppressor tRNAs, and some synthetases have a proofreading function that hydrolyses off non-cognate AAs (6). Thus, while efficiencies of single incorporations per protein of unnatural AAs from tRNA mutants are frequently below 50% (7), the cause of these inefficiencies may be due to competing reactions rather than incomplete interchangeability of domains in translation.

In order to overcome these hurdles in testing and exploiting the modularity of AA-tRNAs in translation, we reconstituted from purified components the molecular machinery necessary for 'replication' of peptides containing unnatural AAs (8). AA-tRNA domain swaps were facilitated by chemoenzymatic preparation of non-suppressor AA-tRNA substrates (7,9). Initial studies used three $tRNA^{Asn}$-based tRNAs (termed $tRNA^{AsnB}_{GUU}$, $tRNA^{AsnB}_{GGU}$ and $tRNA^{AsnB}_{GAC}$, where the subscript refers to the anticodon; Figure 1A). As predicted, this system did improve efficiences of single unnatural L-AA incorporation enough to enable the ribosomal synthesis of defined peptides containing three or five straight unnatural AAs. Unexpectedly, the yield of these peptides was only about 55% or 30%, respectively, when compared with peptides from all-natural AA-tRNAs (10). This implied that further optimization of this complex initial system was required and/or changes to individual tRNA domains were detrimental and/or AA-tRNA domains were not fully interchangeable in translation. Subsequent studies using partially purified (11) or purified (12–17) translation systems to synthesize polymers of unnatural AAs have neither focussed on, nor resolved, the issue of low yields under standard translation conditions. Here, our initial system is optimized and expanded to further evaluate the modularity of AA-tRNA substrates in polymerization by the ribosome. The goal is ribosomal synthesis of combinatorial libraries of polymers substituted with unnatural AAs such as *N*-methyl-AAs and α-hydroxy acids. This would allow genetic selections for protease-resistant drug leads by pure translation display (18).

## MATERIALS AND METHODS

### Abbreviations

AA or X, amino acid; U, unnatural AA; x-$tRNA^{y}_{z}$, $x$ = charged AA, $y$ = AA specificity of either the natural isoacceptor or the natural isoacceptor upon which the chemoenzymatic sequence is based, $z$ = either the natural isoacceptor designation or the anticodon sequence

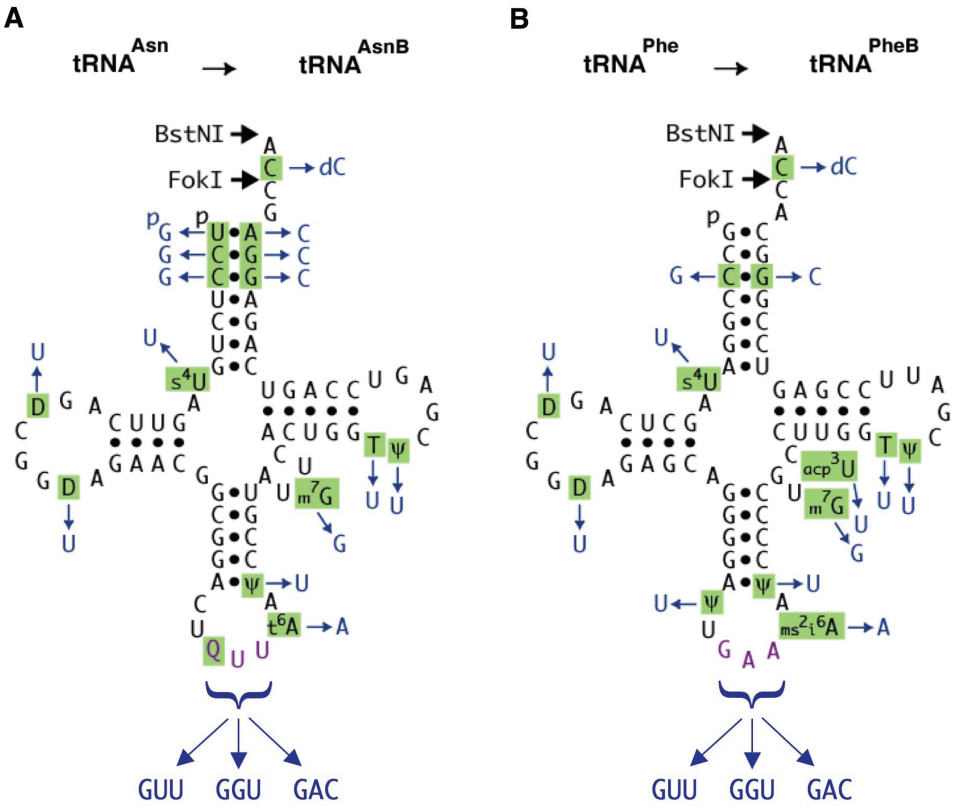


**Figure 1.** Natural *E. coli* $tRNA^{Asn}$ (**A**) and $tRNA^{Phe}$ (**B**) (black), and their seven synthetic counterparts (blue): $tRNA^{AsnB}_{GUU}$ (10), $tRNA^{PheB}_{GAA}$ (20,21) and five anticodon mutants thereof. The anticodons of the natural tRNAs are purple. Substitutions at the 5′ and 3′ termini that maintain the secondary structure of the AA-stems were included to enable efficient transcription initiation at the first nucleotide with GMP by T7 RNA polymerase.

(5′ to 3′) of the chemoenzymatic tRNA sequence. *Formyl*methionine is fM, O-*methyl*serine is mS, and 2-amino-4-pent*en*oic acid and 2-amino-4-pent*yn*oic acid are eU and yU, respectively (eU is also known as allylglycine; structure shown in Figure 2A; structure of peptide containing these 4 AAs in Figure 5A).

## Materials

The preparation of all materials, except those specifically listed below, have been described in detail (8,10,19).

*Messenger RNAs $MT_4E$, $MTNV$, $MTN_2V$, $MTN_5V$ and $MTFV$.* These mRNAs (sequences given in Figures 2A, 3A, 4A and 5A) were prepared by transcription with T7 RNA polymerase of appropriate oligodeoxyribonucleotide templates hybridized to the 18-mer TAATACGACT CACTATAG as illustrated in (8). A second version of mRNA $MTN_5V$ was also prepared by transcription *in vitro* of EcoRI-cut cloned oligos to rule out an artefactually low yield of translation product due to mRNA preparation by direct transcription of a long oligo (result not shown).

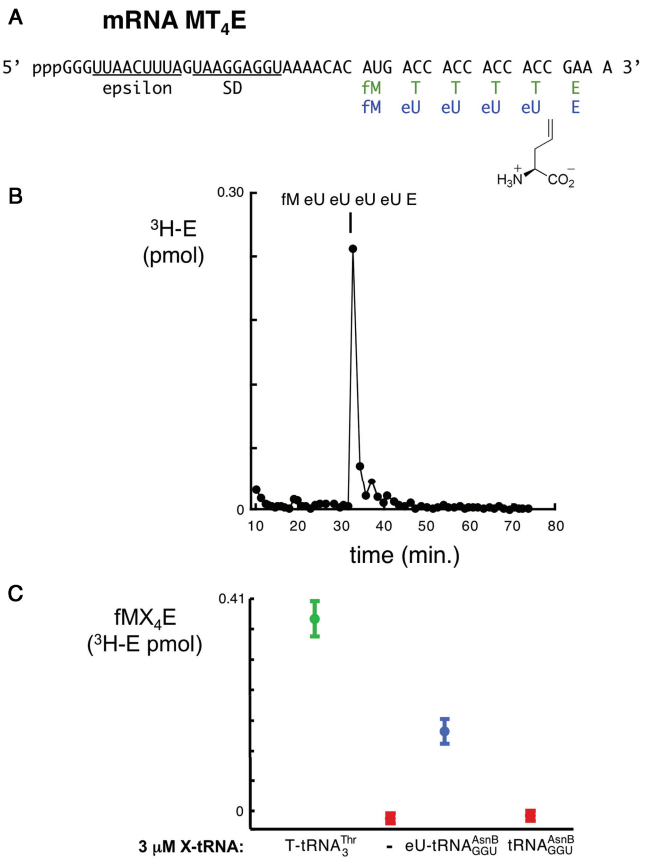


**Figure 2.** Ribosomal incorporation of four adjacent eU AAs into a soluble defined polymer. (**A**) Complete mRNA sequence and encoded translation products using precharged tRNAs including Thr-$tRNA_3^{Thr}$ (green product) or eU-$tRNA^{AsnB}_{GGU}$ (blue product). The epsilon translation enhancer and the Shine-Dalgarno ribosome binding site are underlined. (**B**) HPLC analysis of the blue translation in (A). The elution position of the UV-absorbing, unlabeled, $fM(eU)_4E$ marker peptide mixed with the translation products is indicated above the chromatogram. (**C**) Analysis by cation-exchange, mini-column chromatography of yields for the green and blue translations in (A). The control assay on the right contains uncharged, full-length tRNA (...$CCA_{OH}$ 3′). Background d.p.m. obtained in translations without mRNA were subtracted. Negative controls are red. Standard deviations from the four means are shown.

*Synthetic $tRNA^{Phe}$ genes with mutated anticodons.* Plasmids encoding $tRNA^{PheB}_{GUU}$, $tRNA^{PheB}_{GGU}$ and $tRNA^{PheB}_{GAC}$ were constructed by QuikChange$^R$ oligo-directed mutagenesis of the plasmid encoding $tRNA^{PheB}_{GAA}$ [(20); Figure 1B] by Madhavi Nalam and verified by DNA sequencing. The G-C base pair in the acceptor stem was substituted in comparison with the native isoacceptor to increase the efficiency of synthesis *in vitro* by T7 RNA polymerase (21).

*Synthetic tRNA controls.* Full-length tRNA species were prepared by transcription by T7 RNA polymerase of BstNI-cut plasmids, and then purified by urea polyacrylamide gel electrophoresis and precipitation with ethanol.

*Synthetic AA-tRNAs.* $tRNA^{minusCA}$ species were prepared by transcription by T7 RNA polymerase of FokI-cut plasmids, and then purified by urea polyacrylamide gel electrophoresis and precipitation with ethanol. Each transcript was ligated by T4 RNA ligase to an NVOC-protected pdCA-AA (10,19,22). Concentrations of all unnatural NVOC-aminoacyl-tRNA substrates were estimated by urea polyacrylamide gel electrophoresis at pH 5, and only efficient ligations were kept for later photodeprotection. The efficiency of the photodeprotection procedure was validated by following change in UV absorbance of an NVOC-protected pdCA-AA over time.

*$^3$H-Asn-$tRNA^{Asn}$.* Because of the lack of commercially available, purified $tRNA^{Asn}$ and the difficulty of purifying it, Asn-$tRNA^{Asn}$ was the one AA-tRNA prepared by charging *E. coli* $tRNA^{Total}$. Charging was confined to just the single $tRNA^{Asn}$ isoacceptor in the tRNA mixture by using pure $^3$H-labeled Asn (low specific activity; Moravak Biochemical, CA) and pure His-tagged Asn-tRNA synthetase over-expressed from an available clone (23). The yield of Asn-$tRNA^{Asn}$ was 6× higher with $tRNA^{Total}$ from Roche versus Sigma, so the Roche material was used for large-scale charging. Pilot translations with 0.44 μM Asn-$tRNA^{Asn}$ (including 140 μM $tRNA^{Total}$) prepared from the Roche material and 0.2 μM limiting fMet-$tRNA^{fMet}$ gave single incorporation yields of only ~50% and also inhibited other translations (data not shown). Higher concentations of this Asn-$tRNA^{Asn}$ decreased yields even further, indicating that very high concentrations of uncharged $tRNA^{Total}$ are inhibitory in our system. However, near saturation of Asn incorporation was achieved at a lower concentration of Asn-$tRNA^{Asn}$, 0.2 μM (which necessarily includes 64 μM $tRNA^{Total}$); 0.2 μM charged tRNA is the same concentration as the limiting fMet-$tRNA^{fMet}$.

*C-terminal $^3$H-AA-tRNAs.* $^3$H-Glu-$tRNA^{Glu}$ and $^3$H-Val-$tRNA^{Val}$ were prepared from purified isoacceptors (Sigma

and Subriden RNA) as described (19), with final specific activities ranging from 14 000-22 000 d.p.m./pmol.

*Marker peptides*. fM(eU)$_4$E was synthesized by Zhongping Tan on an Advanced Chemtech peptide synthesizer from commercial reagents (19), purified by HPLC and verified by mass spectrometry.

### Methods

All methods have been described (8,10,19) except the charging and recovery assays below. The translation assay is reiterated for clarification.

*Charging assay for contaminating synthetase activities in the purified translation system*. Our purified translation system uses ribosomes that were washed four times with high salt (8). Although these washes decrease the activity of the ribosomes and are time consuming, the washes were presumed to be necessary for removing contaminating AA-tRNA synthetases. Ribosomes in related purified systems were reported to be highly contaminated with synthetases (24), and a recent purified translation system was found to be contaminated to the extent that its translation products sometimes contain natural AAs where unnatural AAs were encoded (12,13). Thus, the components of our system, including ribosomes washed different numbers of times, were assayed for synthetase contamination. Synthetase charging activities were determined by measuring trichloroacetic acid-precipitated c.p.m. after incubation of the samples with tRNA$^{Total}$ (Sigma), a commercial mixture of 15 different, $^{14}$C-labeled AAs (A, D, E, F, G, H, I, K, L, P, R, S, T, V, Y; New England Nuclear) and ATP at 37°C for 30 min. A DEAE-purified (tRNA-free) crude cell extract (8) was used as a source of the synthetases for the positive control. Supplementary Figure S2 shows that unwashed ribosomes, not translation factors, were indeed highly contaminated with synthetases and that all four of our standard washes were necessary to remove synthetase contaminants completely.

*Translation assays*. To assay synthetic AA-tRNAs in a full translation cycle, translations were always programmed for incorporation of a $^3$H-AA at the C-terminus. Also, to avoid proofreading and *in situ* charging with natural AAs, aminoacyl-tRNA synthetases were omitted from all translations. Translations (10) contained 0.5 μM each of initiation factors 1-3 and elongation factors Ts and G, 2.5 μM elongation factor Tu, 0.25 μM purified ribosomes, 1 μM appropriate mRNA, 0.2 μM (limiting) fMet-tRNA$_i^{fMet}$, 0.5 μM C-terminal, $^3$H-labeled, natural, elongator AA-tRNA, and upstream-encoded, unlabeled or low-specific-activity elongator AA-tRNAs at the following concentrations: 0.2 μM Asn-tRNA$^{Asn}$, 0.5 μM Thr-tRNA$_3^{Thr}$ (except at 3 μM in Figure 2C), 0.5 μM Val-tRNA$_1^{Val}$, and photodeprotected, chemoenzymatic, elongator AA-tRNAs at the concentrations given in the figures (note that higher concentrations of this elongator AA-tRNA were always used when mRNAs programmed multiple, rather than single, incorporations per peptide; see figures). Translations were performed without preincubation (except for Figures S4 and S5) in 5 μl volumes at 37°C for 40 min, then terminated by the addition of NaOH. Maximal yields typically corresponded to half of the limiting 0.2 μM fMet-tRNA$^{fMet}$ incorporated within 40 min into peptide d.p.m. (i.e. 0.5 pmol per 5 μl translation). Substitution of our standard 4× washed ribosomes with a more active preparation (25,26) did not improve the relative yields of unnatural to natural peptide products.

*Quantitation of peptide yields by cation-exchange chromatography*. *N*-formylated peptide products were separated from free AAs by passage through a cation-exchange mini-column (Dowex 50X8-200) in 0.5 M HCl. Recovery from the columns was estimated as follows. Radiolabelled fM(eU)$_4$E and fMT$_4$E were prepared by ribosomal synthesis, passaged through the columns and then quantitated by scintillation counting of aliquots of the eluates (as in Figure 2C). Additional aliquots of these eluates were then re-passaged through the columns (fresh columns) and the eluates counted. Recovery of the re-passaged fM(eU)$_4$E was 92 ± 2% of the loaded eluate, and recovery of the re-passaged fMT$_4$E was 93 ± 2%. Thus, the large difference in yield of the two peptides when synthesized by the ribosome cannot be attributed to differences in recovery.

*Analytical C-18 HPLC*. Radiolabeled translation reaction was treated with NaOH, mixed with authentic unlabeled marker peptide, acidified with acetic acid and filtered through a Microcon 10. The filtrate was analyzed by reversed phase HPLC on a C-18 column using a 9–59% acetonitrile:water gradient containing 0.1% trifluoroacetic acid.

## RESULTS

### Ribosomal synthesis of different polymers of eU from tRNA$^{AsnB}$ bodies

The prior 30% translation yield of the polymer of unnatural eU AAs, fM(eU)$_5$V, compared with the polymer of natural AAs, fMT$_5$V (10), might be due to lower recovery rather than lower synthesis. Consistent with this idea is that the unnatural peptide is much less soluble than the natural peptide in the highly acidic solutions used for analysis (Supplementary Figure 3), and incorporation of a single eU-tRNA$^{AsnB}_{GGU}$ was 100% efficient (10). Thus, a different poly(eU) target was chosen that is theoretically more hydrophilic: fM(eU)$_4$E (Supplementary Figure 3). Marker peptide fM(eU)$_4$E was synthesized chemically and indeed found to be much more soluble than fM(eU)$_5$V. Translation of mRNA MT$_4$E (Figure 2A) with eU-tRNA$^{AsnB}_{GGU}$ (3 μM) yielded a product that comigrated with marker fM(eU)$_4$E on HPLC as expected (Figure 2B). However, the yield of this product measured by cation-exchange mini-columns was still similar to fM(eU)$_5$V and still low (mean = 42%) when compared to the control reaction using the same mRNA template and all-natural AA-tRNAs, in this case producing fMT$_4$E (Figure 2C). The recovery of both peptides from cation-exchange columns was measured at >90% (see

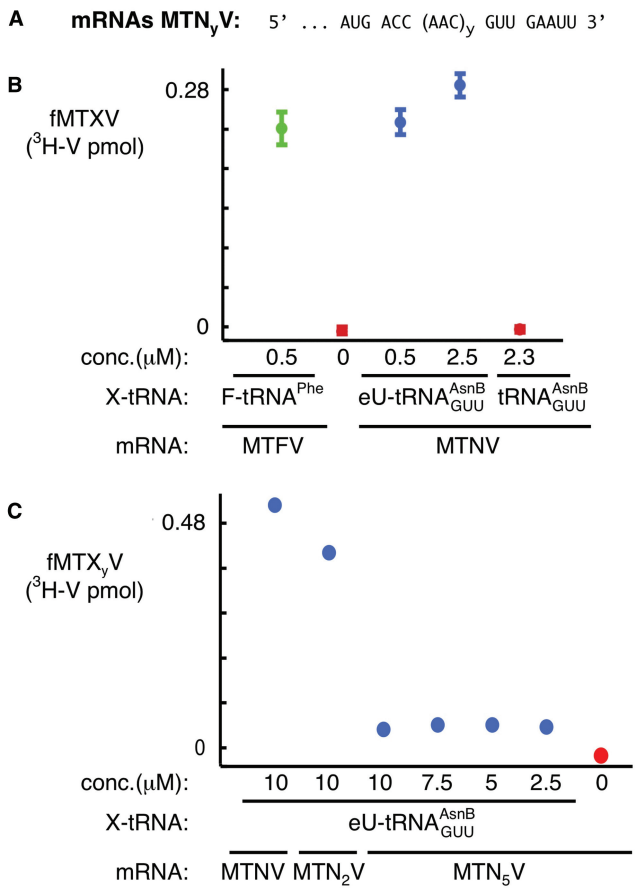


**Figure 3.** Ribosomal incorporation of 1, 2 and 5 eU AAs from a tRNA$^{AsnB}$ with a wild-type anticodon-loop sequence. (**A**) Test mRNA sequences. $y = 1$, 2 or 5. For the GAA codon, no cognate AA-tRNA was supplied. (**B**) Incorporation of a single eU from tRNA$^{AsnB}_{GUU}$ (blue). The positive control translations (green) used all natural AA-tRNAs and mRNA MTFV. Negative controls are red; standard deviations are shown. (**C**) Incorporation of 1, 2 and 5 eUs from tRNA$^{AsnB}_{GUU}$ at various concentrations. (B) and (C) Background d.p.m. obtained in translations without mRNA was essentially the same as the signal from the red translations, and these background d.p.m. were subtracted from all translations.

Materials and Methods section), so the lower yield of eU polymer compared with the natural peptide is due to a difference in synthesis, not recovery.

The translations in Figure 2 were performed for 40 min without preincubation and without release factors so that initiation and recycling should be rate-limiting (8), thereby preventing exhaustion of elongator substrates and premature terminations. To test whether or not these long incubations caused preferentially lower yields with unnatural substrates, the reactions were reconfigured by separately preinitiating ribosomes and preincubating AA-tRNAs with EF-Tu, each for 10 min, and then combining the mixes for short reaction times (8). The relative product yields were still lower with eU-tRNA$^{AsnB}_{GGU}$ than Thr-tRNA, and much lower at the shortest timepoint (15 s; Supplementary Figure S4), suggesting that polymerization of eU-tRNA$^{AsnB}_{GGU}$ is slower than of the natural AA-tRNA. Thus, longer incubation times can favor higher yields with unnatural AA-tRNAs, and preferential deacylation of eU-tRNA versus Thr-tRNA in the long incubation is apparently not the cause of the low yields. Lower complex formation of eU-tRNA$^{AsnB}_{GGU}$ with EF-Tu is also apparently not problematic because the concentration of EF-Tu was saturating (Supplementary Figure S5).

Another possible cause of the suboptimal substrate activity of eU-tRNA$^{AsnB}_{GGU}$ is that the U→G point mutation in the middle of the anticodon triplet reduces the efficiency of decoding. Although this hypothesis is not supported by the finding that this substrate saturates incorporation of a single eU (10), it remains possible that the mutation may misfold the anticodon loop and this may only be detectable in our assay when two misfolded anticodon loops function in adjacent positions on the ribosome. Furthermore, it is known that mutating the anticodon can decrease function in translation (27,28). In order to test this, a tRNA containing an unmodified, wild-type, anticodon loop sequence, eU-tRNA$^{AsnB}_{GUU}$, was synthesized and used to translate mRNAs MTNV, MTN$_2$V and MTN$_5$V (Figure 3A). In Figure 3B, eU-tRNA$^{AsnB}_{GUU}$ was titrated in assays requiring one incorporation per peptide. The yield of fMTeUV was indistingishable from that of the fMTFV control peptide, and control translations substituted with unacylated full-length tRNA$^{AsnB}_{GUU}$ did not synthesize any full-length peptide, confirming that a single eU incorporated very efficiently into product. Surprisingly, the yield of fMT(eU)$_5$V from mMTN$_5$V using the same eU-tRNA$^{AsnB}_{GUU}$ was only 5% compared with the control translation with all-natural AA-tRNAs (Figure 3C), significantly lower than the 30% yield of fM(eU)$_5$V from eU-tRNA$^{AsnB}_{GGU}$ (10). This very low yield cannot be accounted for by insufficient substrate or insufficient photo-deprotection because the yields were independent of substrate concentration (Figure 3C). Nor could the yield be accounted for by general inhibition of translation by the unnatural AA-tRNA, based on the measured incorporation of $^{14}$C-Thr into products in the same translations [as also noted in (20)]. Thus, a wild-type anticodon loop sequence can be less efficient than an anticodon mutant for ribosomal polymerization of unnatural AAs. Of course the anticodon is still not native because it lacks the queuosine (Q) modification, but the result of this modification in comparison with the unmodified G is apparently a slight *decrease* in the stability of pairing with C (29).

Having demonstrated that inefficient polymer synthesis from eU-tRNA$^{AsnB}_{GGU}$ could not be rescued with a wild-type anticodon loop sequence, this implied that rescue required making the substrate even more like native Asn-tRNA$^{Asn}$ (Figure 1A). However, tRNA$^{Asn}$ is a poor experimental system for such tests because it is difficult to introduce modified nucleosides into tRNA transcripts, natural tRNA$^{Asn}$ is not commercially available and is very difficult to purify, the natural 5′ terminus (pUCC …) cannot be synthesized efficiently *in vitro* with T7 RNA polymerase, charging of tRNA$^{AsnB}_{GUU}$ with Asn by the purified *E. coli* Asn synthetase was found to be very low (result not shown), and chemical rearrangement problems are predicted during pdCpA-Asn-NVOC

synthesis. It was therefore decided to synthesize a tRNA body that is more tractable experimentally, $tRNA^{PheB}$.

### Validation of a different tRNA body, $tRNA^{PheB}$

Effects of absence of post-transcriptional nucleoside modifications on translation kinetics have only been reported for one unmodified tRNA, *E. coli* $tRNA^{PheC3G\text{-}G70C}$ (21,30) termed $tRNA^{PheB}_{GAA}$ (Figure 1B). Provided that $tRNA^{PheB}_{GAA}$ was appropriately renatured, lack of modifications had minimal effects on dipeptide synthesis, but translocation had not been evaluated. Translocation was thus tested under our typical conditions, where fMTFV synthesis is saturated by native Phe-$tRNA^{Phe}$ at $\geq$0.25 μM (as predicted for efficient translation limited by 0.2 μM fMet-$tRNA^{fMet}$; Supplementary Figure S6). $tRNA^{PheB}_{GAA}$ chemically charged by ligation to pdCpA-Phe-NVOC was photo-deprotected and compared at different concentrations with 0.5 μM native Phe-$tRNA^{Phe}$ (Figure 4A and B). Concentrations of Phe-$tRNA^{PheB}_{GAA}$ as low as 0.5 μM saturated translation incorporation and translocation to the same extent as 0.5 μM native Phe-$tRNA^{Phe}$. The same results were obtained by mischarging with an unnatural AA, eU (Figure 4C). This validates this unmodified tRNA body and the chemoenzymatic ligation method for efficient incorporation of a cognate and an unnatural AA.

### Synthesis of unnatural polymers from $tRNA^{PheB}$ bodies

The modularity of $tRNA^{Phe}$ was further tested by translation of mRNA MNTVE into fM-yU-mS-eU-E (Figure 5A). This enabled direct comparison with prior results using three different $tRNA^{AsnB}$ tRNAs (10) and has advantages over the poly(eU) systems of having higher product solubility (10) and the ability to substitute incorporation of any combination of individual unnatural AA-tRNAs with a cognate natural AA-tRNA. To this end, three additional $tRNA^{PheB}$ bodies with mutated anticodons were constructed to form base pairs with the N, T and V codons (Figures 1B and 5A). The three encoded $tRNA^{PheBminusCA}$ transcripts were then synthesized and chemically charged with the appropriate unnatural AAs to give yU-$tRNA^{PheB}_{GUU}$, mS-$tRNA^{PheB}_{GGU}$ and eU-$tRNA^{PheB}_{GAC}$. Translations with these three $tRNA^{PheB}$ bodies did indeed produce fM-yU-mS-eU-E (Figure 5B, blue square; also in Figure 6), based on complete dependence on each one of the three $tRNA^{PheB}$s (Figure 5B, red squares). However, the yield was only 15% in comparison with the MVE positive control (Figure 5B legend). This was surprising, given the high activity of the parent $tRNA^{PheB}$ (Figure 4B and C) and the prior 60% yield of an identical translation reaction except with three $tRNA^{AsnB}$ bodies [(10); blue triangle plotted on Figure 5B for comparison].

### Dissecting the effect on polymer yield of each incorporation of an unnatural AA-tRNA

Why was it only possible to incorporate unnatural AA-tRNAs quantitatively within a peptide at single, not multiple, positions? We previously proposed two alternative hypotheses (19): (i) Given evidence that adjacent

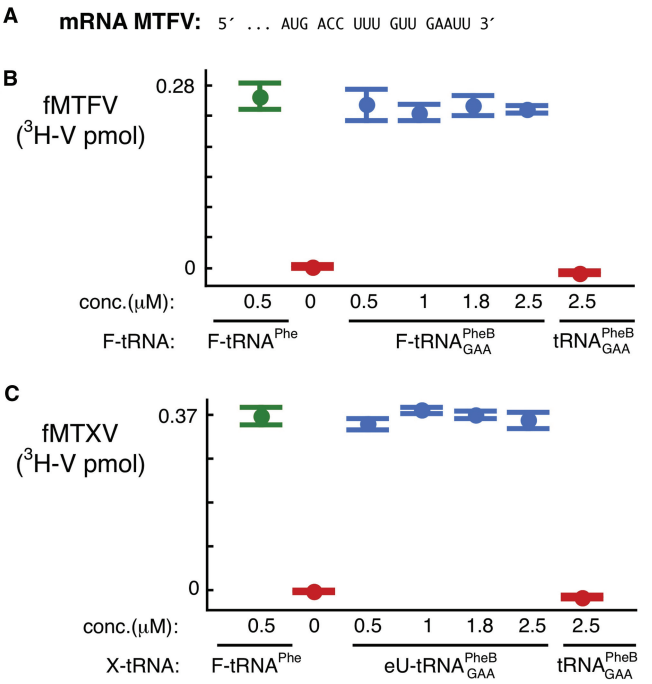


**Figure 4.** Incorporation of Phe and eU and from synthetic $tRNA^{PheB}$ at various concentrations in comparison with natural Phe-$tRNA^{Phe}$. (**A**) mRNA sequence. For the GAA codon, no cognate AA-tRNA was supplied. (**B**) and (**C**) Translation products with the indicated substrates supplied for the F or X AAs. The control assays on the right contain uncharged, full-length tRNA. Standard deviations are shown. Background d.p.m. obtained in translations without mRNA were subtracted.

tRNAs interact on the ribosome [e.g. (31)], such an interaction [either the E and P sites or the P and A sites (Figure S1)] may be inefficient if both AA-tRNAs are unnatural, and (ii) There may be small decreases in yield for single insertions (even for non-neighboring positions) that are difficult to measure, and these decreases become easier to measure when combined for multiple AA insertions. To test these two hypotheses, each unnatural AA-tRNA was substituted with the cognate natural AA-tRNA in all 14 possible combinations for the two blue translations yielding fM-yU-mS-eU-E in Figure 5B. The results (Figure 6) were again unexpected, disproving both hypotheses (i) and (ii).

In the eight translations incorporating two or three unnatural AA-tRNAs per peptide (Figure 6, bottom half), the lowest yield for the $tRNA^{AsnB}$s and the lowest yield for the $tRNA^{PheB}$s were both for fM-yU-T-eU-E synthesis, the only product *lacking* adjacent unnatural AA-tRNAs. Further, the decreases in yield for multiple insertions (Figure 6, bottom half) were *not* simply due to multiplication of small decreases in yield for each single insertion (Figure 6, second row of bars). Rather, all the lowest yielding translations included yU for both the $tRNA^{AsnB}$s and the $tRNA^{PheB}$s. This was not simply just due to some major problem with yU or the AAC codon because fM-yU-mS-eU-E was synthesized from the three $tRNA^{AsnB}$s at 79% yield, and, surprisingly, because yU-$tRNA^{PheB}$ translations gave higher yields if

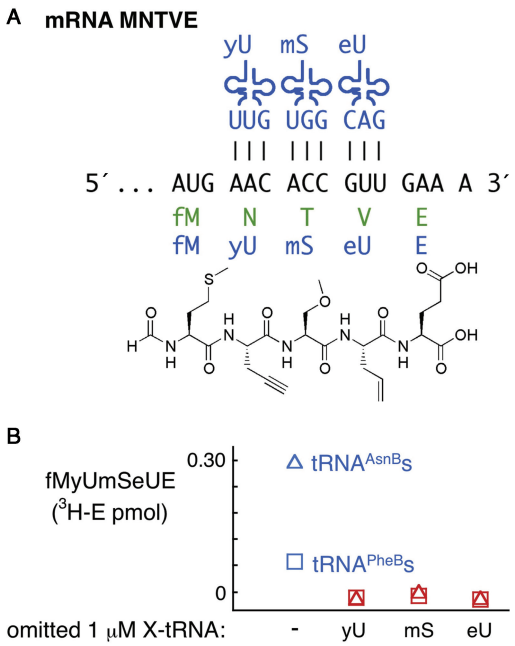


**Figure 5.** Comparison between tRNA$^{PheB}$s and tRNA$^{AsnB}$s for incorporating three adjacent, different unnatural AAs. (**A**) Rational design of three anticodon mutants of tRNA$^{PheB}_{GAA}$ (Figure 2B) for synthesis of a peptide containing three different unnatural AAs. The encoded natural and unnatural translation products are green and blue, respectively, and the structure of the latter is shown. (**B**) Dependence on each unnatural AA-tRNA for synthesis of fM-yU-mS-eU-E. The positive control translations used all natural AA-tRNAs and mRNA MVE (8) to give fMVE with a mean yield of 0.49 pmol (not plotted). The test syntheses (blue) used mRNA MNTVE. As stringent negative controls for the specificity of incorporation (red), each of the unnatural AA-tRNAs was omitted individually. tRNA$^{PheB}$s, open squares. tRNA$^{AsnB}$s, open triangles [taken from ref. (10)]. Background d.p.m. obtained in translations without mRNA were subtracted.

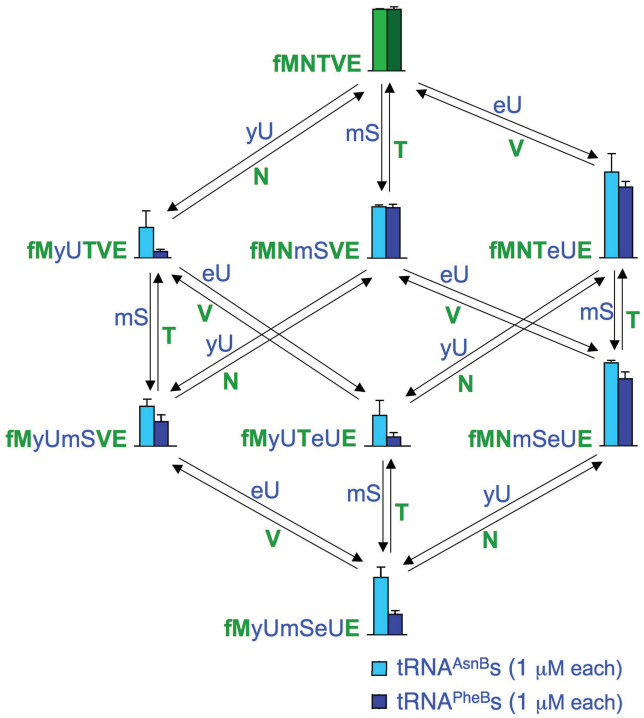


**Figure 6.** Effect on yield of all combinations of three different unnatural AA incorporations into fMNTVE from tRNA$^{AsnB}$s and tRNA$^{PheB}$s. The sequences of the eight product peptides are shown to the left of the bars. The positive control translations with all natural AA-tRNAs gave fMNTVE with mean yield of 0.33 pmol (light green bar done in parallel with tRNA$^{AsnB}$ translations) and 0.39 pmol (dark green bar done in parallel with tRNA$^{PheB}$ translations) and were normalized to 100%. The heights of the other (blue) bars are represented as relative to their all-natural peptide controls (green) performed on the same day. Background d.p.m. obtained in translations without mRNA were subtracted, and standard deviations for each of the 15 different types of translations are shown.

the downstream AA came from the unnatural mS-tRNA$^{PheB}$ instead of the natural Thr-tRNA$^{Thr}$ (Figure 6). Thus, although the tRNA$^{AsnB}$ translations were generally higher yielding than the tRNA$^{PheB}$ translations, as predicted from Figure 5B results, the major differences in yields were *not* due to problems with all three tRNA$^{PheB}$s but rather just due to the poorly-incorporating yU-tRNA$^{PheB}_{GUU}$.

Finally, another variable that might affect AA-tRNA efficiency in translation is considered: the stability of anticodon-codon base pairing (29). For the two tRNA$^{AsnB}$ bodies in the synthesis of poly(eU) (Figures 2C and 3C), the yields correlate with the relative, theoretical, codon-anticodon stabilities:

$$eU - tRNA^{AsnB}_{GUU} < eU - tRNA^{AsnB}_{GGU}$$

Figure 5A shows the anticodon-codon base pairing: from left to right, anticodon GUU forms 2 UA + 1 GC bps; anticodon GGU forms 1 UA + 2 GC bps. However, Figure 6 shows that yields for the six tRNA$^{AsnB}$ and tRNA$^{PheB}$ bodies increased in general as follows:

$$yU - tRNA_{GUU} < mS - tRNA_{GGU} < eU - tRNA_{GAC}$$

Though the efficiencies of these four tRNA$_{GUU}$ and tRNA$_{GGU}$ AA-tRNAs again correlate with the theoretical relative stabilities, the two eU-tRNA$_{GAC}$s do not. The two highest yielding unnatural AA-tRNAs, both with anticodon GAC, form 1CG + 1 AU + 1 GU bps, the *lowest* theoretical stability of the three anticodon-codon pairings shown in Figure 5A.

## DISCUSSION

These synthetic biology experiments, optimizations and controls confirm and extend our prior initial study (10) on unnatural peptide synthesis in a purified translation system. It is clear that low yields in that study cannot be explained by any single potential problem tested here, such as low solubility of products, differential recoveries from the columns, preferential deacylation of unnatural AA-tRNA, a sub-saturating concentration of EF-Tu, an unnatural anticodon loop sequence, tRNA$^{AsnB}$-specific problems, an inability to synthesize an unnatural AA-tRNA completely *in vitro* that incorporates efficiently at a single site, inefficient polymerization of adjacent unnatural aminoacyl-tRNAs, regular decrease in yield with each

unnatural incorporation, and codon-anticodon interactions of low theoretical stability. Figures 2–6 showed that the most stringent test for substrate efficiency in our purified system is measuring multiple, not single, incorporations of unnatural AAs per product. Multiple, as opposed to single, AA incorporations have also proved superior for assaying AA-tRNA substrate efficiency in crude and in vivo systems (31–34). The complexity of the different AA-tRNA structures and their assay results preclude a simple structural explanation for all the inefficiencies. Indeed, it may be naive to expect a single structural explanation, as different AA-tRNAs may be inefficient for different reasons. However, the data do demonstrate the incomplete modularity of AA-tRNAs in polymerization by the ribosome. This has implications for the universality of the genetic code. Constraints in the code are apparently imposed not only by the AA-tRNA synthetases and the proteome, but also by the difficulty in changing the AA specificity of a tRNA without substantially decreasing its activity in translation (28).

Changes in AA-tRNA modules presumably affect translation efficiency by altering structural recognition elements for EF-Tu (5) and/or the ribosome/mRNA complex. Though kinetic studies are needed to define the mechanism(s), some discussion of potential mechanisms consistent with the results is warranted. EF-Tu was used in excess and was apparently saturating (Supplementary Figure S5), and preferential deacylation of unnatural substrates was apparently not problematic, so there was no indication that delivery of L-AA-tRNAs was limiting. Because the translation system is highly purified (Supplementary Figure S2), there should not be any natural substrates or release factors competing with unnatural substrates at the A site. However, the slow synthesis of a polymer of eU-tRNA$^{AsnB}_{GGU}$ (Supplementary Figure S4) opens up the possibility of pausing at each unnatural incorporation for long enough to prevent completion of synthesis of full-length product during the incubation and/or to allow peptidyl-tRNA drop-off from the ribosome before chain completion. Indeed, even single incorporations of more radical unnatural AAs, the D-AA, α-hydroxy acid and *N*-methyl-AA backbone analogs, are dramatically slow (30,35,36). Slow incorporation might explain why product yields were generally lower when multiple, as opposed to single, unnatural incorporations were required (Figures 2–6). Peptidyl-tRNA drop-off from the ribosome competes significantly with elongation during translation of N-terminal codons using natural substrates (37,38), let alone unnatural substrates. Given that drop-off is thought to be much slower at downstream codons due to stronger binding of longer peptidyl sequences to the ribosome tunnel, it is possible that incorporation of unnatural AAs could be improved by incorporation downstream of a long leader peptide.

Incomplete modularity of AA-tRNAs has implications for pure translation display with unnatural AAs, a method for the genetic synthesis and selection of unnatural peptide ligands attached to their mRNAs via the ribosome (18). Attempted synthesis of libraries in a combinatorial manner, and synthesis of longer products (e.g. 10-mers), would be expected to yield only a subset of the desired encoded library products. Though some losses would be acceptable, it would be important to verify that losses did not represent too great a proportion of the encoded library. The preponderance of unexpected results here means that it is difficult to predict the translation activity of any individual unnatural AA-tRNA *a priori*. The synthetic scope might be extended in the future by synthesizing and testing additional tRNA bodies *in vitro* that are more closely related to their wild-type versions. It has also been reported recently that using unnatural AA-tRNAs at extremely high concentrations (0.2 mM each) facilitates polymerization of unnatural L-AAs, *N*-methyl-AAs, *N*-alkyl-glycines and α-hydroxy acids (15–17). Mutation of the translation apparatus can also improve its tolerance for unnatural substrates (39).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

This work began in the Department of Pathology, Brigham and Women's Hospital, Harvard Medical School under the exceptional mentorship of Dr Stephen Blacklow. I am grateful to Drs Madhavi Nalam, Zhongping Tan, Virginia Cornish and Måns Ehrenberg for advice and materials.

## FUNDING

National Institutes of Health; American Cancer Society. Funding for open access charge: National Institutes of Health grant R01-AI072453.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Muramatsu,T., Nishikawa,K., Nemoto,F., Kuchino,Y., Nishimura,S., Miyazawa,T. and Yokoyama,S. (1988) Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature*, **336**, 179–181.
2. Saks,M.E., Sampson,J.R. and Abelson,J. (1998) Evolution of a transfer RNA gene through a point mutation in the anticodon. *Science*, **279**, 1665–1670.
3. Altman,S. (1971) Isolation of tyrosine tRNA precursor molecules. *Nat. New Biol.*, **229**, 19–21.
4. Yarus,M. (1982) Translational efficiency of transfer RNA's: uses of an extended anticodon. *Science*, **218**, 646–652.
5. LaRiviere,F.J., Wolfson,A.D. and Uhlenbeck,O.C. (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science*, **294**, 165–168.
6. Giege,R., Sissler,M. and Florentz,C. (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.*, **26**, 5017–5035.
7. Cornish,V.W., Mendel,D. and Schultz,P.G. (1995) Probing protein structure and function with an expanded genetic code. *Angew. Chem. Int. Ed. Engl.*, **34**, 621–633.
8. Forster,A.C., Weissbach,H. and Blacklow,S.C. (2001) A simplified reconstitution of mRNA-directed peptide synthesis: activity of the epsilon enhancer and an unnatural amino acid. *Anal. Biochem.*, **297**, 60–70.

9. Hecht,S.M., Alford,B.L., Kuroda,Y. and Kitano,S. (1978) "Chemical aminoacylation" of tRNA's. *J. Biol. Chem.*, **253**, 4517–4520.
10. Forster,A.C., Tan,Z., Nalam,M.N.L., Lin,H., Qu,H., Cornish,V.W. and Blacklow,S.C. (2003) Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc. Natl Acad. Sci. USA*, **100**, 6353–6357.
11. Frankel,A., Millward,S.W. and Roberts,R.W. (2003) Encodamers: unnatural peptide oligomers encoded in RNA. *Chem. Biol.*, **10**, 1043–1050.
12. Hartman,M.C.T., Josephson,K., Lin,C.W. and Szostak,J.W. (2007) An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS ONE*, **2**, e972.
13. Josephson,K., Hartman,M.C.T. and Szostak,J.W. (2005) Ribosomal synthesis of unnatural peptides. *J. Am. Chem. Soc.*, **127**, 11727–11735.
14. Murakami,H., Ohta,A., Ashigai,H. and Suga,H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods*, **3**, 357–359.
15. Ohta,A., Murakami,H., Higashimura,E. and Suga,H. (2007) Synthesis of polyester by means of genetic code reprogramming. *Chem. Biol.*, **14**, 1315–1322.
16. Kawakami,T., Murakami,H. and Suga,H. (2008) Messenger RNA-programmed incorporation of multiple N-methyl-amino acids into linear and cyclic peptides. *Chem. Biol.*, **15**, 32–42.
17. Kawakami,T., Murakami,H. and Suga,H. (2008) Ribosomal synthesis of polypeptoids and peptoid-peptide hybrids. *J. Am. Chem. Soc.*, **130**, 16861–16863.
18. Forster,A.C., Cornish,V.W. and Blacklow,S.C. (2004) Pure translation display. *Anal. Biochem.*, **333**, 358–364.
19. Tan,Z., Blacklow,S.C., Cornish,V.W. and Forster,A.C. (2005) De novo genetic codes and pure translation display. *Methods*, **36**, 279–290.
20. Zhang,B., Tan,Z., Dickson,L.G., Nalam,M.N.L., Cornish,V.W. and Forster,A.C. (2007) Specificity of translation for N-alkyl amino acids. *J. Am. Chem. Soc.*, **129**, 11316–11317.
21. Harrington,K.M., Nazarenko,I.A., Dix,D.B., Thompson,R.C. and Uhlenbeck,O.C. (1993) In vitro analysis of translational rate and accuracy with an unmodified tRNA. *Biochemistry*, **32**, 7617–7622.
22. Tan,Z., Forster,A.C., Blacklow,S.C. and Cornish,V.W. (2004) Amino acid backbone specificity of the Escherichia coli translation machinery. *J. Am. Chem. Soc.*, **126**, 12752–12753.
23. Shimizu,Y., Inoue,A., Tomari,Y., Suzuki,T., Yokogawa,T., Nishikawa,K. and Ueda,T. (2001) Cell-free translation reconstituted with purified components. *Nat. Biotechnol.*, **19**, 751–755.
24. Ganoza,M.C., Aoki,H., Burkhardt,N. and Murphy,B.J. (1996) The ribosome as "affinity matrix": efficient purification scheme for translation factors. *Biochimie*, **78**, 51–61.
25. Rodnina,M.V. and Wintermeyer,W. (1995) GTP consumption of elongation factor Tu during translation of heteropolymeric mRNAs. *Proc. Natl Acad. Sci. USA*, **92**, 1945–1949.
26. Zavialov,A.V., Buckingham,R.H. and Ehrenberg,M. (2001) A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3. *Cell*, **107**, 115–124.
27. Olejniczak,M., Dale,T., Fahlman,R.P. and Uhlenbeck,O.C. (2005) Idiosyncratic tuning of tRNAs to achieve uniform ribosome binding. *Nat. Struct. Mol. Biol.*, **12**, 788–793.
28. Raftery,L.A. and Yarus,M. (1987) Systematic alterations in the anticodon arm make tRNAGlu-Suoc a more efficient suppressor. *EMBO J.*, **6**, 1499–1506.
29. Curran,J.F. (1998) Modified Nucleosides in Translation. In Grosjean,H. and Benne,R. (eds), *Modification and editing of RNA*. ASM Press, Washington, DC, pp. 493–516.
30. Pavlov,M.Y., Watts,R.E., Tan,Z., Cornish,V.W., Ehrenberg,M. and Forster,A.C. (2009) Slow peptide bond formation by proline and other N-alkyl amino acids in translation. *Proc. Natl Acad. Sci. USA*, **106**, 50–54.
31. Smith,D. and Yarus,M. (1989) tRNA-tRNA interactions within cellular ribosomes. *Proc. Natl Acad. Sci. USA*, **86**, 4397–4401.
32. Bonekamp,F., Dalboge,H., Christensen,T. and Jensen,K.F. (1989) Translation rates of individual codons are not correlated with tRNA abundances or with frequencies of utilization in Escherichia coli. *J. Bacteriol.*, **171**, 5812–5816.
33. Sorensen,M.A. and Pedersen,S. (1991) Absolute in vivo translation rates of individual codons in Escherichia coli. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J. Mol. Biol.*, **222**, 265–280.
34. Hohsaka,T., Ashizuka,Y., Sasaki,H., Murakami,H. and Sisido,M. (1999) Incorporation of two different nonnatural amino acids independently into a single protein through extension of the genetic code. *J. Am. Chem. Soc.*, **121**, 12194–12195.
35. Yamane,T., Miller,D.L. and Hopfield,J.J. (1981) Discrimination between D- and L-tyrosyl transfer ribonucleic acids in peptide chain elongation. *Biochemistry*, **20**, 7059–7064.
36. Bieling,P., Beringer,M., Adio,S. and Rodnina,M.V. (2006) Peptide bond formation does not involve acid-base catalysis by ribosomal residues. *Nat. Struct. Mol. Biol.*, **13**, 423–428.
37. Karimi,R. and Ehrenberg,M. (1996) Dissociation rates of peptidyl-tRNA from the P-site of E. coli ribosomes. *EMBO J.*, **15**, 1149–1154.
38. Karimi,R., Pavlov,M.Y., Heurgue-Hamard,V., Buckingham,R.H. and Ehrenberg,M. (1998) Initiation factors IF1 and IF2 synergistically remove peptidyl-tRNAs with short polypeptides from the P-site of translating Escherichia coli ribosomes. *J. Mol. Biol.*, **281**, 241–252.
39. Dedkova,L.M., Fahmi,N.E., Golovine,S.Y. and Hecht,S.M. (2003) Enhanced D-amino acid incorporation into protein by modified ribosomes. *J. Am. Chem. Soc.*, **125**, 6616–6617.