frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 19 April 2024
DOI 10.3389/fgene.2024.1386299

Check for updates

OPEN ACCESS

EDITED BY
Patrick O'Donoghue,
Western University, Canada

REVIEWED BY
Jeffery M. Tharp,
Indiana University Bloomington, United States
Noah Reynolds,
University of Illinois at Springfield, United States

*CORRESPONDENCE
John D. Fisk,
john.fisk@ucdenver.edu

RECEIVED 14 February 2024
ACCEPTED 30 March 2024
PUBLISHED 19 April 2024

CITATION
Schmitt MA, Tittle JM and Fisk JD (2024), Codon decoding by orthogonal tRNAs interrogates the *in vivo* preferences of unmodified adenosine in the wobble position.
Front. Genet. 15:1386299.
doi: 10.3389/fgene.2024.1386299

COPYRIGHT
© 2024 Schmitt, Tittle and Fisk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Codon decoding by orthogonal tRNAs interrogates the *in vivo* preferences of unmodified adenosine in the wobble position

Margaret A. Schmitt, Jillyn M. Tittle and John D. Fisk*

Department of Chemistry, University of Colorado Denver, Denver, CO, United States

The *in vivo* codon decoding preferences of tRNAs with an authentic adenosine residue at position 34 of the anticodon, the wobble position, are largely unexplored because very few unmodified A34 tRNA genes exist across the three domains of life. The expanded wobble rules suggest that unmodified adenosine pairs most strongly with uracil, modestly with cytosine, and weakly with guanosine and adenosine. Inosine, a modified adenosine, on the other hand, pairs strongly with both uracil and cytosine and to a lesser extent adenosine. Orthogonal pair directed sense codon reassignment experiments offer a tool with which to interrogate the translational activity of A34 tRNAs because the introduced tRNA can be engineered with any anticodon. Our fluorescence-based screen utilizes the absolute requirement of tyrosine at position 66 of superfolder GFP for autocatalytic fluorophore formation. The introduced orthogonal tRNA competes with the endogenous translation machinery to incorporate tyrosine in response to a codon typically assigned another meaning in the genetic code. We evaluated the codon reassignment efficiencies of 15 of the 16 possible orthogonal tRNAs with A34 anticodons. We examined the Sanger sequencing chromatograms for cDNAs from each of the reverse transcribed tRNAs for evidence of inosine modification. Despite several A34 tRNAs decoding closely-related C-ending codons, partial inosine modification was detected for only three species. These experiments employ a single tRNA body with a single attached amino acid to interrogate the behavior of different anticodons in the background of *in vivo E. coli* translation and greatly expand the set of experimental measurements of the *in vivo* function of A34 tRNAs in translation. For the most part, unmodified A34 tRNAs largely pair with only U3 codons as the original wobble rules suggest. In instances with GC pairs in the first two codon positions, unmodified A34 tRNAs decode the C- and G-ending codons as well as the expected U-ending codon. These observations support the "two-out-of-three" and "strong and weak" codon hypotheses.

KEYWORDS

tRNA modification, inosine, orthogonal translation machinery, sense codon reassignment, wobble, translation, genetic code

1 Introduction

Organisms utilize essentially the same, universal genetic code to convert the information in their DNA genomes into functional proteins. With few exceptions, each of the 64 three nucleotide codons communicates the same signal to the protein translation machinery across the three domains of life. Through the course of evolution, alternative strategies to decode sense codons have developed (Knight et al., 2001). Organisms rely upon complements of tRNA molecules with disparate sets of anticodons and anticodon modifications in order to translate the same mRNA codons. Although the different tRNA complements translate a single mRNA sequence into a protein with a single primary sequence, the energetics of the tRNA/mRNA/ribosome interactions that are required to facilitate translation differ. Thus, the alternative sets of tRNA species lead to differing sets of "preferred" codons across organisms (Andersson and Kurland, 1990). Codon usage is, in part, related to the efficiency with which preferred codons are decoded (dos Reis et al., 2003). The major differences in decoding strategies follows the tripartite division of extant organisms into bacterial, archaeal, and eukaryotic domains (Woese and Fox, 1977).

One prominent difference between the strategies employed to read the genetic code relates to the manner by which the third (wobble) position of codons are decoded (Bjork and Hagervall, 2014; Agris et al., 2018). The wobble hypothesis postulated that some tRNA anticodons could decode multiple codons to explain how fewer than 61 tRNA species were able to translate the full suite of sense codons. The structure of the genetic code, together with the apparent activity of tRNA species, suggested that the first two bases of a codon were read in canonical Watson-Crick fashion and that some "wobble" was allowed in the third codon position (Crick, 1966). At the time when the wobble hypothesis was formulated, the genetic code had not yet been elucidated fully and information on tRNA sequences, modifications, and their interactions with codons was incomplete.

Inosine modification is the prototypical example of the way in which a modified anticodon base in a tRNA might be capable of decoding multiple codons. Inosine (abbreviated as "I") is a modified adenosine base, produced by deamination of the exocyclic N6 amine. When the wobble hypothesis was presented, inosine modification had been found at position 34 in the anticodons of several tRNA species, the location that interacts with the third codon position (Ingram and Sjoquist, 1963; Holley et al., 1965; Söll et al., 1966). According to the original elaboration of the "wobble rules", adenosine was expected to pair exclusively with uracil, while inosine was expected to pair with uracil, cytosine, and adenosine. Modern genomic surveys have shown that tRNAs encoded as A34 are nearly universally converted to inosine (Chan and Lowe, 2016). Unmodified A34 tRNAs are exceedingly rare.

The distribution and prevalence of inosine modification is divergent across the three domains of life (Ehrlich et al., 2021; Srinivasan et al., 2021). In archaea, genes for tRNAs with A34 anticodons, the precursor to I34 anticodons, are absent. In the bacterial domain, typically only a single tRNA gene with an A34 anticodon exists. This tRNA is modified to I34 and reads the arginine CGU, CGC, and CGA codons. A handful of A34 tRNAs, some of which are inosine modified, have recently been identified in other codon four boxes in a small number of bacterial species. Even with these findings, inosine usage remains highly restricted in the bacterial domain. Inosine modification is much more widely employed in eukaryotes, where each of the amino acids degenerately encoded by three, four or six codons, with the exception of glycine, typically utilize I34 tRNAs. The U-, C-, and A-ending codons in the TAPSLIVR family (where "TAPSLIVR" is the acronym containing the one letter amino acid codes for threonine, alanine, proline, serine, leucine, isoleucine, valine, and arginine) are all decoded by a major tRNA species with an I34 anticodon (Rafels-Ybern et al., 2015; Rafels-Ybern et al. 2018; Rafels-Ybern et al. 2019).

Inosine is introduced into tRNAs post-transcriptionally by adenine deaminase acting on tRNA (ADAT) enzymes (Gerber and Keller, 1999; Torres et al., 2014). Each inosine-modified tRNA is genomically encoded and initially transcribed as an A34 tRNA, followed by subsequent quantitative deamination to inosine. In bacteria, the ADAT enzyme encoded by the *tadA* gene is homodimeric and typically has only the single aforementioned arginine tRNA as its substrate (Wolf et al., 2002). In eukaryotes, inosine is introduced into the TAPSLIVR tRNAs post-transcriptionally by a single, heterodimeric ADAT enzyme that recognizes the multiple substrate tRNAs (Gerber and Keller, 1999; Torres et al., 2014).

Previously, we and others have explored the extent to which orthogonal tRNA/aminoacyl tRNA synthetase (aaRS) pairs can be employed to break the degeneracy of the genetic code in *E. coli* (Kwon et al., 2003; Bohlke and Budisa, 2014; Zeng et al., 2014; Biddle et al., 2015; Lee et al., 2015; Mukai et al., 2015; Ho et al., 2016; Kwon and Choi, 2016). Canonical amino acids that are encoded by more than a single codon are potential targets for reassignment since one (or more) of the sense codons can be left to direct incorporation of the natively-encoded amino acid and one (or more) of the codons can be given a reassigned meaning. Breaking the degeneracy of the genetic code requires the identification or engineering of pairs of tRNAs and the aaRSs that recognize them which are "orthogonal" to the complement of tRNA/aaRS pairs naturally employed by the target organism (Liu and Schultz, 2010; Acevedo-Rocha and Budisa, 2016; Melnikov and Söll, 2019). The anticodon of the orthogonal tRNA is engineered to decode a sense codon typically assigned another meaning in the genetic code. In order to expand translation beyond the set of canonical amino acids, the orthogonal aaRS must recognize and aminoacylate its cognate tRNA with a noncanonical amino acid (ncAA). Even partial decoding of a targeted sense codon by the supplied orthogonal machinery can lead to incorporation of the ncAA.

The first demonstration of breaking the degeneracy of the genetic code utilized an orthogonal yeast tRNA species with an AAA anticodon to reassign the UUU phenylalanine codon in *E. coli* (Kwon et al., 2003). In an expansion of the concept demonstrated in this work, we have mapped the extent to which anticodon variants of the orthogonal *Methanocaldococcus jannaschii* (*M. jannaschii*) tyrosine tRNA/aaRS pair can be used to break the degeneracy of the genetic code for 17 of the 18 multi-codon canonical amino acids in *E. coli* (Biddle et al., 2015; Schmitt et al., 2018; Schwark et al., 2020). The *M. jannaschii* tyrosine pair is one of the two orthogonal pairs most commonly used to introduce ncAAs into proteins in response to stop codons. Variants of the *M. jannaschii* aaRS have

been engineered to aminoacylate >100 different ncAAs onto the cognate tRNA (Dumas et al., 2015). Our system for broad exploration of codon reassignment exploits the native function of the *M. jannaschii* orthogonal pair: incorporation of tyrosine. We developed a gain of function screen to quantify introduction of tyrosine by engineered variants of the *M. jannaschii* tRNA in response to non-tyrosine codons at a critical position in green fluorescent protein.

The effectiveness of a given orthogonal pair for reassigning the meaning of a particular sense codon depends on multiple factors, including the efficiency of the aaRS charging its anticodon-engineered cognate tRNA, the competition mounted by the complement of native tRNAs, and the energetics of the codon-anticodon interaction itself. *E. coli* wobble codons are promising locations at which to infiltrate and expand the genetic code because the majority are decoded by endogenous tRNAs via U3/G34 or G3/U34 interactions. The anticodons of the introduced orthogonal tRNAs are engineered to Watson-Crick base pair with the targeted codon and potentially take advantage of favorable codon/anticodon energetics at the wobble position to improve reassignment efficiency.

The extent to which a particular engineered tRNA is able to discriminate between the targeted codon and closely-related codons is an important consideration in choosing an orthogonal system to precisely reassign a given codon. tRNAs with A34 anticodons engineered to target the U3 wobble codons were expected to readily discriminate between the U-ending and non-targeted C-ending codons. Our initial evaluation of the reassignment efficiencies of four sense codons, including His CAU, by the *M. jannaschii* pair revealed unexpected, nearly equal reassignment of the CAU and CAC codons (Biddle et al., 2016). Through sequencing of reverse transcribed tRNA species, we showed that the orthogonal tRNA engineered to have an AUG anticodon was partially modified to inosine by *E. coli* TadA. Typically, "orthogonality" of introduced translation machinery has been interpreted as a lack of crosstalk between the orthogonal tRNAs and host aaRSs and between host tRNAs and the orthogonal aaRSs (Furter, 1998; Wang et al., 2001; Cervettini et al., 2020). Crosstalk between the orthogonal translation components, particularly the tRNA, and other host enzymes, e.g., tRNA modifying enzymes, has not been broadly evaluated. Given the prevalence of tRNA modifications and their myriad roles in the conversion of genotype to phenotype, modification of orthogonal tRNAs by endogenous enzymes is not surprising. Understanding this space of interactions may inform the combination of orthogonal pair and reassignable codon chosen as systems with increasingly expanded genetic codes are engineered (Mukai et al., 2017; Dunkelmann et al., 2021).

In a previous publication, we quantified the codon reassignment efficiency of *M. jannaschii* tRNAs with 14 of the 16 possible ANN anticodons in *E. coli* (Schmitt et al., 2018). In each case, we evaluated the system for reassignment of targeted and non-targeted codons. Some of the introduced orthogonal A34 tRNAs behave as postulated through the wobble hypothesis, decoding the U3 codon nearly exclusively. Another group of introduced orthogonal A34 tRNAs translate other codons within the box with substantial efficiency compared to the efficiency of reassignment of the targeted codon. Decoding of C-ending codons by A34 orthogonal tRNAs suggested that additional anticodons in the context of the *M. jannaschii* tRNA body may be substrates for *E. coli* TadA. Here, we analyze Sanger sequencing of the cDNA products of reverse transcribed *M. jannaschii* tRNA species with 15 A34 anticodons as well as present the reassignment preferences of *M. jannaschii* $tRNA_{ACG}$ targeting Arg CGU and an expanded investigation of translation of NNA and NNG codons by orthogonal A34 tRNAs.

Detection of inosine modification via Sanger sequencing does not map onto the extent that engineered orthogonal tRNAs discriminate between the targeted U-ending codon and closely-related, non-U-ending codons; the majority of *M. jannaschii* tRNAs with A34 anticodons are not inosine modified. Experimental characterization of the decoding preferences of authentic A34 anticodons either *in vivo* or *in vitro* is sparse, in part because of the rarity of unmodified A34 tRNAs across the domains of life. This set of directed sense codon reassignment experiments utilizes a single tRNA body with a single attached amino acid to interrogate the behavior of authentic A34 anticodons in the background of *in vivo E. coli* translation.

2 Methods and materials

DNA vector composition and construction have been described previously (Biddle et al., 2015; Biddle et al., 2016; Schmitt et al., 2018; Schwark et al., 2020). Orthogonal translation machinery vectors utilized in this study differ only at the identity of the nucleotides that correspond to positions 34–36 of the tRNA gene. GFP reporter vectors differ only at the identity of the nucleotides that specify the codon corresponding to position 66 of the superfolding GFP variant. The superfolder GFP mRNA sequence was designed to minimize usage of U3 codons for amino acids read via wobble interactions in *E. coli*. The gain of function, fluorescence-based screen for quantifying codon reassignment has been described (Biddle et al., 2015). Brief summaries of vectors and the screen are provided as Supplementary Material. The Supplementary Material also includes details on cell strains, oligonucleotide primer sequences, numbers of biological replicates evaluated for each codon reassignment measurement, and additional sequencing traces of reverse transcribed tRNAs.

2.1 Reverse transcription of *M. jannaschii* tRNAs for Sanger sequencing

RNAse Away was used routinely throughout the process to keep the work area free of RNAses, and all consumables used were certified RNAse free. The complement of tRNA molecules was extracted from NEB 5-alpha cells harboring an orthogonal translation machinery vector. Briefly, 10^9 cells were pelleted from overnight cultures by spinning sufficient volumes at 17,000 x*g* for 2 min at room temperature. Cell supernatant was removed via pipette. The cell pellets were frozen overnight at −20°C. Following thawing on ice, 200 μL B-PER lysis reagent (Thermo Scientific) was added to each pellet. Cell lysis proceeded at room temperature for 30 min.

A 1:1 volume of 0.3 M NaOAc, 10 mM EDTA, pH 4.5 was added to lower the pH of the solution. 450 μL of water-buffered phenol:chloroform (5:1 ratio) was added to the lysate and vortexed three times for 60 s, 60 s, and 30 s, with lysates held

	codon	orthog anti codon	sense codon reassign efficiency	U3:C3 disc ratio[b]	AA	codon	orthog anti codon	sense codon reassign efficiency	U3:C3 disc ratio	AA	codon	orthog anti codon	sense codon reassign efficiency	U3:C3 disc ratio	AA	codon	orthog anti codon	sense codon reassign efficiency	U3:C3 disc ratio	AA	
U	UUU	AAA	3.2 ± 0.3%	≥94:6	F	UCU	AGA	19.7 ± 1.0%	93:7	S	UAU	AUA		N/A	Y	UGU	ACA	3.6 ± 0.1%	≥95:5	C	U
	UUC		B.D.[a]			UCC		1.4 ± 0.2%			UAC					UGC		B.D.			C
	UUA				L	UCA					UAA				oc	UGA				op	A
	UUG					UCG					UAG				am	UGG				W	G
C	CUU	AAG	4.3 ± 0.4%	74:26	L	CCU	AGG	12.9 ± 1.1%	56:44	P	CAU	AUG	6.1 ± 0.4%	68:32	H	CGU	ACG	7.6 ± 0.5%	55:45	R	U
	CUC		1.5 ± 0.1%			CCC		10.3 ± 0.5%			CAC		2.9 ± 0.1%			CGC		6.1 ± 0.3%			C
	CUA					CCA					CAA				Q	CGA					A
	CUG					CCG					CAG					CGG					G
A	AUU	AAU	4.5 ± 0.5%	91:9	I	ACU	AGU	10.7 ± 0.5%	89:11	T	AAU	AUU	7.2 ± 0.4%	≥97:3	N	AGU	ACU	9.4 ± 0.8%	≥98:2	S	U
	AUC		0.43 ± 0.1%			ACC		1.4 ± 0.2%			AAC		B.D.			AGC		B.D.			C
	AUA					ACA					AAA				K	AGA				R	A
	AUG				M	ACG					AAG					AGG					G
G	GUU	AAC	0.84 ± 0.1%	59:41	V	GCU	AGC	1.3 ± 0.1%	52:48	A	GAU	AUC	3.5 ± 0.5%	≥95:5	D	GGU	ACC	1.3 ± 0.3%	76:24	G	U
	GUC		0.59 ± 0.04%			GCC		1.2 ± 0.1%			GAC		B.D.			GGC		0.4 ± 0.1%			C
	GUA					GCA					GAA				E	GGA					A
	GUG					GCG					GAG					GGG					G
	U					C					A					G					

FIGURE 1
Codon reassignment efficiencies for U- and C-ending codons and U3:C3 discrimination ratios for 15 *M. jannaschii* orthogonal pair tRNA variants with ANN anticodons. [a] B.D. indicates that the reassignment efficiency of the particular codon/anticodon combination was below the limit of detection for the in cell assay (0.2%). The limit of detection is largely determined by the residual fluorescence of cells and growth media. [b] U3:C3 discrimination ratios are calculated as "[RE (reassignment efficiency) U-ending codon]/[RE U-ending codon + RE C-ending codon]". For cases in which the C-ending codon was not reassigned above the limit of detection, the discrimination ratio is calculated as "[RE U-ending codon]/[RE U-ending codon +0.2]". Shaded discrimination ratios indicate orthogonal tRNA anticodons with a notable tendency to decode the C-ending codon as well as the Watson-Crick base pairing U-ending codon. Pale pink shading indicates the three anticodons for which at least partial inosine modification was detected by Sanger sequencing of cDNA products from reverse transcribed tRNAs. Pale aqua shading indicates that no inosine modification was detected despite the tRNA discriminating between the U- and C-ending codons poorly.

on ice for 60 s between each step. Layers were separated by centrifugation at 15,000 x*g* at 4°C for 15 min. The aqueous, RNA-containing layer was transferred to 2.5 volumes of 100% ethanol to precipitate the nucleic acids. Following a brief vortex, the solution was stored overnight at −20°C.

Nucleic acids were pelleted by centrifugation at 15,000 x*g* at 4°C for 30 min. The supernatant was aspirated, and the pellet was allowed to air dry. Nucleic acids were resuspended in 100 µL of 1x DNAseI buffer (Thermo Scientific). One unit of DNaseI was added, and incubation proceeded at 37°C for 20 min. DNAseI was inactivated and removed by phenol:chloroform extraction and ethanol precipitation as described above. RNA was resuspended in 20 µL of 0.3 M NaOAc, 10 mM EDTA, pH 4.5 and quantified using a BioTek Synergy H1 or Synergy Neo2S plate reader. Each RNA extraction yielded 12–20 µg of RNA.

Reverse transcription was carried out using SuperScript IV reverse transcriptase (Thermo Scientific) according to the manufacturer's instructions and primer Mj-RevTrans (sequence in Supplementary Table S1). Duplicate reactions without enzyme were included as negative controls to ensure subsequent amplification resulted from tRNA molecules as opposed to carried-through DNA. cDNA products from reverse transcription were amplified with Taq polymerase (New England Biolabs) using primers cDNA-amp-fwd and cDNA-amp-rev (sequences in Supplementary Table S1). PCR products were cleaned up using the GeneJET PCR spin kit (Thermo Scientific) and analyzed by gel electrophoresis. Products were sequenced with primer cDNA-amp-fwd by Azenta (formerly Genewiz, LLC). Sanger sequencing .ab1 files were converted to .xml files using abi2xml.exe (freely available at https://github.com/eriksjolund/abi2xml) Figures were generated by importing data lines from the .xml files into Microsoft Excel.

3 Results

The fluorescence-based screen for sense codon reassignment exploits the absolute requirement of tyrosine at position 66 in superfolder GFP for fluorophore formation (Zacharias et al., 2006). Reporter vectors with another codon at that position are co-transformed with the *M. jannaschii* tyrosyl aminoacyl tRNA synthetase and a variant of its cognate tRNA. Both the engineered orthogonal tRNA and endogenous tRNAs compete to decode the evaluated codon. Incorporation of tyrosine by the orthogonal tRNA in response to the codon at position 66 in the GFP reporter leads to production of fluorescent protein. Incorporation of a different canonical amino acid by the endogenous translation machinery leads to production of non-fluorescent protein. The mean per cell fluorescence of the system under evaluation is bracketed between a 100% fluorescence reference system in which the codon at position 66 encodes tyrosine and a 0% fluorescence reference system in which the codon at position 66 encodes another amino acid. Stated reassignment efficiencies represent the mean and standard deviation of measurements of the number of biological replicates reported in Supplementary Table S2.

We have evaluated the codon reassignment efficiencies of 15 of the 16 possible orthogonal tRNAs with ANN anticodons (Figure 1) (Schmitt et al., 2018). Our screen cannot be used to evaluate tyrosine codons. For $tRNA_{AUA}$ that targets Tyr codons, decoding of the reporters by either the endogenous or orthogonal tRNAs produces a fluorescent protein. Of these tRNAs, six of 15 differ from the inosine-modified *E. coli* $tRNA^{Arg2}$ at only one anticodon position, those with ANG or ACN anticodons. The eight remaining species differ from *E. coli* $tRNA^{Arg2}$ at both position 35 and 36. Figure 1 is arranged to mirror the standard genetic code table and also presents the calculated U3:C3 discrimination ratios. Several of these tRNAs are capable of decoding both U- and C-ending codons.

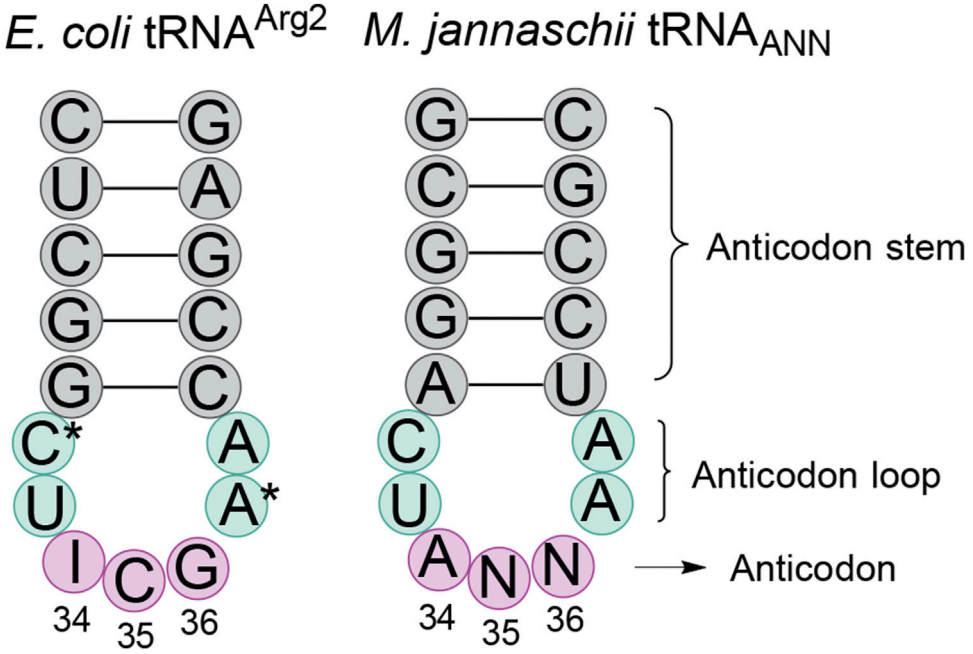

FIGURE 2
Comparison of the anticodon stem loop sequences of *E. coli* $tRNA^{Arg2}$ and *M. jannaschii* $tRNA^{Tyr}{}_{ANN}$. The anticodon loop sequences (aqua shading) of the only *E. coli* tRNA transcribed as A34 and the commonly used *M. jannaschii* orthogonal tRNA are identical, save known nucleotide modifications in the *E. coli* tRNA. In $tRNA^{Arg2}$, position 32 is modified to 2-thiocytidine and position 37 is modified to 2-methylthio-N6-isopentenyladenosine (ms2i6A). The anticodon stem (gray shading) is the same length in the two tRNAs. Depending on the anticodon (pink shading) of the orthogonal tRNA, the sequence may be 0, 1, or 2 nucleotides different than that of $tRNA^{Arg2}$. The three anticodons that permit modification of A34 to inosine in the orthogonal tRNA are the only known modifications to the *M. jannaschii* tRNA when expressed in *E. coli*. Other modifications may exist.

After detecting partial inosine modification of the *M. jannaschii* $tRNA_{AUG}$, we hypothesized that other orthogonal *M. jannaschii* tRNAs, specifically those capable of reassigning C-ending codons, might also be substrates for *E. coli* TadA adenine deaminase (Biddle et al., 2016). The anticodon loop sequences of *E. coli* $tRNA^{Arg2}$ and the orthogonal *M. jannaschii* tRNA are identical, apart from known nucleotide modifications in the *E. coli* tRNA (Figure 2). While the identity elements for TadA recognition have not been definitively mapped, characterization of a limited set of minihelix substrates indicated that anticodon loop sequence was important for TadA activity. Structural evidence suggests that the sequence of the anticodon loop is exclusively involved in recognition (Elias and Huang, 2005; Kuratani et al., 2005; Losey et al., 2006).

We have previously reported the targeted codon reassignment efficiencies for the majority of these orthogonal tRNAs (Schmitt et al., 2018). Here, we characterize the reassignment preferences of the 15th A34 tRNA, *M. jannaschii* $tRNA_{ACG}$ and provide additional quantification of decoding non-U-ending codons by other A34 orthogonal tRNAs (Table 1). We report the observed extent of *in vivo* inosine modification, determined via sequencing of reverse transcribed cDNAs, for the 15 A34 tRNAs (Figure 3). We analyze the expanded suite of data in the context of recognition by *E. coli* TadA. In light of the absence of broad inosine modification, we describe the behavior of a wide set of tRNAs with authentic A34 bases in *E. coli* translation.

The discrimination ratios can be thought of as "number of times out of 100 incorporation events a given tRNA selects the U-ending vs. C-ending codon". Discrimination ratios were calculated by dividing the reassignment efficiency at the targeted, U-ending codon by the sum of the reassignment efficiencies for the U- and C-ending codons. For combinations of A34 tRNA/C3 codon reporters with efficiencies below the limit of detection (0.2%), the discrimination ratio is reported as "greater than or equal to [the quotient of reassignment of the U3 codon divided by reassignment of the U3 codon plus 0.2% (the limit of detection)".

Discrimination ratios provide a normalized metric for how well individual orthogonal tRNAs decode U- and C-ending codons regardless of the absolute value of reassignment efficiencies observed. Orthogonal tRNA performance is dependent upon the effective concentration of the tRNA, which may change as aminoacylation efficiency by the aaRS changes in response to anticodon sequence, and the competition against the endogenous translation machinery. Discrimination ratios cancel the differences in aminoacylation level and competition with endogenous tRNAs inherent to each individual orthogonal tRNA and report how each single tRNA species decodes two codons, allowing normalized comparison across orthogonal tRNA species. For example, an orthogonal tRNA that decodes the U-ending codon with 50% efficiency and the C-ending codon with 5% efficiency has a discrimination ratio of 91:9. An orthogonal tRNA that decodes the U-ending codon with 4% efficiency and the C-ending codon with 0.4% efficiency also has a discrimination ratio of 91:9. Despite the disparate performance of each tRNA within the entirety of the endogenous translation machinery, both tRNA species incorporate their charged amino acid in response to a U-ending codon 91 times out of 100 orthogonal tRNA-directed incorporation events.

TABLE 1 Efficiency of codon reassignment to tyrosine by variants of the *M. jannaschii* orthogonal translation machinery with ANN anticodons.

tRNA anticodon	Codon targeted	Reassignment at PQU[a] (%)	Reassignment at PQC (%)	Reassignment at PQA (%)	Reassignment at PQG (%)
AAA	Phe UUU	**3.2 ± 0.3**	B.D.[b]	———[c]	———
AAG	Leu CUU	**4.3 ± 0.4**	1.5 ± 0.1	B.D.	B.D.
AAU	Ile AUU	**4.5 ± 0.5**	0.43 ± 0.1	B.D.	0.27 ± 0.03
AAC	Val GUU	**0.84 ± 0.1**	0.59 ± 0.04	B.D.	0.35 ± 0.06
AGA	Ser UCU	**19.7 ± 1.0**	1.4 ± 0.2	B.D.	B.D.
AGG	Pro CCU	**12.9 ± 1.1**	10.3 ± 0.5	0.92 ± 0.1	1.0 ± 0.04
AGU	Thr ACU	**10.7 ± 0.5**	1.4 ± 0.2	B.D.	B.D.
AGC	Ala GCU	**1.3 ± 0.1**	1.2 ± 0.1	B.D.	B.D.
AUG	His CAU	**6.1 ± 0.4**	2.9 ± 0.1	B.D.	———
AUU	Asn AAU	**7.2 ± 0.4**	B.D.	B.D.	B.D.
AUC	Asp GAU	**3.5 ± 0.5**	B.D.	B.D.	B.D.
ACA	Cys UGU	**3.6 ± 0.1**	B.D.	———	———
ACU	Arg CGU	**7.6 ± 0.5**	6.1 ± 0.3	0.91 ± 0.1	B.D.
ACU	Ser AGU	**9.4 ± 0.8**	B.D.	B.D.	0.36 ± 0.04
ACC	Gly GGU	**1.3 ± 0.3**	0.40 ± 0.1	B.D.	0.33 ± 0.01

Reassignment efficiencies in bold are those for the codon targeted for reassignment, which fully base pairs to the orthogonal tRNA via Watson-Crick interactions.
Reported reassignment efficiencies represent the mean and standard deviation of multiple biological replicates (typically >12) of each system. A complete list of the number of biological replicates that comprise each reported efficiency is provided in Supplementary Table S2.
[a]For continuity, "P" and "Q" are used to represent the first two positions of the codon in the column headers. The sense codons evaluated with a single tRNA in each row are determined by substituting the complement of the nucleobases in the second and third anticodon positions for P and Q. The third codon position is specified in the column headers. For example, row 1 shows sense codon reassignment by the *M. jannaschii* $tRNA_{AAA}$. Data in the PQU column are for reassignment at the UUU codon, with full Watson-Crick base pairing possible between codon and anticodon. Data in the PQC column are for reassignment at the UUC codon.
[b]B.D. indicates that the codon was evaluated with the specified tRNA, and the measurement was below the detection limit of the in cell assay (0.2%).
[c]"———" indicates that the codon was not evaluated for reassignment by the specified tRNA.

The degree to which U3:C3 codon discrimination may indicate inosine modification varies depending on the incorporation efficiency of the orthogonal tRNA and the detection limit of the assay. As targeted codon reassignment efficiency decreases, a higher percentage of the anticodon must be modified to inosine to reassign the C-ending codon above the limit of detection. Based on the strength of the I34/C3 pairing and expected weakness of an unmodified A34/C3 pairing, as little as 2% inosine modification of the tRNA molecules would be detected as "anomalous" reading of the C-ending codon by an A34 anticodon that reassigns the targeted U-ending codon at 20%. In the case of A34 anticodons that reassign their targeted codons with low efficiency (e.g., 1%), 40% inosine modification of the tRNA molecules would be required for detection as anomalous C-ending codon reassignment.

Our determination of the extent to which each of the tRNAs employed for sense codon reassignment were modified to inosine relied on Sanger sequencing of reverse transcribed tRNAs (Wolf et al., 2002). The tRNA substrates for reverse transcription were isolated using phenol:chloroform extraction. Reverse transcriptase recognizes inosine as guanine and triggers incorporation of cytosine. Unmodified adenosine, on the other hand, triggers incorporation of only thymidine (Wolf et al., 2002). The extent of inosine modification was determined by the ratio of T:C signal corresponding to position 34. Sequencing has been the standard method for detection of inosine modification; the method is semi-quantitative and reasonably sensitive. Other direct detection methods have been developed, but sequencing remains the most commonly used method of inosine detection (Vik et al., 2013; Torres et al., 2018; Knutson et al., 2020; Srinivasan et al., 2021). The observed baseline signals in the sequencing chromatograms suggest that ~5% inosine modification would be clearly evident in chromatograms. Conclusions that individual tRNAs are not modified are based on sequencing traces where incorporation of cytosine, indicative of inosine modification, corresponding to position 34 is indistinguishable from baseline (Figure 3).

Figure 1 indicates the inosine modification status of each of the A34 tRNAs. Despite observation of a wide range of U3:C3 discrimination ratios, inosine modification is detected in only three A34 anticodons on the *M. jannaschii* tRNA body. Orthogonal tRNAs with three of four possible ANG anticodons: ACG (targeting Arg CGU), AUG (targeting His CAU), and AAG (targeting Leu CUU) showed indications of partial inosine modification. Figure 3 presents the relevant section of the chromatogram for Sanger sequencing the cDNA from reverse transcription of the 15 orthogonal tRNAs, arranged to mirror to the standard genetic code table. Only the chromatogram from sequencing *M. jannaschii* $tRNA_{AUG}$ has been reported previously.

In some cases, the reassignment efficiencies reported in Figure 1 and Table 1 differ slightly from those previously reported. As additional biological replicates are added for any given system,

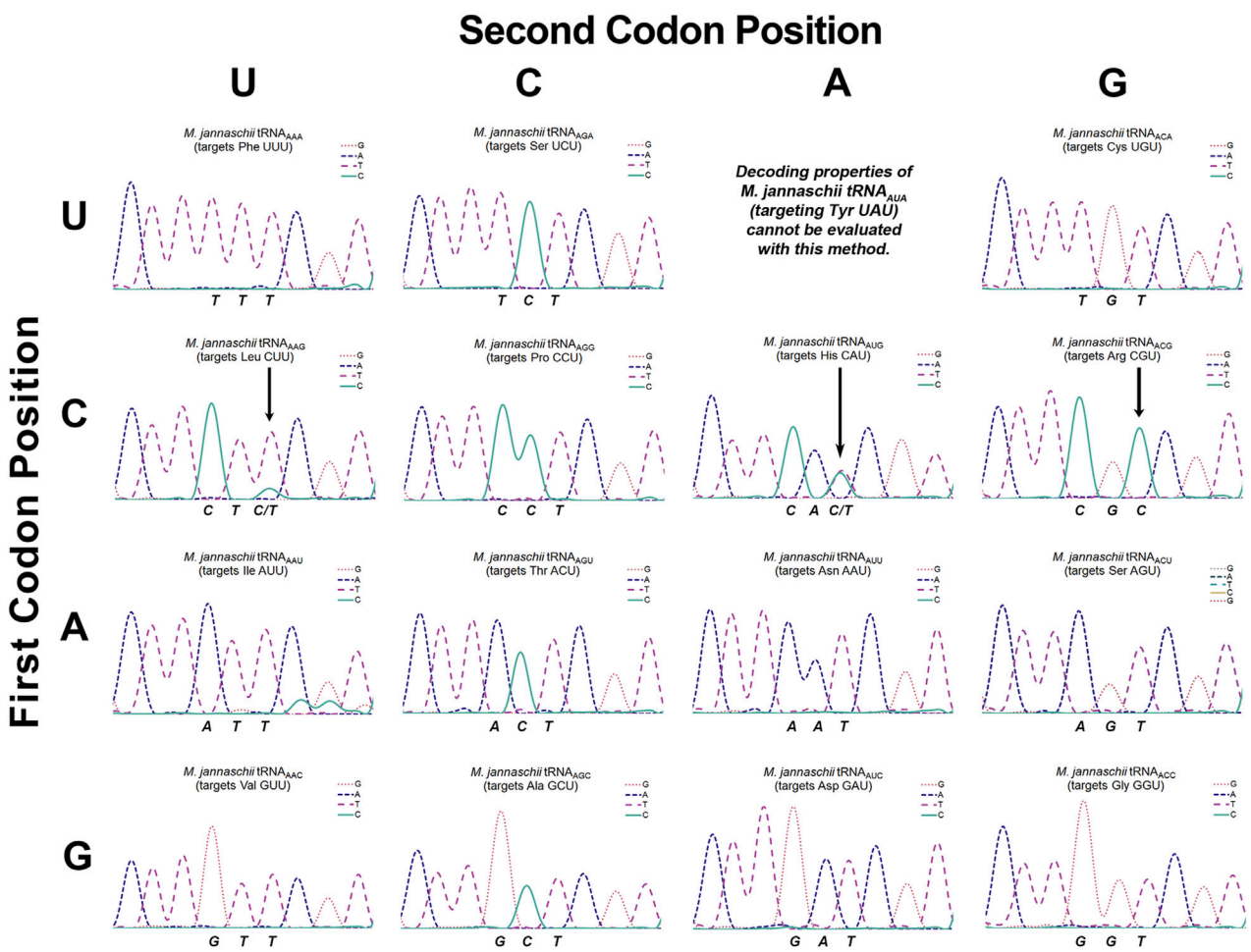

FIGURE 3
Sanger sequencing chromatogram traces of cDNAs reverse transcribed from isolated *M. jannaschii* tRNAs with ANN anticodons. The figure is arranged to mirror the standard genetic code table. Sequencing traces are placed according to the codon targeted for reassignment by the A34 tRNA. The sequenced strand corresponds to the reverse complement of the tRNA. The three nucleotides that specify the codon targeted for reassignment correspond to the anticodon of the tRNA and are stated below each trace. For example, the sequencing output for $tRNA_{AAA}$ (targets Phe UUU) is TTT. Three arrows in traces for the CNU codon row indicate detectable inosine modification of the adenosine at position 34.

we update the stated reassignment efficiency to reflect the accumulated data at the time of publication. Data gathered from multiple experiments across multiple days are readily comparable, demonstrating the robustness and stability of the screen. The consistency of observed reassignment efficiencies across experiments (e.g., batch-to-batch media preparation, person performing the experiment, fluorescence plate reader employed) is evident in the small standard deviations of the measurements.

3.1 Decoding preferences of orthogonal tRNAs with ANG anticodons

The family of A34 tRNAs with ANG anticodon sequences differ from the *E. coli* $tRNA^{Arg2}{}_{ACG}$ anticodon sequence at position 35, which base pairs to the second position of the mRNA codon. The engineered ANG anticodons target the CNU codons found in the same row of the genetic code table as the arginine four box (CGN codons). $tRNA_{AAG}$ targets the Leu CUU codon, $tRNA_{AGG}$ targets the Pro CCU codon, and $tRNA_{AUG}$ targets the His CAU codon. Orthogonal $tRNA_{ACG}$ targets the Arg CGU codon and, on the *M. jannaschii* tRNA body, has a transcribed anticodon loop sequence identical to that of *E. coli* $tRNA^{Arg2}$ (Figure 2) (Chin et al., 2002). The *E. coli* $tRNA^{Arg2}$ is modified at positions 32 and 37 in the anticodon loop. The modification state of the orthogonal *M. jannaschii* tRNAs has not been characterized beyond the evaluation of inosine modification at position 34.

The *M. jannaschii* $tRNA_{ACG}$ is the only A34 tRNA variant for which we have not previously reported decoding efficiencies for the U- and C-ending codons. This orthogonal tRNA competes against the most abundant *E. coli* tRNA, and the CGU and CGC codons are not efficiently reassigned at 7.6% ± 0.5% and 6.1% ± 0.3%, respectively. The U3:C3 discrimination ratio for the *M. jannaschii* $tRNA_{ACG}$ is 55:45. Sanger sequencing of the cDNA of the reverse transcribed orthogonal tRNA showed only C incorporated opposite position 34 of the anticodon, indicating a high level of inosine modification. The degree of modification evident in Sanger sequencing of the orthogonal $tRNA_{ACG}$ was

indistinguishable from the extent of modification determined from sequencing reverse transcribed *E. coli* $tRNA^{Arg2}$, which is quantitatively modified.

As reported previously, a nearly 1:1 mixture of C and T is observed upon sequencing the reverse transcript of the *M. jannaschii* $tRNA_{AUG}$ (Biddle et al., 2016). The 68:32 U3:C3 discrimination ratio for this tRNA suggests that the unmodified A34 version of the tRNA is selective for reading the U-ending codon, while the inosine-modified form is largely responsible for reading the C-ending codon. We identified several anticodon loop sequences that completely abrogate inosine modification from a small library of position 32, 37, and 38 variants. $tRNA_{AUG}$ anticodon loop variants that authentically contain unmodified adenosine at position 34 display U3:C3 discrimination ratios greater than 95:5.

M. jannaschii $tRNA_{AAG}$ targets the Leu CUU codon and displays a U3:C3 discrimination ratio of 76:24. Sanger sequencing of the cDNA showed a mixture of nucleotides corresponding to the A34 position, primarily T with a small amount of C, indicating partial modification to inosine (Figure 3). Relative peak areas suggest that approximately 15% of the anticodons are inosine modified. This degree of inosine modification coupled with the U3:C3 discrimination ratio suggests that the unmodified A34 tRNA modestly discriminates between U- and C- ending codons, in contrast to what is observed for histidine. Inosine modification was not detected by sequencing the *M. jannaschii* $tRNA_{AAG}$ with an A37G mutation. Discrimination by *M. jannaschii* $tRNA_{AAG\text{-}A37G}$ improved to 87:13 (Supplementary Figure S1).

The U3:C3 discrimination ratio for $tRNA_{AGG}$ targeting Pro CCU, 56:44, is nearly identical to that of $tRNA_{ACG}$ (arginine). Given the trends observed for other ANG anticodon tRNAs, we expected to see evidence of inosine modification in *M. jannaschii* $tRNA_{AGG}$. Much to our surprise, no inosine modification was detected via Sanger sequencing of several cDNA preparations of reverse transcribed tRNAs (Figure 3 and additional traces in Supplementary Figure S2). The nearly-equal reassignment efficiency of the Pro CCU and CCC codons is the result of a mechanism(s) other than inosine modification. The apparent lack of discrimination between CCU and CCC codons in *Salmonella* by a mutant proline tRNA with an A34 anticodon, also without inosine modification, has been previously reported (Chen et al., 2002).

3.2 Decoding preferences of orthogonal tRNAs with ACN anticodons

The ACN anticodon family of tRNAs differ from the arginine sequence at position 36, which base pairs with the first position of the mRNA codon. The ACN anticodons target the NGU codons found in the same column of the genetic code table as the arginine four box (CGN codons). $tRNA_{ACA}$ targets the Cys UGU codon, $tRNA_{ACU}$ targets the Ser AGU codon, and $tRNA_{ACC}$ targets the Gly GGU codon. For both $tRNA_{ACA}$ and $tRNA_{ACU}$, reassignment of the corresponding C-ending codon is below the limit of detection for our in cell assay. The calculated U3:C3 discrimination ratios are at least 95:5, depending on the efficiency of decoding the targeted U-ending codon (Figure 1). In both cases, only T is incorporated in response to the nucleotide corresponding to A34 in the Sanger sequencing of the reverse transcribed tRNAs (Figure 3). The U3:C3 codon discrimination ratios do not suggest partial inosine modification, and none is observed. In contrast, fluorescent GFP is produced at a level above that of the limit of detection when the *M. jannaschii* $tRNA_{ACC}$ is evaluated for reassignment of the Gly GGC codon, leading to a calculated U3:C3 discrimination ratio of 76:24. This discrimination ratio is similar to that of the *M. jannaschii* $tRNA_{AAG}$ (Leu CUU), which also differs from the ACG anticodon by a single nucleotide and for which partial inosine modification was detected. The chromatogram from Sanger sequencing of the *M. jannaschii* $tRNA_{ACC}$ does not indicate inosine modification (Figure 3).

3.3 Codon discrimination preferences of A34 tRNAs differing at both the first and second codon positions

The remaining nine tRNAs with ADH anticodons differ from the arginine ACG anticodon at both positions 35 and 36 (The degenerate nucleotide D indicates "not C", and the degenerate nucleotide H indicates "not G.") We hypothesized that these tRNAs would be the least likely substrates for *E. coli* TadA. Indeed, sequencing chromatograms for the eight tRNAs we evaluated shows only T incorporated in response to the nucleotide corresponding to position 34, indicating an absence of inosine modification. The U3:C3 discrimination ratios for this group of tRNAs suggest a strong bias for U3 codons (Figure 1). Six of the eight tRNAs have U3:C3 codon discrimination ratios greater than or equal to 10 to 1.

Both the *M. jannaschii* $tRNA_{AAC}$ (Val GUU) and *M. jannaschii* $tRNA_{AGC}$ (Ala GCU) decode the corresponding U- and C-ending codons with nearly equal efficiency. Despite neither orthogonal tRNA reassigning the targeted U-ending codon efficiently, both around 1%, reassignment of the C-ending codon is well above the limit of detection. The U3:C3 discrimination ratios are 59:41 and 52:48, respectively.

M. jannaschii $tRNA_{AGA}$ (Ser UCU), *M. jannaschii* $tRNA_{AGU}$ (Thr ACU), and *M. jannaschii* $tRNA_{AAU}$ (Ile AUU) also decode their corresponding C-ending codons above the limit of detection. However, given the efficiency with which each of these tRNAs decodes their targeted U-ending codon, U3:C3 discrimination ratios remain quite good, greater than or equal to approximately 10:1. The remaining three *M. jannaschii* A34 tRNAs ($tRNA_{AAA}$ targeting Phe UUU; $tRNA_{AUU}$ targeting Asn AAU; $tRNA_{AUC}$ targeting Asp GAU) do not reassign the corresponding C-ending codon above the limit of detection. In each case, the tRNA reassigns the targeted U-ending codon with greater efficiency than either $tRNA_{AAC}$ (Val GUU) or $tRNA_{AGC}$ (Ala GCU), suggesting a strong bias for decoding U-ending codons within this group of tRNAs.

4 Discussion

4.1 Insight into the specificity of *E. coli* TadA

The precise determinants of adenosine deaminase TadA specificity have not been mapped, in part because bacteria typically only encode a single tRNA with a modifiable

A34 anticodon. The original report describing the function of TadA as a tRNA deaminase relied on a small set of tRNA minisubstrates to evaluate the residues important for tRNA recognition (Wolf et al., 2002). Evaluation of 10 minihelix substrates suggested that the anticodon sequence was the major determinant of TadA recognition. The determinants of eukaryotic ADAT enzymes have been more fully elucidated (Frigolé et al., 2019). Our experiments employ a full tRNA body which presents its anticodon in an effective orientation for recognition by *E. coli* TadA, given that when the anticodon is AGC presumably full inosine modification is detected. With the exception of possible differences in modification state at positions 32 and 37, the anticodon loop of *M. jannaschii* $tRNA_{ACG}$ is identical to that of *E. coli* $tRNA^{Arg2}$ (Bjork and Hagervall, 2014). We examined 15 of 16 possible A34 anticodon variants; the extent of inosine modification of these tRNAs should more precisely map the anticodon elements important for TadA recognition.

cDNA sequencing revealed indications of inosine modification in only three out of 15 tRNAs evaluated. Just two of the three modified tRNA species are highly modified. Sequencing suggests that the TadA deaminase enzyme is very specific for the sequence of the *E. coli* $tRNA^{Arg2}$ anticodon and anticodon loop. For both the AAG (Leu) and AUG (His) anticodons, each mutation to positions 32, 37 and/or 38 that was evaluated was sufficient to block recognition by TadA. Our results are consistent with studies employing minihelix substrates and recent evaluations of inosine tRNA modification in *Oenococcus oeni*, a bacterial strain that encodes four tRNAs with A34 anticodons, only two of which are inosine modified (Wolf et al., 2002; Rafels-Ybern et al., 2018; Rafels-Ybern et al., 2019).

4.2 Extent of A3 and G3 reading by A34 orthogonal tRNAs

The expanded wobble rules suggest that unmodified A34 tRNAs are capable of decoding any nucleotide at position three of the codon (Bjork and Hagervall, 2014; Agris et al., 2018). As we considered the combination of U3:C3 discrimination ratios and detectable inosine modification, we expanded our set of evaluated codons and measured the decoding properties of *M. jannaschii* $tRNA_{ANN}$ species to the corresponding A- and G-ending codons, regardless of whether the codons were synonymous with the U- and C-ending codons. Table 1 presents a nearly complete evaluation of the reassignment efficiencies of *M. jannaschii* A34 tRNAs within their respective codon boxes. Measurements shaded in pale pink have not been previously reported.

Across the 12 A34 tRNAs which have been fully evaluated, six tRNAs show no reassignment of either the A- or G-ending codons. A single tRNA, *M. jannaschii* $tRNA_{AGG}$ (Pro CCU), decodes all four codons above the limit of detection. This observation adds to the conclusion that factors other than inosine modification are influencing the behavior of this tRNA in *E. coli* translation. The CCA and CCG codons are each reassigned approximately 1% of the time. A single tRNA, *M. jannaschii* $tRNA_{ACG}$ (Arg CGU), reassigns the CGA codon above the limit of detection (0.91% ± 0.1%) and does not reassign the CGG codon. This observation is unsurprising given the decoding preferences of tRNAs with inosine at position 34.

$tRNA_{ACU}$ (Ser AGU) represents the only instance of an *M. jannaschii* A34 tRNA decoding only the U- and G-ending codons. Reassignment of the C- and A-ending codons is below the limit of detection. The AGG codon is the least commonly used sense codon in *E. coli* and is decoded by the least abundant endogenous tRNA species (Dong et al., 1996). The AGA codon is slightly more frequently used in *E. coli*, and an additional endogenous tRNA is utilized to decode it. This favorable competition environment to decode AGG, relative to even AGA, may explain the quantified reassignment preferences of the introduced $tRNA_{ACU}$ and is one reason this codon has been explored for genetic code expansion (Zeng et al., 2014; Lee et al., 2015; Mukai et al., 2015; Biddle et al., 2022).

The remaining three tRNAs, $tRNA_{AAU}$ (Ile AUU), $tRNA_{AAC}$ (Val GUU), and $tRNA_{ACC}$ (Gly GGU), decode both the C- and G-ending codons above the limit of detection and not the A-ending codon. The relative order of codon preference observed in an *in vitro* study was that A34 tRNAs generally preferred U > C > G > A (Boren et al., 1993).

4.3 Function of unmodified A34 tRNAs in translation

We observed a wide range of U3:C3 discrimination ratios; sequencing indicated inosine modification in only three out of 15 tRNAs evaluated. Our data describe a wider examination of the function of authentic A34 tRNAs than has been previously performed. The measurements of A34 tRNA decoding specificity speak to two related and relatively unstudied questions. 1. How do A34 tRNAs function in translation (Chen et al., 2002)? 2. Why are unmodified A34 tRNAs essentially absent from the tRNA complements of extant organisms (Ehrlich et al., 2021)? The wide range of observed U3:C3 discrimination ratios suggests that the functions of adenosine wobble interactions in codon recognition are idiosyncratic. Significant evidence that the effects of individual tRNA modifications also seem to be tRNA and codon context dependent exists (Phizicky and Alfonzo, 2010).

The set of directed sense codon reassignment experiments by *M. jannaschii* $tRNA_{ANN}$ variants employ a single tRNA body with a single attached amino acid to interrogate the behavior of different anticodons in the background of *in vivo E. coli* translation. The major variables in our evaluations are the identity of the codon/anticodon pair measured and the subset of *E. coli* tRNAs against which the orthogonal tRNA competes. Our results greatly expand the set of experimental measurements of the function of A34 tRNAs in *in vivo* translation systems (Boren et al., 1993; Inagaki et al., 1995; Takai et al., 1999; Chen et al., 2002; Agris et al., 2018). The pattern of discrimination appears to be consistent with the "two-out-of-three" and "strong and weak" codon decoding hypotheses (Lagerkvist, 1978; Grosjean and Westhof, 2016).

Reading of codons by A34 tRNAs was addressed only in passing in Crick's original discussion of the wobble hypothesis. Unmodified A34 anticodons were expected to read only U-ending codons (Crick, 1966). A major component of the argument supporting the wobble hypothesis was that modification of A34 tRNAs to inosine expanded the set of codons read by a single tRNA and thus reduced the set of tRNAs needed to read the genetic code. Within the anticodon

sequences available at the time, inosine was present in several tRNA species. Those tRNAs were capable of pairing with U-, C-, and A-ending codons. As more tRNA sequences were determined, the facts that unmodified A34 tRNAs were extremely rare and that eukaryotes, in particular, largely utilized I34 tRNAs to read three and four codon boxes became increasingly clear.

While broadly correct in outline, the wobble hypothesis has been updated. The first adjustments reconciled decoding mechanisms in *Mycoplasma* and mitochondria, both of which utilize highly abbreviated tRNA complements. Identification of the large number of modified tRNA nucleotides in the anticodon stem loop and appreciation for their impact on codon/anticodon pairing has further refined the original hypothesis (Agris, 1991; Yokoyama and Nishimura, 1995; Grosjean and Westhof, 2016; Agris et al., 2018).

Unmodified A34 tRNAs are extremely rare, and their decoding properties have only been studied in a few instances (Rafels-Ybern et al., 2018; Rafels-Ybern et al., 2019). The majority of tRNAs that are encoded with an A34 anticodon are modified to inosine. In eukaryotes, each of the TAPSLIVR amino acids are primarily introduced by tRNAs with inosine at position 34 (Rafels-Ybern et al., 2018; Rafels-Ybern et al., 2019). In bacteria, only a single tRNA with an inosine-modified A34 is commonly employed to introduce arginine in the CGN codon box. Archaea do not typically utilize any A34 tRNAs (Phillips and de Crécy-Lagard, 2011). A recent analysis of the encoded tRNA sets from over 1,000 bacterial and eukaryotic genomes representing all major phylogenetic taxa confirmed that A34 tRNAs are extremely rare in bacteria and archaea. A handful of instances of bacterial genomes encoding additional A34 tRNAs beyond an arginine-incorporating variant are evident (Rafels-Ybern et al., 2018; Rafels-Ybern et al., 2019). While some bacterial taxa have expanded their repertoire of encoded A34 tRNAs, other bacteria have lost the inosine modification and modifying enzyme completely (Yokobori et al., 2013; Rafels-Ybern et al., 2018; Rafels-Ybern et al., 2019).

Until the 2018 genomic analysis study, unmodified A34 tRNAs had only been observed in organisms and organelles with highly reduced genomes or studied as isolated or prepared mutant tRNAs. The translation of organellar and *Mycoplasma* genomes employ greatly reduced tRNA complements relative to complements in typical bacteria and eukaryotic systems. The *Mycoplasma* and organellar-like interactions have been referred to as "super-wobble" interactions (Rogalski et al., 2008). Yeast mitochondria encode a single unmodified arginine tRNA with an ACG anticodon as the only tRNA capable of decoding the CGN codon box. Unlike bacteria, yeast mitochondria rarely use CGN codons to encode arginine. The function of this tRNA species has not been experimentally evaluated. *Mycoplasma* species employ an authentic A34 threonine tRNA (Andachi et al., 1987; Samuelsson et al., 1987). The function of this tRNA has been evaluated in *in vitro* translation experiments employing *Mycoplasma* S30 extracts. These experiments show that the A34 threonine tRNA is capable of decoding ACU, ACC, and ACG codons (Inagaki et al., 1995).

The function of A34 tRNAs have been evaluated in the context of glycine and serine tRNA backgrounds in *E. coli*-derived *in vitro* translation systems. An overexpressed and purified *E. coli* glycyl $tRNA_{ACC}$ strongly decoded GGU, weakly decoded GGC, and surprisingly competed with glycyl $tRNA_{UCC}$ to decode both GGA and GGG (Boren et al., 1993). A closely-related *in vitro* translation experiment examining the properties of A34 tRNAs in the context of a Ser tRNA backbone showed that A34 anticodons strongly discriminated and read predominantly U-ending codons, with minor reading of C-ending codons and no apparent reading of A- or G-ending codons in the Ser UCN codon box (Takai et al., 1999).

For 10 of the 15 codons evaluated, authentic orthogonal A34 tRNAs show a strong preference for U3 as opposed to C3 codons. Discrimination ratios are greater than or equal to approximately 10 to one for *M. jannaschii* A34 tRNAs targeting codons for Phe, Cys, Ile, Thr, Asn, and Ser (both possible A34 anticodons). Orthogonal A34 tRNAs for Leu and His exhibit high discrimination ratios in variants that are not inosine modified. No orthogonal tRNA with an ACG anticodon, targeting the Arg CGU codon, that was not predominantly inosine modified was identified. Generation of an orthogonal $tRNA_{ACG}$ with an authentic A34 is of particular interest as it would allow characterization of the competition between orthogonal $tRNA_{ACG}$ and endogenous $tRNA_{ICG}$.

The codons for proline, alanine, and glycine include GC pairs in the first two positions of the codon. That the third position of interaction in codon/anticodon recognition is less important as a result of the strength of the G/C interactions describes the expectation of the "two-out-of-three" and "strong and weak" codon hypotheses (Lagerkvist, 1978; Grosjean and Westhof, 2016). The low U3:C3 discrimination ratios as well as detectable decoding of the G-ending codons for each of these authentic A34 orthogonal tRNAs support these interpretations of genetic code reading. In the cases of proline and alanine, the A34 tRNAs essentially do not discriminate between C- and U-ending codons. In the case of glycine, the U-ending codon is slightly preferred. The arginine four box is the fourth codon box with 2 GC pairs at the first two codon positions. As mentioned, a non-inosine modified version of this tRNA anticodon could not be identified.

5 Conclusion

We set out to explore the extent to which *M. jannaschii* tyrosyl tRNAs with ANN anticodons were modified to inosine by *E. coli* TadA. Given that apparent quantitative A to I modification was observed in the orthogonal *M. jannaschii* $tRNA_{ACG}$, which possesses an anticodon identical to the only *E. coli* tRNA species that undergoes A to I modification ($tRNA^{Arg2}$), and that the closely-related orthogonal $tRNA_{AUG}$ was highly modified, we hypothesized that orthogonal *M. jannaschii* tRNAs with anticodons similar to ACG (i.e., ANG or ACN) may also be substrates for *E. coli* TadA. We also recognized that the heterodimeric ADATs which are responsible for modifying 8 A34 tRNAs to inosine in eukaryotes evolved from the bacterial TadA, and that *E. coli* TadA could have broader substrate recognition than had been previously observed. We expected the degree to which an orthogonal A34 tRNA was able to distinguish between U- and C-ending codons would correlate with the extent of A-to-I modification, as unmodified A34 tRNAs are expected to have a strong preference for pairing with U in the third codon position.

The expectation that A34 tRNAs show a strong preference for U-ending codons was based on a limited set of sometimes conflicting experimental data on the function of A34 tRNAs in translation as well as our previous observation that "anomalous" decoding of histidine CAC codons by the *M. jannaschii* $tRNA_{AUG}$ was due to inosine modification (Boren et al., 1993; Inagaki et al., 1995; Takai et al., 1999; Chen et al., 2002; Biddle et al., 2016; Agris et al., 2018). Anticodon loop sequence variants of $tRNA_{AUG}$ that abrogated inosine modification displayed high level discrimination between U- and C-ending codons. We anticipated that U3:C3 discrimination ratios might provide insight on inosine modification states of corresponding A34 tRNAs.

We observed a wide range of U3:C3 discrimination ratios, yet sequencing indicated inosine modification in only three out of fifteen tRNAs evaluated. Thus, orthogonal pair directed sense codon reassignment using the *M. jannaschii* tRNA/aaRS pair enabled a wider examination of the function of authentic A34 tRNAs in *in vivo* translation than has previously been performed. For the most part, unmodified A34 tRNAs largely pair with only U3 codons as the original wobble rules suggest. In the instances with GC pairs in the first two positions of the codon (i.e., proline, alanine, and glycine), unmodified A34 tRNAs readily decode the corresponding C- and G-ending codons. In only one instance is decoding of an A-ending codon by an authentic A34 tRNA observed. $tRNA_{AGG}$ decoding each codon in the proline four box may represent a unique situation pertaining to proline as opposed to an indication of broader A34/A3 pairing tendencies. Inosine modification explains the other A-ending codon reassigned with an efficiency above the limit of detection.

This evaluation demonstrates that authentic A34 tRNAs perform effectively in the context of the *E. coli* translation apparatus (Chen et al., 2002). Our data offer no apparent functional reason why A34 tRNAs did not evolve to translate codon four boxes. The absence of unmodified A34 tRNAs across the domains of life suggests that more than just pressure for reduced tRNA complements is responsible for the absence of A34 tRNAs, but a clear functional deficiency is not evident in *E. coli*.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

MS: Writing–review and editing, Writing–original draft, Visualization, Methodology, Investigation, Funding acquisition, Conceptualization. JT: Writing–review and editing, Methodology, Investigation. JF: Writing–review and editing, Supervision, Funding acquisition, Conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by the National Science Foundation, grant number CHE-1507055 to JF. The APC was funded by University of Colorado Denver.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1386299/full#supplementary-material

References

Acevedo-Rocha, C. G., and Budisa, N. (2016). Xenomicrobiology: a roadmap for genetic code engineering. *Microb. Biotechnol.* 9 (5), 666–676. doi:10.1111/1751-7915.12398

Agris, P. F. (1991). Wobble position modified nucleosides evolved to select transfer-RNA codon recognition - a modified-wobble hypothesis. *Biochimie* 73 (11), 1345–1349. doi:10.1016/0300-9084(91)90163-u

Agris, P. F., Eruysal, E. R., Narendran, A., Väre, V. Y. P., Vangaveti, S., and Ranganathan, S. V. (2018). Celebrating wobble decoding: half a century and still much is new. *RNA Biol.* 15 (4-5), 537–553. doi:10.1080/15476286.2017.1356562

Andachi, Y., Yamao, F., Iwami, M., Muto, A., and Osawa, S. (1987). Occurrence of unmodified adenine and uracil at the first position of anticodon in threonine tRNAs in Mycoplasma capricolum. *Proc. Natl. Acad. Sci. U.S.A.* 84 (21), 7398–7402. doi:10.1073/pnas.84.21.7398

Andersson, S. G. E., and Kurland, C. G. (1990). Codon preferences in free-living microorganisms. *Microbiol. Rev.* 54 (2), 198–210. doi:10.1128/mmbr.54.2.198-210.1990

Biddle, W., Schmitt, M. A., and Fisk, J. D. (2015). Evaluating sense codon reassignment with a simple fluorescence screen. *Biochemistry* 54 (50), 7355–7364. doi:10.1021/acs.biochem.5b00870

Biddle, W., Schmitt, M. A., and Fisk, J. D. (2016). Modification of orthogonal tRNAs: unexpected consequences for sense codon reassignment. *Nucleic Acids Res.* 44 (21), 10042–10050. doi:10.1093/nar/gkw948

Biddle, W., Schwark, D. G., Schmitt, M. A., and Fisk, J. D. (2022). Directed evolution pipeline for the improvement of orthogonal translation machinery for genetic code expansion at sense codons. *Front. Chem.* 10, 815788. doi:10.3389/fchem.2022.815788

Bjork, G. R., and Hagervall, T. G. (2014). Transfer RNA modification: presence, synthesis, and function. *EcoSal Plus* 6 (1). doi:10.1128/ecosalplus.ESP-0007-2013

Bohlke, N., and Budisa, N. (2014). Sense codon emancipation for proteome-wide incorporation of noncanonical amino acids: rare isoleucine codon AUA as a target for genetic code expansion. *FEMS Microbiol. Lett.* 351 (2), 133–144. doi:10.1111/1574-6968.12371

Boren, T., Elias, P., Samuelsson, T., Claesson, C., Barciszewska, M., Gehrke, C. W., et al. (1993). Undiscriminating codon reading with adenosine in the wobble position. *J. Mol. Biol.* 230 (3), 739–749. doi:10.1006/jmbi.1993.1196

Cervettini, D., Tang, S., Fried, S. D., Willis, J. C. W., Funke, L. F. H., Colwell, L. J., et al. (2020). Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase-tRNA pairs. *Nat. Biotechnol.* 38 (8), 989–999. doi:10.1038/s41587-020-0479-2

Chan, P. P., and Lowe, T. M. (2016). Gtrnadb 2.0: an expanded database of transfer rna genes identified in complete and draft genomes. *Nucleic Acids Res.* 44 (D1), D184–D189. doi:10.1093/nar/gkv1309

Chen, P., Qian, Q., Zhang, S. P., Isaksson, L. A., and Bjork, G. R. (2002). A cytosolic tRNA with an unmodified adenosine in the wobble position reads a codon ending with the non-complementary nucleoside cytidine. *J. Mol. Biol.* 317 (4), 481–492. doi:10.1006/jmbi.2002.5435

Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002). Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 99 (17), 11020–11024. doi:10.1073/pnas.172226299

Crick, F. H. C. (1966). Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* 19 (2), 548–555. doi:10.1016/s0022-2836(66)80022-0

Dong, H. J., Nilsson, L., and Kurland, C. G. (1996). Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* 260 (5), 649–663. doi:10.1006/jmbi.1996.0428

dos Reis, M., Wernisch, L., and Savva, R. (2003). Unexpected correlations between gene expression and codon usage bias from microarray data for the whole *Escherichia coli* K-12 genome. *Nucleic Acids Res.* 31 (23), 6976–6985. doi:10.1093/nar/gkg897

Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015). Designing logical codon reassignment - expanding the chemistry in biology. *Chem. Sci.* 6 (1), 50–69. doi:10.1039/c4sc01534g

Dunkelmann, D. L., Oehm, S. B., Beattie, A. T., and Chin, J. W. (2021). A 68-codon genetic code to incorporate four distinct non-canonical amino acids enabled by automated orthogonal mRNA design. *Nat. Chem.* 13 (11), 1110–1117. doi:10.1038/s41557-021-00764-5

Ehrlich, R., Davyt, M., López, I., Chalar, C., and Marín, M. (2021). On the track of the missing tRNA genes: a source of non-canonical functions? *Front. Mol. Biosci.* 8, 643701. doi:10.3389/fmolb.2021.643701

Elias, Y., and Huang, R. H. (2005). Biochemical and structural studies of a-to-I editing by tRNA:A34 deaminases at the wobble position of transfer RNA. *Biochemistry* 44 (36), 12057–12065. doi:10.1021/bi050499f

Frigolé, H. R., Camacho, N., Coma, M. C., Fernández-Lozano, C., García-Lema, J., Rafels-Ybern, A., et al. (2019). tRNA deamination by ADAT requires substrate-specific recognition mechanisms and can Be inhibited by tRFs. *RNA* 25 (5), 607–619. doi:10.1261/rna.068189.118

Furter, R. (1998). Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in *Escherichia coli*. *Protein Sci.* 7 (2), 419–426. doi:10.1002/pro.5560070223

Gerber, A. P., and Keller, W. (1999). An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* 286 (5442), 1146–1149. doi:10.1126/science.286.5442.1146

Grosjean, H., and Westhof, E. (2016). An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* 44 (17), 8020–8040. doi:10.1093/nar/gkw608

Ho, J. M., Reynolds, N. M., Rivera, K., Connolly, M., Guo, L.-T., Ling, J., et al. (2016). Efficient reassignment of a frequent serine codon in wild-type *Escherichia coli*. *ACS Synth. Biol.* 5 (2), 163–171. doi:10.1021/acssynbio.5b00197

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., et al. (1965). Structure of a ribonucleic acid. *Science* 147 (3664), 1462–1465. doi:10.1126/science.147.3664.1462

Inagaki, Y., Kojima, A., Bessho, Y., Hori, H., Ohama, T., and Osawa, S. (1995). Translation of synonymous codons in family boxes by Mycoplasma capricolum tRNAs with unmodified uridine or adenosine at the first anticodon position. *J. Mol. Biol.* 251 (4), 486–492. doi:10.1006/jmbi.1995.0450

Ingram, V. M., and Sjoquist, J. A. (1963). Studies on the structure of purified alanine and valine transfer RNA from yeast. In: synthesis and structure of macromolecules. *Cold Spring Harb. Symp. Quant. Biol.* 28, 133–138. doi:10.1101/sqb.1963.028.01.025

Knight, R. D., Freeland, S. J., and Landweber, L. F. (2001). Rewiring the keyboard evolvability of the genetic code. *Nat. Rev. Genet.* 2 (1), 49–58. doi:10.1038/35047500

Knutson, S. D., Korn, M. M., Johnson, R. P., Monteleone, L. R., Dailey, D. M., Swenson, C. S., et al. (2020). Chemical profiling of A-to-I RNA editing using a click-compatible phenylacrylamide. *Chem. Eur. J.* 26 (44), 9874–9878. doi:10.1002/chem.202001667

Kuratani, M., Ishii, R., Bessho, Y., Fukunaga, R., Sengoku, T., Shirouzu, M., et al. (2005). Crystal structure of tRNA adenosine deaminase (TadA) from *aquifex aeolicus*. *J. Biol. Chem.* 280 (16), 16002–16008. doi:10.1074/jbc.M414541200

Kwon, I., and Choi, E. S. (2016). Forced ambiguity of the leucine codons for multiple-site-specific incorporation of a noncanonical amino acid. *PLOS ONE* 11 (3), e0152826. doi:10.1371/journal.pone.0152826

Kwon, I., Kirshenbaum, K., and Tirrell, D. A. (2003). Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc.* 125 (25), 7512–7513. doi:10.1021/ja0350076

Lagerkvist, U. (1978). "Two out of three": an alternative method for codon reading. *Proc. Natl. Acad. Sci. U.S.A.* 75 (4), 1759–1762. doi:10.1073/pnas.75.4.1759

Lee, B. S., Shin, S., Jeon, J. Y., Jang, K.-S., Lee, B. Y., Choi, S., et al. (2015). Incorporation of unnatural amino acids in response to the agg codon. *ACS Chem. Biol.* 10 (7), 1648–1653. doi:10.1021/acschembio.5b00230

Liu, C. C., and Schultz, P. G. (2010). Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* 79, 413–444. doi:10.1146/annurev.biochem.052308.105824

Losey, H. C., Ruthenburg, A. J., and Verdine, G. L. (2006). Crystal structure of *Staphylococcus aureus* tRNA adenosine deaminase TadA in complex with RNA. *Nat. Struct. Mol. Biol.* 13 (2), 153–159. doi:10.1038/nsmb1047

Melnikov, S. V., and Söll, D. (2019). Aminoacyl-tRNA synthetases and tRNAs for an expanded genetic code: what makes them orthogonal? *IJMS* 20 (8), 1929. doi:10.3390/ijms20081929

Mukai, T., Lajoie, M. J., Englert, M., and Söll, D. (2017). Rewriting the genetic code. *Annu. Rev. Microbiol. S. Gottesman.* 71, 557–577. doi:10.1146/annurev-micro-090816-093247

Mukai, T., Yamaguchi, A., Ohtake, K., Takahashi, M., Hayashi, A., Iraha, F., et al. (2015). Reassignment of a rare sense codon to a non-canonical amino acid in *Escherichia coli*. *Nucleic Acids Res.* 43 (16), 8111–8122. doi:10.1093/nar/gkv787

Phillips, G., and de Crécy-Lagard, V. (2011). Biosynthesis and function of tRNA modifications in archaea. *Curr. Opin. Microbiol.* 14 (3), 335–341. doi:10.1016/j.mib.2011.03.001

Phizicky, E. M., and Alfonzo, J. D. (2010). Do all modifications benefit all tRNAs? *FEBS Lett.* 584 (2), 265–271. doi:10.1016/j.febslet.2009.11.049

Rafels-Ybern, A., Attolini, C. S. O., and de Pouplana, L. R. (2015). Distribution of adat-dependent codons in the human transcriptome. *IJMS* 16 (8), 17303–17314. doi:10.3390/ijms160817303

Rafels-Ybern, A., Torres, A. G., Camacho, N., Herencia-Ropero, A., Frigolé, H. R., Wulff, T. F., et al. (2019). The expansion of inosine at the wobble position of tRNAs, and its role in the evolution of proteomes. *Mol. Biol. Evol.* 36 (4), 650–662. doi:10.1093/molbev/msy245

Rafels-Ybern, A., Torres, A. G., Grau-Bove, X., Ruiz-Trillo, I., and de Pouplana, L. R. (2018). Codon adaptation to tRNAs with inosine modification at position 34 is widespread among eukaryotes and present in two bacterial phyla. *RNA Biol.* 15 (4-5), 500–507. doi:10.1080/15476286.2017.1358348

Rogalski, M., Karcher, D., and Bock, R. (2008). Superwobbling facilitates translation with reduced tRNA sets. *Nat. Struct. Mol. Biol.* 15 (2), 192–198. doi:10.1038/nsmb.1370

Samuelsson, T., Guindy, Y. S., Lustig, F., Boren, T., and Lagerkvist, U. (1987). Apparent lack of discrimination in the reading of certain codons in mycoplasma-mycoides. *Proc. Natl. Acad. Sci. U.S.A.* 84 (10), 3166–3170. doi:10.1073/pnas.84.10.3166

Schmitt, M. A., Biddle, W., and Fisk, J. D. (2018). Mapping the plasticity of the *Escherichia coli* genetic code with orthogonal pair-directed sense codon reassignment. *Biochemistry* 57 (19), 2762–2774. doi:10.1021/acs.biochem.8b00177

Schwark, D. G., Schmitt, M. A., Biddle, W., and Fisk, J. D. (2020). The influence of competing tRNA abundance on translation: quantifying the efficiency of sense codon reassignment at rarely used codons. *ChemBioChem* 21 (16), 2274–2286. doi:10.1002/cbic.202000052

Söll, D., Jones, D. S., Ohtsuka, E., Faulkner, R. D., Lohrmann, R., Hayatsu, H., et al. (1966). Specificity of sRNA for recognition of codons as studied by the ribosomal binding technique. *J. Mol. Biol.* 19 (2), 556–573. doi:10.1016/s0022-2836(66)80023-2

Srinivasan, S., Torres, A. G., and de Pouplana, L. R. (2021). Inosine in biology and disease. *Genes* 12 (4), 600. doi:10.3390/genes12040600

Takai, K., Takaku, H., and Yokoyama, S. (1999). *In vitro* codon-reading specificities of unmodified tRNA molecules with different anticodons on the sequence background of *Escherichia coli* $tRNA_1^{Ser}$. *Biochem. Biophys. Res. Commun.* 257 (3), 662–667. doi:10.1006/bbrc.1999.0538

Torres, A. G., Piñeyro, D., Filonava, L., Stracker, T. H., Batlle, E., and de Pouplana, L. R. (2014). A-to-I editing on tRNAs: biochemical, biological and evolutionary implications. *FEBS Lett.* 588 (23), 4279–4286. doi:10.1016/j.febslet.2014.09.025

Torres, A. G., Wulff, T. F., Rodriguez-Escriba, M., Camacho, N., and de Pouplana, L. R. (2018). Detection of inosine on transfer RNAs without a reverse transcription reaction. *Biochemistry* 57 (39), 5641–5647. doi:10.1021/acs.biochem.8b00718

Vik, E. S., Nawaz, M. S., Andersen, P. S., Fladeby, C., Bjorås, M., Dalhus, B., et al. (2013). Endonuclease V cleaves at inosines in RNA. *Nat. Commun.* 4, 2271. doi:10.1038/ncomms3271

Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001). Expanding the genetic code of *Escherichia coli*. *Science* 292 (5516), 498–500. doi:10.1126/science.1060077

Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74 (11), 5088–5090. doi:10.1073/pnas.74.11.5088

Wolf, J., Gerber, A. P., and Keller, W. (2002). TadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J.* 21 (14), 3841–3851. doi:10.1093/emboj/cdf362

Yokobori, S., Kitamura, A., Grosjean, H., and Bessho, Y. (2013). Life without $tRNA^{Arg}$ adenosine deaminase TadA: evolutionary consequences of decoding the four CGN codons as arginine in mycoplasmas and other mollicutes. *Nucleic Acids Res.* 41 (13), 6531–6543. doi:10.1093/nar/gkt356

Yokoyama, S., and Nishimura, S. (1995). "Modified nucleosides and codon recognition. tRNA: structure, biosynthesis, and function," in *tRNA: structure, biosynthesis, and function* (Washington DC: Wiley).

Zacharias, D. A., and Tsien, R. Y. (2006). "Molecular biology and mutation of green fluorescent protein," in *Green fluorescent protein: properties, applications, and protocols, 2nd edition*. Editors M. Chalfie and S. R. Kain (Spinger), 83–120.

Zeng, Y., Wang, W., and Liu, W. S. R. (2014). Towards reassigning the rare AGG codon in *Escherichia coli*. *ChemBioChem* 15 (12), 1750–1754. doi:10.1002/cbic.201400075