**Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Ha An Nguyen Date

# Molecular mechanisms of bacterial translation regulation

By

Ha An Nguyen
Doctor of Philosophy

Chemistry

_____
Christine Dunham
Advisor

_____
Vincent Conticello
Committee Member

_____
Khalid Salaita
Committee Member

Accepted:

_____

Kimberly Jacob Arriola, Ph.D.

Dean of the James T. Laney School of Graduate Studies

_____
Date

# Molecular mechanisms of bacterial translation regulation

By

Ha An Nguyen
B.S. University of Richmond, 2016

Advisor: Christine M. Dunham, Ph. D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2021

## ABSTRACT

### Molecular mechanisms of bacterial translation regulation

By Ha An Nguyen

Rapid and accurate translation to produce properly folded and functional proteins is essential to cell growth and survival. The ribosome is a complex macromolecular machine that directs the conversion of a nucleotide-based code (i.e. genes encoded in messenger RNA (mRNA)) into amino acids, the code of proteins. To accomplish this task, transfer RNAs (tRNAs) must accurately “decode” the genetic information on mRNA by pairing their anticodon sequences with mRNA codons, while carrying with them the specific corresponding amino acid. While the correct base pairing between the tRNA anticodon and mRNA codon is the fundamental driving force behind the fidelity of protein synthesis, it is also known that other elements of the tRNA as well as the ribosome also act as regulatory elements during protein synthesis. These additional layers of regulation are needed because the ribosome needs to quickly select the correct tRNA from the large pool of chemically and structurally similar tRNA molecules.

This dissertation analyzes how the ribosome maintains the three nucleotide mRNA frame and decoding fidelity. First, I investigated how the $m^1G37$ modification and insertions of either G37.5 or A37.5 in $tRNA^{Pro}$ sequence can circumvent recognition and proofreading by the ribosome enabling misreading and +1 frameshifting. After that, I studied structures of mismatched tRNA-mRNA pairs in both the ribosomal A and the P sites to uncover mechanisms with which the ribosome can recognize correct from incorrect tRNAs. Using $tRNA^{Ala}$ bound to either a cognate or a near-cognate codon, I found that disrupting the evolutionarily correlated 32-38 pairing in the tRNA renders the ribosomal rRNA nucleotide A1913 unable to recognize the correct tRNA. Finally, I discovered an A-site mRNA positioning mechanism with which mismatches in the P site can trigger post-peptidyl transfer proofreading. Altogether, the biochemical and structural data presented here elucidate previously unappreciated mechanisms with which the integrity of the tRNA-mRNA pairing is regulated to ensure the accuracy and efficiency of gene expression.

# Molecular mechanisms of bacterial translation regulation

By

Ha An Nguyen
B.S. University of Richmond, 2016

Advisor: Christine M. Dunham, Ph. D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2021

# ACKNOWLEDGMENTS

Throughout this PhD, I have received a great deal of support and assistance. I am deeply grateful for everyone who has helped me along the way, even if I am unable to exhaustively list everyone.

First and foremost, I would like to express my deepest appreciation to my advisor, Dr. Christine Dunham, for accepting me into her group as an inexperienced chemistry student, and for going above and beyond in her role as a PI and mentor to get me to where I am today. Her sincere enthusiasm for science inspired me even on the worst of days and she pushed me to achieve so much more than I thought I could.

I'm extremely thankful to my committee, Dr. Vincent Conticello and Dr. Khalid Salaita, for letting me rotate in their labs, for serving on my committee and supervising the progress of my degree, and for all their insightful comments and suggestions. In addition, I also thank Dr. Graeme Conn and Dr. Homa Ghalei for their continued scientific and technical support, as well as supporting me for various applications.

I would also like to mention my first scientific mentor, Dr. Carrie Wu, for taking in an overconfident teenager and equipped me exceptionally well for my PhD with scientific as well as life skills.

A great thank you to all the past and present members of the Dunham lab for their scientific contributions as well as day to day support. I literally could not have done this without such an amazing group of people. Dr. Eric Hoffer taught me all the necessary scientific knowledge and technical skills I needed in the early years of my PhD. I thank Dr. Dongxue Wang and John Shanks for all their wisdom and for purifying proteins for me. Special thanks to my current lab mates: Ian Pavelich for all the pep talks and biting

humor; Pooja Srinivas for her scientific fearlessness and insistence on everyone being healthy; Dr. Alexandra Kuzmishin Nagy for her extreme niceness in all aspects of life; Jacob Mattingly and Evelyn Kimbrough for keeping my projects alive; Julia Tanquary, Eliseo Salas, and Tiffany Trieu for being great new teammates and continuing the Dunham lab legacy.

Special thanks to the Conn lab being so kind to me and making their lab a space where I could walk in at any time and feel like it is also my lab. In addition, I thank the neighboring labs of Ortlund, Kahn, and Ghalei for their generosity with sharing their time and expertise, maintaining shared instruments, lending reagents, free food, and their general kindness whenever I had a crisis.

I would like to express my sincere gratitude to all my external collaborators: Drs. Kurt Fredrick, Ruben Gonzalez, Kenneth Keiler, Julia Bandow, Ya Ming Hou for sharing their projects with me and sending me materials I needed.

I also thank all the synchrotron beamline staff members at SER-CAT and NE-CAT, who helped me at all times of day and night. For EM, I thank Dr. Puneet Juneja and Dr. Ricardo Guerrero-Ferreira at Emory for freezing my grids, collecting my data, teaching me RELION and trying to prevent me from crashing Glacier, as well as national centers NCCAT and PNCC for collecting massive EM datasets for me.

I am grateful for the assistance of all the administrators and staff who has been instrumental in making sure I had everything in order and actually get my degree.

Finally, and most importantly, I thank my husband Will and my cat Rafi for being my everything.

# TABLE OF CONTENTS

# LIST OF FIGURES

## Chapter 1



## Chapter 2



## Chapter 3

**Chapter 4**

**Chapter 5**

# LIST OF TABLES

## List of Abbreviations

| | |
|---|---|
| μM | micromolar |
| aa-tRNA | aminoacyl-tRNA |
| aaRS | aminoacyl-tRNA synthetase |
| ASL | anticodon stem-loop |
| ATP | adenosine triphosphate |
| $cmo^5U34$ | uridine-5-oxyacetic acid |
| DNA | deoxyribonucleic acid |
| *E. coli* | *Escherichia coli* |
| EF-G | elongation factor G |
| EF-P | elongation factor P |
| EF-Tu | elongation factor thermo unstable |
| fs | frameshift |
| GTP | guanosine-5'-triphosphate |
| LB | Luria Broth |
| $m^1G$ | 1-methylguanosine |
| $m^6A$ | *N*6-methyladenosine |
| MDa | megadalton |
| mM | milimolar |
| $mnm^5s^2$ | 5-methylaminomethyl-2-thiouridine |
| $mnm^5U$ | uridine 5-oxyacetic acid |
| MPD | 2-methyl-2,4-pentanediol |
| mRNA | messenger RNA |

| | |
|---|---|
| NE-CAT | Northeastern Collaborative Access Team |
| nM | nanomolar |
| PDB | protein data bank |
| PEG | polyethylene glycol |
| PTC | peptidyl transferase center |
| RF1 | release factor 1 |
| RF2 | release factor 2 |
| RF3 | release factor 3 |
| RNA | ribonucleic acid |
| rRNA | ribosomal RNA |
| SAM | *S*-adenosyl-*L*-methionine |
| SER-CAT | SouthEast Regional Collaborative Access Team |
| SRL | sarcin-ricin loop |
| $t^6A$ | *N*6-threonylcarbamoyladenosine |
| tRNA | transfer RNA |
| w/v | weight/volume |
| β-Me | β-mercaptoethanol |

# Chapter 1.

# Introduction

## 1.1. Overview of translation

The central dogma defines a sequential transfer of information through biological systems from DNA to RNA to protein such that genetic information encoded in DNA results in fully functional organisms. Translation is the last step of this information flow where the ribosome decodes nucleic acid information in messenger RNA (mRNA) into amino acid sequences to synthesize polypeptide chains. Proteins perform essential roles in the cell and account for 60-80% of the cellular biomass so it is critical that functional proteins are produced in a timely manner. The ribosome is one of the most conserved macromolecular machines in all cells that consists of two asymmetric subunits (the larger 50S subunit the smaller 30S subunit in bacteria) that join together to form the 2.5 megadalton (MDa) 70S ribosome (1, 2) (**Figure 1.1**).

Protein synthesis is a complex and delicately coordinated process involving the massive RNA-protein complex called the ribosome and many translation factors (3, 4) (**Figure 1.2**). Translation has four defined stages: initiation, elongation, termination, and recycling. During this process, the ribosome guides the binding of template mRNA, protein translation factors and transfer tRNAs (tRNAs) to catalyze the synthesis of a polypeptide chain. Initiation controls the frequency with which an mRNA is translated mainly based on the availability of the ribosome binding site and the strength of the special recognition sequence called the Shine-Dalgarno sequence (5). The reading frame is defined here by the formyl initiator $tRNA^{fMet}$ binding the start codon, AUG in the peptidyl (P) site. Elongation is when tRNAs deliver amino acids to be added to the newly synthesized nascent chain in the aminoacyl (A) site. Three nucleotides of the mRNA

codon are read or decoded by corresponding anticodon nucleotides of tRNAs until a stop codon is recognized by release factors and translation is terminated. Factors such as mRNA secondary structure, rare codons, or variations in tRNA levels impede efficient movement and cause pausing. This pausing can be due to the unavailability of a cognate tRNA, or unfavorable interactions of the nascent chain with the exit tunnel serving as regulatory mechanisms to control programmed frameshifting or to slow down the translation speed to allow for proteins to fold while on the ribosome (6-10). Biochemical and genetic studies have provided critical insights into the mechanism of protein synthesis, however, the molecular basis for how the ribosome maintains the three-nucleotide mRNA frame, one of the fundamental aspects of translation, is still under investigation. This dissertation focuses on the mechanisms of frame maintenance during elongation and how mismatches in the codon-anticodon interaction are recognized by the ribosome.

### 1.2. The ribosome undergoes large conformational changes during elongation

Translation is a highly complicated process where a lot of factors need to come together correctly to maintain the speed (~20 amino acids per second or 50 ms per codon (11, 12)) and accuracy (error rate of 1 in $10^3$ to $10^4$ (13)) of protein synthesis (14) (**Figure 1.2**). The ribosome samples a large pool of tRNAs in the cell to select the correct one corresponding to the mRNA codon presented in the A site. The degeneracy of the genetic code is partially explained by how the tRNA binds in the decoding center of the ribosome. The first two positions of the codon-anticodon interaction are heavily monitored by extensive ribosomal contacts with the minor groove of the codon-anticodon duplex (1, 15). While the shape and position of the helix is important for recognition of the ribosome, the identity of the third codon-anticodon interaction (position 34 of the tRNA with the 3$^{rd}$

and last nucleotide of the mRNA triplet codon, **Figure 1.3**) is not as strictly constrained by the ribosome. This is why the first and second nucleotides of the three-nucleotide codon mostly determines the amino acid the codon codes for, while only a few changes in the third nucleotide identity changes the amino acid the codon corresponds to (16, 17). For example, alanine is encoded by four codons with third position degeneracy: GCU, GCC, GCA, and GCG. There are more codons (three nucleotides each with four possible bases result in 64 possible codons) than encoded amino acids (20 essential and 1 stop), so the genetic code is able to tolerate some point mutations since the final protein sequence is not affected.

During elongation, the ribosome is moved along the mRNA to add amino acids delivered by aminoacyl-tRNAs (aa-tRNAs) on to the growing nascent chain in cycles of decoding, peptide bond formation and translocation (18). The aa-tRNA is delivered to the ribosome by elongation factor thermo unstable (EF-Tu) in a ternary complex with GTP and interacts with the ribosomal proteins L7/L12 that form a multioligomeric stalk (19). If the codon-anticodon interaction in the decoding center on the 30S is deemed to be correct, EF-Tu hydrolyzes GTP quickly which releases the aa-tRNA to be accommodated into the A site (20-22). If the interaction is not correct, the ternary complex can quickly dissociate and GTP hydrolysis is much slower. The ribosomal decoding center in the A site monitors the codon-anticodon pairing extensively through interactions between the ribosomal RNA (rRNA) and the tRNA-mRNA base pairing. 16S rRNA nucleotides A1492, A1493, A1913, G530, and C1054 directly contact the base pairs of the codon-anticodon interaction. The ribosome undergoes conformational changes when the tRNA is correct and accepted. 16S rRNA base A1492 flips out of helix 44 (h44), 23S rRNA base A1913

also flips out of Helix 69 (H69), and 16S rRNA base G530 switches from a *syn* to an *anti* conformation to interrogate the interaction (**Figure 1.3**) (21, 23-26). These rRNA changes influence the overall conformation of the 30S region that encompasses the A site called the head domain to close (**Figure 1.3**). Once the tRNA is accommodated on the 50S where peptidyl transfer occurs, the ribosome positions the CCA ends of the tRNA in the peptidyl transferase center on the 50S such that the amino group on the A-site aa-tRNA performs a nucleophilic attack on the ester carbonyl of the P-site tRNA to form a peptide bond. This reaction transfers the nascent chain with the addition of the new amino acid on to the A-site tRNA (27). After the peptide bond is formed, the ribosomal subunits move relative to each other from the nonrotated state (with the tRNAs in the A/A and P/P position (tRNA positions are denoted corresponding to their binding sites on 30S/50S) where the A-site tRNA is bound to the A site and the P-site tRNA is bound to the P site on both small and large subunits) to the so-called hybrid states (where the tRNAs are bound in hybrid P/E and A/P states meaning the anticodon ends of the tRNA and mRNA are bound in the 30S P and A sites but the CCA ends of the tRNAs have moved to the 50S E and P sites) (28-30). The hybrid states are driven by the peptidyl transfer where deacylated tRNAs preferentially bind to the E site, and peptidyl tRNAs favored the A/P state (31, 32). Elongation factor G (EF-G) then translocates the tRNAs and the mRNA forward through the ribosome on the 30S in a series of intermediate states by hydrolyzing GTP (**Figure 1.2**) (30, 33-35). Now the P-site tRNA is fully moved to the E site (E/E) and the A-site tRNA is fully moved to the P site (P/P) leaving an empty A site. The 30S head domain swivels in a counterclockwise rotation to dissociate E-site tRNA and EF-G, and for the ribosome to return to a nonrotated state with a P/P tRNA and an empty A site

ready for the next aa-tRNA. The ribosome undergoes repeated cycles of elongation, reading the mRNA three nucleotides at a time, until it encounters a stop codon in the mRNA.

### 1.3. tRNA selection on the ribosome

Translation is the most energy-intensive process of the cell, and it is the major factor limiting growth of rapidly dividing bacterial cells (36, 37). In particular, about 60% of a cell's energy (ATP) is spent on ribosome biogenesis and maintenance alone (38). It is then not surprising that translation fidelity is carefully regulated to ensure cellular resources are not wasted on ribosomes that are mistranslating to produce nonsense peptides or are stalled and hence not actively translating proteins. In general, the ribosome maintains high levels of accuracy and is able to discriminates against most mismatches during translation (39).

Biochemical characterization of translation rates and structural information indicates the ribosome employs several kinetic discrimination checkpoints to efficiently select the correct tRNA from the pool of chemically and structurally similar cellular tRNAs (3, 15, 20, 40-42). Generally, the ribosome samples the tRNA and irreversibly commits to the correct tRNA to ensure the processivity of the reaction at two phases: initial selection and proofreading. As discussed previously, the ribosome samples the tRNAs to select for the correct mRNA codon-tRNA anticodon pairing. This is the first kinetic checkpoint where the ribosome discriminates between correct and incorrect tRNAs, where the binding of cognate tRNA binding to the A site is stabilized and incorrect tRNAs rapidly dissociates (140 $s^{-1}$ as compared to 0.2 $s^{-1}$ for cognate ternary complexes) from the ribosome (20, 41). This stabilization is not only due to the hydrogen bonds formed from the correct Watson-Crick base pairing but is also due to the structural recognition by the ribosome.

This is the first step of commitment the ribosome makes to accept a tRNA as correct based on the complementarity of the codon-anticodon to ensure the rapid and more frequent movement of the correct tRNA towards the fully accommodated state. There also appears to be a tradeoff between the accuracy of cognate codon reading and efficiency of selecting the correct tRNA during initial selection such that there is a variation in the fidelity of this kinetic discrimination step across the possible combinations of codon-anticodon pairings (43, 44).

In addition to the kinetic discrimination steps, the ribosome also employs an induced-fit mechanism to drive forward the acceptance of the correct tRNA. The head domain closure of the ribosomal 30S small subunit around the correct A-site codon-anticodon interaction is driven by the conserved monitoring rRNA nucleotides (A1492, A1493, and G530) (45). This 30S head movement causes the ternary complex (aa-tRNA•EF-Tu•GTP) and specific regions within EF-Tu to move: the tRNA is bent in the A/T state (the anticodon is bound to the codon in the 30S A site but the CCA-end is bent towards the outside of the ribosome in the ternary (T) state) and EF-Tu His84 residue in the Switch II region and the Asp21 residue in the P loop contacts the sarcin-ricin loop (SRL) of the 23S rRNA in the large subunit (21, 22, 46). This contact activates the GTPase activity of EF-Tu leading to the irreversible GTP hydrolysis by EF-Tu to release the aa-tRNA to the ribosome (20). This movement also changes the tRNA conformation from the A/T state and favors the acceptance of the tRNA fully into the A site (A/A). This movement accelerating the acceptance of the correct tRNA correlates with kinetic studies showing that the GTPase activation is rapid for cognate aa-tRNAs (190 $s^{-1}$) but is very slow for near-cognate aa-tRNAs (0.6 $s^{-1}$) (3, 47, 48).

After the codon recognition during initial selection, the tRNA undergoes further discrimination during proofreading where the incorrect tRNA substrates are discarded (26). After EF-Tu is activated, it hydrolyses GTP and rearranges into the GDP-bound conformation (22, 49). This allows the 3' CCA end of the A-site tRNA carrying the amino acid to be positioned in the peptidyl transferase center (PTC) of the ribosome ready for peptide bond formation. Kinetic studies showed this accommodation step is slower for complexes containing mismatches than that of the cognate complex, presenting the opportunity for the ribosome to reject incorrect tRNAs even after the irreversible GTP hydrolysis (42, 50). The rate of aa-tRNA accommodation is 22 $s^{-1}$ for cognate aa-tRNAs in contrast to 0.06 $s^{-1}$ for near-cognate aa-tRNAs. Subsequently, near-cognate aa-tRNA is released from the ribosome post-GTP hydrolysis is much faster (0.84 $s^{-1}$) than cognate aa-tRNAs (0.1 $s^{-1}$) (20, 41, 42, 51, 52).

### 1.4. tRNA elements contribute to accurate tRNA selection

tRNAs are essential 'adaptor' molecules that transports the amino acid corresponding to the mRNA triplet codon to be incorporated into the polypeptide chain. tRNAs are essential in all three kingdoms of life and also have roles beyond translation (53, 54). In *Escherichia coli* (*E. coli),* there are 2 initiator and 43 elongator tRNAs to decode for every amino acid (55). While tRNAs have great diversity in nucleotide sequence, they have highly conserved cloverleaf secondary structure that consist of ~70-90 nucleotides forming 5 helical stems with some ending in unpaired bases to form loops (**Figure 1.4A**). The acceptor stem is where the aminoacyl moiety is attached to the 3' CCA end and there is a 5'-terminal phosphate group. The TΨC arm contains a highly conserved pseudouridine modification and forms important tertiary interactions with the D-loop to form the 3D structure of tRNAs. The variable loop is typically 4-5 nucleotides in length

with a few exceptions where they can be 10 nucleotides or more. The anticodon stem is consistently made of 5 base pairs ending in 7-nucleotide anticodon loop containing the 3 anticodon nucleotides (54, 56). tRNAs adopt very similar L-shaped 3D structures with the TΨC and D arms coming together to form the bend (**Figure 1.4B**). The identity of the tRNA is determined by the anticodon sequence (the three bases at positions 34, 35, 36) but other tRNA elements are also critical for recognition by aminoacyl synthetases, EF-Tu and also the ribosome (40, 46, 57, 58). In addition to the primary nucleotide sequence, tRNAs are also heavily modified post-transcriptionally with over 100 chemical moieties identified to date (**Figure 1.5**). These modifications have many additional regulatory functions such as changing the levels of aminoacylation or maintaining the translational accuracy and so disruptions in their biosynthetic pathways leading to deficiencies in modification levels often have severe consequences for the cell (59-66).

The anticodon loop, specifically nucleotides 34 and 37, (**Figure 1.5**) is very highly modified with a wide variety of chemical modifications (**Figure 1.6**). Modifications at these two positions are so critical that 54 of the 61 sense codons in *E. coli* are read by tRNAs that have modifications at positions 34 and 37 (67-69). Nucleotide 34 is a part of the anticodon and modifications at position 34 often expand the base pairing capability of the anticodon at the third codon-anticodon pairing known as the 'wobble' base pairing position (70). The plasticity in the codon-anticodon interaction is important to expand the degeneracy of the genetic code. Most common wobble base pairs are the non-Watson-Crick base pairs G•U, I•U, I•C, and I•A (I = inosine, a modified adenosine). Structurally, these base pairs are able to adopt Watson-Crick-like conformations in the decoding center of the ribosome (17, 71-78). Since nucleotide 34 directly pairs with the 3$^{rd}$ mRNA

codon nucleotide, modifications at position 34 directly impact decoding. For example, $tRNA^{Lys}_{UUU}$ requires both the uridine 5-oxyacetic acid ($mnm^5U$) modification at position 34 and the *N*6-threonylcarbamoyladenosine ($t^6A$) at position 37 to bind to its lysine codons (AAA and AAG) in the ribosomal A site. These modifications also allow for efficient translocation to the P site but the reasons for this are unclear (79, 80).

Nucleotide 37 is 3′ adjacent to the anticodon and is also highly modified (**Figure 1.4**). In all sequenced bacterial tRNAs, position 37 is modified ~75% of the time (59). Out of the modifications, t6A, *N*6-methyladenosine ($m^6A$), and 1-methylguanosine ($m^1G$) are frequently implicated with important biological functions (**Figure 1.5**). tRNAs lacking modifications at position 37 lose their ability to accurately decode the mRNA triplet codon and this can have detrimental effects on the cell (81-86). There has been a special interest on the impact of the 37 modification on the $tRNA^{Pro}$ decoding and frame maintenance since it was discovered over three decades ago (87, 88). Proline codons (CCN where N can be any nucleotide, so 4 codons encode for proline) are especially frameshift-prone. They cause the ribosome to lose the three-nucleotide mRNA reading frame during the slow rate of peptide bond formation with a proline (89). In fact, translating consecutive proline codons is so challenging for the ribosome that it induces ribosomal stalling and disrupts translation (90, 91). Bacteria contain a specialized translation factor called elongation factor P (EF-P) specifically to facilitate the successful movement of ribosomes through the stretches of polyprolines (81, 90, 92-94). In *E. coli* and other bacteria, all $tRNA^{Pro}$ isoacceptors are naturally modified with $m^1G37$. It is now known that the $m^1G37$ modification in $tRNA^{Pro}$ is critical for frame maintenance and the lack of this modification impacts binding, translocation, tRNA stability, and P-site gripping causing ribosome

stalling (81, 86, 88, 95-99). In bacteria, the *S*-adenosyl methionine dependent methyltransferase TrmD methylates the N1 position of the G37 base in tRNA and is essential for growth. Other tRNAs also contain $m^1G37$ modifications such as $tRNA^{Leu}_{GAG}$, $tRNA^{Leu}_{CAG}$, and $tRNA^{Arg}_{CCG}$, and so a decrease in $m^1G37$ has global effects on the cell. Recently, it was discovered that modification deficient $tRNA^{Pro}$ was not aminoacylated efficiently and this might be the cause for TrmD deletion being lethal (85). Recently, it was found that a deficiency in the $m^1G37$ modification affects the synthesis of membrane proteins. The decrease in $m^1G37$ levels in *E. coli* and *Salmonella* disrupted membrane integrity and sensitized these bacteria to antibiotics providing a link between the $m^1G37$ tRNA modification and bacterial antibiotic resistance (100, 101). The $m^1G37$ is also conserved in all three domains of life, with Trm5 being the eukaryotic and archaeal counterpart for TrmD (102).

Suppressor tRNAs were originally identified as mutants of tRNAs that can alter the decoding the three-nucleotide genetic code (23, 103-109). Most of these suppressor tRNAs have a nucleotide insertion and have noncanonical translational properties to decode two or four nucleotide codons or incorporate amino acids at stop codons. Suppressor tRNAs were originally discovered in genetic screens and were found to compensate for the insertions or deletions in the genetic code where they 'suppressed' the mutation in the mRNA and reverted the reading frame (103, 104). Stop codon suppressors are especially interesting because of their ability to readthrough a stop codon and can be exploited to incorporate an unnatural amino acid to expand the genetic code and enabling novel protein chemistries (110-114). The first sequenced suppressor tRNA was $tRNA^{SufD}$ which is a $tRNA^{Gly}$ variant with a cytosine insertion immediately 5' of the

anticodon between positions 33 and 34 of the tRNA (115). $tRNA^{SufD}$ could decode a four-nucleotide GGG-G codon. Another well-known suppressor is the Hirsh suppressor, a $tRNA^{Trp}$ variant with a single guanosine to adenosine mutation at position 24 of the tRNA distant from the anticodon, that can read the near-cognate UGA stop codon in addition to its cognate UGG codon (116). Structure of the Hirsh suppressor bound to the tRNA discovered an additional internal hydrogen bond that forms with the G24A mutation stabilizing the bent form of $tRNA^{Trp}$ allowing it to be more readily accepted by the ribosome even when the tRNA-mRNA interaction contains a mismatch (117).

### 1.5. Modifications and insertions in $tRNA^{Pro}$ affect decoding and frame maintenance

Proline codons present a special challenge to the ribosome during protein synthesis causing mistranslation. Across the three domains of life, there is only a single conserved stretch of three prolines in ValS, the aminoacyl-tRNA synthetase for $tRNA^{Val}$ (118, 119). Reporter assays investigating the propensity of proline codons to frameshift showed that the effect is most pronounced when the proline codon is in the second position of the reading frame, right after the start codon (81).

In *E. coli,* there are 3 $tRNA^{Pro}$ isoacceptors decoding proline codons: $tRNA^{Pro}_{CGG}$, $tRNA^{Pro}_{GGG}$ and $tRNA^{Pro}_{UGG}$, also known by their corresponding gene names *proK*, *proL* and *proM* respectively (**Figure 1.7**). Imbalances in the concentrations of the tRNA isoacceptors in the cell promote frameshifting (120), and p*roM* is naturally frameshift-prone (81, 121). In bacterial, all three isoacceptors contain the $m^1G37$ modification and decrease in this modification using a TrmD knockdown strain (TrmD knockout is lethal) causes an increase in +1 frameshifting and ribosome stalling (81, 86, 122). Among the three isoacceptors, $tRNA^{Pro}_{CGG}$ (*proK*) is the major isoacceptor decoding the CCG mRNA

codon. Previous work by the Dunham laboratory showed that the $m^1G37$ modification in $tRNA^{Pro}_{CGG}$ stabilizes its interaction with the cognate codon and maintains the mRNA frame (97).

There are two known frameshift (fs) suppressors of $tRNA^{Pro}$, *sufA6* and *sufB2*, that were identified during genetic suppression assays in *Salmonella* (107, 122). Both of these suppressors contain insertions of a single guanosine nucleotide at the same position between nucleotides 37 and 38 (called G37.5). $tRNA^{SufA6}$ is a variant of $tRNA^{Pro}_{CGG}$ and $tRNA^{SufB2}$ is a variant of $tRNA^{Pro}_{GGG}$ (**Figure 1.7**). Both of the fs tRNAs also contain the $m^1G37$ modification (123) and they decoded four-nucleotide codons causing a +1 shift in the reading frame of the mRNA.

Chapter 2 of this dissertation investigates how the $m^1G37$ modification as well as the G37.5 insertion in $tRNA^{Pro}_{CGG}$ alters the binding of the tRNA to the ribosomal A site to distinguish between correct and incorrect tRNAs. When $tRNA^{Pro}_{CGG}$ lacks $m^1G37$, it binds to both cognate and slippery near-cognate codons with similar affinity instead of binding tightly to the cognate codon only. The G37.5 insertion in $tRNA^{SufA6}$ also had this analogous effect of stabilizing the tRNA binding on the frameshift-prone near-cognate codon. This loss of recognition in binding affinities indicates the ribosome is unable to distinguish between the correct and incorrect tRNAs both when the $m^1G37$ modification is deficient and when there is an insertion in the anticodon stem loop of $tRNA^{Pro}_{CGG}$. Previous work in the Dunham lab showed that the G37.5 insertion destabilized the anticodon loop through disruption of the evolutionarily important 32-38 base pair. We then found that changing the G37.5 insertion into an A can restore the stabilization of the tRNA using both biochemical assays and X-ray crystallography (95). This work contributes an

additional layer of regulation to the known mechanisms of the tRNA-mediated frameshifting.

### 1.6. The 32-38 base pair in tRNAs is an important recognition element

Since the ribosome relies heavily on kinetic proofreading to quickly accept correct tRNAs and reject incorrect tRNAs, tRNAs have evolved to have remarkably similar binding affinities to the A and P sites on the ribosome (124). There are 45 tRNAs decoding the 61 sense codons in *E. coli* with a large diversity in sequence and modifications (68). tRNAs undergo distinct pathways of recognition, aminoacylation, and editing by their cognate aaRSs. After that, aa-tRNAs undergo the same reaction pathways during translation and are treated as analogous substrates by EF-Tu and the ribosome (124, 125). Despite their similar 3D structures (**Figure 1.4**), the ribosome has to quickly sample the tRNAs and determine if it is correct or not. tRNAs have evolved features to be able to compensate for the incredibly diverse chemical properties of the esterified amino acids (e.g., glycine vs histidine) and also the inherently different strengths of the codon-anticodon interactions (e.g., three C-G would be stronger than three A-U base pairs) (126). The 32-38 pairing is the last base pair in the ASL of the tRNA right before the anticodon stem (**Figure 1.4**) is one of these evolutionarily features that have evolved to compensate for the strength of the codon-anticodon interaction. Examination of the bacterial tRNA sequences indicate that the rest of the tRNA has evolved to correspond with the anticodon nucleotides themselves such that even the isoacceptors decoding the same amino acid have different consensus sequences (127, 128).

Extensive work by the Uhlenbeck laboratory has shown that the identity of the 32-28 base pair ensures correct codon recognition (124, 127, 129, 130). Changing the 32 and 38 nucleotides in stop codon suppressor tRNAs can modulate their efficiency (131, 132)

while overexpressing a $tRNA^{Gly}$ U32C mutant increased levels of frameshifting (133) indicating a direct role of these two nucleotides in tRNA function. Previous research in the Dunham lab and my work has also shown that the G37.5 insertion in $tRNA^{SufA6}$ disrupts the 32-38 pairing and changing the insertion into an A37.5 restored the U32-A37.5 interaction resulting in stabilized binding and tRNA structure (95, 97). In particular, the 32-28 pairing in $tRNA^{Ala}_{GGC}$ is shown to be critical for accurate decoding, and disruptions to this base pair resulted in near-cognate codons being translated as efficiently as cognate codons (130). This resulted in misincorporations in a reconstituted translation systems and is toxic to the cell in an overexpression system (134). Changing the A32-U38 pairing in $tRNA^{Ala}_{GGC}$ to C•A, U•A, or U•U (• denotes a non-Watson-Crick base pair) was so detrimental that overexpressing these tRNA variants overwhelmed canonical translation and caused cell death (134).

Chapter 3 of this dissertation provides a structural mechanism with which the ribosome uses the 32-38 pairing to recognize a correct vs. incorrect tRNA-mRNA interaction. Examining structures of $tRNA^{Ala}_{GGC}$ bound to cognate and near-cognate codons as compared to a $tRNA^{Ala}_{GGC}$ variant with the 32-38 nucleotides reversed bound to cognate and near-cognate codons, we found two mechanisms with which the 32-38 pair contributes to correct tRNA selection on the ribosome. When wild-type $tRNA^{Ala}_{GGC}$ binds to a near-cognate codon containing a G•A mismatch at the third position, the 32-38 pairing is disordered and a ribosomal RNA nucleotide (23S A1913) adopts an alternate position. These recognition signals are lost when the 32-38 identities are reversed, so the cognate and near-cognate codon-anticodon interactions are treated similarly in a loss of fidelity (135).

### 1.7. Post peptidyl transfer quality control

Previous chapters have focused on translation fidelity maintained at the decoding level when the tRNA first arrives to the ribosome, but there are post peptidyl transfer quality control mechanisms beyond the A site. It was initially noted that incorrect P-site tRNAs induced +1 frameshifting in the A site in *Saccharomyces cerevisiae* (136, 137). The effects were more pronounced for certain codon-anticodon mismatches than others with frameshift efficiency increased as much as 30-fold when all 64 possible codons were tested or when certain frameshift-prone tRNAs were overexpressed. More recent biochemical data showed that mismatches between the codon and the anticodon in the P site increased the rate of peptide release on sense codons and decreased the fidelity of tRNA selection in the A site (13, 138). Once an amino acid is incorporate into the nascent chain, there is no editing mechanism with which an error can be corrected unlike the arsenal of proofreading and repair procedures in DNA (139).

The mismatches in the P site indicate a possible erroneous amino acid (some mismatches do not impact amino acid identity) has already been added to the growing polypeptide chain. This mismatch, in a mechanism still unknown, signals to the A site to reduce fidelity. This reduced fidelity can result in another incorrect tRNA being accepted by the ribosome or can recruit release factors that normally recognize stop codons to bind to sense codons in the A site and trigger premature termination and peptide release (termination stage in **Figure 1.2**). This post peptidyl transfer quality control (post PT QC) mechanism allows the incorrect peptide chain to be hydrolyzed and degraded before a potentially toxic protein is produced. Also this mechanism allows the translation machinery to be recycled before even more energy is expended on an incorrect translation product (140).

Chapter 4 of this dissertation proposes a mechanism with which mismatches in the P site cause mispositioning of the mRNA in the A site disrupting the A-site tRNA-mRNA interactions. This disruption could decrease the fidelity at the A site and serve as a signal to trigger the quality control mechanisms resulting in premature termination.

## 1.8. Figures

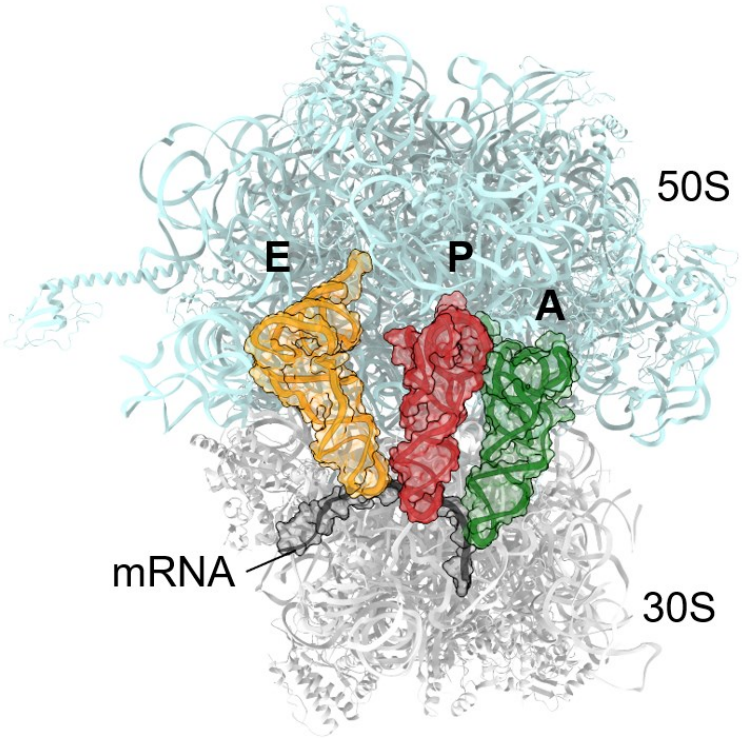


**Figure 1.1. Overview of the 70S ribosome.**

The bacterial ribosome has two subunits: the larger 50S (cyan) and the smaller 30S (gray). The tRNAs and the mRNA are overlaid on top of the 70S indicating their binding sites. tRNAs span both the subunits in three positions: aminoacyl (A, green), peptidyl (P, red), exit (E, orange). The A site is where aminoacyl-tRNAs are first delivered to the ribosome and also where release factors bind during termination. The tRNA then moves to the P site once peptide bond formation occurs, and finally it exits the ribosome through the E site.

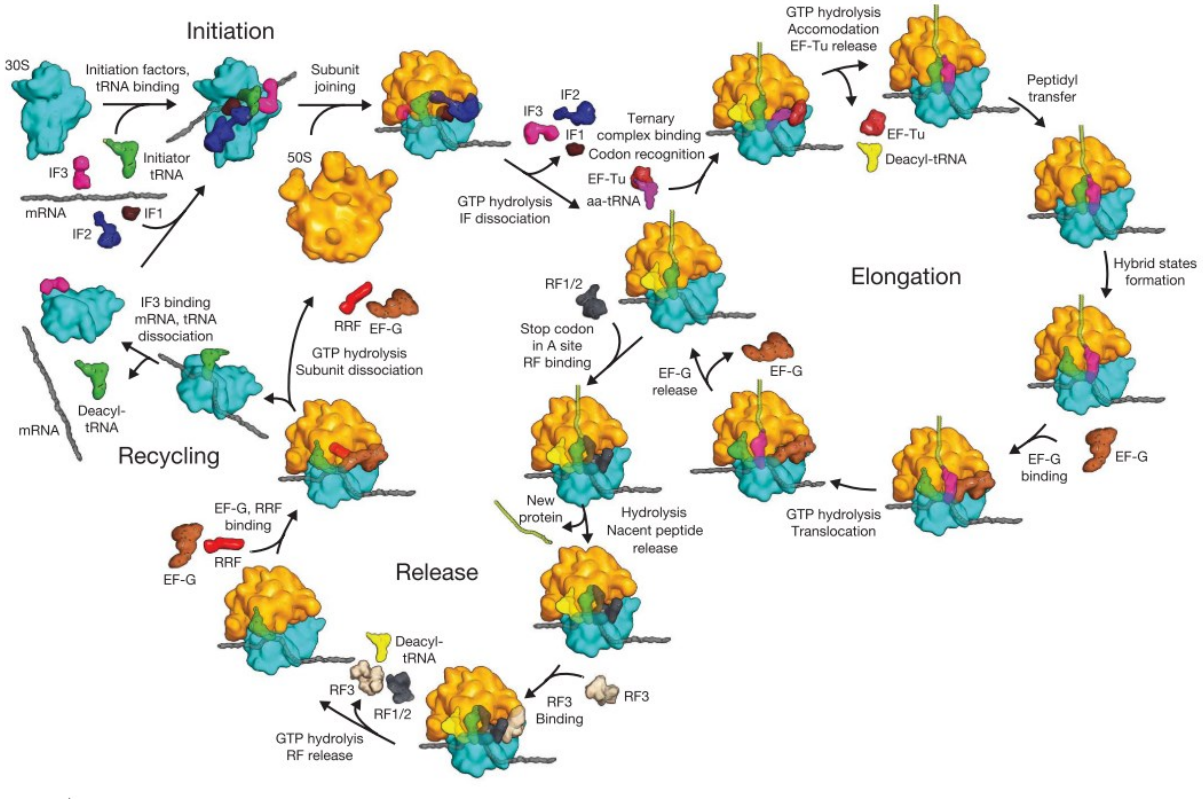


**Figure 1.2. Overview of bacterial translation.**

Translation is a highly dynamic process involving many translation factors consisting of four main steps: initiation, elongation, termination, and recycling. During initiation, the 30S small subunit forms a pre-initiation complex with the initiator $tRNA^{fMet}$ and the mRNA start codon is positioned. The 50S subunit then joins to form the 70S initiation complex to be ready for elongation. The ribosome then iteratively moves through the mRNA during elongation with elongator aa-tRNAs delivering aminoacyl groups in cycles of decoding, peptide bond formation and translocation until the ribosome encounters a stop codon on the mRNA. Translation then terminates to release the synthesized peptide. The ribosome then undergoes recycling where the subunits are split to allow them to participate in the next round of translation. Figure from reference (2).

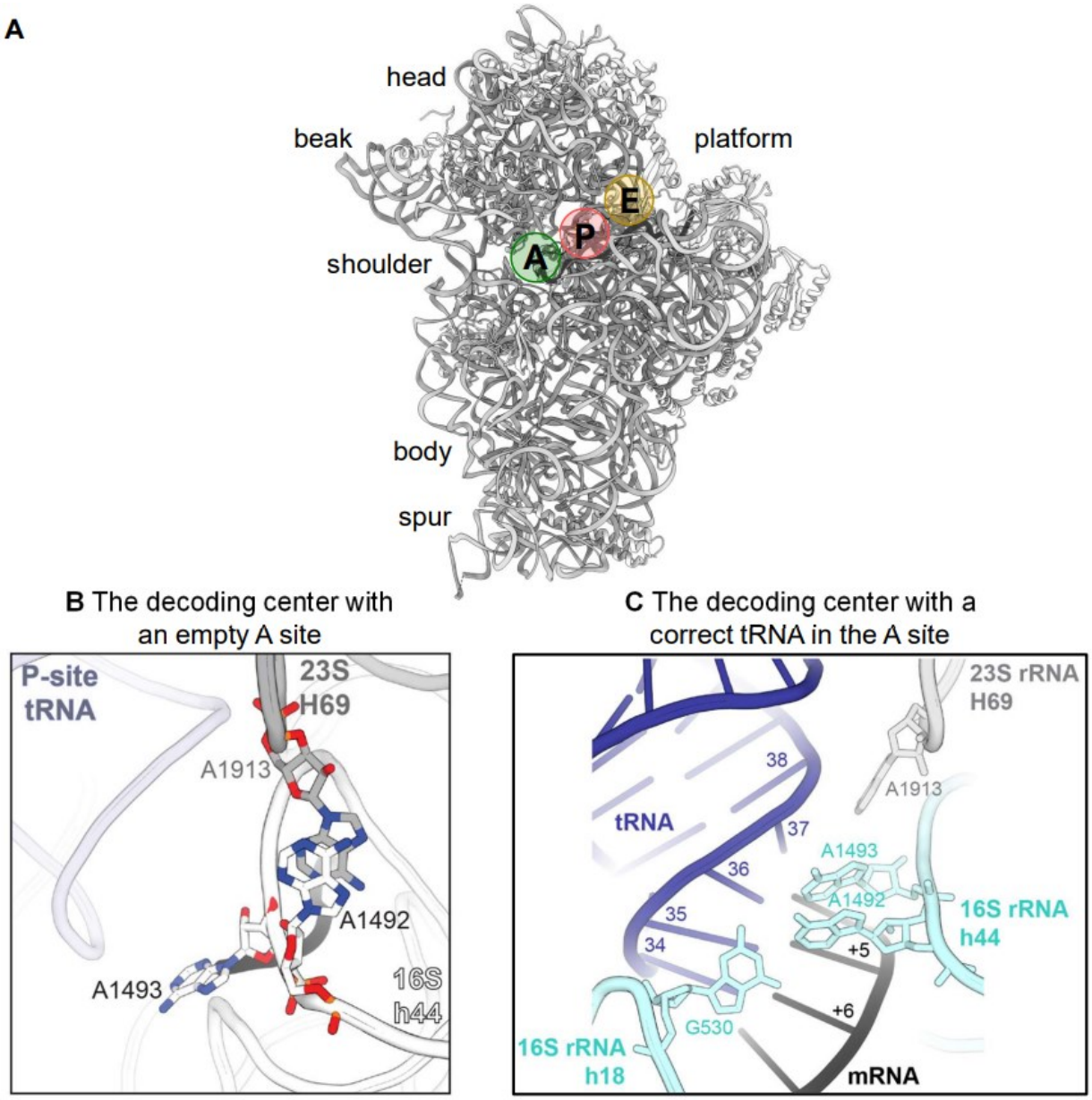


**Figure 1.3. The decoding center undergoes conformational changes in response to a correct tRNA binding to the A site.**

**A**. An overview of the 30S showing the locations of the head, beak, shoulder, body, spur, platform and the tRNA binding sites A (green), P (red), and E (yellow) on the small subunit relative to the mRNA (in black). **B**. When there is no tRNA bound in the A site, 16S rRNA nucleotide A1492 (white) is 'flipped-in' in h44 to base stack with 23S rRNA

nucleotide A1913 in H69 (gray). The P-site tRNA is shown in light purple with the mRNA in black (it does not extend into the A site) **C**. When a correct tRNA is bound to the A site (blue), H69 base A1913 moves towards the tRNA to form a hydrogen bond with the 2' OH of position 37 of the tRNA and h44 bases A1493 and A1492 directly interacts with the codon-anticodon base pairs. A1492 base pairs with h18 base G530 to form a gate stabilizing the mRNA-tRNA codon-anticodon helix in the decoding center. The mRNA numbering refers to the first nucleotide position in the P site as +1, with A-site nucleotides denoted as +4, +5, and +6.

Figure adapted from reference (135).

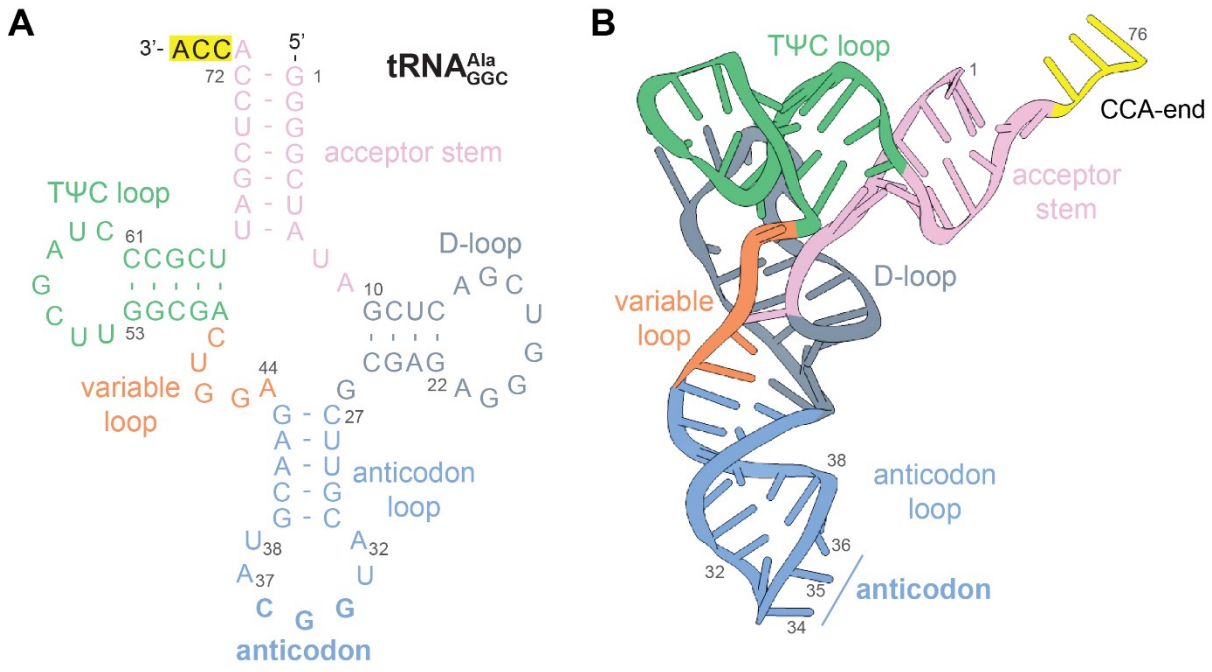


**Figure 1.4. The structure of a transfer RNA.**

**A**. tRNAs adopt a highly conserved cloverleaf secondary structure with distinct domains: acceptor stem, TΨC loop, D-loop, variable loop, and anticodon loop. The 3' CCA end is where the aminoacyl group is attached to the terminal A76 nucleotide. The anticodon (positions 34, 35, 36) is the primary determinant for the identity of the tRNA. **B**. The tRNA tertiary structure adopts a highly conserved L shape in solution, with the domains color coded corresponding to the secondary structure.

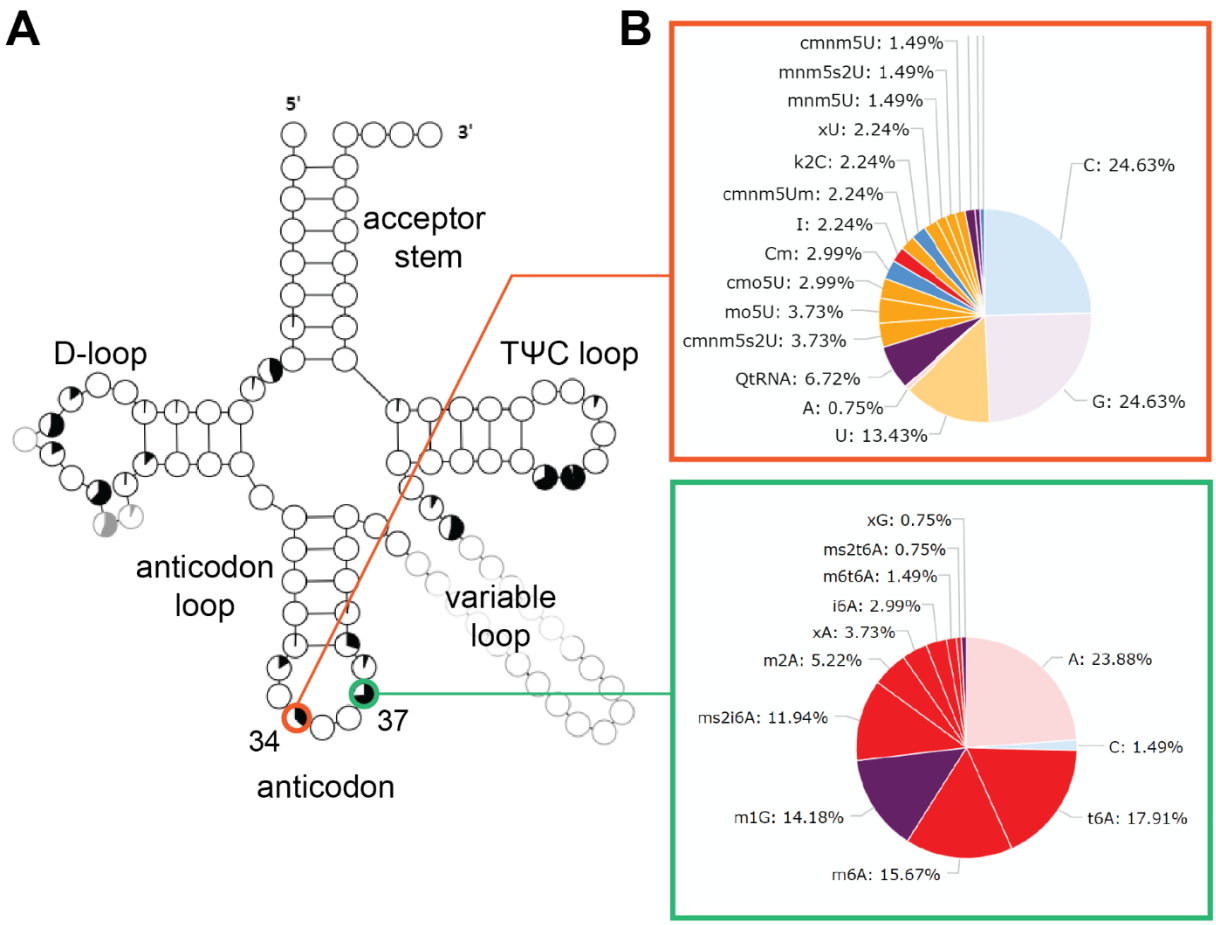


**Figure 1.5. Positions 34 and 37 in the anticodon loop are highly modified.**

**A**. The distribution and frequency of modified nucleosides at all positions of the tRNA from all known bacterial tRNA sequences (134 total sequences). The black wedge in the circle represents the proportion of modifications found at that position. Positions 34 (orange) in the anticodon and 37 (green) adjacent to the anticodon are highly modified in bacteria. **B.** Detailed breakdown of modifications found at positions 34 (orange) and 37 (green) showing the frequency of unmodified nucleosides and different types of modified nucleosides. Figures generated using the tRNAmodviz server to plot the distribution of modifications using bacterial tRNA sequences (http://genesilico.pl/trnamodviz/) (59).

The abbreviations for each RNA modification are as follows: cmnm5U, 5-carboxymethylaminomethyluridine; mnm5s2U, 5-methylaminomethyl-2-thiouridine; mnm5U, 5-methylaminomethyluridine; xU, unknown modified uridine; k2C, 2-lysidine; cmnm5Um, 5-carboxymethylaminomethyl-2′-O-methyluridine; I, inosine; Cm, 2′-O-methylcytidine; cmo5U, uridine 5-oxyacetic acid; mo5U, 5-methoxyuridine; cmnm5s2U, 5-carboxymethylaminomethyl-2-thiouridine; QtRNA, queuosine; xG, unknown modified guanosine; ms2t6A, 2-methylthio-N6-threonylcarbamoyladenosine; m6t6A, N6-methyl-N6-threonylcarbamoyladenosine; i6A, N6-isopentenyladenosine; xA, unknown modified adenosine; m2A, 2-methyladenosine; ms2i6A, 2-methylthio-N6-isopentenyladenosine; m1G, 1-methylguanosine; m6A, N6-methyladenosine; t6A, N6-threonylcarbamoyladenosine.

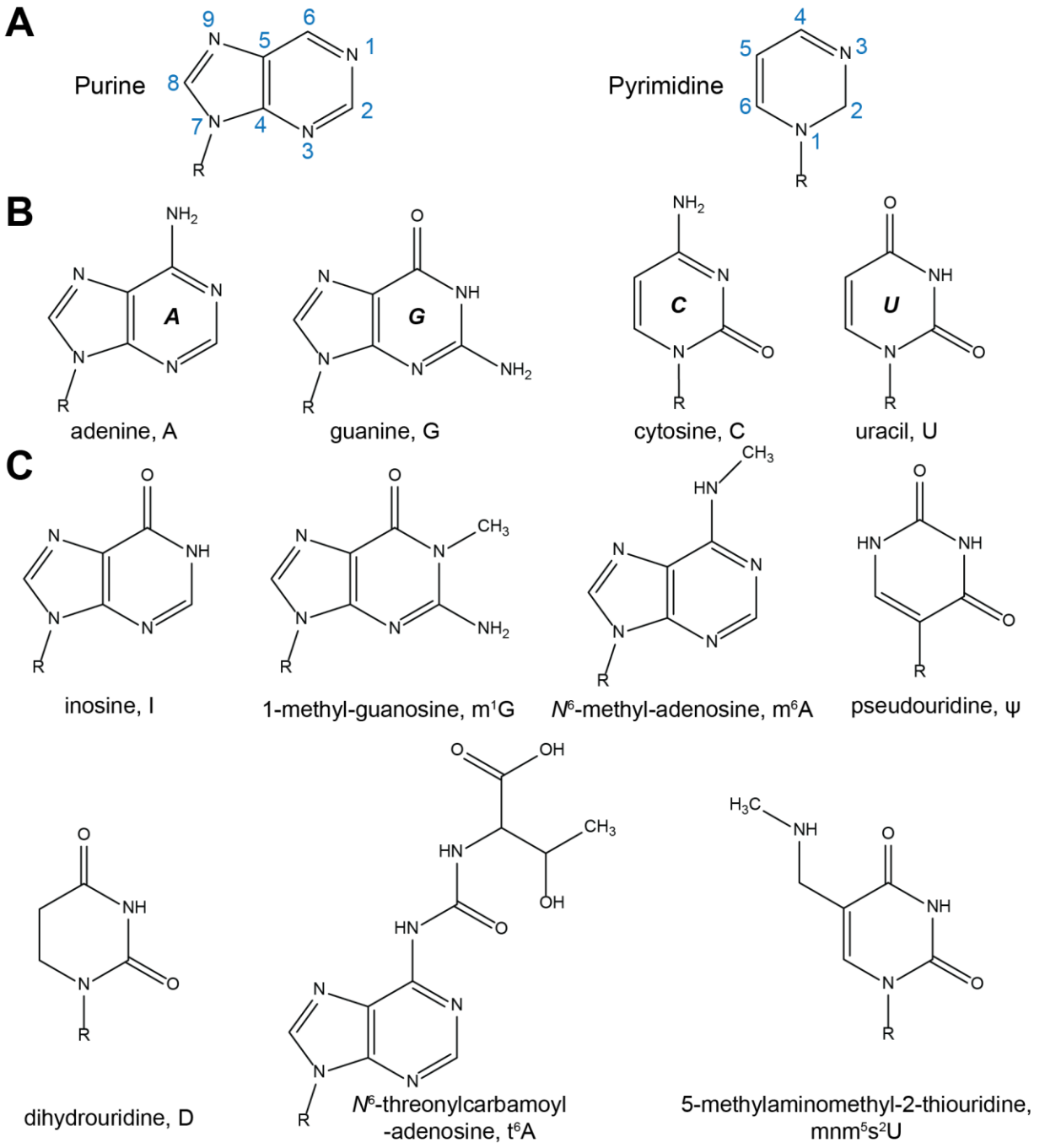


**Figure 1.6. Chemical structures of tRNA bases.**

**A**. General structures of the two general classes of nucleic acid bases with their numbering conventions. Purines are larger with two heterocyclic rings, while pyrimidines only have one 6-membered ring. **B.** The four major nitrogenous bases of RNA are

adenine, guanine, cytosine, and uracil. **C.** There is a wide variety of modified RNA bases. R represents the connection with the ribose.

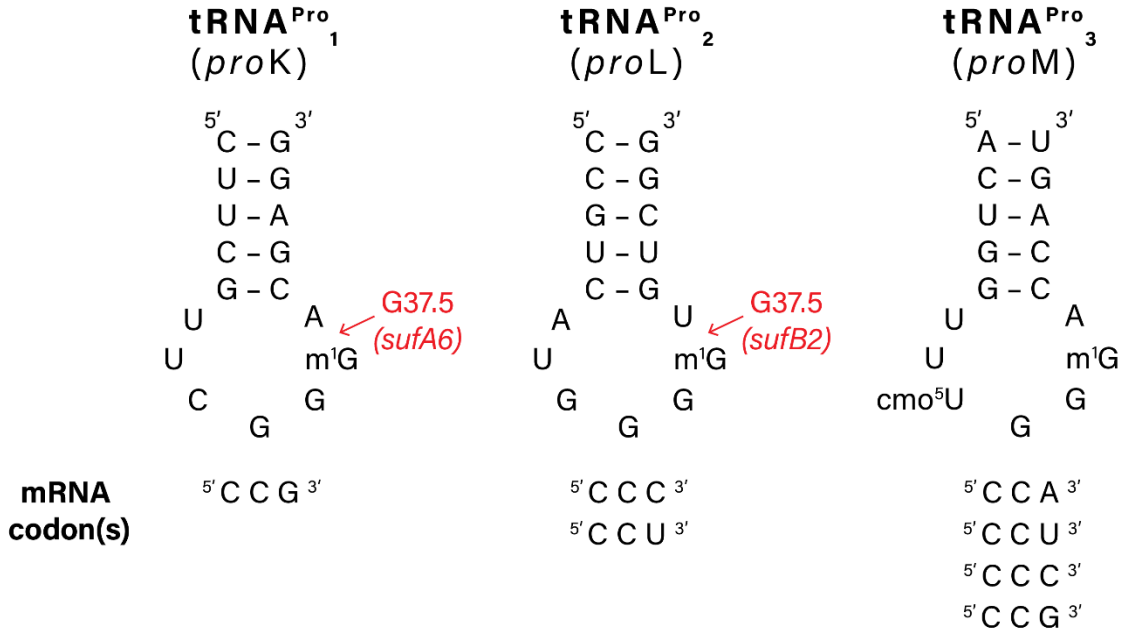


**Figure 1.7. The anticodon stem loops of the three isoacceptors of proline.**

The proline codons (CCA, CCU, CCC, CCG) are decoded by 3 $tRNA^{Pro}$ isoacceptors: $tRNA^{Pro}_{CGG}$ (*proK*), $tRNA^{Pro}_{GGG}$ (*proL*) and $tRNA^{Pro}_{UGG}$ (*proM*). The secondary sequence of their anticodon stem loops is shown with the mRNA codons that they decode below. In red are the nucleotide insertions in the frameshift suppressor variants of *proK* and *proL* called *sufA6* and *sufB2* respectively. Modifications $m^1G$ is 1-methylguanosine, and $cmo^5U$ is uridine 5-oxyacetic acid.

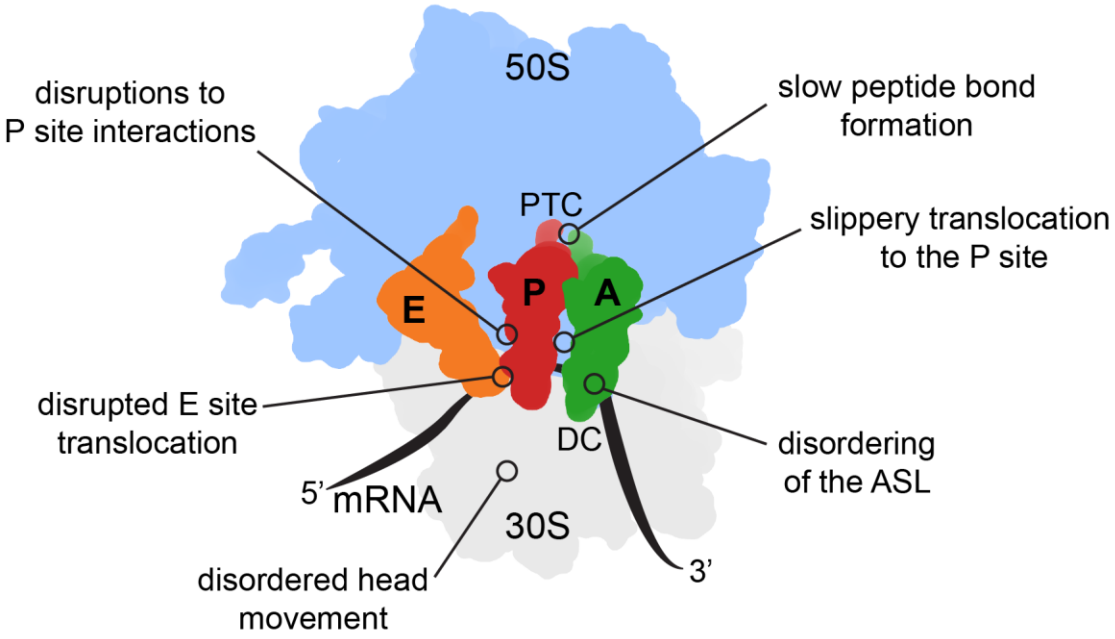


**Figure 1.8. Mechanisms of tRNA-mediated regulation of translation fidelity.**

Overview of the known mechanisms of frameshifting mediated by tRNAs. The large subunit 50S is in blue, small subunit 30S in gray, the E, P, A tRNAs are shown in orange, red, and green respectively. PTC = peptidyl transferase center, DC = decoding center.

## 1.9. References


1. M. Selmer *et al.*, Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942 (2006).
2. T. M. Schmeing, V. Ramakrishnan, What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**, 1234-1242 (2009).
3. M. V. Rodnina, Translation in Prokaryotes. *Cold Spring Harb Perspect Biol* 10.1101/cshperspect.a032664 (2018).
4. J. M. Ogle, A. P. Carter, V. Ramakrishnan, Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem Sci* **28**, 259-266 (2003).
5. C. O. Gualerzi, C. L. Pon, Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cellular and Molecular Life Sciences* **72**, 4341-4367 (2015).
6. H. Girstmair *et al.*, Depletion of cognate charged transfer RNA causes translational frameshifting within the expanded CAG stretch in huntingtin. *Cell Rep* **3**, 148-159 (2013).
7. J. E. Wilusz, Controlling translation via modulation of tRNA levels. *Wiley Interdiscip Rev RNA* **6**, 453-470 (2015).
8. N. Korniy, E. Samatova, M. M. Anokhina, F. Peske, M. V. Rodnina, Mechanisms and biomedical implications of -1 programmed ribosome frameshifting on viral and bacterial mRNAs. *FEBS Lett* 10.1002/1873-3468.13478 (2019).
9. N. inCaliskan *et al.*, Conditional Switch between Frameshifting Regimes upon Translation of dnaX mRNA. *Mol Cell* **66**, 558-567 e554 (2017).
10. V. Chandrasekaran *et al.*, Mechanism of ribosome stalling during translation of a poly(A) tail. *Nat Struct Mol Biol* **26**, 1132-1140 (2019).
11. M. Zhu, X. Dai, Y.-P. Wang, Real time determination of bacterialin vivoribosome translation elongation speed based on LacZα complementation system. *Nucleic Acids Research* 10.1093/nar/gkw698, gkw698 (2016).
12. I. L. Volkov, M. Johansson, Single-molecule tracking approaches to protein synthesis kinetics in living cells. *Biochemistry* 10.1021/acs.biochem.8b00917 (2018).
13. H. S. Zaher, R. Green, Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161-166 (2009).
14. H. S. Zaher, R. Green, Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**, 746-762 (2009).
15. A. B. Loveland, G. Demo, N. Grigorieff, A. A. Korostelev, Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546**, 113-117 (2017).

16. T. P. Hoernes *et al.*, Translation of non-standard codon nucleotides reveals minimal requirements for codon-anticodon interactions. *Nat Commun* **9**, 4865 (2018).

17. A. Rozov, E. Westhof, M. Yusupov, G. Yusupova, The ribosome prohibits the G*U wobble geometry at the first position of the codon-anticodon helix. *Nucleic Acids Res* **44**, 6434-6441 (2016).

18. H. F. Noller, L. Lancaster, J. Zhou, S. Mohan, The ribosome moves: RNA mechanics and translocation. *Nat Struct Mol Biol* **24**, 1021-1027 (2017).

19. U. Kothe, H.-J. Wieden, D. Mohr, M. V. Rodnina, Interaction of Helix D of Elongation Factor Tu with Helices 4 and 5 of Protein L7/12 on the Ribosome. *Journal of Molecular Biology* **336**, 1011-1021 (2004).

20. K. B. Gromadski, M. V. Rodnina, Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* **13**, 191-200 (2004).

21. T. M. Schmeing *et al.*, The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688-694 (2009).

22. J. C. Schuette *et al.*, GTPase activation of elongation factor EF-Tu by the ribosome during decoding. *EMBO J* **28**, 755-765 (2009).

23. D. Hughes *et al.*, Genetic characterization of frameshift suppressors with new decoding properties. *J Bacteriol* **171**, 1028-1034 (1989).

24. X. Agirrezabala *et al.*, Structural insights into cognate versus near-cognate discrimination during decoding. *EMBO J* **30**, 1497-1507 (2011).

25. X. Zeng, J. Chugh, A. Casiano-Negroni, H. M. Al-Hashimi, C. L. Brooks, 3rd, Flipping of the ribosomal A-site adenines provides a basis for tRNA selection. *J Mol Biol* **426**, 3201-3213 (2014).

26. A. B. Loveland, G. Demo, A. A. Korostelev, Cryo-EM of elongating ribosome with EF-Tu•GTP elucidates tRNA proofreading. *Nature* 10.1038/s41586-020-2447-x (2020).

27. Y. S. Polikanov, T. A. Steitz, C. A. Innis, A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nat Struct Mol Biol* **21**, 787-793 (2014).

28. A. Korostelev, D. N. Ermolenko, H. F. Noller, Structural dynamics of the ribosome. *Curr Opin Chem Biol* **12**, 674-683 (2008).

29. W. Zhang, J. A. Dunkle, J. H. Cate, Structures of the ribosome in intermediate states of ratcheting. *Science* **325**, 1014-1017 (2009).

30. J. Zhou, L. Lancaster, J. P. Donohue, H. F. Noller, Crystal structures of EF-G-ribosome complexes trapped in intermediate states of translocation. *Science* **340**, 1236086 (2013).

31. D. Sharma, EF-G-independent reactivity of a pre-translocation-state ribosome complex with the aminoacyl tRNA substrate puromycin supports an intermediate (hybrid) state of tRNA binding. *RNA* **10**, 102-113 (2004).

32. Y. P. Semenkov, M. V. Rodnina, W. Wintermeyer, Energetic contribution of tRNA hybrid state formation to translocation catalysis on the ribosome. *Nat Struct Biol* **7**, 1027-1031 (2000).

33. Y. G. Gao *et al.*, The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* **326**, 694-699 (2009).

34. J. Zhou, L. Lancaster, J. P. Donohue, H. F. Noller, How the ribosome hands the A-site tRNA to the P site during EF-G-catalyzed translocation. *Science* **345**, 1188-1191 (2014).

35. J. Lin, M. G. Gagnon, D. Bulkley, T. A. Steitz, Conformational changes of elongation factor G on the ribosome during tRNA translocation. *Cell* **160**, 219-227 (2015).

36. J. B. Russell, G. M. Cook, Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol Rev* **59**, 48-62 (1995).

37. H. Bremer, P. P. Dennis, Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus* **3** (2008).

38. X. Zhou, W.-J. Liao, J.-M. Liao, P. Liao, H. Lu, Ribosomal proteins: functions beyond the ribosome. *Journal of Molecular Cell Biology* **7**, 92-104 (2015).

39. N. Manickam, N. Nag, A. Abbasi, K. Patel, P. J. Farabaugh, Studies of translational misreading in vivo show that the ribosome very efficiently discriminates against most potential errors. *RNA* **20**, 9-15 (2014).

40. M. V. Rodnina, W. Wintermeyer, Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem* **70**, 415-435 (2001).

41. K. B. Gromadski, T. Daviter, M. V. Rodnina, A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. *Mol Cell* **21**, 369-377 (2006).

42. M. V. Rodnina, N. Fischer, C. Maracci, H. Stark, Ribosome dynamics during decoding. *Philos Trans R Soc Lond B Biol Sci* **372** (2017).

43. M. Johansson, J. Zhang, M. Ehrenberg, Genetic code translation displays a linear trade-off between efficiency and accuracy of tRNA selection. *Proceedings of the National Academy of Sciences* **109**, 131-136 (2012).

44. J. Zhang, K.-W. Ieong, M. Johansson, M. Ehrenberg, Accuracy of initial codon selection by aminoacyl-tRNAs on the mRNA-programmed bacterial ribosome. *Proceedings of the National Academy of Sciences* **112**, 9602-9607 (2015).

45. J. M. Ogle, F. V. Murphy, M. J. Tarry, V. Ramakrishnan, Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721-732 (2002).

46. I. M. Krab, A. Parmeggiani, "Mechanisms of EF-Tu, a pioneer GTPase" in Progress in Nucleic Acid Research and Molecular Biology. (Elsevier, 2002), 10.1016/s0079-6603(02)71050-7, pp. 513-551.

47. M. V. Rodnina, T. Pape, R. Fricke, L. Kuhn, W. Wintermeyer, Initial Binding of the Elongation Factor Tu·GTP·Aminoacyl-tRNA Complex Preceding Codon Recognition on the Ribosome. *Journal of Biological Chemistry* **271**, 646-652 (1996).

48. T. Pape, Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E.coli ribosome. *The EMBO Journal* **17**, 7490-7497 (1998).

49. K. Abel, M. D. Yoder, R. Hilgenfeld, F. Jurnak, An α to β conformational switch in EF-Tu. *Structure* **4**, 1153-1159 (1996).

50. K.-W. Ieong, Ü. Uzun, M. Selmer, M. Ehrenberg, Two proofreading steps amplify the accuracy of genetic code translation. *Proceedings of the National Academy of Sciences* **113**, 13744-13749 (2016).

51. S. C. Blanchard, R. L. Gonzalez, H. D. Kim, S. Chu, J. D. Puglisi, tRNA selection and kinetic proofreading in translation. *Nature Structural & Molecular Biology* **11**, 1008-1014 (2004).

52. S. C. Blanchard, H. D. Kim, R. L. Gonzalez, Jr., J. D. Puglisi, S. Chu, tRNA dynamics on the ribosome during translation. *Proc Natl Acad Sci U S A* **101**, 12893-12898 (2004).

53. J. Shepherd, M. Ibba, Bacterial transfer RNAs. *FEMS Microbiology Reviews* **39**, 280-300 (2015).

54. A. K. Hopper, tRNA transfers to the limelight. *Genes & Development* **17**, 162-180 (2003).

55. S. Rudorf, R. Lipowsky, Protein Synthesis in E. coli: Dependence of Codon-Specific Elongation on tRNA Concentration and Codon Usage. *PLOS ONE* **10**, e0134994 (2015).

56. N. Krahn, J. T. Fischer, D. Soll, Naturally Occurring tRNAs With Non-canonical Structures. *Front Microbiol* **11**, 596914 (2020).

57. K. Gabriel, J. Schneider, W. H. McClain, Functional evidence for indirect recognition of G.U in tRNA(Ala) by alanyl-tRNA synthetase. *Science* **271**, 195-197 (1996).

58. Y. Goldgur *et al.*, The crystal structure of phenylalanyl-tRNA synthetase from thermus thermophilus complexed with cognate tRNAPhe. *Structure* **5**, 59-68 (1997).

59. P. Boccaletto *et al.*, MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* **46**, D303-D307 (2018).

60. F. Tuorto, F. Lyko, Genome recoding by tRNA modifications. *Open Biol* **6** (2016).

61. J. Urbonavicius, Q. Qian, J. M. Durand, T. G. Hagervall, G. R. Bjork, Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J* **20**, 4863-4873 (2001).

62. C. Yarian *et al.*, Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem* **277**, 16391-16395 (2002).

63. E. M. Gustilo, F. A. Vendeix, P. F. Agris, tRNA's modifications bring order to gene expression. *Curr Opin Microbiol* **11**, 134-140 (2008).

64. T. Pan, Modifications and functional genomics of human transfer RNA. *Cell Res* **28**, 395-404 (2018).

65. T. Suzuki, The expanding world of tRNA modifications and their disease relevance. *Nature Reviews Molecular Cell Biology* **22**, 375-392 (2021).

66. V. de Crecy-Lagard, M. Jaroch, Functions of Bacterial tRNA Modifications: From Ubiquity to Diversity. *Trends Microbiol* **29**, 41-53 (2021).

67. P. F. Agris, F. A. Vendeix, W. D. Graham, tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol* **366**, 1-13 (2007).

68. M. Sprinzl, C. Horn, M. Brown, A. Ioudovitch, S. Steinberg, Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* **26**, 148-153 (1998).

69. H. Grosjean, M. Sprinzl, S. Steinberg, Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie* **77**, 139-141 (1995).

70. U. Lagerkvist, "Two out of three": an alternative method for codon reading. *Proceedings of the National Academy of Sciences* **75**, 1759-1762 (1978).

71. L. B. Jenner, N. Demeshkina, G. Yusupova, M. Yusupov, Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat Struct Mol Biol* **17**, 555-560 (2010).

72. N. Demeshkina, L. Jenner, E. Westhof, M. Yusupov, G. Yusupova, A new understanding of the decoding principle on the ribosome. *Nature* **484**, 256-259 (2012).

73. N. Demeshkina, L. Jenner, E. Westhof, M. Yusupov, G. Yusupova, New structural insights into the decoding mechanism: translation infidelity via a G.U pair with Watson-Crick geometry. *FEBS Lett* **587**, 1848-1857 (2013).

74. A. Rozov *et al.*, Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat Commun* **7**, 10457 (2016).

75. A. Rozov *et al.*, Tautomeric G*U pairs within the molecular ribosomal grip and fidelity of decoding in bacteria. *Nucleic Acids Res* 10.1093/nar/gky547 (2018).

76. F. V. t. Murphy, V. Ramakrishnan, A. Malkiewicz, P. F. Agris, The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat Struct Mol Biol* **11**, 1186-1191 (2004).

77. F. V. t. Murphy, V. Ramakrishnan, Structure of a purine-purine wobble base pair in the decoding center of the ribosome. *Nat Struct Mol Biol* **11**, 1251-1252 (2004).

78. A. Weixlbaumer *et al.*, Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nat Struct Mol Biol* **14**, 498-502 (2007).

79. C. Yarian *et al.*, Modified Nucleoside Dependent Watson−Crick and Wobble Codon Binding by tRNALysUUU Species. *Biochemistry* **39**, 13390-13395 (2000).

80. S. S. Phelps, O. Jerinic, S. Joseph, Universally Conserved Interactions between the Ribosome and the Anticodon Stem-Loop of A Site tRNA Important for Translocation. *Molecular Cell* **10**, 799-807 (2002).

81. H. B. Gamper, I. Masuda, M. Frenkel-Morgenstern, Y. M. Hou, Maintenance of protein synthesis reading frame by EF-P and m(1)G37-tRNA. *Nat Commun* **6**, 7226 (2015).

82. Y. M. Hou, R. Matsubara, R. Takase, I. Masuda, J. I. Sulkowska, TrmD: A Methyl Transferase for tRNA Methylation With m(1)G37. *Enzymes* **41**, 89-115 (2017).

83. H. Hori, Transfer RNA methyltransferases with a SpoU-TrmD (SPOUT) fold and their modified nucleosides in tRNA. *Biomolecules* **7** (2017).

84. Y. M. Hou, I. Masuda, H. Gamper, Codon-Specific Translation by m(1)G37 Methylation of tRNA. *Front Genet* **9**, 713 (2018).

85. B. E. Clifton, M. A. Fariz, G.-I. Uechi, P. Laurino (2021) Evolutionary repair reveals an unexpected role of the tRNA modification m1G37 in aminoacylation. (Cold Spring Harbor Laboratory).

86. I. Masuda *et al.*, Loss of N1-methylation of G37 in tRNA induces ribosome stalling and reprograms gene expression. *eLife* **10** (2021).

87. G. R. Bjork, P. M. Wikstrom, A. S. Bystrom, Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science* **244**, 986-989 (1989).

88. T. G. Hagervall, T. M. Tuohy, J. F. Atkins, G. R. Bjork, Deficiency of 1-methylguanosine in tRNA from Salmonella typhimurium induces frameshifting by quadruplet translocation. *J Mol Biol* **232**, 756-765 (1993).

89. M. Y. Pavlov *et al.*, Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc Natl Acad Sci U S A* **106**, 50-54 (2009).

90. L. Peil *et al.*, Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proc Natl Acad Sci U S A* **110**, 15265-15270 (2013).

91. C. J. Woolstenhulme, N. R. Guydosh, R. Green, A. R. Buskirk, High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep* **11**, 13-21 (2015).

92. A. Louie, N. S. Ribeiro, B. R. Reid, F. Jurnak, Relative affinities of all Escherichia coli aminoacyl-tRNAs for elongation factor Tu-GTP. *J Biol Chem* **259**, 5010-5016 (1984).

93. S. Ude *et al.*, Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* **339**, 82-85 (2013).

94. R. Tollerson, 2nd, A. Witzky, M. Ibba, Elongation factor P is required to maintain proteome homeostasis at high growth rate. *Proc Natl Acad Sci U S A* **115**, 11072-11077 (2018).

95. H. A. Nguyen, E. D. Hoffer, C. M. Dunham, Importance of a tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in tRNACGGPro for decoding. *Journal of Biological Chemistry* **294**, 5281-5291 (2019).

96. S. Hong *et al.*, Mechanism of tRNA-mediated +1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* 10.1073/pnas.1809319115 (2018).

97. T. Maehigashi, J. A. Dunkle, S. J. Miles, C. M. Dunham, Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. *Proc Natl Acad Sci U S A* **111**, 12740-12745 (2014).

98. E. D. Hoffer *et al.*, Structural insights into mRNA reading frame regulation by tRNA modification and slippery codon–anticodon pairing. *eLife* **9** (2020).

99. G. Jager, K. Nilsson, G. R. Bjork, The phenotype of many independently isolated +1 frameshift suppressor mutants supports a pivotal role of the P-site in reading frame maintenance. *PLoS One* **8**, e60246 (2013).

100. I. Masuda *et al.*, tRNA Methylation Is a Global Determinant of Bacterial Multi-drug Resistance. *Cell Syst* 10.1016/j.cels.2019.03.008 (2019).

101. Y. M. Hou, I. Masuda, L. J. Foster, tRNA methylation: An unexpected link to bacterial resistance and persistence to antibiotics and beyond. *WIREs RNA* **11** (2020).

102. G. R. Bjork *et al.*, A primordial tRNA modification required for the evolution of life? *EMBO J* **20**, 231-239 (2001).

103. D. L. Riddle, J. R. Roth, Suppressors of frameshift mutations in Salmonella typhimurium. *J Mol Biol* **54**, 131-144 (1970).

104. D. L. Riddle, J. R. Roth, Frameshift suppressors. 3. Effects of suppressor mutations on transfer RNA. *J Mol Biol* **66**, 495-506 (1972).

105. C. M. Cummins, M. R. Culbertson, G. Knapp, Frameshift suppressor mutations outside the anticodon in yeast proline tRNAs containing an intervening sequence. *Mol Cell Biol* **5**, 1760-1771 (1985).

106. J. F. Atkins, R. B. Weiss, S. Thompson, R. F. Gesteland, Towards a genetic dissection of the basis of triplet decoding, and its natural subversion: programmed reading frame shifts and hops. *Annu Rev Genet* **25**, 201-228 (1991).

107. G. E. Sroga, F. Nemoto, Y. Kuchino, G. R. Bjork, Insertion (sufB) in the anticodon loop or base substitution (sufC) in the anticodon stem of tRNA(Pro)2 from Salmonella typhimurium induces suppression of frameshift mutations. *Nucleic Acids Res* **20**, 3463-3469 (1992).

108. T. M. Tuohy, S. Thompson, R. F. Gesteland, J. F. Atkins, Seven, eight and nine-membered anticodon loop mutants of tRNA(2Arg) which cause +1 frameshifting. Tolerance of DHU arm and other secondary mutations. *J Mol Biol* **228**, 1042-1054 (1992).

109. T. J. Magliery, J. C. Anderson, P. G. Schultz, Expanding the genetic code: selection of efficient suppressors of four-base codons and identification of "shifty" four-base codons with a library approach in Escherichia coli. *J Mol Biol* **307**, 755-769 (2001).

110. V. Bourdeau *et al.*, Amber suppression in Escherichia coli by unusual mitochondria-like transfer RNAs. *Proc Natl Acad Sci U S A* **95**, 1375-1380 (1998).

111. A. K. Kowal, C. Kohrer, U. L. RajBhandary, Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria. *Proc Natl Acad Sci U S A* **98**, 2268-2273 (2001).

112. H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464**, 441-444 (2010).

113. Y. Zhang *et al.*, A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **551**, 644-647 (2017).

114. V. T. Dien, S. E. Morris, R. J. Karadeema, F. E. Romesberg, Expansion of the genetic code via expansion of the genetic alphabet. *Curr Opin Chem Biol* **46**, 196-202 (2018).

115. D. L. Riddle, J. Carbon, Frameshift suppression: a nucleotide addition in the anticodon of a glycine transfer RNA. *Nat New Biol* **242**, 230-234 (1973).

116. D. Hirsh, Tryptophan transfer RNA as the UGA suppressor. *J Mol Biol* **58**, 439-458 (1971).

117. T. M. Schmeing, R. M. Voorhees, A. C. Kelley, V. Ramakrishnan, How mutations in tRNA distant from the anticodon affect the fidelity of decoding. *Nature Structural & Molecular Biology* **18**, 432-436 (2011).

118. A. L. Starosta *et al.*, A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. *Cell Rep* **9**, 476-483 (2014).

119. P. Huter *et al.*, Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. *Mol Cell* **68**, 515-527 e516 (2017).

120. M. O'Connor, Imbalance of tRNA(Pro) isoacceptors induces +1 frameshifting at near-cognate codons. *Nucleic Acids Res* **30**, 759-765 (2002).

121. H. B. Gamper, I. Masuda, M. Frenkel-Morgenstern, Y. M. Hou, The UGG Isoacceptor of tRNAPro Is Naturally Prone to Frameshifts. *Int J Mol Sci* **16**, 14866-14883 (2015).

122. Q. Qian, G. R. Bjork, Structural alterations far from the anticodon of the tRNAProGGG of Salmonella typhimurium induce +1 frameshifting at the peptidyl-site. *J Mol Biol* **273**, 978-992 (1997).

123. Q. Qian *et al.*, A new model for phenotypic suppression of frameshift mutations by mutant tRNAs. *Mol Cell* **1**, 471-482 (1998).

124. S. Ledoux, O. C. Uhlenbeck, Different aa-tRNAs Are Selected Uniformly on the Ribosome. *Molecular Cell* **31**, 114-123 (2008).

125. J. M. Schrader, S. J. Chapman, O. C. Uhlenbeck, Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proceedings of the National Academy of Sciences* **108**, 5215-5220 (2011).

126. N. Manickam, K. Joshi, M. J. Bhatt, P. J. Farabaugh, Effects of tRNA modification on translational accuracy depend on intrinsic codon-anticodon strength. *Nucleic Acids Res* **44**, 1871-1881 (2016).

127. M. Olejniczak, O. C. Uhlenbeck, tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. *Biochimie* **88**, 943-950 (2006).

128. K. Pernod *et al.*, The nature of the purine at position 34 in tRNAs of 4-codon boxes is correlated with nucleotides at positions 32 and 38 to maintain decoding fidelity. *Nucleic Acids Res* 10.1093/nar/gkaa221 (2020).

129. M. Olejniczak, T. Dale, R. P. Fahlman, O. C. Uhlenbeck, Idiosyncratic tuning of tRNAs to achieve uniform ribosome binding. *Nat Struct Mol Biol* **12**, 788-793 (2005).

130. S. Ledoux, M. Olejniczak, O. C. Uhlenbeck, A sequence element that tunes Escherichia coli tRNA(Ala)(GGC) to ensure accurate decoding. *Nat Struct Mol Biol* **16**, 359-364 (2009).

131. Y. M. Hou, P. Schimmel, A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* **333**, 140-145 (1988).

132. L. G. Kleina, J. M. Masson, J. Normanly, J. Abelson, J. H. Miller, Construction of Escherichia coli amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of suppressor efficiency. *J Mol Biol* **213**, 705-717 (1990).

133. F. Tsai, J. F. Curran, tRNA2 Gln mutants that translate the CGA arginine codon as glutamine in Escherichia coli. *RNA* **4**, 1514-1522 (1998).

134. H. Murakami, A. Ohta, H. Suga, Bases in the anticodon loop of tRNA(Ala)(GGC) prevent misreading. *Nat Struct Mol Biol* **16**, 353-358 (2009).

135. H. A. Nguyen, S. Sunita, C. M. Dunham, Disruption of evolutionarily correlated tRNA elements impairs accurate decoding. *Proceedings of the National Academy of Sciences* **117**, 16333-16338 (2020).

136. A. Vimaladithan, P. J. Farabaugh, Special peptidyl-tRNA molecules can promote translational frameshifting without slippage. *Mol Cell Biol* **14**, 8107-8116 (1994).

137. S. Pande, A. Vimaladithan, H. Zhao, P. J. Farabaugh, Pulling the ribosome out of frame by +1 at a programmed frameshift site by cognate binding of aminoacyl-tRNA. *Mol Cell Biol* **15**, 298-304 (1995).

138. H. S. Zaher, R. Green, Kinetic basis for global loss of fidelity arising from mismatches in the P-site codon:anticodon helix. *RNA* **16**, 1980-1989 (2010).

139. A. Bębenek, I. Ziuzia-Graczyk, Fidelity of DNA replication—a matter of proofreading. *Current Genetics* **64**, 985-996 (2018).

140. S. C. Blanchard, When starting over makes more sense. *ACS Chem Biol* **4**, 89-92 (2009).

## Chapter 2.

## Importance of tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in $tRNA^{Pro}_{CGG}$ for decoding

Ha An Nguyen, Eric D. Hoffer, Christine M. Dunham.

This research was originally published in the Journal of Biological Chemistry. H. A. Nguyen, E. D. Hoffer, C. M. Dunham, Importance of a tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in $tRNA^{Pro}_{CGG}$ for decoding. *Journal of Biological Chemistry* **294**, 5281-5291 (2019).

**Author contributions**: H. A. N., E. D. H., and C. M. D. conceptualization, data curation, formal analysis, validation, investigation, methodology, writing-original draft, project administration, writing-review and editing, and supervision; C. M. D. resources and funding acquisition.

## 2.1. Abstract

Modification of anticodon nucleotides allows tRNAs to decode multiple codons, expanding the genetic code. Additionally, modifications located in the anticodon loop, but outside the anticodon itself, stabilize tRNA–codon interactions, increasing decoding fidelity. Anticodon loop nucleotide 37 is 3′ to the anticodon and, in $tRNA^{Pro}_{CGG}$, is methylated at the N1 position in its nucleobase ($m^1G37$). The $m^1G37$ modification in $tRNA^{Pro}_{CGG}$ stabilizes its interaction with the codon and maintains the mRNA frame. However, it is unclear how $m^1G37$ affects binding at the decoding center to both cognate and +1 slippery codons. Here, we show that the $tRNA^{Pro}_{CGG}$ $m^1G37$ modification is important for the association step during binding to a cognate CCG codon. In contrast, $m^1G37$ prevented association with a slippery CCC-U or +1 codon. Similar analyses of frameshift suppressor $tRNA^{SufA6}$, a $tRNA^{Pro}_{CGG}$ derivative containing an extra nucleotide in its anticodon loop that undergoes +1 frameshifting, reveal that $m^1G37$ destabilizes interactions with both the cognate CCG and slippery codons. One reason for this destabilization is the disruption of a conserved U32·A38 nucleotide pairing in the anticodon stem through insertion of G37.5. Restoring the $tRNA^{SufA6}$ U32·A37.5 pairing results in a high-affinity association on the slippery CCC-U codon. Further, an X-ray crystal structure of the 70S ribosome bound to $tRNA^{SufA6}$ U32·A37.5 at 3.6 Å resolution shows a reordering of the anticodon loop consistent with the findings from the high-affinity measurements. Our results reveal how the tRNA modification at nucleotide 37 stabilizes interactions with the mRNA codon to preserve the mRNA frame.

## 2.2. Introduction

Protein synthesis is performed by the ribosome, a conserved protein–RNA macromolecular machine where mRNA, tRNAs, and translation factors read the genetic information as presented on mRNA into proteins. There are four defined stages of protein synthesis: initiation, elongation, termination, and recycling (reviewed in Ref. 1). During elongation, three nucleotides of the mRNA codon are read (or decoded) by three anticodon nucleotides of a tRNA in the ribosomal aminoacyl site (A site) 2 on the small 30S subunit. The three-nucleotide code on the mRNA defines a single amino acid delivered by the corresponding tRNA. The regulation of the mRNA frame is critically important to maintain the correct sequential addition of amino acids to the nascent chain (2). Despite the importance of accurate protein expression for cell viability, the molecular basis for how the ribosome maintains this three-nucleotide mRNA frame is not well-understood.

Since tRNAs decode mRNAs, these RNA molecules probably play a role in mRNA frame maintenance. tRNAs are ~76–90 nucleotides in length and adopt an L-shaped tertiary structure allowing them to fit into ribosome-binding sites that span both subunits (Fig. 2.1). tRNAs undergo extensive post-transcriptional modifications important for the correct tertiary fold of the tRNA, including the conformation of the anticodon stem-loop (ASL) (3). RNA modifications that are located in the anticodon and neighboring nucleotides in the ASL contribute to the accuracy and speed of translation (4, 5) by stabilizing the interactions between the anticodon and codon (3, 6, 78). After decoding, these tRNA modifications are also important during translocation of the mRNA-tRNA pairs (9) and have also been implicated in mRNA frame maintenance (6).

The selection of the correct tRNA for each mRNA codon relies on the formation of Watson-Crick base pairs between the first two nucleotides of the codon and nucleotides 36 and 35 of the anticodon (Fig. 2.1*A*). The interaction between the third nucleotide of the codon and anticodon nucleotide 34 is not required to be Watson-Crick. Instead, a G•U wobble pair or a modified anticodon nucleotide 34 - codon nucleotide pair can form. The modification of nucleotide 34 enables non-Watson-Crick interactions with the third nucleotide position of the mRNA that is accepted as cognate by the ribosome. The increased flexibility in codons that each tRNA can decode allows for the degeneracy of the genetic code where the 61 codons are decoded by fewer tRNAs (10). Therefore, tRNA modifications at nucleotide 34 have an important and essential role in the process of decoding.

Nucleotide 34 contains many diverse modifications that are typically required for accurate translation (11). Two examples include the uridine-5-oxyacetic acid ($cmo^5U34$) in $tRNA^{Ala}_{CGU}$ and the 5-methylaminomethyl-2-thiouridine ($mnm^5s^2$) U34 modification in $tRNA^{Lys}_{UUU}$. The $cmo^5U34$ modification in $tRNA^{Ala}$ stabilizes its interaction with U6 at the wobble position (12). The $mnm^5s^2$ U34 modification in $tRNA^{Lys}_{UUU}$ allows for pairing with UUC/U/G or A codons (13,14). However, a 6-threonylcarbamoyladenosine ($t^6A$) or a 2-methylthio derivative ($ms^2t^6A$) at nucleotide 37 is required for recognition of AAA or AAG codons (4,7). Although all these codon-anticodon pairings should be recognized by the ribosome, the instability of the anticodon loop of $tRNA^{Lys}_{UUU}$ and thus, its interactions with the AAA codon, require both modifications at nucleotides 34 and 37. Nucleotide 37 of the tRNA is located 3' to the anticodon, adjacent to the first position of the Watson-Crick base pair between the codon and anticodon nucleotide 36 (Fig. 2.1*A*). The codon-anticodon

pairing between $tRNA^{Lys}_{UUU}$ and its codon is weak in the absence of the modifications because of the three A-U base pairs and poor stacking of the UUU anticodon (15). Both the $t^6A$ and $ms^2t^6A$ modifications contain planar heterocycle moieties that promote cross-strand stacking interactions between A38 of the tRNA and the first base in the mRNA codon to stabilize the codon-anticodon pairing (16). Since decoding relies on both the high affinity binding of cognate tRNAs to the decoding center and conformational changes of the 30S known as domain closure (17-19), these modifications influence both aspects of decoding.

Nucleotide 37 is modified in >70% of all tRNAs and is typically a purine (20). Among the modified nucleosides, the $t^6A$ and methylated guanosine ($m^1G$) are the most common (11). In the absence of the modification at nucleotide 37, the anticodon loops of human $tRNA^{Lys}$ and yeast $tRNA^{Asp}$ lack structural rigidity compared to their modified forms (21,22). Likewise, the modification at nucleotide 37 in *E. coli* $tRNA^{Lys}$ and $tRNA^{Phe}$ stabilizes the canonical uridine turn (U turn) in the ASL which is required for high affinity binding to the A site (23-25).

Approximately 63% of tRNAs containing the $m^1G37$ modification decode CNN codons (where N indicates any nucleotide) including Leu, Pro, His, Gln and Arg codons (11). The $m^1G37$ modification is present in ~95% of all known sequences of proline tRNAs (26). In bacteria, the essential methyltransferase TrmD (Trm5 in eukaryotes and archaea) catalyzes the N1-methylation of G37 in tRNAs. Furthermore, the $m^1G37$ modification stabilizes the anticodon of the tRNA to prevent +1 ribosomal frameshift errors while mutations in *trmD* also cause growth defects (6,26-28). $tRNA^{Pro}_{CGG}$, the major isoacceptor for proline, decodes the CCG codon and in the absence of the $m^1G37$ modification,

causes +1 frameshifting on CCC-N codons. It was previously thought that $tRNA^{Pro}_{CGG}$ lacking the $m^1G37$ modification would allow for a four-nucleotide interaction between the anticodon and the mRNA codon with G37 interacting with the mRNA codon (29). However, biochemical and structural studies of $ASL^{Pro}_{CCG}$ lacking the modification revealed that this four-nucleotide interaction does not occur in the A site during decoding (30,31). Additionally, the mRNA is positioned in the unshifted or zero frame indicating that the frameshift event occurs post-decoding consistent with recent structures (32). Interestingly, the absence of the methylation at G37 causes a distortion of the tRNA on the opposite side of the anticodon loop at nucleotide U32 (31), leading to the disruption of interactions with A38. Collectively, these results suggest a previously unappreciated role of the stabilization of the 32•38 pairing in $tRNA^{Pro}_{CGG}$ predicted to maintain the correct mRNA frame (31).

Frameshift-suppressor tRNAs derived from $tRNA^{Pro}$ contain an insertion between anticodon loop nucleotides 37 and 38 (referred to as 37.5) decode CCC-N codons as proline (26,33-36) (Fig. 2.1). These mutant tRNAs are genetic suppressors that perform noncanonical reading of the genetic code to restore the reading frame (37,38). In this case, frameshift suppressor $tRNA^{SufA6}$, isolated from *Salmonella enterica serova* Typhimurium, contains an eight-nucleotide anticodon loop by the addition of G37.5 that causes +1 frameshifting. The structure of 70S-$tRNA^{SufA6}$ bound to the decoding center to CCC-A/U/C codons that undergoes +1 frameshifting revealed similarities to the structure of 70S-$tRNA^{Pro}_{CGG}$ lacking the $m^1G37$ bound to a near-cognate codon that also promotes +1 frameshifting (31). Both tRNAs decode the mRNA in the unshifted or zero frame indicating the shift into the new frame occurred post-decoding. Moreover, the inserted

37.5 nucleotide and the lack of $m^1G37$ both cause destabilization of nucleotides on the opposite side of the anticodon loop that ablates a conserved, non-Watson-Crick U32•A38 pairing. The 32•38 pairing was restored in both the $tRNA^{Pro}$ and $tRNA^{SufA6}$ in the context of recognizing a cognate, three-nucleotide codon. The disruption of the 32•38 pairing is particularly notable due to its universal significance in tuning the ribosomal binding across tRNAs (39). These results provide insight into how tRNA modifications and the 32•38 pairing in the anticodon loop together lead to mRNA frame maintenance.

Here, we tested how the $m^1G37$ modification in $tRNA^{Pro}_{CGG}$ and $tRNA^{SufA6}$ impacts binding at the decoding center to cognate and slippery +1 codons. Further, we engineer $tRNA^{SufA6}$ to contain a conserved U32•A38 pairing to attempt to restore high-affinity binding to decoding center. Lastly, a 3.6 Å X-ray crystal structure of $tRNA^{SufA6}$ containing this engineered 32•38 pairing bound to the 70S ribosomal A site reveals a reordering of the 32•38 pair required for decoding.

## 2.3. Results

### *The $m^1G37$ modification in $tRNA^{Pro}_{CGG}$ stabilizes binding to the A site.*

To assess the importance of the $m^1G37$ modification in $tRNA^{Pro}_{CGG}$ in decoding, we used established filter binding assays to determine binding kinetics to the A site (39). *E. coli* 70S ribosomes were programmed with mRNA containing a peptidyl(P)-site AUG start codon and an A-site proline CCG codon, and P-site *E. coli* $tRNA^{fMet}$. We used a chemically synthesized ASL containing 18 nucleotides of $tRNA^{Pro}_{CGG}$ and a $m^1G37$ modification to ensure the RNA was completely modified (Table 2.3). $ASL^{Pro}_{CCG}$ with the $m^1G37$ modification binds to a cognate CCG codon in the A site with an equilibrium dissociation constant ($K_d$) of 284 nM (Fig. 2.2*A*, Table 2.4). This affinity is within the range of reported

values for ASLs binding to the A site (33-500 nM) (9,40). Removal of the $m^1G37$ modification ($\Delta m^1G37$) significantly reduced binding, by ~6.5 fold (1.8 μM; Fig. 2.2*A*, Table 2.4).

Although the data could be fit with reasonable confidence (Fig. 2.3, Table 2.4), the low maximum binding was concerning, if not unprecedented (39-41). Furthermore, impractical 70S concentrations required to reach maximum binding for weaker interactions prevented us from attempting to continue with this approach. Therefore, we instead performed competition binding assays which allows for the calculation of the equilibrium dissociation constant ($K_d$) based on measured association ($k_{on}$) and dissociation ($k_{off}$) rates (Fig. 2.2*B*)(42,43). Unlike measuring the experimental $K_d$ to observe equilibrium binding, the $k_{on}$ and $k_{off}$ measurements provide information regarding the influence of the $m^1G37$ modification at each step of the binding event separately, that is, the association to and dissociation from the ribosomal A site. Using this approach, we found a modest difference in $k_{off}$ between $ASL^{Pro}_{CCG}$ and $ASL^{Pro}_{CCG}$ $\Delta m^1G37$ implying the modification does not stabilize tRNA binding to the ribosomes. Instead, the $m^1G37$ modification is important for initial binding as shown by the 2.6-fold slower association rate of $ASL^{Pro}_{CCG}$ $\Delta m^1G37$ ($k_{on}$ = 0.012 $\mu M^{-1}$ $min^{-1}$) as compared to $ASL^{Pro}_{CCG}$ ($k_{on}$ = 0.031 $\mu M^{-1}$ $min^{-1}$) to the CCG A-site codon (Fig. 2.2*C,D*; Table 2.1). These data are consistent with 70S structures demonstrating conformational distortion of the anticodon loop in the absence of the $m^1G37$ modification (31). Furthermore, we found that the calculated dissociation constant ($K_d$) of $ASL^{Pro}_{CCG}$ $m^1G37$ with the cognate CCG codon is 420 nM (Table 2.1) whereas $ASL^{Pro}_{CCG}$ $\Delta m^1G37$ binds to the A-site CCG codon with a calculated $K_d$ of 1.4 μM, consistent with directly measured $K_d$ values (Fig. 2.2*A*, Table 2.4).

***Recognition of a +1 slippery CCC-U codon is enhanced by the lack of the $m^1$G37 modification in $tRNA^{Pro}_{CGG}$.***

$tRNA^{Pro}_{CGG}$ lacking the $m^1$G37 modification undergoes high levels of +1 frameshifting on CCC-N codons (26,28). Previous 70S structures of $tRNA^{Pro}_{CGG}$ $\Delta m^1$G37 bound to A-site CCC-N codons revealed the tRNA decodes in the unshifted or zero frame (31). The three tRNA anticodon nucleotides C34-G35-G36 form three interactions with the C4-C5-C6 mRNA codon, respectively (where the first nucleotide of the P-site mRNA codon is denoted as 1 and the A-site nucleotides are 4, 5 and 6) (Fig. 2.1*A*). The interaction between the anticodon and the codon is near cognate as defined by a single mismatch between C34 and C6 (Fig. 2.4*A*). To test the impact of the $m^1$G37 modification on the ability of $tRNA^{Pro}_{CGG}$ to form a stable complex with the slippery CCC-U codon, we again measured binding kinetics. In the context of $ASL^{Pro}_{CGG}$ binding to a slippery codon, the $m^1$G37 modification influences binding but in the opposite manner to binding to the cognate CCG codon (Fig. 2.2). $ASL^{Pro}_{CGG}$ lacking the $m^1$G37 modification associates ~4-fold faster to the slippery codon ($k_{on}$ = 0.017 versus 0.0039 $\mu M^{-1}min^{-1}$); however, lack of the modification only has a moderate impact on $k_{off}$ (0.009 versus 0.014 $min^{-1}$) (Fig. 2.4*B*,*C*; Table 2.1). Calculated $K_d$ measurements of 3.62 μM and 0.41 μM for $ASL^{Pro}_{CGG}$ and $ASL^{Pro}_{CGG}$ $\Delta m^1$G37, respectively, indicate a 9-fold difference in binding affinity. In summary, the stabilizing effect observed in the cognate CCG context is reversed on a slippery CCC-U codon: the presence of the $m^1$G37 modification actually impairs the association of $ASL^{Pro}_{CGG}$ to the A site programmed with a slippery CCC-U codon.

***Nucleotide insertion in the ASL of $tRNA^{SufA6}$ counteracts the stabilization exerted by $m^1$G37.***

Frameshift suppressor $tRNA^{SufA6}$ undergoes +1 frameshifting similar to $tRNA^{Pro}_{CGG}$ lacking the $m^1G37$ modification (26,31). Both $tRNA^{SufA6}$ and $tRNA^{Pro}_{CGG}$ contain the $m^1G37$ modification and, in the case of $tRNA^{SufA6}$, this modification is located adjacent to the inserted nucleotide (30) (Fig. 2.5*A*). we next tested the importance of the $m^1G37$ modification in the context of an eight-nucleotide anticodon loop in $ASL^{SufA6}$, using the same kinetic binding assays as before. In contrast to the stabilizing effect observed with the $m^1G37$ modification in $ASL^{Pro}_{CGG}$ on a cognate CCG codon, $ASL^{SufA6}$ associates with the CCG codon 10-fold faster in the absence of the modification (0.053 versus 0.0053 $\mu M^{-1}$ $min^{-1}$) (Fig. 2.5*B*, Table 2.1). In contrast, the dissociation of $ASL^{SufA6}$ was essentially unaffected by the absence or presence of the modification ($k_{off}$ = 0.013 and 0.014 $min^{-1}$; Fig. 2.5*C*). The 10-fold difference in the calculated $K_d$ between $ASL^{SufA6}$ containing the $m^1G37$ modification (2.5 μM) and lacking the $m^1G37$ modification (0.26 μM) is thus reflective of the large changes in the tRNA association with the A site. One interpretation of these observations could be that although the $m^1G37$ modification imparts a stabilizing effect in anticodon loops of canonical seven nucleotides (10,44), increasing the anticodon loop to eight nucleotides as seen in $ASL^{SufA6}$, ablates any stabilization from the modification. Additionally, the overall trends of the $ASL^{SufA6}$ association rates are similar to the rates seen with $ASL^{Pro}_{CGG}$ on the slippery CCC-U codon.

Next, we tested the binding of $ASL^{SufA6}$ to the slippery CCC-U codon. We found that the influence of the modification status of $ASL^{SufA6}$ follows similar trends regardless of whether $ASL^{SufA6}$ is recognizing a cognate CCG or a slippery CCC-U codon (Table 2.1). The association rate of $ASL^{SufA6}$ for a slippery CCC-U codon is 0.0041 $\mu M^{-1}$ $min^{-1}$ as opposed to 0.0053 $\mu M^{-1}$ $min^{-1}$ for binding to the CCG codon in the presence of $m^1G37$ in

the ASL. In the absence of the $m^1G37$ modification, $ASL^{SufA6}$ has a 10-fold greater association rate for both the CCG and slippery CCC-U codons (0.053 $\mu M^{-1}$ $min^{-1}$ and 0.062 $\mu M^{-1}$ $min^{-1}$, respectively). The dissociation of $ASL^{SufA6}$ from a cognate CCG or a slippery CCC-U codon are all very similar regardless of the G37 modification status (0.013-0.019 $min^{-1}$). Together these data indicate that the inserted G37.5 nucleotide in $ASL^{SufA6}$ removes the dependency on the $m^1G37$ modification required for tight association to the ribosome for the parent $tRNA^{Pro}_{CGG}$. Further, the G37.5 nucleotide also prevents the ribosome from distinguishing between cognate and near-cognate, slippery codons as evidenced by the similar calculated $K_d$ values in the absence of the $m^1G37$ modification (0.26 and 0.31 μM, respectively).

***Engineering of the U32•A37.5 pairing in $ASL^{SufA6}$ allows for tight association to the A site.***

Our affinity assays show that frameshift suppressor $ASL^{SufA6}$ is unable to bind with high affinity to a cognate CCG codon despite containing the same GGC anticodon as $tRNA^{Pro}_{CGG}$ (Fig. 2.1). Therefore, the $m^1G37$ modification in $ASL^{SufA6}$ has a very different role in stabilizing the interactions between the anticodon and codon in contrast to $ASL^{Pro}_{CCG}$. Although $tRNA^{SufA6}$ undergoes +1 frameshifting, it does so with low efficiencies because of its poor association to the slippery CCC-U codon (Fig. 2.4*B*) (30). These data lead us to question whether it is the inserted anticodon loop nucleotide alone that causes reduced binding affinity as seen with other frameshift-suppressor ASL binding studies (40). In the case of both $ASL^{Pro}_{CCG}$ and $ASL^{SufA6}$ that bind poorly to the A site, 70S structures of these same tRNA-mRNA pairs bound have been solved (Fig. 2.6*A*-*C*)(31). In the case of $ASL^{Pro}_{CCG}$ $\Delta m^1G37$ bound to a cognate CCG codon, electron density is missing for

nucleotide U32 which is located on the opposite side of the anticodon loop from A38 (Fig. 2.6*B*). The destabilization of the ASL is likely due to the apparent flexibility of the 5' stem that, in turn, disrupts the conserved U32•A38 interaction located at the base of the RNA stem. The 32•38 disruption is noteworthy because the identity of these nucleotides is universally important in fine-tuning tRNA affinity and therefore translation fidelity (39,45). The same structural phenomenon is also observed in 70S structures containing $ASL^{SufA6}$; $ASL^{SufA6}$ binding to a slippery CCC-U codon results in local distortion of the 5' stem disrupting the U32•A38 pairing (Fig. 2.6*C*). In both cases, the tRNA-mRNA pair undergoes +1 frameshifting. Therefore, we postulated that the frameshift event was directly influenced by the destabilization of the 32•38 pairing after tRNA selection but before movement to the P site.

The G37.5 insertion in $tRNA^{SufA6}$ changes the potential base pairing interaction of U32•A38 to U32•G37.5 (Fig. 2.6*C*). In this context, the U32•G37.5 pairing should render the ribosome unable to distinguish a cognate from noncognate interaction as the 32-38 nucleotide identity is directly correlated to the anticodon sequence (39,45). Indeed, $ASL^{SufA6}$ binds to cognate CCG and near-cognate (*i.e.* slippery) CCC-U codon with similar affinities (calc. $K_d$ of 0.31 μM and 0.26 μM, respectively, in the absence of the $m^1G37$ modification; Table 2.1). We next tested whether changing G37.5 to A37.5 could restore high affinity A-site binding due to the possible formation of a new U32•A37.5 pair. We found that potentially restoring the U32•A37.5 base pair in does not restore high-affinity binding to a cognate CCG codon in the absence or presence of the $m^1G37$ modification (calc. $K_d$ of 7.4 μM and 15 μM, respectively; Fig. 2.6*D*,*E*, Table 2.1). Notably, in contrast to $ASL^{Pro}_{CCG}$, $ASL^{SufA6}$A37.5 displays similar association rates both in the presence (0.0057

$\mu M^{-1}min^{-1}$) or absence (0.0046 $\mu M^{-1}min^{-1}$) of $m^1G37$, but $k_{off}$ is reduced ~3-fold (0.09 $min^{-1}$ and 0.034 $min^{-1}$, respectively) (Fig. 2.6*D*,*E*).

In binding to the slippery CCC-U codon, $ASL^{SufA6}$ A37.5 has a ~two-fold higher affinity (calc. $K_d$ = 1.8 μM) than $ASL^{SufA6}$ containing G37.5 (calc. $K_d$ = 3.7 μM) in the presence of the $m^1G37$ modification (Table 2.1). Removal of $m^1G37$ results in $ASL^{SufA6}$ A37.5 binding with high affinity, similar to $ASL^{SufA6}$ G37.5 (calc. $K_d$ = 0.31 μM for $ASL^{SufA6}$ G37.5 and 0.45 μM for $ASL^{SufA6}$ A37.5). The $K_d$ for $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ binding to a slippery CCC-U codon is comparable to modified $ASL^{Pro}_{CCG}$ binding to a cognate CCG codon (calc. $K_d$ = 0.42 μM). For the $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ binding to a slippery CCC-U codon, both the $k_{on}$ and $k_{off}$ rates are higher than those of other ASLs ($k_{on}$ = 0.146 $\mu M^{-1}min^{-1}$, $k_{off}$ = 0.06 $min^{-1}$), implying the recognition mechanism is altered. The increase in affinity implies that in the case of $ASL^{SufA6}$ with the restored U32-A37.5, the lack of $m^1G37$ enables high affinity binding and recognition of the ASL when the slippery CCC-U codon is presented in the A site. Most importantly, we demonstrate that by changing the identity of the base insertion and controlling the modification at position 37 in the anticodon stem loop, we can tune the affinity of the ASL to the ribosomal A site. This has significant implications for understanding how the ribosome interacts with rationally engineered tRNAs.

***Engineering the 32•38 pairing in $ASL^{SufA6}$*** to ***U32•A37.5 reorders 5' stem of the anticodon loop.***

To determine if the engineered $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ does indeed reorders the ASL as suggested from the binding kinetics, we solved a 3.6 Å resolution X-ray crystal structure of $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ bound to the *Thermus thermophilus* 70S ribosome (Fig. 2.6*F*). The ASL and mRNA density are well-ordered and unambiguously

demonstrate a change in the anticodon loop (Fig. 2.6*F*) as compared to other $ASL^{SufA6}$ structures bound to the ribosome (Fig. 2.6*C*)(31). The $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ has good density for U32 in contrast to the previous structures that showed distortion of the 5' stem of the ASL (Fig. 2.6*B,C*). The phosphate backbone of nucleotide A37.5 shifts by 2.8 Å as compared to wild-type $ASL^{Pro}_{CCG}$ bound to its cognate CCG codon, and by 6.2 Å when compared to $ASL^{SufA6}$ with the G37.5 (Fig. 2.6*F, 2.7*). This movement places A37.5 across from U32 allowing the possible formation of a single hydrogen bond similar to the 32•38 orientation observed in other tRNAs (Fig. 2.6*A,* 2.8) (31,46,47). Overall, the A37.5 insertion in $ASL^{SufA6}$ seems to orient the ASL to a conformation more similar to that of $ASL^{Pro}_{CCG}$ than that of wild-type $ASL^{SufA6}$ (Fig. 2.7). 16S rRNA nucleotides A1492 and A1493 flip from their internal position in helix 44 and G530 is positioned close to A1492 demonstrating recognition by the ribosome (Fig. 2.8).

**2.4. Discussion**

Modification of tRNAs adds an important layer of regulation during translation. These modifications are so functionally important that more genes are devoted to tRNA modification pathways than to the expression of tRNAs themselves (48). Modifications in the ASLs of tRNAs are critical given that only 7 of the 61 sense codons are decoded by tRNAs that lack modifications at nucleotide 34 or 37 in *E. coli* (49,50). In this work, we determine that the $m^1G37$ modification in $ASL^{Pro}_{CGG}$ is required for high affinity binding to a cognate CCG codon in the decoding center (Fig. 2.2). The absence of the modification results in low affinity binding and specifically, the association ($k_{on}$) is reduced almost 3-fold while $k_{off}$ is unaffected (Fig. 2.2). These results indicate that the $m^1G37$ modification in $tRNA^{Pro}_{CGG}$ provides stability in association with the decoding center rather than causing

A-site drop-off. Consistent with these data are our previous structural studies that showed destabilization of the anticodon loop when $ASL^{Pro}_{CGG}$ lacks the $m^1G37$ modification and interacts with a cognate CCG codon (31) (Fig. 2.6).

Both $tRNA^{Pro}_{CGG}$ and $tRNA^{Pro}_{GGG}$ isoacceptors lacking the $m^1G37$ modification undergo +1 frameshifting on CCC-N codons (26,29,51). Although $ASL^{Pro}_{CGG}$ containing the $m^1G37$ modification significantly impairs binding to a slippery CCC-U codon (3.62 μM; Fig. 2.4*A*), its removal causes a ~4-fold enhancement in $ASL^{Pro}_{CGG}$ association with the slippery codon (Fig. 2.4*B*). This association results in tighter binding (calc. $K_d$ of 0.41 μM) that is comparable to binding of wild-type $ASL^{Pro}_{CGG}$ to a cognate CCG codon (0.42 μM). These data suggest that the additional stability that the $m^1G37$ modification imparts in binding to a cognate CCG codon is lost in the context of a non-Watson-Crick C34-C4 pair at the third or wobble position. Collectively, these results show the $m^1G37$ modification in $ASL^{Pro}_{CGG}$ stabilizes high-affinity interactions in the cognate case but prevents recognition of slippery codons that would result in +1 frameshifting.

tRNAs containing expanded anticodon stem loops can cause ‘slipping’ on mRNA codons resulting in frameshifts (37). An extra nucleotide insertion in the anticodon loop of $tRNA^{Pro}_{CGG}$ was identified in a frameshift suppressor tRNA (named $tRNA^{SufA6}$) that reverts a +1 frameshift. Primer extension analyses revealed that $tRNA^{SufA6}$ was also modified at nucleotide 37 similar to all three $tRNA^{Pro}$ isoacceptors (30,52) but the extent of modification was not determined. It is unclear what role, if any, the $m^1G37$ modification affects $tRNA^{SufA6}$-mediated +1 frameshifting. We find that the presence of the $m^1G37$ modification renders $ASL^{SufA6}$ unable to bind to both cognate CCG and slippery CCC-U codons (Fig. 2.5*B*). In contrast, $ASL^{SufA6}$ lacking $m^1G37$ binds with high affinity to both a

cognate CCG or a slippery CCC-U codon. These data support the notion that the $m^1G37$ modification and the inserted G37.5 nucleotide likely stabilize the anticodon loop in similar ways. In support of the functional similarities of $m^1G37$ and G37.5, structures of $tRNA^{Pro}_{CGG}$ $\Delta m^1G37$ or $tRNA^{SufA6}$ decoding codons that allow for +1 frameshifting reveal structural similarities. The 3' stem of the ASL, in particular nucleotides 30-32, on the opposite side of G37.5/G37 are conformationally dynamic in both structures strongly suggesting that +1 frameshifts induced by these two tRNAs occur by a similar mechanism (Fig. 2.6*B,C*) (31).

In $tRNA^{Pro}_{CGG}$, nucleotide U32 normally forms a single hydrogen bond with A38 and thus is not a Watson-Crick base pair (Fig. 2.6*A*, 2.8). The nucleotide identity of the 32•38 pairing in all tRNAs is inversely correlated to the strength of the codon-anticodon interaction (39,41,45). For example, the anticodon of *E. coli* $tRNA^{Ala}_{GGC}$ is considered strong because of the three GC pairs between the codon and the anticodon. Therefore, in this strong case, the 32•38 pairing needs to be correspondingly weak to counterbalance the strength of the codon-anticodon. Changing the 32•38 pairing in $tRNA^{Ala}_{GGC}$ from a weak, conserved U32•A38 pair to a strong pair such as C32•A38, prevents the ribosome from being able to distinguishing correct from incorrect tRNA-mRNA pairs (41,53). In the context of $tRNA^{SufA6}$, the inserted G37.5 displaces A38 preventing a U32•A38 pairing (31). Binding of $ASL^{SufA6}$ to a cognate CCG or a +1 slippery CCC-U codon is extremely weak as indicated by both the $k_{on}$ and $k_{off}$ rates (Fig. 2.5). We attempted to restore the wild-type U32•A38 found in $tRNA^{Pro}_{CGG}$ by changing G37.5 in $tRNA^{SufA6}$ to an adenosine. $ASL^{SufA6}$ A37.5 binds poorly to a CCG codon regardless of the $m^1G37$ modification status (Fig. 2.5). Interestingly, the A37.5 mutant bound tightly to the slippery CCC-U codon but only in the absence of the $m^1G37$ modification, similar to $ASL^{SufA6}$. An X-ray crystal structure

of the ribosome with an A-site $ASL^{SufA6}$ A37.5 (lacking the $m^1G37$ modification) bound to a slippery CCC-U codon reveals a reordering of the 3' stem such that U32 regains rigidity as assessed by its electron density (Fig. 2.6*F*). We predict this engineered $tRNA^{SufA6}$ does not undergo +1 frameshifting because of ordering of the ASL but further studies are required to test this.

The studies here demonstrate that the $m^1G37$ modification of $tRNA^{Pro}$ influences recognition of both cognate and near-cognate, slippery codons. $tRNA^{Pro}_{CGG}$ lacking the $m^1G37$ modification undergoes +1 frameshifting but our previous structures, along with other structures of extended ASLs that frameshift, show that the shift into the new frame does not occur in the decoding center (31,54,55). At what stage of elongation does $tRNA^{Pro}_{CGG}$ lacking the $m^1G37$ modification cause a +1 frameshift? Kinetic analyses of $tRNA^{Pro}_{GGG}$ movement through the ribosome reveal the shift can occur at two distinct stages: a fast mechanism during translocation of the tRNA-mRNA pairs to the P site and a slower mechanism when $tRNA^{Pro}_{GGG}$ is stalled in the P site while waiting for A-site tRNA delivery (28). A recent structure of $ASL^{SufA6}$ bound to a +1 codon in the P site demonstrates the ASL *alone* is sufficient for the +1 mRNA frameshift consistent with the slow mechanism presented above (32). Another possibility is what is also observed in the structure of $ASL^{Pro}_{CGG}$ lacking the $m^1G37$ modification in the A site (31). The 3' anticodon stem is destabilized which, in turn, may influence how elongation factor G (EF-G) recognizes the tRNA in the A site to initiate translocation. This second possibility is consistent with the fast mechanism as observed in the kinetic analyses.

Although the data presented focus on the impact of the $m^1G37$ modification on decoding of single proline codons, poly-proline sequences in protein-coding genes are

prone to +1 frameshifts (56-60). The unique nature of proline where it is both a poor donor and acceptor during the peptidyl-transferase reaction results in a slow rate of peptide bond formation. Therefore, consecutive prolines cause ribosome stalling (60,61). Additionally, the nucleotide repeats in the proline codons presents the same codon-anticodon interactions regardless if the tRNA binds in the zero or +1 frame (62). These events can collectively lead to the shifty nature of $tRNA^{Pro}_{CGG}$ but are counterbalanced by the action of elongation factor P (EF-P) that helps stabilize peptidyl-$tRNA^{Pro}$ located in the P site of the ribosome (28,60,61,63). EF-P binds in the exit (E) site of the ribosome on both the small and large ribosomal subunits and abuts against P-site peptidyl-$tRNA^{Pro}$ (63). A modified Lys residue of EF-P protrudes into the 50S P site and stabilizes peptidyl-$tRNA^{Pro}$ to help resume protein synthesis stalled at a stretch of poly-prolines (60,64,65). Cryo-EM structures of ribosomes bound to peptidyl-$tRNA^{Pro}$ reveal the flexibility of the peptidyl-$tRNA^{Pro}$ in the absence of EF-P. EF-P binding orders peptidyl-$tRNA^{Pro}$ to facilitate efficient peptide bond formation of the cyclic proline moiety (63). In addition to this function, EF-P also can suppress +1 frameshifts suggesting an previously unappreciated role in helping to maintain the mRNA frame (28). Conserved EF-P residues that interact with anticodon stem nucleotides 41 and 42 are essential for function (63). In the absence of the $m^1G37$ modification, the interaction between EF-P and $tRNA^{Pro}$ may be destabilized due to the flexibility of the 3' stem (31). The interplay between the $m^1G37$ modification in $tRNA^{Pro}$ and EF-P suggests that this elongation factor has synergistic roles in translational fidelity dependent on tRNA metabolism.

## 2.5. Experimental procedures

### *Ribosome purification.*

*E. coli* 70S ribosomes were purified as previously described (66). *E. coli* MRE600 cells were grown to an $OD_{600}$ ~0.7 in Luria Broth (LB) media at 37 °C then cooled on ice for 20 min. All centrifugation steps were performed at 4°C. Cells were pelleted by centrifugation and washed with buffer 1 (10 mM HEPES-KOH, pH 7.6, 10 mM $MgCl_2$, 1 M $NH_4Cl$, 6 mM β-mercaptoethanol (β-Me)) twice then resuspended in buffer 2 (10 mM HEPES-KOH, pH 7.6, 10 mM $MgCl_2$, 100 mM $NH_4Cl$, 6 mM β-Me). Cells were lysed using an EmulsiFlex-C5 high pressure homogenizer (Avestin) and cell debris were removed by centrifuging at 13,000 *x g* for 15 min. The lysate was further centrifuged at 27,000 *x g* for 30 min to obtain the S30 fraction. Ribosomes were pelleted by centrifuging at 42,000 *x g* for 17 hr. The pellets were resuspended in Buffer 2 and ribosomes were further purified over a 10-40% sucrose gradient in Buffer 2 at 70,000 *x g* for 12 hrs. 70S ribosomes were separated from polysomes and subunits using a Brandel gradient fractionator. The 70S fractions were pooled, pelleted, resuspended in buffer 2 and stored at -80°C.

### *70S complex formation.*

ASLs and mRNAs were chemically synthesized (Integrated DNA Technologies) and purified *E. coli* $tRNA^{fMet}$ was purchased from Chemical Block (Table 2.3). mRNAs contain either an CCG or CCC-U in the A site. The *E. coli* 70S ribosome complex was formed by incubation with two-fold molar excess of mRNA for 5 min followed by two-fold molar excess of $tRNA^{fMet}$ for 30 min at 37°C. A-site ASLs were 5'-labeled with [γ-$^{32}$P]-ATP (PerkinElmer) using T4 PNK enzyme (NEB).

***Kinetic binding assays.***

A modified 96-well Bio-rad dot-blot apparatus with two membranes was used to study binding kinetics of tRNAs to ribosomes (67). An upper nitrocellulose membrane and a lower nylon membrane (Amersham Hybond-N+, GE Healthcare) were pre-equilibrated in cold buffer 3 (5 mM HEPES-KOH, pH 7.5, 50 mM KCl, 10 mM $NH_4Cl$, 10 mM $Mg(CH_3COO)_2$, 6 mM β-Me). All binding experiments were performed in buffer G. Non-specific binding was controlled for by having a [$^{32}$P]-ASL only sample with each experiment. Reactions were filtered by vacuum and immediately washed using 100 µl cold Buffer 3. After incubations, membranes were dried and exposed to a phosphorimager screen (GE Healthcare) and imaged on a Typhoon FLA 7000. Quantification was performed using ImageQuantTL software and analyzed using GraphPad Prism. The fraction of A-site ASL bound was calculated as the ratio between nitrocellulose counts and the total counts on both membranes after correcting for nonspecific binding.

***Measuring dissociation constants ($K_d$).***

Serial dilutions of the ribosome complex (70S, mRNA, P-site $tRNA^{fMet}$) were performed to generate a series of 70S concentrations ranging from 0.98 nM to 1 µM. [$^{32}$P]-ASL was added and incubated for 3 hr at 25°C. Reaction volumes of 10 µl were applied to the filters, washed and then quantified as described above. The ASL fraction bound was fit using a one site specific binding equation in GraphPad Prism as previously described (Fraction bound=$\frac{B_{max}.[70S]}{K_d+[70S]}$) (39,68).

***Measuring association rates ($k_{on}$).***

Association to the 70S A site was measured as previously described (42,43). Briefly, 15 µl of 4.5 nM $^{32}$P-labeled ASL was added to 15 µl of increasing concentrations of 70S complex programmed with mRNA and tRNA$^{fMet}$. Three µl aliquots were removed at different times (0.5, 1, 2, 3, 4, 5, 6, 7 min) immediately filtered and washed with 100 µl of cold buffer 3. Initial association rates were obtained using different concentrations of the 70S complex (final concentrations of 12.5, 25, 50, 100, and 200 nM). $k_{on}$ was derived as the slope of the linear regression performed on the initial rates vs. [70S] plot.

***Measuring dissociation constants ($k_{off}$).***

Dissociation of the ASL from the 70S A site was initiated by a 1:100 dilution of the equilibrium binding reaction (1 µM 70S, 3 µM mRNA, 5 µM tRNA$^{fMet}$, 0.05 µM $^{32}$P-ASL) in buffer 3 containing 0.3 µM unlabeled ASL. At 5 min intervals, 10 µl aliquots of the reaction were removed, filtered and washed. The ASL fraction bound was normalized to t = 0. The natural log of the normalized fraction bound was fitted with a linear regression vs. time, and $k_{off}$ was derived as the negative of the slope.

***Crystallization, X-ray data collection, and structural determination.***

Purification of *Thermus thermophilus* 70S ribosomes, formation of complexes with mRNA and tRNAs, and initial screening conditions followed previously established protocols (46,68). Two µl of the ribosome complex (4.4 µM 70S, 8.8 µM mRNA, 11 µM tRNA$^{fMet}$, 22 µM ASL$^{SufA6}$, 11 µM CC-puromycin (Dharmacon) and 2.8 µM deoxy BigCHAP (Hampton Research) was mixed with 2.4 µl of reservoir condition (0.1 M Tris-HOAc, pH 7.0, 0.2 M KSCN, 4.5-5.5% (w/v) PEG 20K, 4.5-5.5% (w/v) PEG 550 MME, 10 mM $Mg(OAc)_2$). Crystals were grown by sitting drop at 20°C in two weeks. Crystals were cryoprotected using increasing amounts of PEG 550 MME to final concentration of 35%, with the final

solution containing 22 μM $ASL^{SufA6}$ and 11 μM CC-puromycin. The crystals were screened at the SER-CAT 22-ID and NE-CAT 24ID-C/E beamlines and datasets were collected at the SER-CAT 22-ID beamline, all at the Advanced Photon Source, Argonne National Laboratory (Table 2.2). Datasets were integrated and scaled using XDS (69) and the structure was solved by molecular replacement using 70S coordinates from PDB code 4Y4O. Crystallographic refinements were performed with PHENIX (70) followed by manual model building in Coot (71). Figures were generated in PyMOL (72).

**Data availability:** The atomic coordinates and structure factors (code 6NDK) have been deposited in the Protein Data Bank (http://wwpdb.org/).

## 2.6. Figures and tables

### FIGURES

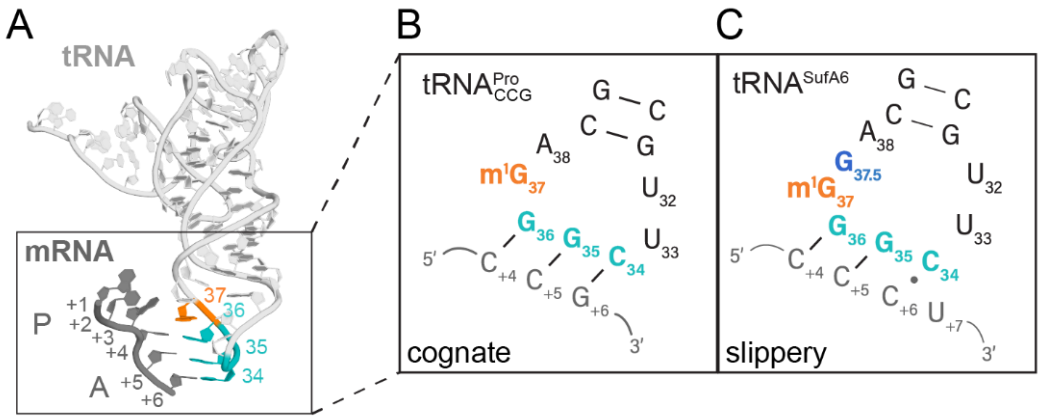


**Figure 2.1. Frameshift suppressor tRNA$^{SufA6}$ is a derivative of tRNA$^{Pro}_{CCG}$.**

*A,* Tertiary structure of tRNA$^{Pro}$ with its anticodon depicted in cyan, anticodon loop nucleotide 37 in orange, the mRNA in gray with the first peptidyl(P)-site nucleotide denoted as +1, P-site codon nucleotides listed as +1, +2 and +3, and aminoacyl(A)-site codon nucleotides listed as +3-+4, and +5. *B,* Secondary structure of the anticodon stem loop of tRNA$^{Pro}$ bound to a cognate CCG codon containing three Watson-Crick base pairs. *C,* Secondary structure of the anticodon stem loop of tRNA$^{SufA6}$ bound to a four-nucleotide slippery CCC-U codon and containing a C6 - C34 wobble interaction. tRNA$^{SufA6}$ contains an extra guanosine between positions 37 and 38 (G37.5, blue).

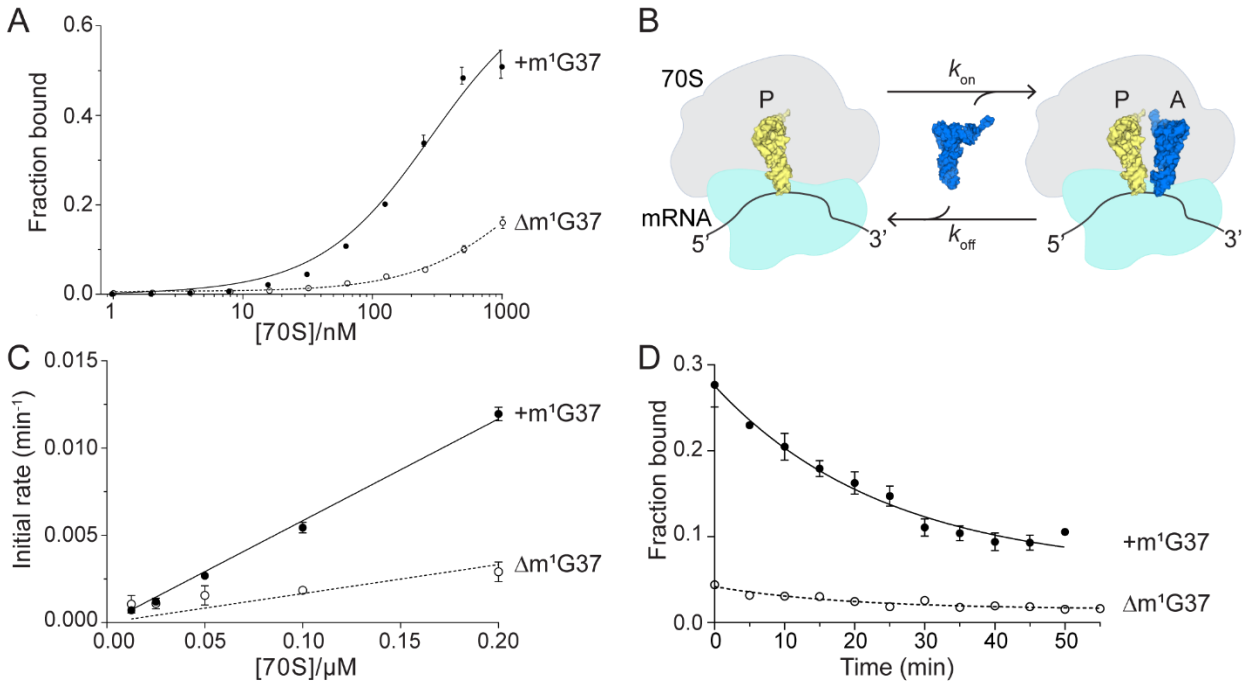


**Figure 2.2. The m$^1$G37 modification in tRNA$^{Pro}_{CCG}$ is important for binding to a cognate CCG codon.**

*A*, Equilibrium binding of ASL$^{Pro}_{CGG}$ with (+m$^1$G37) or without (Δm$^1$G37) the modification to a programmed 70S containing an A-site CCG codon. *B*, Schematic of the association and dissociation of tRNA from the ribosomal A site. *C*, Association ($k_{on}$) and *D*, dissociation ($k_{off}$) rates of ASL$^{Pro}_{CCG}$ with (+m$^1$G37) or without (Δm$^1$G37) the modification to the 70S containing a cognate CCG codon in the A site. Data is the mean ± s.e. of at least five independent experiments.

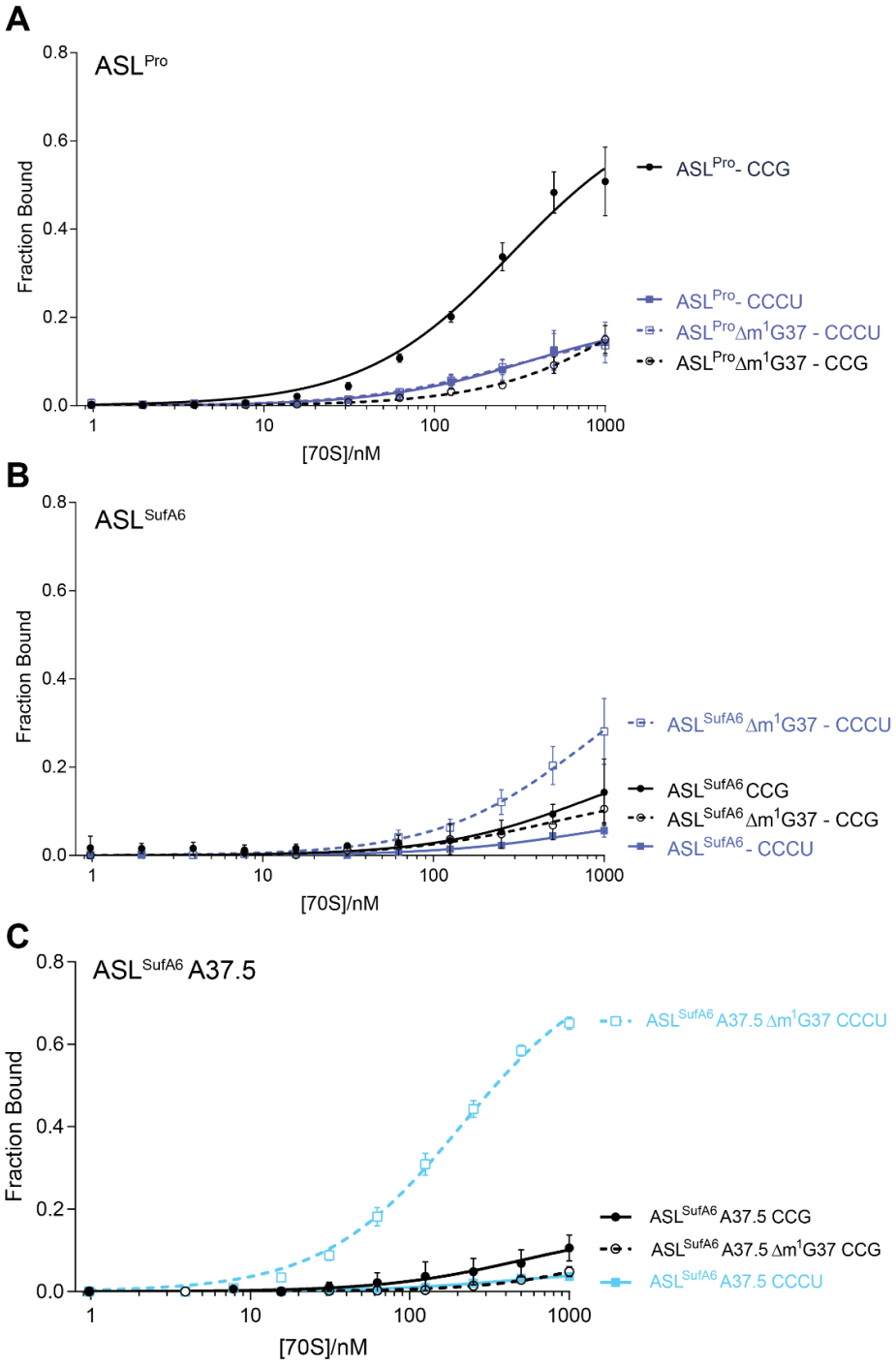


**Figure 2.3. Equilibrium binding curves $ASL^{Pro}$, $ASL^{SufA6}$, and $ASL^{SufA6}$ A37.5.**

Binding constants were obtained by fitting the data from 6 replicates to a one site specific binding model using GraphPad Prism.

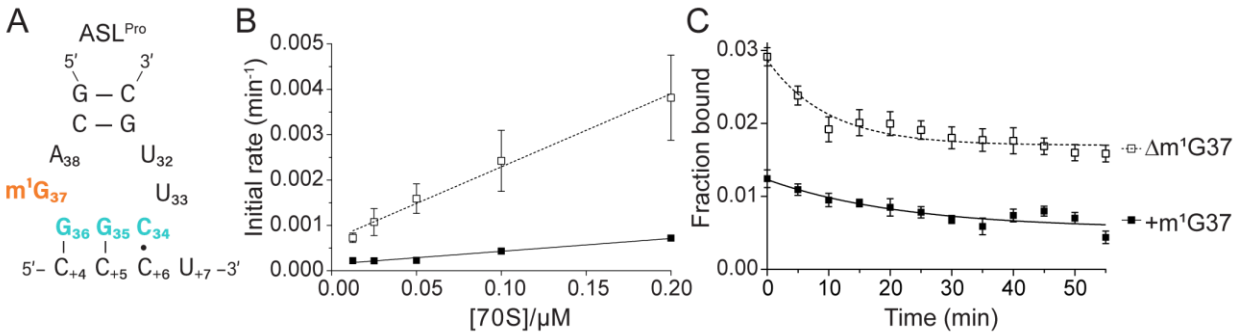


**Figure 2.4. The m$^1$G37 modification prevents tRNA$^{Pro}_{CCG}$ binding to the slippery CCC-U codon.**

*A,* Secondary structure of ASL$^{Pro}_{CCG}$ (same coloring scheme as in Fig. 1) shown bound to its slippery CCC-U codon. *B*, Association ($k_{on}$) and *C*, dissociation ($k_{off}$) rates of ASL$^{Pro}_{CCG}$ with (+m$^1$G37) or without (Δm$^1$G37) the modification to a slippery CCC-U codon in the ribosomal A site. Data is the mean ± s.e. of at least five independent experiments.

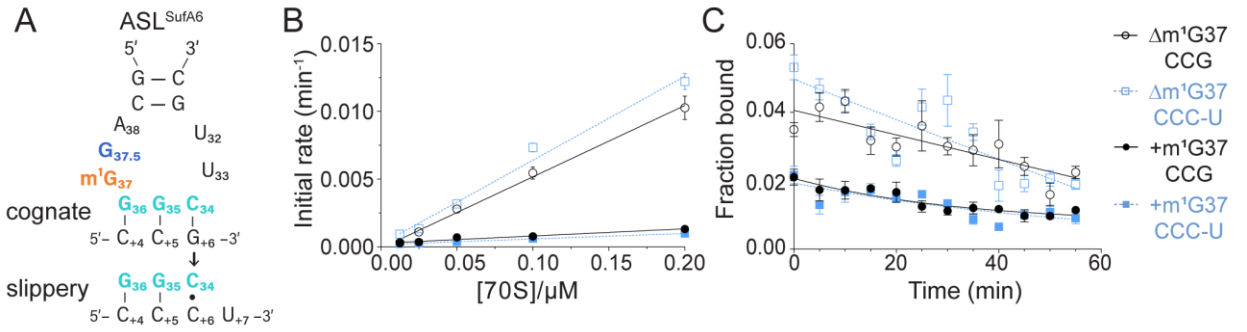


**Figure 2.5. The m$^1$G37 modification in ASL$^{SufA6}$ impairs binding to a cognate CCG or slippery CCC-U codon.**

*A,* Secondary structure of ASL$^{SufA6}$ (same coloring scheme as in Fig. 1) with either a cognate or slippery codon- anticodon interaction. *B*, Association ($k_{on}$) and *C*, dissociation ($k_{off}$) rates of ASL$^{SufA6}$ with (+m$^1$G37) or without (Δm$^1$G37) the modification bound to either the cognate CCG (black) or slippery CCC-U (blue) codon in the ribosomal A site. Data is mean ± s.e. of at least five independent experiments.

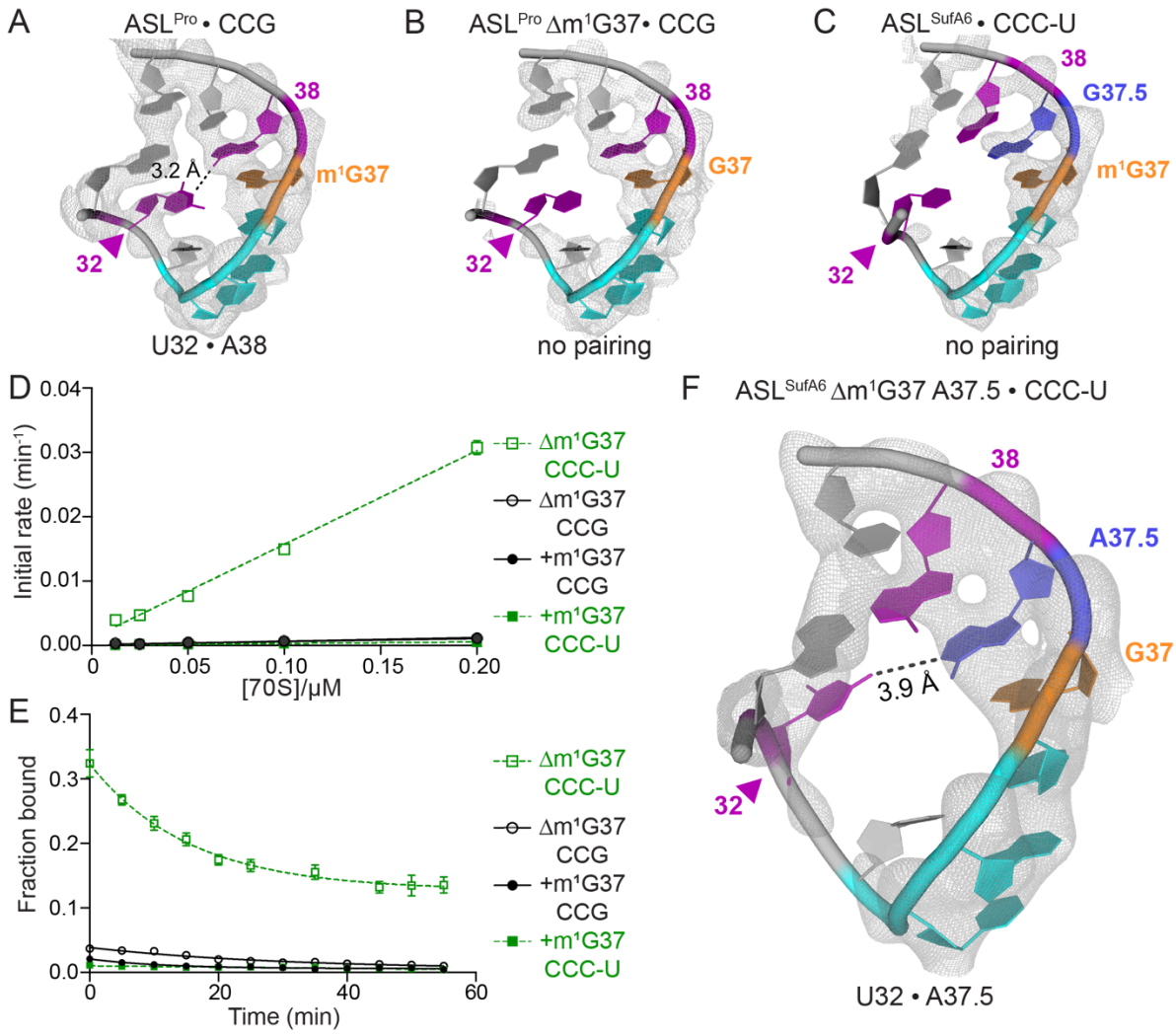


**Figure 2.6. Reordering of the 32-38 pairing allows for high affinity binding of $ASL^{SufA6}$ to a slippery CCC-U codon.**

*A,* $2F_o$-$F_c$ electron density maps from a structure containing the 70S ribosome with $ASL^{Pro}_{CCG}$ decoding a cognate CCG codon in the A site (PDB code 4LSK; color scheme is the same as Fig. 1); *B*, 70S ribosome with $ASL^{Pro}_{CCG}$ $\Delta m^1G37$ decoding a cognate CCG codon (PDB code 4P70); and *C*, 70S ribosome with $ASL^{SufA6}$ decoding a slippery CCC-U codon (PDB code 4L47). These structures demonstrate that the $m^1G37$ modification stabilizes the U32•A38 interaction in $ASL^{Pro}_{CCG}$ on a cognate CCG codon (*A*) whereas the lack of the $m^1G37$ modification results in disorder of the 3' region of $ASL^{Pro}_{CCG}$ (*B*). A similar

disordering is seen when $ASL^{SufA6}$ containing an inserted nucleotide in its anticodon loop (G37.5) decodes a slippery CCC-U codon (*C*). *D,* Association ($k_{on}$) and *E*, dissociation ($k_{off}$) rates of $ASL^{SufA6}$ with a mutated A37.5 with ($+m^1G37$) or without ($\Delta m^1G37$) the modification bound to either the cognate CCG (black) or slippery CCC-U (green) codon. Data is the mean ± s.e. of at least five independent experiments. *F*, $2F_o$-$F_c$ electron density maps from a structure containing the 70S ribosome with $ASL^{SufA6}$ A37.5 bound to an A-site slippery CCC-U codon. Mutation of G37.5 to A37.5 reorders the 3' stem of the ASL, specifically nucleotides 31 and 32. $2F_o$-$F_c$ electron density maps are contoured at 1.5σ.

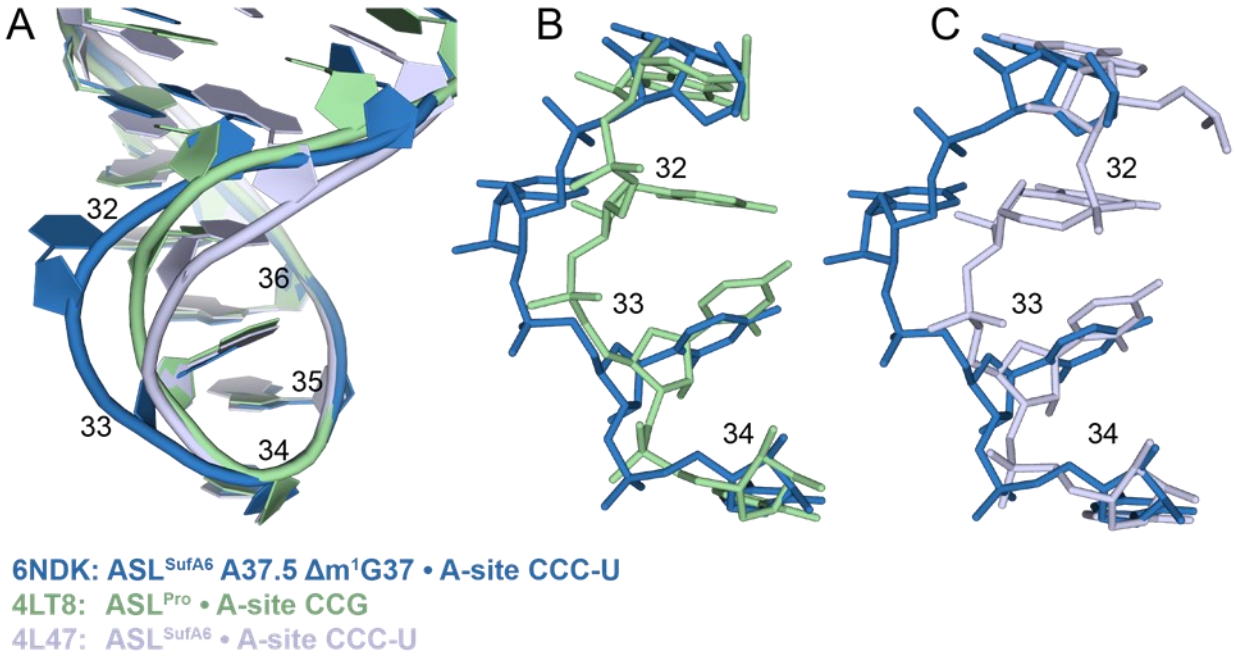


**Figure 2.7. The phosphate backbone of the ASL$^{SufA6}$ is widened in the case of the A37.5 insertion.**

*A*, Overlay of ASL$^{SufA6}$ A37.5 Δm$^{1}$G37 bound to CCC-U codon (blue, PDB code 6NDK(97)), ASL$^{Pro}$ bound to CCG (green, PDB code 4LT8, and ASL$^{SufA6}$ bound to CCC-U (gray, PDB code 4L47) showing the reordering of the ASL on the opposite side of the insertion.

*B-C,* ASL$^{SufA6}$ A37.5 Δm$^{1}$G37 bound to A-site CCC-U (blue) codon adopts a conformation more similar to ASL$^{Pro}$ on CCG (green) than that of ASL$^{SufA6}$ bound to CCC-U (gray).

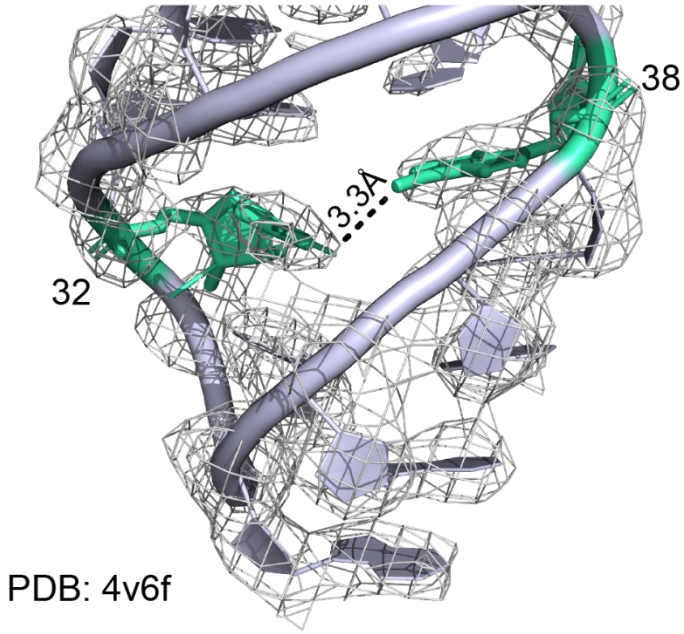


**Figure 2.8. The 32-38 pairing in $tRNA^{Phe}$ shows a conserved hydrogen bond (PDB code 4V6F) similar to that of modified $tRNA^{Pro}$ binding to a cognate codon.**

$2F_o$-$F_c$ electron density map contoured at 1.5σ is shown in gray.

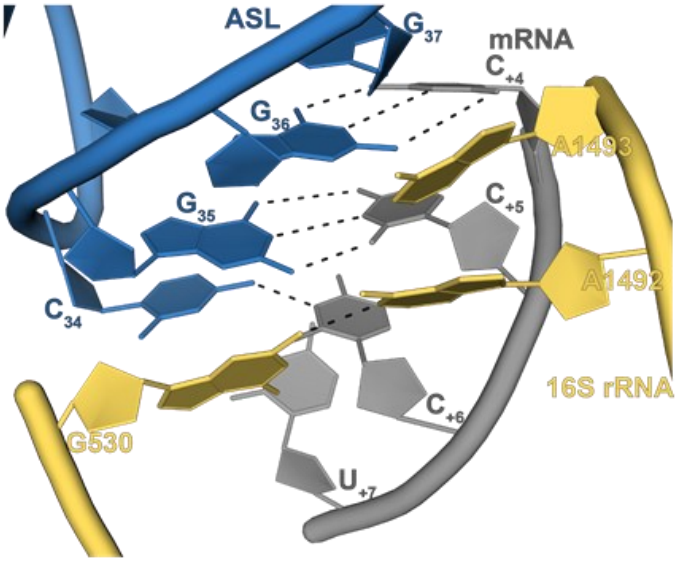


**Figure 2.9. The ASL$^{SufA6}$ A37.5 is accepted by the ribosome.**

16S rRNA nucleotides A1492 and A1493 flip from the internal helix 44 and G530 is positioned close to A1492 demonstrating recognition by the ribosome. Paromomycin was not added to the crystallization complex formation.

## TABLES

**Table 2.1. $k_{on}$, $k_{off}$, and calculated $K_d$ values (best fit ± SE) for $ASL^{Pro}_{CCG}$, $ASL^{SufA6}$, and $ASL^{SufA6}$ A37.5.**

Data from at least five replicates were fit in GraphPad Prism. Calculated $K_d = k_{off}/k_{on}$.

| | m$^1$G37 | A-site codon | $k_{on}$ (μM$^{-1}$ min$^{-1}$) | $k_{off}$ (min$^{-1}$) | calc. $K_d$ (μM) | Fold change in $K_d$ |
|---|---|---|---|---|---|---|
| $ASL^{Pro}_{CCG}$ | + | CCG | 0.031 ± 0.004 | 0.013 ± 0.002 | 0.42 | 3 |
| | - | CCG | 0.012 ± 0.002 | 0.017 ± 0.002 | 1.4 | |
| | + | CCC-U | 0.0039 ± 0.0005 | 0.014 ± 0.003 | 3.62 | 9 |
| | - | CCC-U | 0.017 ± 0.003 | 0.009 ± 0.001 | 0.41 | |
| $ASL^{SufA6}$ | + | CCG | 0.0053 ± 0.0007 | 0.013 ± 0.002 | 2.5 | 10 |
| | - | CCG | 0.053 ± 0.002 | 0.014 ± 0.003 | 0.26 | |
| | + | CCC-U | 0.0041 ± 0.0001 | 0.015 ± 0.004 | 3.7 | 12 |
| | - | CCC-U | 0.062 ± 0.004 | 0.019 ± 0.004 | 0.31 | |
| $ASL^{SufA6}$ A37.5 | + | CCG | 0.0057 ± 0.0004 | 0.09 ± 0.01 | 15 | 2 |
| | - | CCG | 0.0046 ± 0.0004 | 0.034 ± 0.007 | 7.4 | |
| | + | CCC-U | 0.003 ± 0.001 | 0.005 ± 0.018 | 1.8 | 4 |
| | - | CCC-U | 0.146 ± 0.006 | 0.06 ± 0.01 | 0.45 | |

**Table 2.2. Data collection and refinement statistics**

| | $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ |
|---|---|
| **Data collection** | |
| Space group | $P2_12_12_1$ |
| Wavelength (Å) | 1.00000 |
| Cell dimensions | |
| a, b, c (Å) | 208.91, 445.91, 617.31 |
| α, β, γ (°) | 90, 90, 90 |
| Resolution (Å) | 49.17 - 3.64 (3.77 - 3.64) |
| $R_{pim}$ (%) | 11.5 (62.8) |
| I/σI | 4.77 (1.14) |
| Completeness (%) | 93.31 (87.06) |
| Redundancy | 5.0 (4.2) |
| CC1/2 | 0.986 (0.395) |
| **Refinement** | |
| Reflections | 596,260 (55,238) |
| $R_{work}/R_{free}$ (%) | 21.5/26.2 |
| No. atoms | 291,830 |
| B-factors (Å$^2$) | |
| Overall | 150.84 |
| Macromolecule | 151.28 |
| Ligand/ion | 94.30 |
| R.m.s deviations | |
| Bond lengths (Å) | 0.006 |
| Bond angles (°) | 0.89 |

*Highest resolution shell is shown in parentheses.

**Table 2.3. Sequences of RNA used in this study.**

Insertions are bolded, the anticodon is underlined, and the N1 methylation at position 37 is noted as $m^1G$ for ASLs.

| | |
|---|---|
| $ASL^{SufA6}$ | GCU CGU U<u>CG G</u>($m^1G$)**G** ACG AGC |
| $ASL^{SufA6}$ $\Delta m^1G37$ | GCU CGU U<u>CG G</u>G**G** ACG AGC |
| $ASL^{SufA6}$ A37.5 | GCU CGU U<u>CG G</u>($m^1G$)**A** ACG AGC |
| $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ | GCU CGU U<u>CG G</u>G**A** ACG AGC |
| $ASL^{Pro}$ | GCU CGU U<u>CG G</u>($m^1G$)A CGA GC |
| $ASL^{Pro}$ $\Delta m^1G37$ | GCU CGU U<u>CG G</u>GA CGA GC |
| mRNA_Asite_CCG | GGC AAG GAG GUA AAA AUG CCG UAC CA |
| mRNA_Asite_CCCU | GGC AAG GAG GUA AAA AUG CCCU ACCA |

**Table 2.4. $K_d$ and $B_{max}$ values (best fit ± SE) for $ASL^{Pro}$, $ASL^{SufA6}$, and $ASL^{SufA6}$ A37.5 as determined by filter binding experiments at equilibrium.**

Binding constants were obtained by fitting the data from 6 replicates to a one site specific binding model using GraphPad Prism.

| | mRNA | | | |
|---|---|---|---|---|
| | CCG | | CCCU | |
| | $K_d$ (nM) | $B_{max}$ | $K_d$ (nM) | $B_{max}$ |
| $ASL^{Pro}$ | 284 ± 27 | 0.69 ± 0.03 | 351 ± 79 | 0.20 ± 0.02 |
| $ASL^{Pro}$ $\Delta m^1G37$ | 1795 ± 475 | 0.42 ± 0.08 | 274 ± 46 | 0.18 ± 0.01 |
| $ASL^{SufA6}$ | 781 ± 333 | 0.25 ± 0.06 | 767 ± 173 | 0.10 ± 0.01 |
| $ASL^{SufA6}$ $\Delta m^1G37$ | 497 ± 187 | 0.15 ± 0.03 | 791 ± 180 | 0.51 ± 0.06 |
| $ASL^{SufA6}$ A37.5 $\Delta m^1G37$ | 4558 ± 4067 | 0.27 ± 0.21 | 212 ± 9 | 0.81 ± 0.01 |
| $ASL^{SufA6}$ A37.5 | 1084 ± 188 | 0.19 ± 0.02 | 444 ± 98 | 0.055 ± 0.006 |

## 2.7. Acknowledgments

We thank Dunham laboratory members for advice and encouragement, Dr. Kurt Fredrick for advice on kinetic binding assays, Stacey Miles for technical assistance, Dr. Tatsuya Maehigashi for crystallization advice, and Dr. Graeme Conn for comments on the manuscript. Crystals were screened and X-ray crystallography data sets were collected at the NE-CAT beamlines (funded by NIGMS, National Institutes of Health , Grant P30 GM124165 ), using a Pilatus detector (RR029205) and an Eiger detector (OD021527), and at the SER-CAT beamlines (funded by its member institutions and National Institutes of Health Equipment Grants RR25528 and RR028976 ). This research used resources of the APS, a United States Department of Energy Office of Science User Facility operated by Argonne National Laboratory under Contracts DE-AC02-06CH11357 (NE-CAT) and W-31-109-Eng-38 (SER-CAT).

## 2.8. References

1. Schmeing, T. M., and Ramakrishnan, V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**, 1234-1242
2. Dunkle, J. A., and Dunham, C. M. (2015) Mechanisms of mRNA frame maintenance and its subversion during translation of the genetic code. *Biochimie* **114**, 90-96
3. Agris, P. F., Eruysal, E. R., Narendran, A., Vare, V. Y. P., Vangaveti, S., and Ranganathan, S. V. (2018) Celebrating wobble decoding: Half a century and still much is new. *RNA biology* **15**, 537-553
4. Yarian, C., Townsend, H., Czestkowski, W., Sochacka, E., Malkiewicz, A. J., Guenther, R., Miskiewicz, A., and Agris, P. F. (2002) Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem* **277**, 16391-16395
5. Kruger, M. K., Pedersen, S., Hagervall, T. G., and Sorensen, M. A. (1998) The modification of the wobble base of tRNAGlu modulates the translation rate of glutamic acid codons in vivo. *Journal of molecular biology* **284**, 621-631

6. Urbonavicius, J., Qian, Q., Durand, J. M., Hagervall, T. G., and Bjork, G. R. (2001) Improvement of reading frame maintenance is a common function for several tRNA modifications. *Embo J* **20**, 4863-4873

7. Yarian, C., Marszalek, M., Sochacka, E., Malkiewicz, A., Guenther, R., Miskiewicz, A., and Agris, P. F. (2000) Modified nucleoside dependent Watson-Crick and wobble codon binding by tRNALysUUU species. *Biochemistry* **39**, 13390-13395

8. Weixlbaumer, A., Murphy, F. V. t., Dziergowska, A., Malkiewicz, A., Vendeix, F. A., Agris, P. F., and Ramakrishnan, V. (2007) Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nat Struct Mol Biol* **14**, 498-502

9. Phelps, S. S., Jerinic, O., and Joseph, S. (2002) Universally conserved interactions between the ribosome and the anticodon stem-loop of A site tRNA important for translocation. *Mol Cell* **10**, 799-807

10. Hou, Y. M., Gamper, H., and Yang, W. (2015) Post-transcriptional modifications to tRNA--a response to the genetic code degeneracy. *RNA* **21**, 642-644

11. Boccaletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., de Crecy-Lagard, V., Ross, R., Limbach, P. A., Kotter, A., Helm, M., and Bujnicki, J. M. (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* **46**, D303-D307

12. Kothe, U., and Rodnina, M. V. (2007) Codon reading by tRNAAla with modified uridine in the wobble position. *Mol Cell* **25**, 167-174

13. Murphy, F. V. t., and Ramakrishnan, V. (2004) Structure of a purine-purine wobble base pair in the decoding center of the ribosome. *Nat Struct Mol Biol* **11**, 1251-1252

14. Rozov, A., Demeshkina, N., Khusainov, I., Westhof, E., Yusupov, M., and Yusupova, G. (2016) Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat Commun* **7**, 10457

15. Agris, P. F. (1991) Wobble position modified nucleosides evolved to select transfer RNA codon recognition: a modified-wobble hypothesis. *Biochimie* **73**, 1345-1349

16. Murphy, F. V. t., Ramakrishnan, V., Malkiewicz, A., and Agris, P. F. (2004) The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat Struct Mol Biol* **11**, 1186-1191

17. Ogle, J. M., Murphy, F. V., Tarry, M. J., and Ramakrishnan, V. (2002) Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721-732

18. Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M., and Yusupova, G. (2012) A new understanding of the decoding principle on the ribosome. *Nature* **484**, 256-259

19. Hoffer, E. D., Maehigashi, T., Fredrick, K., and Dunham, C. M. (2019) Ribosomal ambiguity (ram) mutations promote the open (off) to closed (on) transition and thereby increase miscoding. *Nucleic Acids Res* 47, 1557-1563

20. Machnicka, M. A., Olchowik, A., Grosjean, H., and Bujnicki, J. M. (2014) Distribution and frequencies of post-transcriptional modifications in tRNAs. *RNA Biol* **11**, 1619-1629

21. Durant, P. C., and Davis, D. R. (1999) Stabilization of the anticodon stem-loop of tRNALys,3 by an A+-C base-pair and by pseudouridine. *J Mol Biol* **285**, 115-131

22. Perret, V., Garcia, A., Puglisi, J., Grosjean, H., Ebel, J. P., Florentz, C., and Giege, R. (1990) Conformation in solution of yeast tRNA(Asp) transcripts deprived of modified nucleotides. *Biochimie* **72**, 735-743

23. Ashraf, S. S., Ansari, G., Guenther, R., Sochacka, E., Malkiewicz, A., and Agris, P. F. (1999) The uridine in "U-turn": contributions to tRNA-ribosomal binding. *RNA* **5**, 503-511

24. Sundaram, M., Durant, P. C., and Davis, D. R. (2000) Hypermodified nucleosides in the anticodon of tRNA(Lys) stabilize a canonical U-turn structure. *Biochemistry* **39**, 15652

25. Cabello-Villegas, J., Winkler, M. E., and Nikonowicz, E. P. (2002) Solution conformations of unmodified and A(37)N(6)-dimethylallyl modified anticodon stem-loops of Escherichia coli tRNA(Phe). *J Mol Biol* **319**, 1015-1034

26. Bjork, G. R., Wikstrom, P. M., and Bystrom, A. S. (1989) Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science* **244**, 986-989

27. Li, J., Esberg, B., Curran, J. F., and Bjork, G. R. (1997) Three modified nucleosides present in the anticodon stem and loop influence the in vivo aa-tRNA selection in a tRNA-dependent manner. *J Mol Biol* **271**, 209-221

28. Gamper, H. B., Masuda, I., Frenkel-Morgenstern, M., and Hou, Y. M. (2015) Maintenance of protein synthesis reading frame by EF-P and m(1)G37-tRNA. *Nature communications* **6**, 7226

29. Hagervall, T. G., Tuohy, T. M., Atkins, J. F., and Bjork, G. R. (1993) Deficiency of 1-methylguanosine in tRNA from Salmonella typhimurium induces frameshifting by quadruplet translocation. *J Mol Biol* **232**, 756-765

30. Qian, Q., Li, J. N., Zhao, H., Hagervall, T. G., Farabaugh, P. J., and Bjork, G. R. (1998) A new model for phenotypic suppression of frameshift mutations by mutant tRNAs. *Mol Cell* **1**, 471-482

31. Maehigashi, T., Dunkle, J. A., Miles, S. J., and Dunham, C. M. (2014) Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. *Proc Natl Acad Sci U S A* **111**, 12740-12745

32. Hong, S., Sunita, S., Maehigashi, T., Hoffer, E. D., Dunkle, J. A., and Dunham, C. M. (2018) Mechanism of tRNA-mediated +1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* **115**, 11226-11231

33. Riddle, D. L., and Roth, J. R. (1970) Suppressors of frameshift mutations in Salmonella typhimurium. *J Mol Biol* **54**, 131-144

34. Yourno, J. (1972) Externally suppressible +1 "glycine" frameshift: possible quadruplet isomers for glycine and proline. *Nat New Biol* **239**, 219-221

35. Sroga, G. E., Nemoto, F., Kuchino, Y., and Bjork, G. R. (1992) Insertion (sufB) in the anticodon loop or base substitution (sufC) in the anticodon stem of tRNA(Pro)2 from Salmonella typhimurium induces suppression of frameshift mutations. *Nucleic Acids Res* **20**, 3463-3469

36. O'Connor, M. (2002) Imbalance of tRNA(Pro) isoacceptors induces +1 frameshifting at near-cognate codons. *Nucleic Acids Res* **30**, 759-765

37. Atkins, J. F., and Bjork, G. R. (2009) A gripping tale of ribosomal frameshifting: extragenic suppressors of frameshift mutations spotlight P-site realignment. *Microbiol Mol Biol Rev* **73**, 178-210

38. Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E., and Baranov, P. V. (2016) Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res* **44**, 7007-7078

39. 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

40. Walker, S. E., and Fredrick, K. (2006) Recognition and positioning of mRNA in the ribosome by tRNAs with expanded anticodons. *J Mol Biol* **360**, 599-609

41. Ledoux, S., Olejniczak, M., and Uhlenbeck, O. C. (2009) A sequence element that tunes Escherichia coli tRNA(Ala)(GGC) to ensure accurate decoding. *Nat Struct Mol Biol* **16**, 359-364

42. Fahlman, R. P., and Uhlenbeck, O. C. (2004) Contribution of the esterified amino acid to the binding of aminoacylated tRNAs to the ribosomal P- and A-sites. *Biochemistry* **43**, 7575-7583

43. Shoji, S., Abdi, N. M., Bundschuh, R., and Fredrick, K. (2009) Contribution of ribosomal residues to P-site tRNA binding. *Nucleic Acids Res* **37**, 4033-4042

44. Konevega, A. L., Soboleva, N. G., Makhno, V. I., Semenkov, Y. P., Wintermeyer, W., Rodnina, M. V., and Katunin, V. I. (2004) Purine bases at position 37 of tRNA stabilize codon-anticodon interaction in the ribosomal A site by stacking and Mg2+-dependent interactions. *RNA* **10**, 90-101

45. Olejniczak, M., and Uhlenbeck, O. C. (2006) tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. *Biochimie* **88**, 943-950

46. Selmer, M., Dunham, C. M., Murphy, F. V. t., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942

47. Jenner, L. B., Demeshkina, N., Yusupova, G., and Yusupov, M. (2010) Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat Struct Mol Biol* **17**, 555-560

48. Bjork, G. R., Ericson, J. U., Gustafsson, C. E., Hagervall, T. G., Jonsson, Y. H., and Wikstrom, P. M. (1987) Transfer RNA modification. *Annu Rev Biochem* **56**, 263-287

49. Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* **26**, 148-153

50. Agris, P. F. (2004) Decoding the genome: a modified view. *Nucleic Acids Res* **32**, 223-238

51. Bjork, G. R., Jacobsson, K., Nilsson, K., Johansson, M. J., Bystrom, A. S., and Persson, O. P. (2001) A primordial tRNA modification required for the evolution of life? *EMBO J* **20**, 231-239

52. Kuchino, Y., Yabusaki, Y., Mori, F., and Nishimura, S. (1984) Nucleotide sequences of three proline tRNAs from Salmonella typhimurium. *Nucleic Acids Res* **12**, 1559-1562

53. Murakami, H., Ohta, A., and Suga, H. (2009) Bases in the anticodon loop of tRNA(Ala)(GGC) prevent misreading. *Nat Struct Mol Biol* **16**, 353-358

54. Dunham, C. M., Selmer, M., Phelps, S. S., Kelley, A. C., Suzuki, T., Joseph, S., and Ramakrishnan, V. (2007) Structures of tRNAs with an expanded anticodon loop in the decoding center of the 30S ribosomal subunit. *RNA* **13**, 817-823

55. Fagan, C. E., Dunkle, J. A., Maehigashi, T., Dang, M. N., Devaraj, A., Miles, S. J., Qin, D., Fredrick, K., and Dunham, C. M. (2013) Reorganization of an intersubunit bridge induced by disparate 16S ribosomal ambiguity mutations mimics an EF-Tu-bound state. *Proc Natl Acad Sci U S A* **110**, 9716-9721

56. Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M. V. (2008) Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J Biol Chem* **283**, 32229-32235

57. Muto, H., and Ito, K. (2008) Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin. *Biochem Biophys Res Commun* **366**, 1043-1047

58. 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-alkylamino acids in translation. *Proc Natl Acad Sci U S A* **106**, 50-54

59. Johansson, M., Ieong, K. W., Trobro, S., Strazewski, P., Aqvist, J., Pavlov, M. Y., and Ehrenberg, M. (2011) pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. *Proc Natl Acad Sci U S A* **108**, 79-84

60. Doerfel, L. K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M. V. (2013) EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* **339**, 85-88

61. Ude, S., Lassak, J., Starosta, A. L., Kraxenberger, T., Wilson, D. N., and Jung, K. (2013) Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* **339**, 82-85

62. Yourno, J., and Tanemura, S. (1970) Restoration of in-phase translation by an unlinked suppressor of a frameshift mutation in Salmonella typhimurium. *Nature* **225**, 422-426

63. Huter, P., Arenz, S., Bock, L. V., Graf, M., Frister, J. O., Heuer, A., Peil, L., Starosta, A. L., Wohlgemuth, I., Peske, F., Novacek, J., Berninghausen, O., Grubmuller, H., Tenson, T., Beckmann, R., Rodnina, M. V., Vaiana, A. C., and Wilson, D. N. (2017) Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. *Mol Cell* **68**, 515-527 e516

64. Navarre, W. W., Zou, S. B., Roy, H., Xie, J. L., Savchenko, A., Singer, A., Edvokimova, E., Prost, L. R., Kumar, R., Ibba, M., and Fang, F. C. (2010) PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in Salmonella enterica. *Mol Cell* **39**, 209-221

65. Bullwinkle, T. J., Zou, S. B., Rajkovic, A., Hersch, S. J., Elgamal, S., Robinson, N., Smil, D., Bolshan, Y., Navarre, W. W., and Ibba, M. (2013) (R)-beta-lysine-modified elongation factor P functions in translation elongation. *J Biol Chem* **288**, 4416-4423

66. Zhang, Y., Hong, S., Ruangprasert, A., Skiniotis, G., and Dunham, C. M. (2018) Alternative Mode of E-Site tRNA Binding in the Presence of a Downstream mRNA Stem Loop at the Entrance Channel. *Structure* **26**, 437-445 e433

67. Wong, I., and Lohman, T. M. (1993) A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc Natl Acad Sci U S A* **90**, 5428-5432

68. Fagan, C. E., Maehigashi, T., Dunkle, J. A., Miles, S. J., and Dunham, C. M. (2014) Structural insights into translational recoding by frameshift suppressor tRNASufJ. *RNA* **20**, 1944-1954

69. Kabsch, W. (2010) Xds. *Acta crystallographica* **66**, 125-132

70. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica* **66**, 213-221

71. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica* **60**, 2126-2132

72. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

## Chapter 3.

## Disruption of evolutionarily correlated tRNA elements impairs accurate decoding

Ha An Nguyen, S. Sunita, Christine M. Dunham.

This research was originally published in the Proceedings of the National Academy of Sciences. H. A. Nguyen, S. Sunita, C. M. Dunham, Disruption of evolutionarily correlated tRNA elements impairs accurate decoding. *Proceedings of the National Academy of Sciences* **117**, 16333-16338 (2020).

**Author contributions**: S.S. and C.M.D. designed research; H.A.N. and S.S. performed research; H.A.N., S.S., and C.M.D. analyzed data; and H.A.N., S.S., and C.M.D. wrote the paper.

## 3.1. Abstract

Bacterial tRNAs contain evolutionarily conserved sequences and modifications that ensure uniform binding to the ribosome and optimal translational accuracy despite differences in their aminoacyl attachments and anticodon nucleotide sequences. In the tRNA anticodon stem-loop, the anticodon sequence is correlated to a base pair in the anticodon loop (nucleotides 32-38) that tunes the binding of each tRNA to the decoding center in the ribosome. Disruption of this correlation renders the ribosome unable to distinguish correct from incorrect tRNAs. The molecular basis for how these two coordinated aspects of the tRNA lead to accurate decoding is unclear. Here, we solved structures of the bacterial ribosome containing either wild-type $tRNA^{Ala}_{GGC}$ or $tRNA^{Ala}_{GGC}$ containing a reversed 32-38 pair on cognate and near-cognate codons. Structures of wild-type $tRNA^{Ala}_{GGC}$ bound to the ribosome reveal 23S rRNA nucleotide A1913 positional changes that are dependent on whether the codon-anticodon interaction is cognate or near-cognate. Further, the 32-38 pair is destabilized in the context of a near-cognate, codon-anticodon pair. Reversal of the pairing in $tRNA^{Ala}_{GGC}$ ablates A1913 movement regardless if the interaction is cognate or near-cognate. These results demonstrate that disrupting the 32-38 and anticodon sequences alters interactions with the ribosome that contribute to misreading.

## 3.2. Significance

Accurate gene expression relies on polymerases distinguishing correct from incorrect substrates that are chemically and structurally similar. During protein synthesis, Watson-Crick base pairing between the tRNA anticodon and the mRNA codon is essential for accurate translation. However, other tRNA elements outside the anticodon also contribute

to correct selection by the ribosome but the structural basis for this selectivity is unknown. Here, we determined structures of the bacterial ribosome containing tRNAs with altered nucleotide pairings in the anticodon loop that are known to prevent the ribosome from distinguishing correct from incorrect tRNAs. Collectively, these structures reveal a previously unappreciated role for a ribosomal decoding site nucleotide in sensing the integrity of the tRNA that contributes to decoding fidelity.

### 3.3. Introduction

The ribosome orchestrates the binding of mRNA, protein translation factors and tRNAs in a sequential manner for the synthesis of all cellular proteins. This process is remarkably complicated and involves numerous steps that have been evolutionarily optimized to select correct tRNAs from a pool of structurally and chemically similar tRNAs. tRNAs are the so-called “adaptor” molecules that decode the genetic code on mRNA and carry an aminoacyl group (1). As noted in a recent review (2), the term “adaptor” implies plasticity between the aminoacyl group and anticodon, yet it is clear that other nucleotide sequences and modifications in each tRNA are also evolutionarily tuned to permit comparable binding affinities to both EF-Tu and to the ribosome for optimal accuracy (3-5). This tuning of tRNAs occurs in all bacterial tRNAs (5), yet the structural basis for how optimized sequences contribute to efficient recognition on the ribosome is unclear.

All tRNAs adopt L-shaped structures that traverse both the small 30S and the large 50S subunits when bound to the ribosome. The aminoacyl group is attached to the 3’ end of the tRNA which is located ~90 Å away from the anticodon (Fig. 3.1*A*). EF-Tu interacts extensively with the acceptor arm of tRNAs and surrounds the 3’ aminoacyl group. These aminoacyl groups represent a broad range of chemical diversity yet all tRNAs bind to EF-Tu with the same relative affinities despite these intrinsic physical differences, due to

compensatory binding contributions from the tRNA (2, 3, 6). Likewise, the anticodon region of all tRNAs bind to their cognate mRNA codons on the ribosome with similar affinities, despite diverse codon-anticodon pairings that should exhibit differences in base pairing strengths (4). In both cases, the sequences of each tRNA have evolved to compensate for the chemical diversity of the aminoacyl group or the codon-anticodon strength to achieve similar rates of binding and optimal accuracy.

The evolution of tRNA sequence and modification patterns implicate previously unappreciated roles of specific tRNA regions in translation, including that of the process of decoding. During decoding, the Watson-Crick base pairing between the three nucleotides of the mRNA codon and tRNA anticodon is monitored by the ribosome. Correct pairing causes conformational changes of the 30S subunit and GTPase activation of EF-Tu to release the tRNA and for elongation to proceed (7). Phylogenic and biochemical analyses reveal that a universal feature of all bacterial tRNAs is a correlation between the nucleotide identities of the anticodon (nucleotides 34, 35, and 36) and nucleotides 32 and 38 located in the anticodon loop (5, 8-10) (Fig. 3.1*A*). Specifically, strong G-C codon−anticodon interactions are always balanced by a weaker 32–38 pairing and, conversely, a weak AU-rich codon−anticodon interaction is coupled with a stronger 32–38 pairing. This coordination of nucleotide identities ensures uniform binding affinities of all tRNAs to their cognate codon (4, 11). Further, when the nucleotide identity of the 32-38 pair is disrupted, the ribosome is unable to distinguish correct from incorrect tRNAs, establishing that this correlation is important for translation fidelity (9, 10).

Mutations of the A32-U38 pair in $tRNA^{Ala}_{GGC}$ have provided significant insights into how 32-38 sequence changes alone disrupt the fidelity of translation (5, 8-10, 12). The wild-

type A32-U38 pairing in $tRNA^{Ala}_{GGC}$ is considered a weak interaction that counterbalances the GGC anticodon that binds tightly to the GCC alanine codon (9, 10). Without a weak A32-U38 pair, the GGC anticodon of $tRNA^{Ala}_{GGC}$ binds very tightly to its own cognate codon but also binds tightly to other near-cognate codons (8) (near-cognate is defined as a single mismatch pair between the codon and anticodon). Another consequence of the tRNA binding too tightly to its own cognate codon could be that downstream aspects of decoding may also be impaired. When tRNAs bind too tightly to near-cognate codons, this can result in the incorrect tRNA outcompeting the normal cognate tRNAs resulting in the misincorporation of the tRNA or misreading. changing the 32-38 pairing to U-A, C-G, or C-A in $tRNA^{Ala}_{GGC}$ causes efficient misreading of near-cognate codons resulting in the misincorporation of alanine (8, 12). Specifically, the $tRNA^{Ala}_{GGC}$ U32-A38 variant, which contains a reversed 32-38 pair, decodes the near-cognate GCA codon efficiently. While the initial binding affinity of the EF-Tu•GTP•$tRNA^{Ala}_{GGC}$ U32-A38 ternary complex to the ribosome (either on cognate GCC or near-cognate GCA codons) is nominally decreased as compared to that of wild-type $tRNA^{Ala}_{GGC}$, the GTP hydrolysis rates by EF-Tu and dipeptide formation are comparable to that of the wild-type tRNA decoding a cognate GCC codon (8). Therefore, the U32-A38 reversal circumvents proofreading mechanism that allow for the acceptance of the incorrect tRNA. In these cases, typically all aspects of decoding, that is, EF-Tu•GTP•tRNA-binding, GTP hydrolysis and dipeptide formation, are impaired. In summary, these data strongly implicate bypassing of proofreading as the major impacted step of decoding upon dysregulation of the 32-38 pair and the anticodon sequence coordination (11). Further, while deletion of $tRNA^{Ala}_{GGC}$ only incurs a minor growth defect (13)(1), overexpression of certain $tRNA^{Ala}_{GGC}$ variants, including the $tRNA^{Ala}_{GGC}$

U32-A38 variant, is toxic in *E. coli* (12). Therefore, perturbing the 32-38 base pair in a single tRNA isoacceptor can overwhelm the canonical translation machinery that results in cell death.

Motivated by these compelling biochemical and *in vivo* assays, we solved four x-ray crystal structures of *Thermus thermophilus* ribosomes containing either wild-type $tRNA^{Ala}_{GGC}$ or the $tRNA^{Ala}_{GGC}$ U32-A38 variant in the aminoacyl (A) site decoding either a cognate GCC or near-cognate GCA codon (Fig. 3.1). We find that when wild-type $tRNA^{Ala}_{GGC}$ decodes a near-cognate GCA codon (single mismatch at the third position of the codon-anticodon interaction), interactions between 23S rRNA A1913 and the tRNA are ablated. However, in the case of the $tRNA^{Ala}_{GGC}$ with the reversed 32-38 pair, A1913 maintains interactions with the tRNA regardless if the codon-anticodon interaction is cognate or near-cognate.

### 3.4. Results

**Near-cognate interactions between $tRNA^{Ala}_{GGC}$ and the GCA codon influence the position of A1913.** To determine the structural basis for how $tRNA^{Ala}_{GGC}$ misreads codons upon reversal of its 32-38 pair, we first solved a structure of *Thermus thermophilus* 70S with wild-type $tRNA^{Ala}_{GGC}$ bound to a cognate GCC codon in the A site (Figs. 3.1*A* and 3.1*B* and 3.2*A* and Table 3.1). We formed ribosome complexes with mRNA containing an initiation AUG codon at the peptidyl (P) site and GCC alanine codon at the A site. *E. coli* $tRNA^{fMet}$ and $tRNA^{Ala}_{GGC}$ were added to the P and the A sites, respectively and crystallization trials were initiated according to standard procedures (14). In the structure solved to 3.2 Å, three Watson-Crick base pairs form between the codon-anticodon and is recognized by the A site as a cognate interaction by 16S rRNA nucleotides G530, A1492,

A1493 and 23S rRNA A1913 (Fig. 3.2*B* and Fig. 3.3). These rRNA nucleotides surround the codon-anticodon pair and adopt “ON” positions whereby A1492 and A1493 interact directly with the first two positions of the codon-anticodon, G530 forms interactions with A1492 to lock the 30S in a closed conformation, and A1913 packs against the tRNA backbone of the anticodon stem (Figs. 3.3 and 3.4) (15-17). Although the 32–38 nucleotides are commonly shown as unpaired in secondary structural representations, these two nucleotides often form interactions between their Watson–Crick faces (18) (Fig. 3.5). In this structure, the A32−U38 pair in the anticodon loop adopts a canonical Watson–Crick pair with A1913 stacking against the phosphate backbone of U38 (Figs. 3.2*C* and 3.3*B*).

We next solved a 3.2-Å structure of the *Thermus thermophilus* 70S-wild-type $tRNA^{Ala}_{GGC}$ bound to a near-cognate GCA codon at the A site (Figs. 3.1*B* and 3.6*A* and Table 3.1). Two Watson-Crick base pairs form at the first two positions of the codon-anticodon and a single $A_{+6}$•G34 mismatch forms at the third or wobble position (Fig. 3.6*B*). Both G34 and $A_{+6}$ adopt *anti* conformations with two hydrogen bonds forming between the amino N6 of $A_{+6}$ with the carbonyl O6 of G34, and between the N1 of $A_{+6}$ with the amino N1 of G34 (Fig. 3.6*B*). The C1’-C1’ distance between a Watson Crick base pair is typically 10.5 Å while G•A pairs have a distance of ~12.6 Å (19); the C1’-C1’ distance of the $A_{+6}$•G34 pair in this study is ~12.6 Å consistent with previous studies. To our knowledge, a $A_{+6}$•G34 mismatch has not been observed in the decoding center of the ribosome. At the wobble position, G•U pairs and chemically modified anticodon nucleotides that interact with mismatched codon nucleotides can be decoded as Watson-Crick-like base pairs by the ribosome, a mechanism known as wobble base decoding

(16). By this definition, a $A_{+6}$•G34 mismatch in an anticodon-codon interaction should be rejected by the ribosome resulting in rapid dissociation of the ternary complex after initial selection.

The near-cognate codon-anticodon interaction also causes the nucleobases of the A32-U38 base pair to become disordered as indicated by the lack of electron density (Fig. 3.6*C*). This lack of electron density is notable because in both structures we solved of ribosomes bound to cognate versus near-cognate codon-anticodon interactions are at comparable resolutions (both 3.2 Å). The destabilization of the 32-38 pairing has also been seen when the tRNA-mRNA pairs are near-cognate and cause mRNAs frameshifts (20, 21). These data further emphasize the critical role of this base pair in the accurate decoding of cognate codons.

The ribosome closely monitors the codon-anticodon interaction (Figs. 3.3*A* and 3.4) but the rest of the ASL is minimally inspected in the A site, providing a conundrum in understanding how the correlation between the anticodon and the 32-38 pairing could tune tRNA binding and acceptance by the ribosome (8-10). The closest ribosomal nucleotide or protein to anticodon stem nucleotides 32-38 is 23S rRNA nucleotide A1913 (Figs. 3.4 and 3.7*A*). A1913 is located in the loop of Helix 69 (H69) which is a universally conserved helix that forms an intersubunit bridge, contacts 16S rRNA nucleotide A1493 during decoding (22), and is also important for release factor recognition of stop codons (22, 23). A1913 typically packs against nucleotide 38 of the A-site tRNA and forms a hydrogen bond with the 2'-OH of nucleotide 37 (Fig. 3.3*B*). In the case of wild-type $tRNA^{Ala}_{GGC}$ decoding a near-cognate GCA codon that causes disordering of U32-A38, we find that the nucleobase of A1913 shifts ~7 Å from the tRNA as compared to when

$tRNA^{Ala}_{GGC}$ binds to a cognate codon (as measured between N1 atoms of the nucleobases) (Figs. 3.3*C* and 3.7*B*-*C*).

**Disrupting the 32-38 pair of $tRNA^{Ala}_{GGC}$ renders the ribosome unable to distinguish cognate from near-cognate codon-anticodon pairs.** To understand the structural basis for how reversing the 32-38 pairing in tRNAs leads to miscoding, we solved a structure of the ribosome bound to an A-site $tRNA^{Ala}_{GGC}$ U32-A38 variant decoding either a cognate GCC or a near-cognate GCA codon (Fig. 3.8 and Table 3.1). Three Watson-Crick base pairs form between the $tRNA^{Ala}_{GGC}$ U32-A38 variant and a GCC cognate codon similar to the wild-type $tRNA^{Ala}_{GGC}$-GCC codon interaction (Figs. 3.2*B* and 3.8*A*). Similarly, the near-cognate interaction formed between $tRNA^{Ala}_{GGC}$ U32-A38 and the GCA codon is identical to how wild-type $tRNA^{Ala}_{GGC}$ interacts with near-cognate codon (Figs. 3.6*A* and 3.8*B*). However, in contrast to the disordering of the A32-U38 nucleobases observed in wild-type $tRNA^{Ala}_{GGC}$ decoding a near-cognate GCA codon (Fig. 3.6*C*), all regions of the anticodon stem-loop of the $tRNA^{Ala}_{GGC}$ U32-A38 variant are well-ordered regardless if bound to a cognate GCC or near-cognate GCA codon (Fig. 3.9). The only notable difference in the $tRNA^{Ala}_{GGC}$ U32-A38 variant structures is the position of the mutated U32-A38 nucleobases; U32-A38 swivel ~20° away from each other as compared to the wild-type $tRNA^{Ala}_{GGC}$ (Fig. 3.10). The amino N6 of A38 now forms a single hydrogen bond with the carbonyl O2 of U32 in contrast to a typical Watson-Crick base pair that forms with wild-type $tRNA^{Ala}_{GGC}$ (Figs. 3.2*C* and 3.6*C*). In both structures of $tRNA^{Ala}_{GGC}$ U32-A38 variant decoding either a cognate or a near-cognate codon when the ribosome is unable to distinguish between correct and incorrect tRNAs (8), A1913 remains packed against the

A-site tRNA variant (Fig. 3.11). These results indicate reversing the 32-38 pair results in changes in their pairing that, in turn, influence how A1913 packs with the anticodon stem-loop.

### 3.5. Discussion

To maintain efficient and accurate protein synthesis, tRNAs acquired diverse sequences and chemical modifications that enable their specific recognition by specific aminoacyl-tRNA synthetases and the decoding of cognate mRNA codons (24, 25). In addition, these tRNA elements are evolutionarily optimized to ensure that all tRNAs have similar binding affinities to both EF-Tu and the ribosome, thus preventing potential thermodynamic differences from contributing to the decoding process. As part of the tuning of tRNAs to bind uniformly to the ribosome, the nucleotide identities and strength of the 32–38 base pair and the anticodon nucleotides are correlated (9, 10). Since the ribosome closely monitors the codon−anticodon interaction but the rest of the ASL is minimally inspected in the A site, it was unclear how disrupting this correlation would influence the overall tRNA structure and whether this dysregulation affects how the ribosome interacts with the A-site tRNA. Here, we solved X-ray crystal structures of ribosome complexes containing $tRNA^{Ala}_{GGC}$ or a $tRNA^{Ala}_{GGC}$ U32−A38 variant known to cause high levels of miscoding (8). The U32−A38 variant binds tighter to both cognate and near-cognate codons, resulting in increased misreading as assessed by biochemical assays and cell death in vivo upon overexpression (8, 12). The reversal of U32−A38 does not affect other aspects of tRNA selection, including GTP hydrolysis by EF-Tu or dipeptide formation kinetics (8), establishing $tRNA^{Ala}_{GGC}$ as an ideal system to study these miscoding events.

In this study, we report two observations that may explain the misreading propensity when the identity of the 32–38 pair and the anticodon nucleotides are dysregulated: In the context of the wild-type $tRNA^{Ala}_{GGC}$ decoding a near-cognate codon, the 32–38 pairing becomes disordered, and 23S rRNA A1913 moves away from the backbone of the ASL of the tRNA (Figs. 3.6*C* and 3.7*C*). Normally, when wild-type $tRNA^{Ala}_{GGC}$ decodes a cognate GCC codon, nucleotides 32 and 38 form a stable Watson–Crick base pair with well-defined electron density (Fig. 3.2*C*). However, when the same tRNA is bound to a near-cognate codon, resulting in a $A_{+6}$•G36 mismatch at the wobble position, the electron density of the 32–38 base pair is weakened, signifying that these nucleotides are more mobile (Fig. 3.6*C*). These data provide hints to how noncanonical interactions between the codon and the anticodon may be sensed by other regions of the tRNA. A similar disruption of the 32–38 pairing was previously observed in the context of mutant tRNAs that cause mRNA frameshifts in the +1 direction (20, 21). Notably, in these structures, there is a complete lack of density for nucleotide 32 of the anticodon loop and 5' of the anticodon stem (nucleotides 28 to 31), suggesting this disorder and/or instability may be a common feature of tRNA−mRNA pairs that cause miscoding.

A1913 adopts two different conformations that appear to be dependent on whether the codon−anticodon interaction is cognate or near cognate (Figs. 3.3 and 3.7). In the case of a cognate codon−anticodon interaction, A1913 packs against the backbone of tRNA nucleotide 38 in a position that is observed in most ribosome structures solved to date (Figs. 3.5 and 3.7*B*). In contrast, in the case of a near-cognate, codon−anticodon interaction, A1913 moves ~7 Å away from the tRNA (Fig. 3.7*C*). We propose that A1913 is part of the response of the ribosome to monitor the structural integrity of the A-site tRNA

[previously termed “ON” (17)]. When the 32–38 nucleotides are reversed in $tRNA^{Ala}_{GGC}$, the position of A1913 is always “ON,” packed against the tRNA backbone, regardless of whether the codon−anticodon interaction is cognate or near cognate (Fig. 3.7). These data strongly suggest that, in the case of the $tRNA^{Ala}_{GGC}$ U32−A38 variant, the ribosome recognizes both the cognate and the near-cognate interaction as correct (or “ON”), consistent with previous biochemical analyses demonstrating that $tRNA^{Ala}_{GGC}$ U32-A38 bypasses decoding checkpoints to misread near-cognate codons (8, 9).

To our knowledge, the position of A1913 has never been seen to move away from the tRNA backbone in the context of a mismatched codon−anticodon interaction in the A site (Fig. 3.7). In structures containing single C•A, A•C, A•A, U•G, and G•U mismatches at the first (16, 26) or second (16, 17, 27) position of the codon−anticodon interaction, A1913 packs against the tRNA; in one structure with a U•G mismatch at the second A-site position, A1913 is conformationally dynamic and unresolvable in the electron potential map (17). In these cases, the mismatched interaction was formed by systematically changing the sequence of the mRNA codon using four standard tRNAs in the absence of prior biochemical knowledge of how these pairs impact the decoding process (Fig. 3.12). It is now well appreciated that certain codon−anticodon pairs undergo high levels of misreading in vivo, while other pairings do not (28). Therefore, perhaps the movement of A1913 wasn’t previously identified because the mismatched complexes affect a different stage of initial tRNA selection than what was captured in the structure, or the mismatched codon−anticodon pair does not cause high levels of miscoding. Future studies are required to understand this previously unappreciated role of A1913 in maintaining the fidelity of the decoding process.

## 3.6. Methods

### *in vitro* transcription of $tRNA^{Ala}_{GGC}$.

Two DNA oligos spanning the A32-U38 $tRNA^{Ala}_{GGC}$ variant were annealed, PCR amplified, and subcloned into a linearized pUC18 plasmid. The tRNA sequence was flanked by a T7 RNA polymerase promoter and a BsaI restriction digest site. *E. coli* DH5α were transformed with pUC18-Ala and grown overnight in super broth (3.5% tryptone, 2.0% yeast extract, 0.5% NaCl, 5 mM NaOH) supplemented with 100 µg/ml ampicillin at 37°C. The cell pellet was harvested by centrifugation, plasmid DNA was purified, digested using BsaI, run-off transcription was performed and the RNA purified as previously described (29).

### Crystallization, X-ray data collection, and structural determination.

70S ribosomes were purified from *Thermus thermophilus* using previously established protocols (14). The ribosome complex was formed by incubating 4.4 µM 70S with 8.8 µM mRNA (IDT) in buffer (5 mM HEPES-KOH, pH 7.5, 50 mM KCl, 10 mM $NH_4Cl$, 10 mM $Mg(CH_3COO)_2$, 6 mM β-mercaptoethanol (β-Me) at 55 °C for 5 mins. Then 11 µM $tRNA^{fMet}$ (Chemical Block) and 22 µM $tRNA^{Ala}$ were sequentially incubated at 55° C for 15 mins. The reaction was cooled to 37 °C, and 0.1 mM paromomycin was incubated at 37 °C. After equilibrating at 20°C, a final concentration of 2.8 µM deoxy BigCHAP (Hampton Research) was added to the complex. Crystals grew from either a polyethylene glycol (PEG) condition (0.1 M Tris-HOAc pH 7.0, 0.2 M KSCN, 4-4.5% (v/v) PEG 20K, 4.5-5.5% (v/v) PEG 550MME, 10 mM $Mg(CH_3COO)_2$) or an MPD condition (0.1 M L-arginine HCl, 0.1 M Tris-HCl pH 7.5, 3% PEG 20K, 10-16.5% MPD (2-methyl-2,4-pentanediol), 1 mM β-Me). Data collection was performed at the SER-CAT 22-ID and NE-CAT 24ID-C

beamlines at the Advanced Photon Source. Data were integrated and scaled using XDS (30), molecular replacement performed in PHENIX (31) using coordinates from PDB structure 4Y4O (32). Initial refinement was done using rigid-body restraints in PHENIX, followed by jelly-body refinement in REFMAC5 (33) in the CCP4i2 suite (34) and further iterative rounds of crystallographic refinements were performed in PHENIX. Model building was performed in Coot (35) and figures generated using PyMol (36).

**Data Availability.**

Crystallography, atomic coordinates, and structure factors have been deposited in the PDB, https://www.wwpdb.org/ (PDB ID codes 6OF6, 6OJ2, 6OPE, 6ORD).

### 3.7. Figures and tables

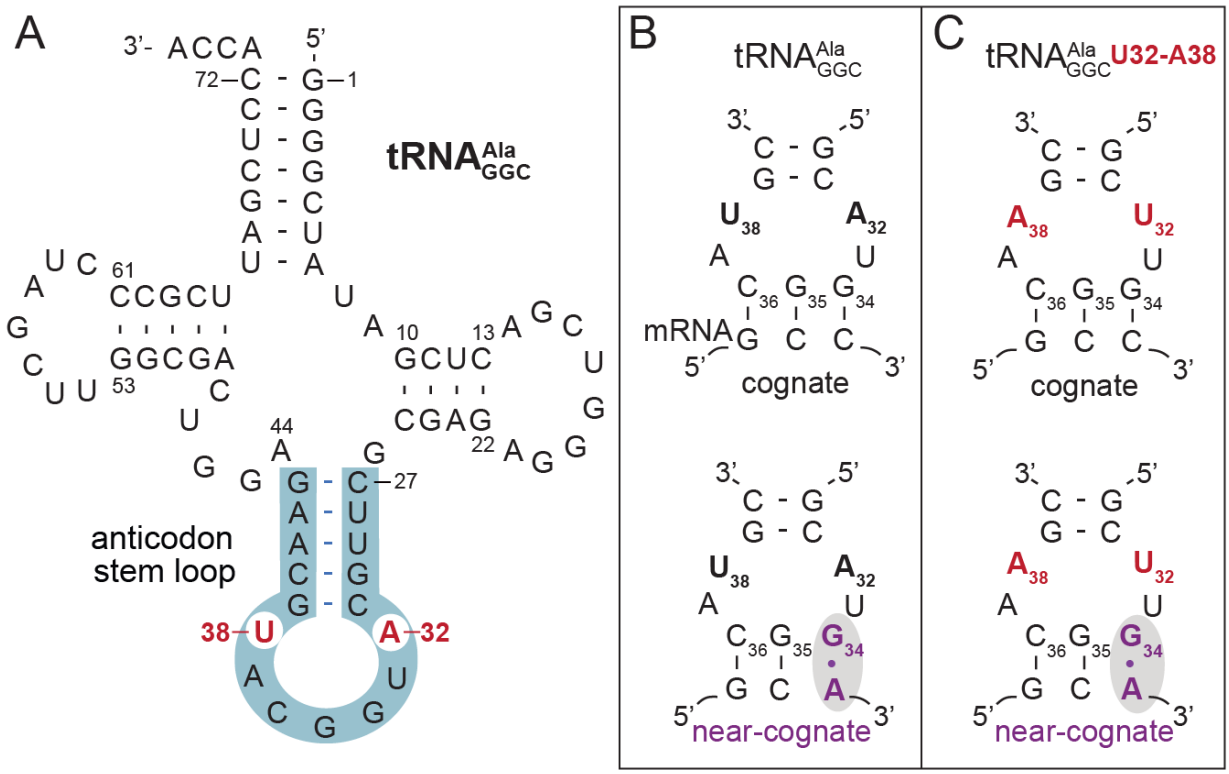


**Figure 3.1. tRNA$^{Ala}_{GGC}$-mRNA complexes used in this study.**

***A***, The secondary structure of tRNA$^{Ala}_{GGC}$ with the anticodon stem-loop highlighted in blue. ***B***, Complexes of wild-type tRNA$^{Ala}_{GGC}$ (A32-U38 in purple) and the ***C***, tRNA$^{Ala}_{GGC}$ variant containing U32-A38 (red) on either a cognate GCC or near-cognate GCA codons.

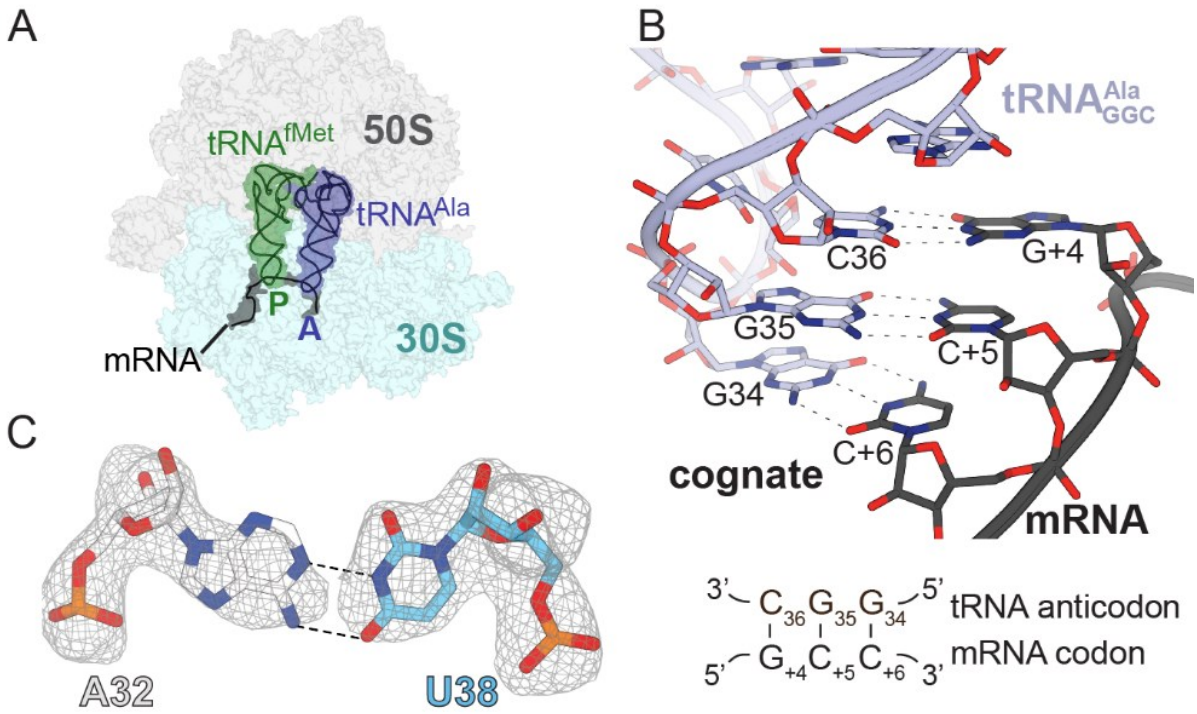


**Figure 3.2. Cognate interactions between $tRNA^{Ala}_{GGC}$ and the GCC mRNA codon.**

***A,*** Overview of the bacterial 70S ribosome containing a P-site $tRNA^{fMet}$ and an A-site $tRNA^{Ala}_{GGC}$ bound to a GCC codon. ***B,*** Zoomed in view of the anticodon stem-loop of $tRNA^{Ala}_{GGC}$ showing three Watson-Crick pairings between the anticodon and the codon. The +1 mRNA numbering starts in the P site with the A-site nucleotide positions labeled as +4, +5 and +6. ***C,*** The $2F_O$-$F_C$ electron density map of the A32-U38 pair is contoured at 1.0σ.

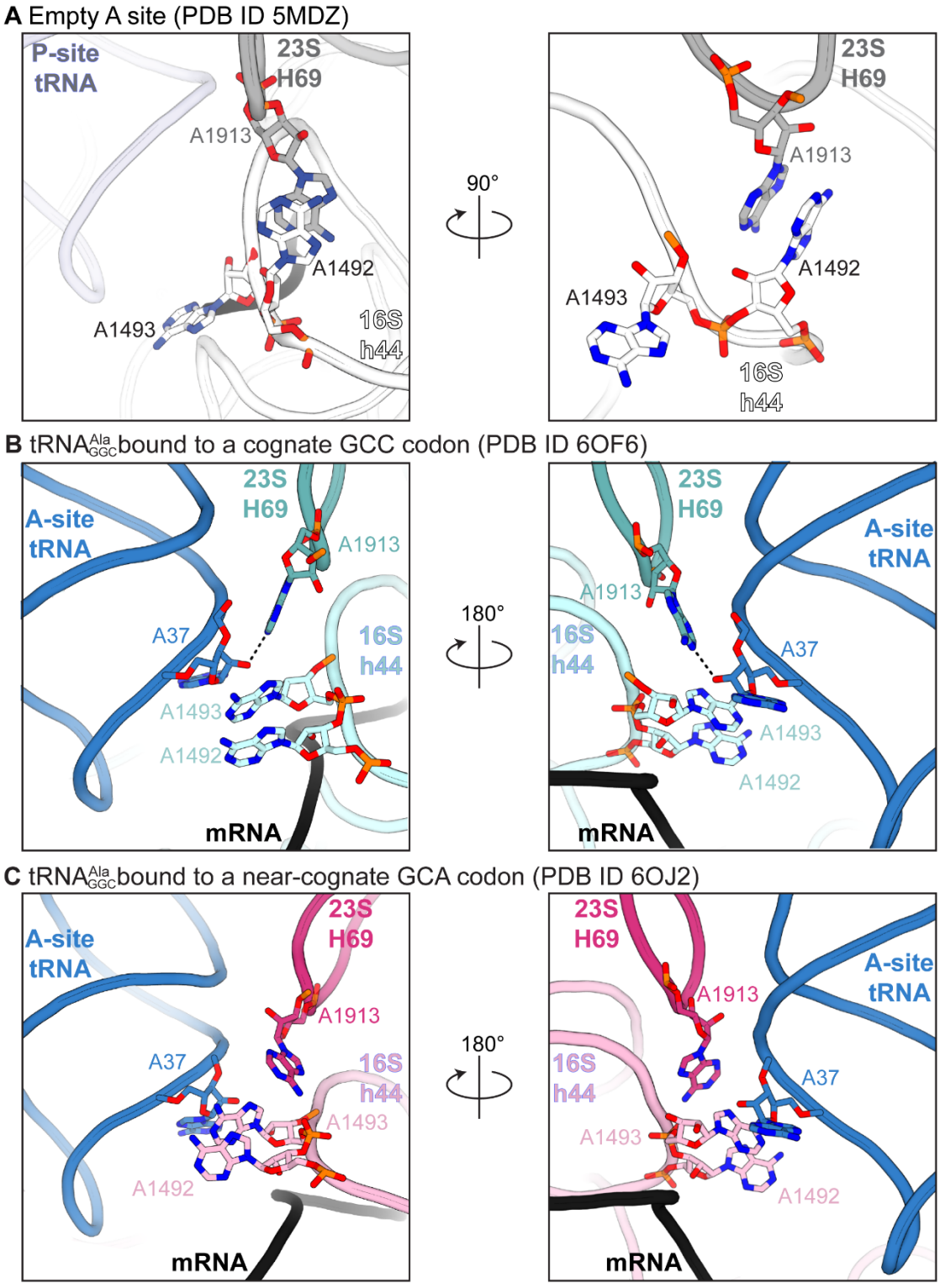


**Figure 3.3. Conformational changes of rRNA nucleotides 23S rRNA A1913 (from Helix 69) and 16S rRNA A1492-1493 (from helix 44) during decoding.**

***A****,* In the absence of tRNA in the A site, A1913 and A1492 form a stacking interaction (PDB code 5MDZ (1)). ***B****,* When $tRNA^{Ala}_{GGC}$ binds to a cognate GCC codon in the A site, A1492 flips to form stacking interactions with A1493 replacing A1492's interaction with A1913.

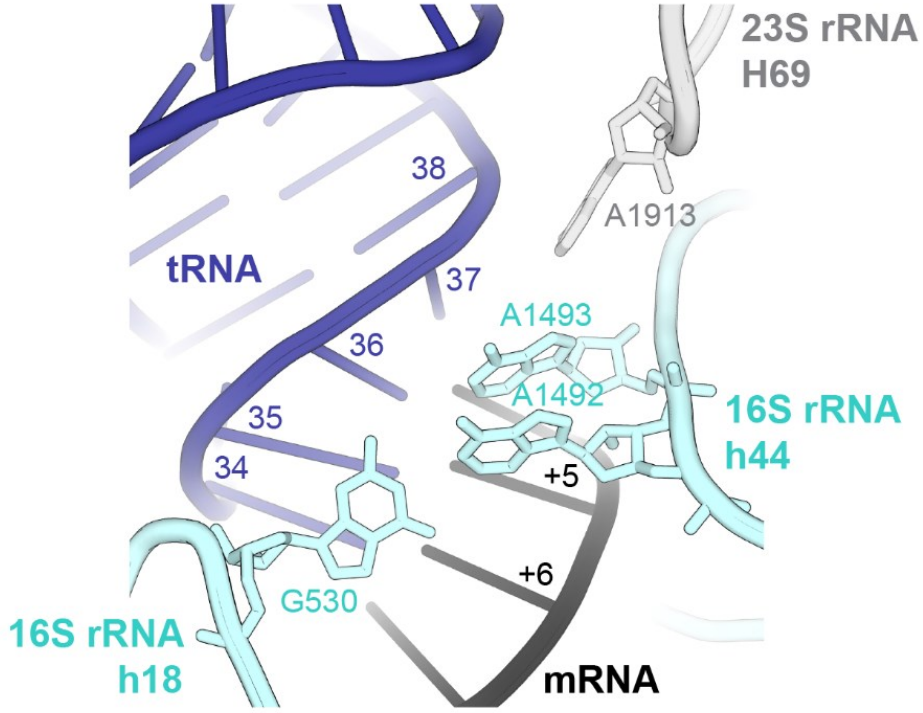


**Figure 3.4. The decoding center of the ribosome.**

16S rRNA nucleotides G530 (from helix 18), A1492 and A1493 (from helix 44) (teal) inspect the A-site tRNA anticodon (blue) and mRNA codon (black) interaction, while 23S rRNA nucleotide A1913 (from Helix 69) (gray) packs against the tRNA phosphate backbone.

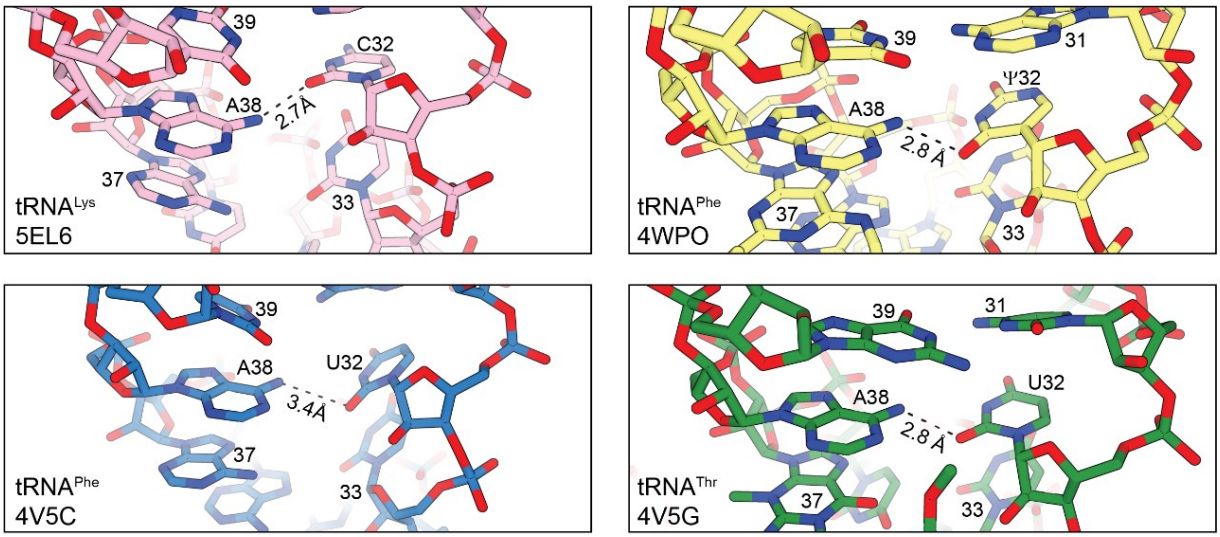


**Figure 3.5. The 32-38 pair in tRNAs bound to the ribosomal A site.**

Four structures of different tRNAs all showing an interaction between nucleotides 32 and 38 (PDB codes 5EL6 (1), 4WPO (2), 4V5C (3), 4V5G (4)).

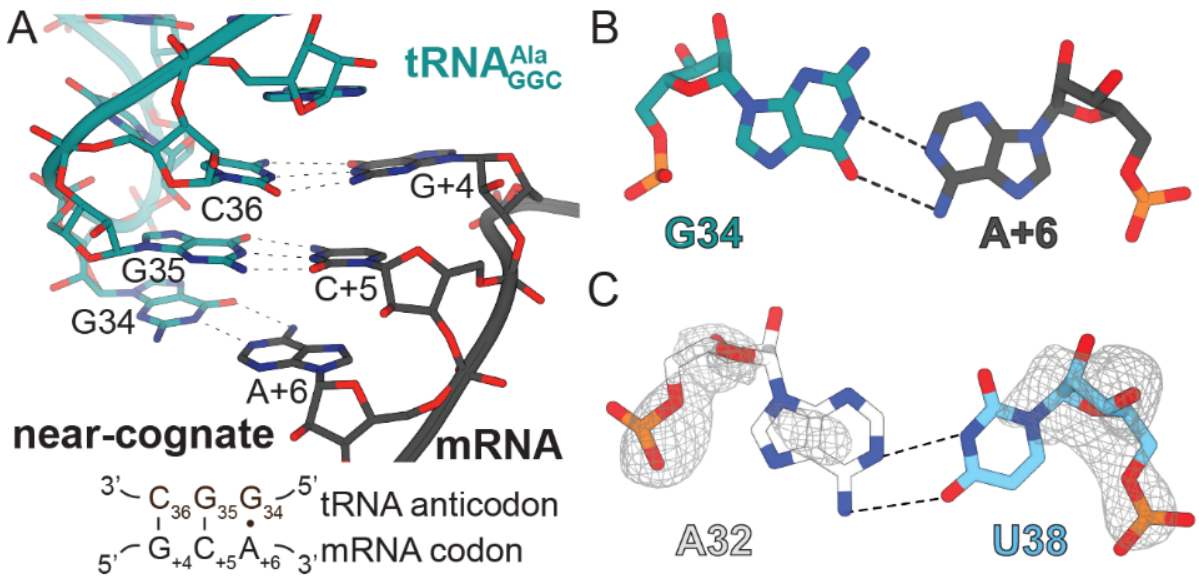


**Figure 3.6. A near-cognate codon-anticodon interaction in tRNA$^{Ala}_{GGC}$ influences the stability of the A32-U38 pair.**

***A*,** Zoomed in view of the interaction between the anticodon stem-loop of tRNA$^{Ala}_{GGC}$ with a near-cognate GCA codon. Two Watson-Crick pairs forms between C36-G$_{+4}$ and G35-C$_{+5}$ and a G34•A$_{+6}$ mismatch forms at the third or wobble position. ***B*,** Two hydrogen bonds form between the *cis* Watson-Crick G34•A$_{+6}$ interaction. ***C,*** The A32-U38 pair is destabilized in a structure of the 70S bound to tRNA$^{Ala}_{GGC}$ in the A site decoding a near-cognate GCA codon.

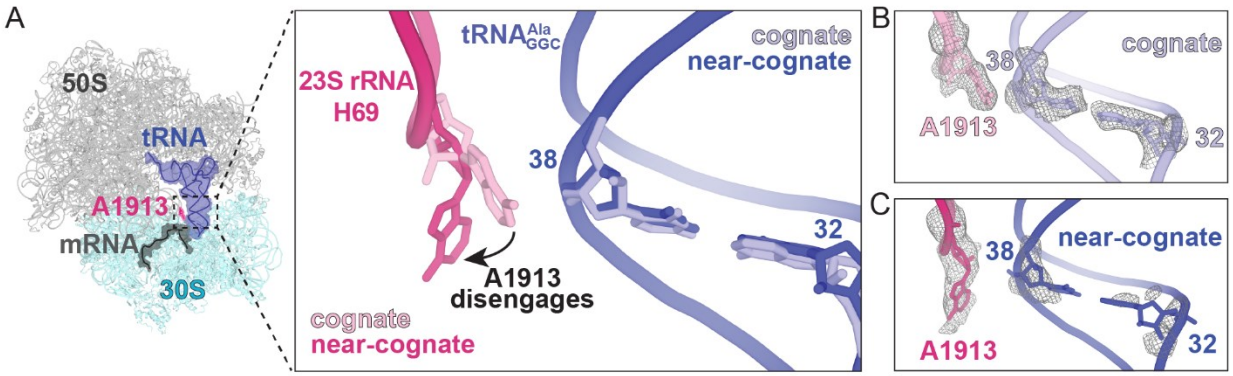


**Figure 3.7. Interaction of 23S rRNA A1913 with tRNA and its ablation when the 32-38 pair is destabilized.**

***A,*** The A-site tRNA (blue) contacts 23S rRNA A1913 in H69 (light pink) that packs against the backbone of U38 (light blue) in the context of a cognate tRNA-mRNA pair. When a near-cognate tRNA-mRNA pair binds at the A site, A1913 (dark pink) moves away from the tRNA (dark blue). ***B,*** Interactions between A1913 showing its nucleobase is proximal to the U38-A32 pair in tRNA when $tRNA^{Ala}_{GGC}$ decodes a cognate CGG codon. A1913 also adopts this conformation in the structures where the $tRNA^{Ala}_{GGC}$ U32-A38 variant is a cognate GCC or near-cognate GCA codon. ***C,*** When $tRNA^{Ala}_{GGC}$ decodes a near-cognate GCA codon, the 32-38 pair becomes disordered and A1913 moves away from the A-site tRNA. $2F_O$-$F_C$ electron density maps are contoured at 1.0σ.

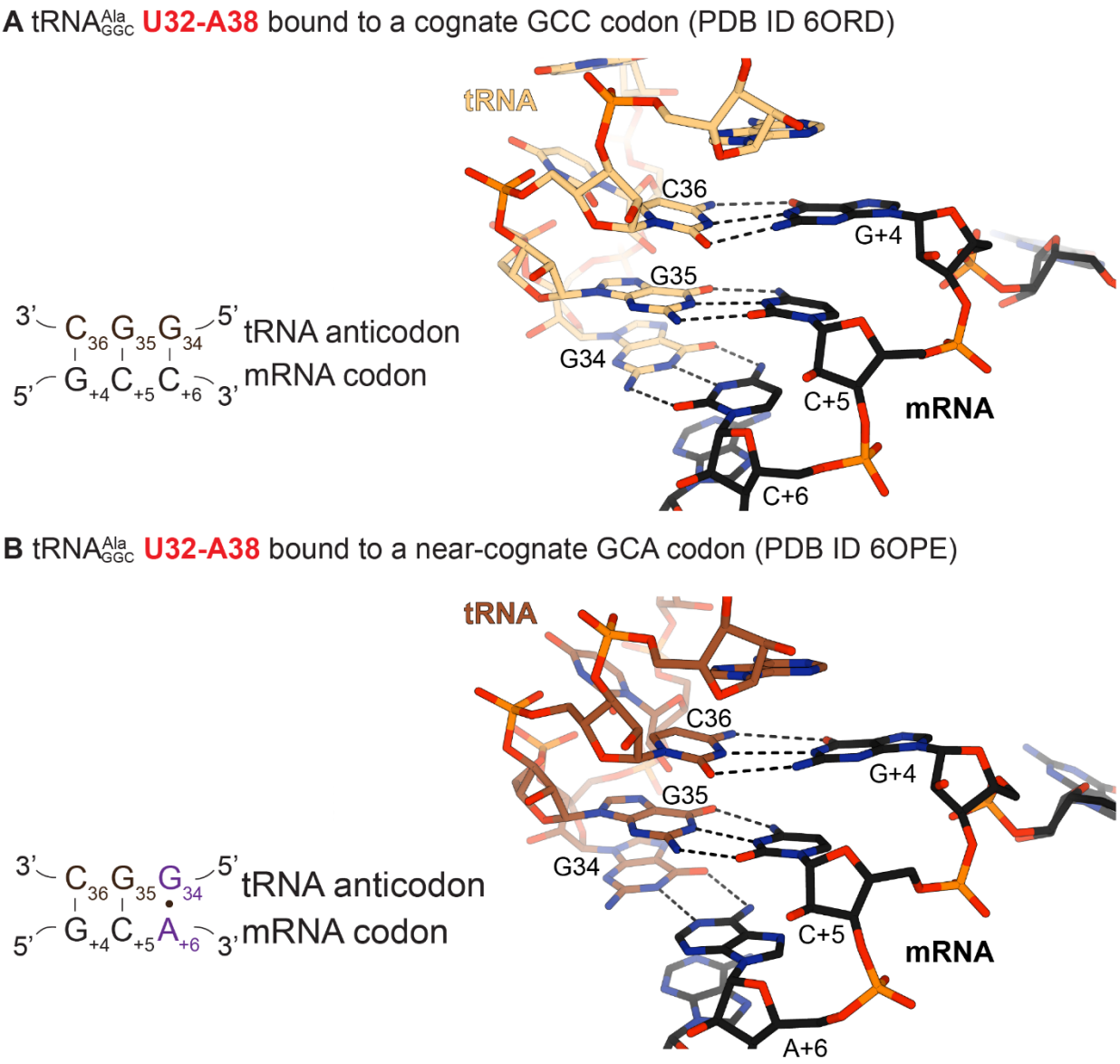


**Figure 3.8. The codon-anticodon interactions in structures of the tRNA$^{Ala}_{GGC}$ U32-A38 mutant bound in the A site.**

The interaction is maintained for the cognate (panel ***A***) and near-cognate (panel ***B***) codons similar to that in the wild-type tRNA.

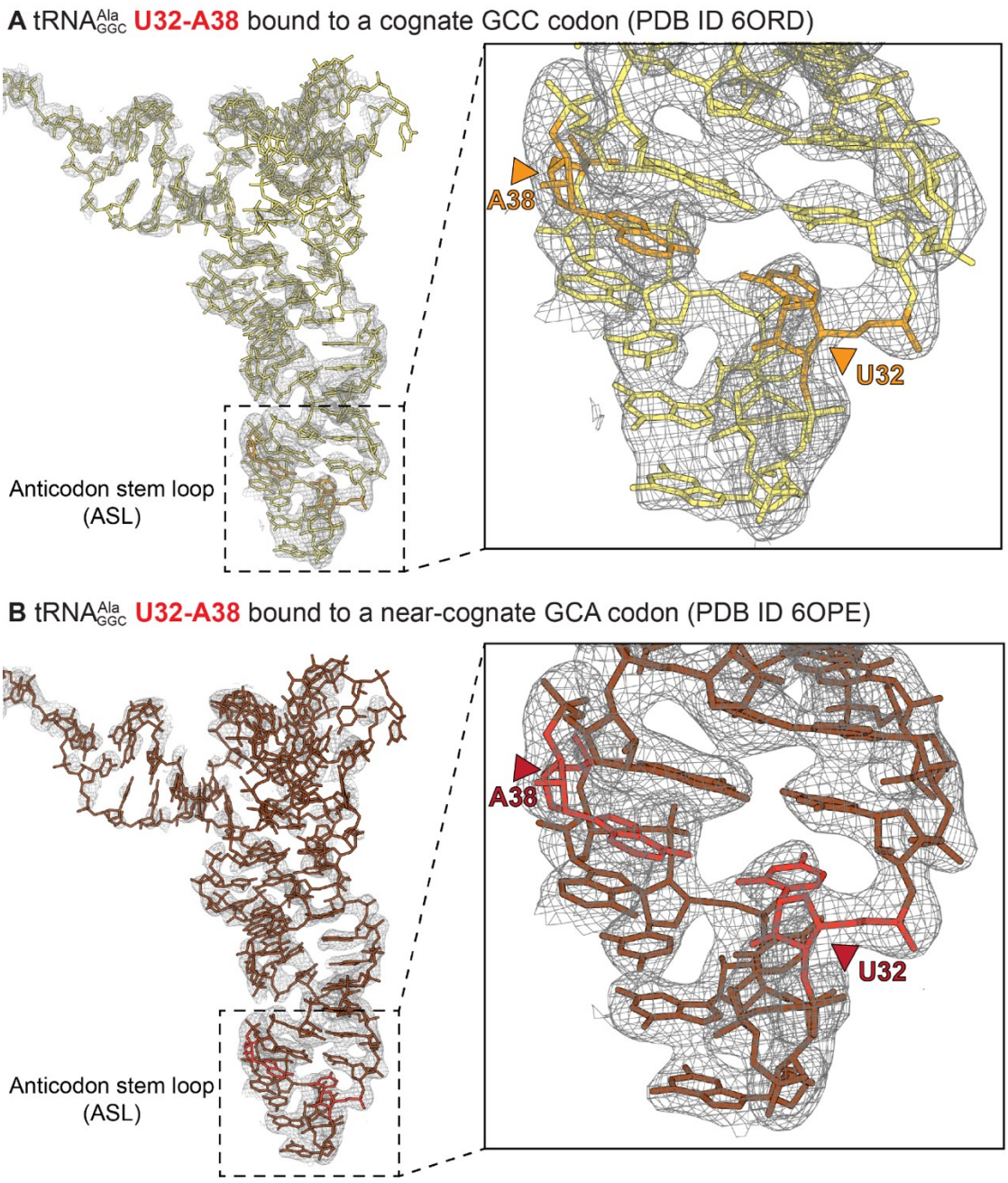


**Figure 3.9. tRNA$^{Ala}_{GGC}$ with the reversed 32-38 pairing shows good electron density of the pairing even when bound to a near-cognate codon.**

***A,*** In the structure of tRNA$^{Ala}_{GGC}$ with the reversed 32-38 pairing bound to a cognate codon in the A site, the tRNA shows good electron density for the whole tRNA and, in particular, for the 32-38 pair (inset). ***B,*** In the structure of 70S-tRNA$^{Ala}_{GGC}$ with the reversed

32-38 pairing bound to a near-cognate codon in the A site, again there is good electron density for the whole tRNA and, in particular, for the 32-38 pair (inset). This is contrast to the structure of 70S with wild-type $tRNA^{Ala}_{GGC}$ where the 32-38 pair shows a lack of electron density in the presence of a near-cognate codon. The $2F_o$-$F_c$ electron density map is contoured at 1σ.

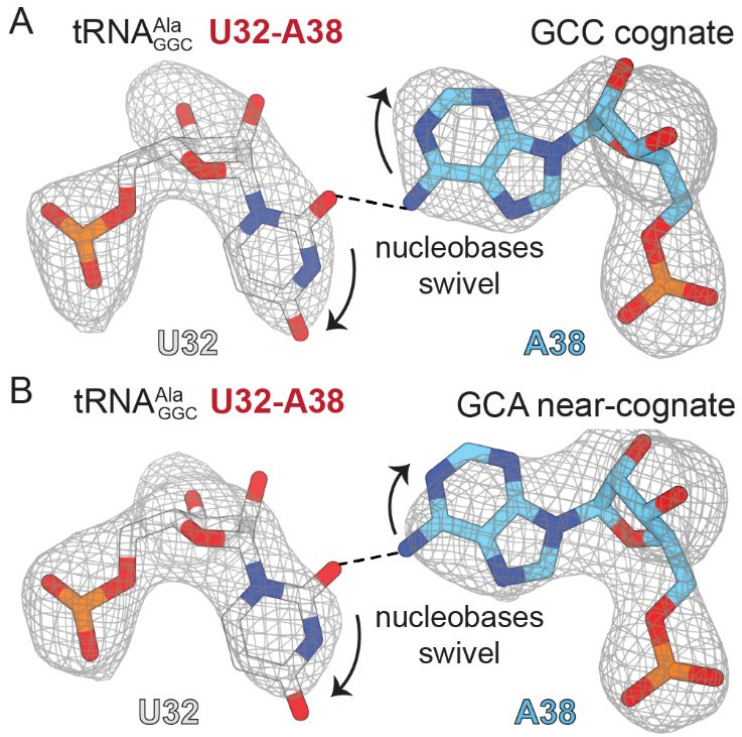


**Figure 3.10. Reversing the 32-38 pair in tRNA$^{Ala}_{GGC}$ changes their orientation.**

***A*,** The structure of 70S-wild-type tRNA$^{Ala}_{GGC}$ A32-U38 interacting with a cognate GCC codon. ***B***, The 70S-tRNA$^{Ala}_{GGC}$ U32-A38 variant interacting with a near-cognate GCA codon. 2F$_O$-F$_C$ electron density maps in gray are contoured at 1.0σ.

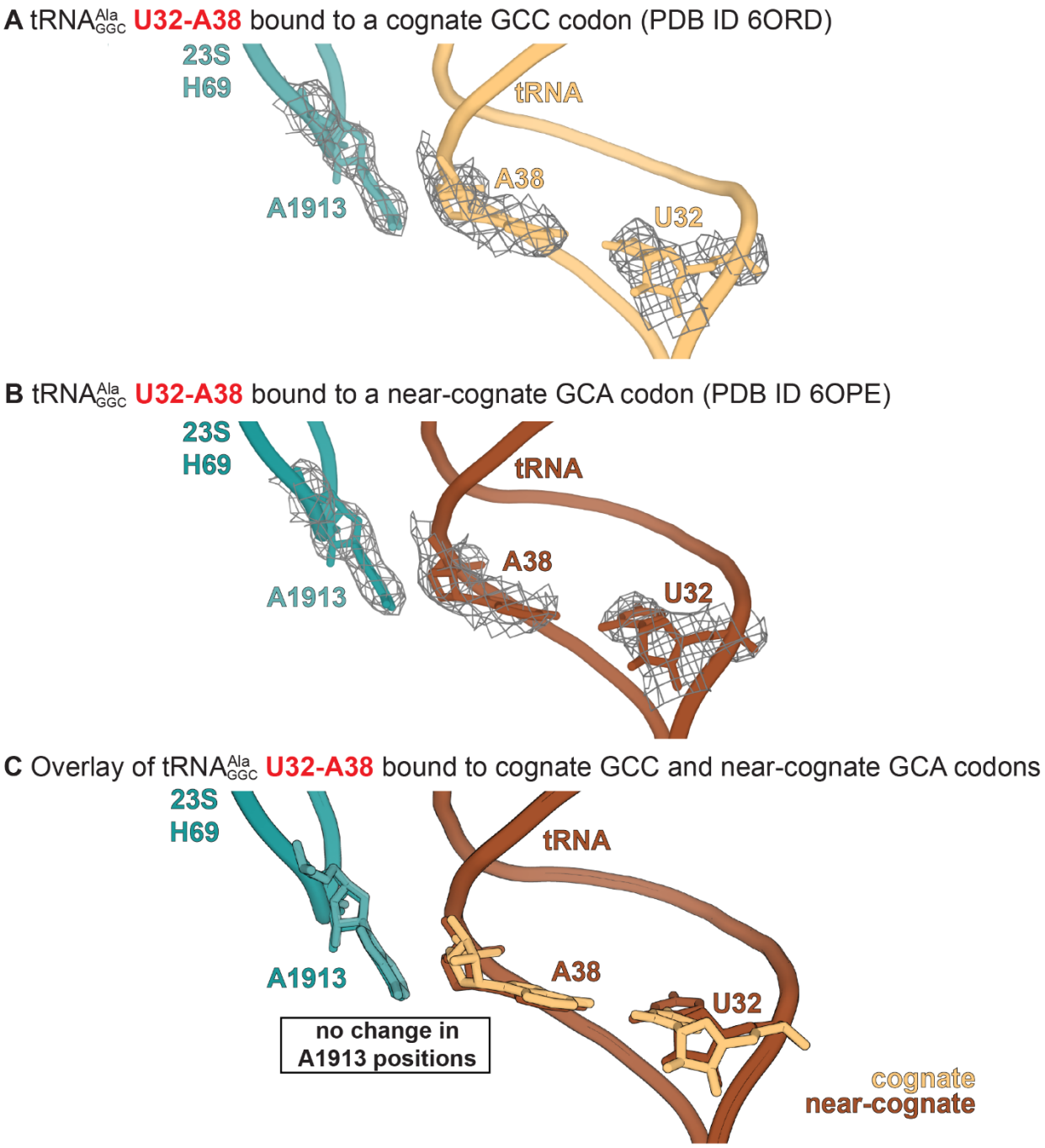


**Figure 3.11. Representative electron density of 23S rRNA A1913 and the reversed 32-38 pairing when bound to a cognate or near-cognate codon.**

***A,*** The 70S-tRNA$^{Ala}_{GGC}$ structure containing the reversed 32-38 pairing shows A1913 packing against the tRNA with good electron density when bound to a cognate codon.

***B,*** When bound to a near-cognate codon, A1913 and the 32-38 pairing also have good electron density. ***C*****,** An overlay of the two structures indicates that the position of A1913 superimpose well. The $2F_o$-$F_c$ electron density map is contoured at 1σ.

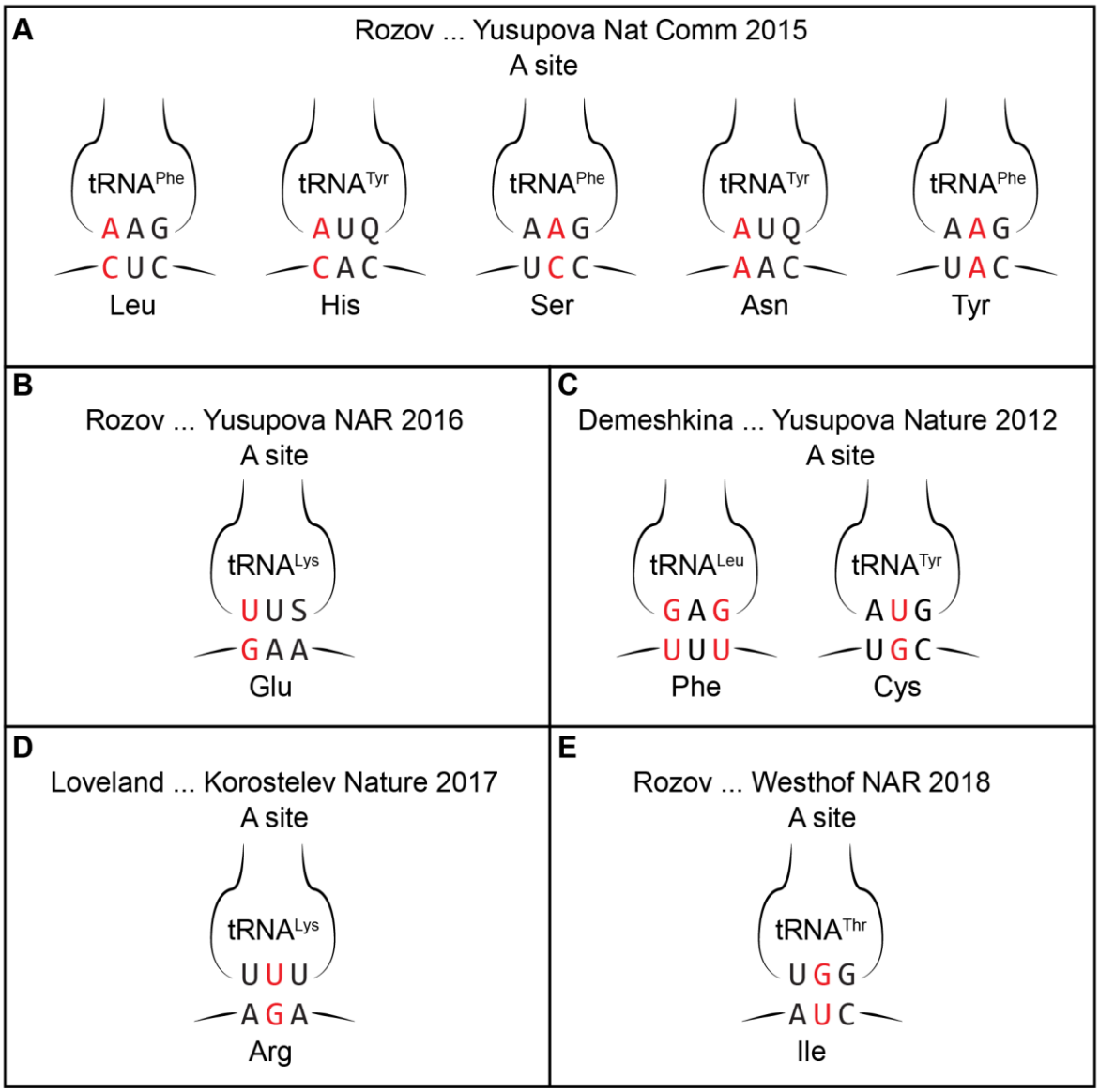


**Figure 3.12. Examples of previously solved structures of anticodon-codon mismatches bound to the ribosome.**

***A,*** A•A and A•C mismatches in the ribosomal A site in the first and second positions do not form stable Watson-Crick base pairs, but H69 A1913 adopts the 'ON' position in all of these cases (5). ***B-E,*** The G•U mismatches have been extensively investigated at the first, second, and third positions in both the P and A sites, and at all positions, a G•U mismatch adopts a Watson-Crick geometry (6-9).

Q: queuosine, S: 5-methylaminomethyl-2-thiouridine, $mnm^5s^2U$

**References**


1. A. Rozov *et al.*, Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat Commun* **7**, 10457 (2016).
2. J. Lin, M. G. Gagnon, D. Bulkley, T. A. Steitz, Conformational changes of elongation factor G on the ribosome during tRNA translocation. *Cell* **160**, 219-227 (2015).
3. R. M. Voorhees, A. Weixlbaumer, D. Loakes, A. C. Kelley, V. Ramakrishnan, Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol* **16**, 528-533 (2009).
4. T. M. Schmeing *et al.*, The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688-694 (2009).
5. A. Rozov, N. Demeshkina, E. Westhof, M. Yusupov, G. Yusupova, Structural insights into the translational infidelity mechanism. *Nat Commun* **6**, 7251 (2015).
6. A. Rozov, E. Westhof, M. Yusupov, G. Yusupova, The ribosome prohibits the G*U wobble geometry at the first position of the codon-anticodon helix. *Nucleic Acids Res* **44**, 6434-6441 (2016).
7. N. Demeshkina, L. Jenner, E. Westhof, M. Yusupov, G. Yusupova, A new understanding of the decoding principle on the ribosome. *Nature* **484**, 256-259 (2012).
8. A. B. Loveland, G. Demo, N. Grigorieff, A. A. Korostelev, Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546**, 113-117 (2017).
9. A. Rozov *et al.*, Tautomeric G*U pairs within the molecular ribosomal grip and fidelity of decoding in bacteria. *Nucleic Acids Res* 10.1093/nar/gky547 (2018).

**Table 3.1. Data collection and refinement statistics**

| | $tRNA^{Ala}_{GGC}$ - cognate codon | $tRNA^{Ala}_{GGC}$ - near-cognate codon | $tRNA^{Ala}_{GGC}$ UA-cognate codon | $tRNA^{Ala}_{GGC}$ UA-near-cognate codon |
|---|---|---|---|---|
| PDB ID | 6OF6 | 6OJ2 | 6ORD | 6OPE |
| **Data collection** | | | | |
| Wavelength (Å) | 1.0000 | 1.0000 | 0.9792 | 0.9791 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell dimensions | | | | |
| *a, b, c (Å)* | 211.3, 452.3, 626.5 | 211.0, 453.5, 625.4 | 209.4, 445.8, 616.0 | 209.4, 449.8, 620.0 |
| *α, β, γ (°)* | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution (Å) | 43.9-3.2 (3.3-3.2) | 50-3.2 (3.3-3.2) | 49.9-3.1 (3.2-3.1) | 49.6-3.1 (3.2-3.1) |
| $R_{pim}$ (%) | 7.4 (34.4) | 11.0 (74.0) | 12.7 (83.7) | 14.1 (77.0) |
| $I/\sigma I$ | 8.26 (2.04) | 4.88 (0.98) | 4.57 (0.71) | 6.29 (1.12) |
| $CC_{1/2}$ | 0.993 (0.645) | 0.995 (0.329) | 0.991 (0.183) | 0.989 (0.387) |
| Completeness (%) | 91.15 (91.01) | 99.52 (99.92) | 95.19 (91.69) | 97.13 (99.13) |
| Redundancy | 6.6 (6.0) | 5.8 (5.6) | 3.0 (2.2) | 3.0 (3.0) |
| **Refinement** | | | | |
| No. reflections | 889,696 | 971,076 | 981,305 | 1,016,419 |
| $R_{work}/R_{free}$ (%) | 19.5/23.2 | 23.6/27.0 | 23.2/26.8 | 22.5/26.6 |
| No. atoms | | | | |
| Macromolecules | 297,356 | 297,318 | 290,854 | 294,030 |
| Ligands | 1,319 | 1,053 | 2,965 | 951 |
| B-factors | | | | |
| Macromolecules | 90.33 | 129.71 | 91.34 | 87.33 |
| Ligands | 56.13 | 130.73 | 60.81 | 59.98 |
| R.m.s. deviations | | | | |
| Bond lengths (Å) | 0.005 | 0.006 | 0.008 | 0.012 |
| Bond angles (°) | 0.87 | 1.07 | 0.99 | 1.25 |

*Values in parentheses are for highest-resolution shell.

### 3.8. Acknowledgments

We thank the Dunham lab members Drs. Jack Dunkle and Eric Hoffer for their technical assistance and helpful scientific insights, Dr. Graeme L. Conn for critical reading of the manuscript, and the staff members of the NE-CAT and SER-CAT beamlines for assistance during data collection. This work was supported by National Institutes of Health Grant R01 GM093278. X-ray crystallography data sets were collected at the NE-CAT beamlines (funded by NIGMS, National Institutes of Health, Grant P30 GM124165), using a Pilatus detector (RR029205) and an Eiger detector (OD021527), and at the SER-CAT beamlines (funded by its member institutions and National Institutes of Health Equipment Grants S10_RR25528 and S10_RR028976). This research used resources of the APS, a United States Department of Energy Office of Science User Facility operated by Argonne National Laboratory under Contracts DE-AC02-06CH11357 (NE-CAT) and W-31-109-Eng-38 (SER-CAT).

### 3.9. References

1. Crick FH (1958) On protein synthesis. Symp Soc Exp Biol 12:138-163.
2. Uhlenbeck OC & Schrader JM (2018) Evolutionary tuning impacts the design of bacterial tRNAs for the incorporation of unnatural amino acids by ribosomes. Curr Opin Chem Biol 46:138-145.
3. Fahlman RP, Dale T, & Uhlenbeck OC (2004) Uniform binding of aminoacylated transfer RNAs to the ribosomal A and P sites. Mol Cell 16(5):799-805.
4. Ledoux S & Uhlenbeck OC (2008) Different aa-tRNAs are selected uniformly on the ribosome. Mol Cell 31(1):114-123.
5. Louie A, Ribeiro NS, Reid BR, & Jurnak F (1984) Relative affinities of all Escherichia coli aminoacyl-tRNAs for elongation factor Tu-GTP. J Biol Chem 259(8):5010-5016.
6. Schuette JC, et al. (2009) GTPase activation of elongation factor EF-Tu by the ribosome during decoding. EMBO J 28(6):755-765.

7. Ledoux S, Olejniczak M, & Uhlenbeck OC (2009) A sequence element that tunes Escherichia coli tRNA(Ala)(GGC) to ensure accurate decoding. Nat Struct Mol Biol 16(4):359-364.

8. Olejniczak M, Dale T, Fahlman RP, & Uhlenbeck OC (2005) Idiosyncratic tuning of tRNAs to achieve uniform ribosome binding. Nat Struct Mol Biol 12(9):788-793.

9. Olejniczak M & Uhlenbeck OC (2006) tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. Biochimie 88(8):943-950.

10. Gromadski KB, Daviter T, & Rodnina MV (2006) A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. Mol Cell 21(3):369-377.

11. Murakami H, Ohta A, & Suga H (2009) Bases in the anticodon loop of tRNA(Ala)(GGC) prevent misreading. Nat Struct Mol Biol 16(4):353-358.

12. Gabriel K, Schneider J, & McClain WH (1996) Functional evidence for indirect recognition of G.U in tRNA(Ala) by alanyl-tRNA synthetase. Science 271(5246):195-197.

13. Selmer M, et al. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313(5795):1935-1942.

14. Ogle JM, Murphy FV, Tarry MJ, & Ramakrishnan V (2002) Selection of tRNA by the ribosome requires a transition from an open to a closed form. Cell 111(5):721-732.

15. Demeshkina N, Jenner L, Westhof E, Yusupov M, & Yusupova G (2012) A new understanding of the decoding principle on the ribosome. Nature 484(7393):256-259.

16. Loveland AB, Demo G, Grigorieff N, & Korostelev AA (2017) Ensemble cryo-EM elucidates the mechanism of translation fidelity. Nature 546(7656):113-117.

17. Auffinger P & Westhof E (1999) Singly and bifurcated hydrogen-bonded base-pairs in tRNA anticodon hairpins and ribozymes. Journal of molecular biology 292(3):467-483.

18. Westhof E & Fritsch V (2000) RNA folding: beyond Watson-Crick pairs. Structure 8(3):R55-65.

19. Maehigashi T, Dunkle JA, Miles SJ, & Dunham CM (2014) Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. Proc Natl Acad Sci U S A 111(35):12740-12745.

20. Fagan CE, Maehigashi T, Dunkle JA, Miles SJ, & Dunham CM (2014) Structural insights into translational recoding by frameshift suppressor tRNASufJ. RNA 20(12):1944-1954.

21. Ortiz-Meoz RF & Green R (2011) Helix 69 is key for uniformity during substrate selection on the ribosome. J Biol Chem 286(29):25604-25610.

22. Ali IK, Lancaster L, Feinberg J, Joseph S, & Noller HF (2006) Deletion of a conserved, central ribosomal intersubunit RNA bridge. Mol Cell 23(6):865-874.

23. Agris PF, Narendran A, Sarachan K, Vare VYP, & Eruysal E (2017) The Importance of Being Modified: The Role of RNA Modifications in Translational Fidelity. Enzymes 41:1-50.

24. Vare VY, Eruysal ER, Narendran A, Sarachan KL, & Agris PF (2017) Chemical and Conformational Diversity of Modified Nucleosides Affects tRNA Structure and Function. Biomolecules 7(1).

25. Rozov A, Westhof E, Yusupov M, & Yusupova G (2016) The ribosome prohibits the G*U wobble geometry at the first position of the codon-anticodon helix. Nucleic Acids Res 44(13):6434-6441.

26. Rozov A, et al. (2016) Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. Nature communications 7:10457.

27. Manickam N, Nag N, Abbasi A, Patel K, & Farabaugh PJ (2014) Studies of translational misreading in vivo show that the ribosome very efficiently discriminates against most potential errors. RNA 20(1):9-15.

28. Linpinsel JL & Conn GL (2012) General protocols for preparation of plasmid DNA template, RNA in vitro transcription, and RNA purification by denaturing PAGE. Methods Mol Biol 941:43-58.

29. Kabsch W (2010) Xds. Acta crystallographica 66(Pt 2):125-132.

30. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica 66(Pt 2):213-221.

31. Polikanov YS, Melnikov SV, Soll D, & Steitz TA (2015) Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. Nat Struct Mol Biol 22(4):342-344.

32. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta crystallographica 67(Pt 4):355-367.

33. Potterton L, et al. (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. Acta Crystallogr D Struct Biol 74(Pt 2):68-84.

34. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta crystallographica 60(Pt 12 Pt 1):2126-2132.

35. Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

# Chapter 4.

## Second position P-site codon-anticodon mismatch mispositions the downstream mRNA

Ha An Nguyen, Eric D. Hoffer, Tatsuya Maehigashi, Crystal E. Fagan
and Christine M. Dunham

This research is unpublished and is currently in preparation for journal submission.

**Author contributions:** T.M., C.E.F., and C.M.D designed research; H.A.N., T.M., and C.E.F. performed the research; all authors analyzed the data; H.A.N and C.M.D. wrote the paper with input from all authors.

## 4.1. Abstract

Rapid and accurate translation is essential in all organisms to produce properly folded and functional proteins. mRNA codons that define the protein coding sequences are decoded by tRNAs on the ribosome in the aminoacyl (A) binding site. The mRNA codon and the tRNA anticodon interaction is extensively monitored by the ribosome to ensure accuracy in tRNA selection. While other polymerases that synthesize DNA and RNA can correct for misincorporations, the ribosome is unable to correct mistakes. Instead, if a misincorporation occurs, the mismatched tRNA-mRNA pair moves to the peptidyl (P) site and from this location, causes a reduction in the fidelity at the A site. This reduced fidelity allows for either additional mismatched tRNAs to be accepted or release factors to recognize sense codons and hydrolyze the aberrant peptide. Here, we solved crystal structures of the ribosome containing a mismatched $tRNA^{Lys}$ in the P site interacting with U·U nucleotide mismatches in the codon. We find that when the mismatch occurs in the second position of the P-site codon-anticodon interaction, the first mRNA nucleotide of the A-site codon flips from the mRNA path to engage 16S rRNA nucleotides that comprise the decoding center. This mRNA nucleotide mispositioning may be the cause for the reduction of fidelity at the A site. Further, this may provide an opportunity for release factors to bind to sense codons triggering premature termination to hydrolyze the erroneous nascent chain before it can disrupt the cellular proteome.

## 4.2. Introduction

The accurate flow of genetic information is vital for cellular life. DNA and RNA polymerases copy nucleic acid templates into complementary nucleic acids and Watson-Crick base pairing between these nucleotide strands guides accuracy. The thermodynamic differences between base pairings alone cannot fully account for the exceptional accuracy of replication (~$10^{-9}$) and transcription (~$10^{-5}$) (1). To accomplish such high accuracy, both DNA and RNA polymerases have the ability to detect misincorporations, excise the incorrect nucleotide and replace the incorrect nucleotide with the correct nucleotide. This proofreading mechanism allows for the continuous replication or transcription without having to discard the current product and restart. In contrast, during protein synthesis, the incorporation of an incorrect amino acid is irreversible because the mRNA codon template specifies an amino acid product and thus, Watson-Crick base pairing cannot be monitored by the ribosome to remove the incorrectly added product. An additional challenge is the large distance of ~70 Å between the mRNA-tRNA base pairing in the decoding center on the small ribosomal subunit and the peptidyl transfer center on the large ribosomal submit where aminoacyl groups attached to tRNAs are added to the nascent chain. This distance prevents a rapid response (2, 3). Collectively, these differences may account for much higher error rates in protein synthesis (4).

Despite having a ~$10^{-3}$ error rate, the bacterial ribosome still maintains sufficient protein synthesis fidelity to maintain a functional proteome. Using both kinetic proofreading and induced-fit mechanisms, the ribosome rapidly selects the correct tRNA substrate from incorrect but structural similar tRNAs (5-9). Ternary complexes containing

aminoacyl-tRNAs (aa-tRNAs), EF-Tu and GTP (aa-tRNA•EF-Tu•GTP) are delivered to the aminoacyl (A) site of the ribosome and encounter two kinetic checkpoints before acceptance. First, Watson-Crick base pairing between the codon and anticodon is inspected during a process called initial codon selection. The ribosomal A site has an extensive monitoring network: ribosomal RNA (rRNA) nucleotides G530, C1054, A1492, A1493 and A1913 (*E. coli* numbering) undergo conformational changes to directly monitor the pairing of the codon-anticodon on the small 30S subunit (Fig. 4.1) (10-12). The first two positions of the codon-anticodon interaction are required to be Watson-Crick (A-U or G-C) due to the constraints by the A site while the third position can either be Watson-Crick or an interaction that resembles the geometry of a Watson-Crick pairing (e.g. G•U pairing or the pairing of a modified anticodon nucleotide with an mRNA nucleotide). The complementarity of the codon-anticodon interaction stabilizes the ternary complex while incorrect aa-tRNAs rapidly dissociate (13, 14). Second, a correct Watson-Crick base pair causes rapid hydrolysis of GTP by EF-Tu while incorrect pairings induce much slower GTP hydrolysis and EF-Tu disassociation (15). Rapid hydrolysis also enables conformational changes of the ternary complexes leading to full loading of aa-tRNAs on the large 50S ribosomal subunit. These steps ensure high accuracy and speed during protein synthesis.

Despite existing proofreading mechanisms during tRNA selection, missense errors still occur *in vivo* at a rate of one in ~3000 amino acid incorporated (16), which is notably lower than the *in vitro* error rate of 1 in 500 codons (1). This discrepancy implies additional quality control processes beyond tRNA selection in the A site contributes to fidelity in protein synthesis. The discovery of a post-peptidyl transfer quality control mechanism

(post-PT QC) revealed that codon-anticodon mismatches that bypass A-site surveillance mechanisms are subsequently detected in the peptidyl (P) site (17). Using the well-known *in vivo* misincorporation event where $tRNA^{Lys}$ (anticodon is UUU, all anticodons and mRNAs are depicted in the 5' to 3' direction) decodes the near-cognate asparagine AAU codon (18), once the mismatched codon-anticodon pairing moved into the P site, a subsequent loss in fidelity at the A site ensued (17, 19, 20) (near cognate is defined as a single mismatch between the codon and the anticodon). This loss of fidelity manifests as a stabilization of either the binding of incorrect tRNAs or release factors (RF) on non-stop or sense codons in the A site (1). Near-cognate tRNA binding in this context leads to accelerated rates of GTP hydrolysis by EF-Tu with the tRNA accommodated at similar rates to those of correct tRNAs (19). RFs recognition of sense codons triggers premature termination that is two orders of magnitude higher after a single misincorporation event (incorrect mRNA-tRNA pairing in the P site) and four orders of magnitude higher after two consecutive misincorporation events (incorrect mRNA-tRNA pairings now located in both the P and E sites) (1). Interestingly, different mismatch types elicit different post PT QC responses, and it is not clear how the ribosome discriminates between these slight differences in mismatch errors. For example, first or second codon position G•U and U•U mismatches activate post PT QC mechanisms that release the nascent chain in a similar manner. However, at the third codon position, a G•U mismatch does not trigger peptidyl release while a U•U mismatch does. In addition, the magnitude of the post PT QC activation and premature termination is also varied with mismatches at the second codon position having the largest increase in the rate of hydrolysis (1).

Unlike in the A site, the codon-anticodon helix in the ribosomal P site is not stringently monitored for Watson-Crick base complementarity. Instead, the P site is optimized for the recognition of initiator $tRNA^{fMet}$ and the gripping of elongator tRNAs to ensure the correct positioning and movement of the tRNA. Various rRNA bases and ribosomal proteins form a network of interactions with the P-site tRNA but there is little inspection of the anticodon. 16S rRNA nucleotides G966 and C1400 form a bridging interaction beneath the third position of the codon-anticodon while other 16S rRNA nucleotides stabilize the P-site tRNA by interacting with the anticodon stem (21). G1338 and A1339 grip base pairs in the anticodon stem to distinguish between initiator and elongator tRNAs and prevent tRNA translocation from the P to the exit (E) site prematurely (22, 23). C-terminal tails of ribosomal proteins uS9 and uS13 also contact the tRNA anticodon stem loop (ASL) backbone (Fig. 4.1). These interactions collectively stabilize already selected tRNAs, but it is not clear how codon-anticodon mismatches are recognized given the minimal interactions of the P site with the codon-anticodon helix.

To decipher how the ribosome recognizes mismatches in the P site, we solved two structures of 70S ribosomes bound with U•U mismatches in the first and second codon-anticodon positions in the P site (Fig. 4.2, Table 4.1). The mismatches were in the codons- we changed the lysine AAA codon to either UAA or AUA, both codons that $tRNA^{Lys}$ miscodes (Kramer Farabaugh). We find that the second U•U mismatch in the codon-anticodon interaction causes the first nucleotide of the A-site mRNA to flip ~90° away from its normal position in the mRNA path. This movement leaves the A-site codon with only two of the three nucleotides properly positioned to interact with either incoming tRNAs or RFs. In contrast, the first position codon-anticodon mismatch in the P site does

not influence the position of the A-site codon. These results are consistent with biochemical studies that demonstrate the second position mismatch triggers post PT QC mechanisms at much higher levels than mismatches at the first or the third position of the P-site codon-anticodon interaction (1). Mispositioning of the A-site codon could serve as a signal to trigger post PT QC mechanisms and initiate RF2-mediated hydrolysis on non-stop or sense codons resulting in the termination of erroneous protein synthesis and recycling of ribosomes required for cell survival.

### 4.3. Results

**Single codon-anticodon nucleotide mismatches minimally impact the architecture of the P site.** We solved two crystal structures of *T. thermophilus* ribosomes containing a U•U mismatch in the P site at the first and second positions of the codon-anticodon interaction to understand how these mismatches trigger different post PT QC responses once they have moved into the P site (Fig. 4.2). These ribosome complexes also contain $ASL^{Phe}$ bound to a cognate codon in the A site. Similar to previous structures of G•U codon-anticodon mismatches in the P site (24), the overall architecture of the P site reveals no apparent changes as observed in both structures. The P site adopts a similar structure to when a cognate mRNA-tRNA pair is present: 16S rRNA nucleotides C1400 and G966 form a bridge and stack with the third base-pair of the codon-anticodon interaction, G1338 and G1339 form A-minor interactions with the tRNA above the anticodon stem loop, and uS9 and uS9 tails extend into the P site (Fig. 4.3). The P-site tRNAs also adopt very similar structures to each other, regardless of the position of the mismatch (Fig. 4.4). Although there is a nonstandard U•U base pairs, both U•U base pairs mimic the geometry of a Watson-Crick base pair (Fig. 4.2). In contrast to the various

structural reorganization observed in the A site in response to mismatches at the first or second codon-anticodon positions (10, 11, 25), the ribosomal P site is not observed to undergo any structural reorganization when a mismatch is present.

**Second nucleotide U•U mismatch in the codon-anticodon located in the P site causes the first nucleotide of the mRNA codon to deviate from the path.** The mRNA path located on the 30S is surrounded by 16S rRNA nucleotides and critical magnesium ions (21, 26). The E, P, and A sites each accommodate one three-nucleotide codon (Fig. 4.5A). The E site has minimal interactions with the mRNA, with non-sequence specific contacts established by nucleotides 16S rRNA G693 and G926 with the first and third positions of the E-site codon. In the P site, C1402, C1403, and U1498 form interactions with the mRNA backbone stabilizing its position and as mentioned, the C1400-G966 bridge engages the third base pair of the codon-anticodon interaction. The mRNA undergoes a $Mg^{2+}$-mediated kink between the P and the A sites and rRNA. Decoding site rRNA nucleotides A1492, A1493, and G530 probe the first and second codon-anticodon base pairs directly to check for Watson-Crick pair, and C1054 stacks with the third base pair. In both structures presented here, the backbone of the mRNA in the P site follows a similar path seen in all previously determined bacterial ribosome structures to date (Fig. 4.5A). In summary, all the 30S contacts with mRNA are maintained in our structures with each codon properly positioned in the E, P, and A sites.

In our 70S structure containing a P-site U•U mismatch at the first position of the codon-anticodon ($U_{+1}$•$U_{36}$), we find that the interaction between the codon-anticodon in the A site contains three Watson-Crick base pairs ($A_{36}$-$U_{+4}$, $A_{35}$-$U_{+5}$, $G_{34}$-$C_{+6}$) (Fig. 4.5B). However, when the U•U mismatch is located at the second position of the codon-

anticodon pairing ($U_{+2}$•$U_{35}$), the first nucleotide of the mRNA codon in the A site ($U_{+4}$) deviates from its normal positioning and is flipped out from the mRNA path (Figs. 4.5C, 4.6). This mispositioning thereby disrupts base pairing with $A_{36}$ of the A-site $tRNA^{Phe}$ (we used the 18-nucleotide anticodon stem loop (ASL) of $tRNA^{Phe}$ and CC-puromycin for technical reasons as described in the Methods section). Specifically, $U_{+4}$ flips ~90° away from $A_{36}$ and instead interacts with 16S nucleotide A1493 of the decoding center: the 2'-OH of $U_{+4}$ forms hydrogen bonds with the 2'-OH and the phosphate oxygen of A1493 (Figs. 4.5, 4.6). This movement of $U_{+4}$ out of the normal mRNA path in the A site is surprising for several reasons. First, the codon-anticodon interaction is cognate at the A site and three Watson-Crick base pairs should form. Second, when there is a $U_{+1}$•$U_{36}$ mismatch in the first position of the P site, the mRNA path in the A site is unaffected (Fig. 4.2B). Biochemical characterization of the post PT QC response reveals a much more robust response to second position mismatches (17) consistent with our structures presented here. The lack of fidelity at the A site due to the second position $U_{+1}$•$U_{36}$ mismatch is likely the result of the mispositioning of the first nucleotide of the A-site codon. This impaired mRNA presentation in the A site may allow binding of any ligand including non-cognate tRNAs and release factors and ultimately, induce termination to stop erroneous protein synthesis.

### 4.4. Discussion

In this study, we examine the structural basis for how P-site U•U mismatches affect the fidelity of the adjacent A site. It was previously shown that when an incorrect codon-anticodon pairing escapes rejection at the A site and its aminoacyl group is added to the growing nascent chain, quality control mechanisms exist to reduce fidelity at the A site

(17, 19). This remarkable reduction in fidelity allows for non- and near-cognate tRNAs to be kinetically accepted as cognate by the ribosome and for RFs to recognize non-stop codons and hydrolyze the nascent chain. Normally, incorrect near-cognate tRNAs rapidly dissociate from the ribosome during the initial codon recognition in the A site (9). However, incorrect mRNA-tRNA pairings in the P site stabilize incorrect tRNAs at the A site leading to higher rates of miscoding (19). Successive misincorporations lead to an accumulation of errors and ultimately premature termination. Hydrolysis of this erroneous peptide is necessary to stop the synthesis of aberrant polypeptides and for the ribosome to be recycled for further rounds of translation.

While strict Watson-Crick base pair complementarity is required at the first two positions of the codon-anticodon pairing, the third or wobble position have some degree of plasticity and can accept either a G•U pair or a pairing between a modified nucleotide of the tRNA anticodon and the mRNA codon (24, 27, 28). In both cases, the pairing adopts a Watson-Crick-like orientation that permits the acceptance of the tRNA by the ribosome and thus expands the number of codons that tRNAs recognize. Much research has focused on $tRNA^{Lys}$ because of its well-known propensity to miscode in vivo on asparagine AAU codons (29). Structures of A-site bound $tRNA^{Lys}$ carrying U•U mismatch at either the first, second, and third position of the codon-anticodon interaction on the ribosome demonstrate that the U•U mismatches adopt Watson-Crick-like conformations (24). Luciferase reporter assays quantifying every possible misreading error by $tRNA^{Lys}$ on a codon revealed that error rates vary widely across the different types of mismatches with no clear trend of misreading propensity based on the position or the type of mismatch, only with certain codon contexts where there is an increased error propensity

by $tRNA^{Lys}$ on arginine, asparagine, and termination codons (29). Expanding the work to other tRNAs suggests that miscoding rates are higher in U•U mismatches than other base pairs (30). Since the ribosomal A site heavily constraints the structure of the ASL, perhaps the smaller pyrimidine-pyrimidine mismatch is able to bypass the A-site proofreading mechanisms better because it does not induce steric clashes or cause unfavorable backbone torsions as a wider pyrimidine-purine or purine-purine mismatch would. Since uridine is the smallest base, the U•U mismatch would be the least destabilizing to the tRNA interactions with the mRNA and the ribosome (31). In our structures of the U•U mismatches in the P site, we also observe that the U•U pairing adopts a wider Watson-Crick-like pairing. Further, U•U mismatches do not cause major perturbations in the shape of the anticodon stem loop suggesting their influence appears to be localized to the codon-anticodon (Figs. 4.3, 4.4, 4.7).

It is not clear why we see mRNA mispositioning in the case of the second position mismatch, but not the first. Both codons used in this study, the UAA ‘ochre’ stop codon and the AUA isoleucine codon are naturally misread by $tRNA^{Lys}$ at a higher frequency than others (29). $tRNA^{Lys}$ misreading on the AUA codon was suggested to be possible because the authors predict steric clashes would be minimal due to the pyrimidine-pyrimidine interaction (31). UAA is a stop codon where there is no competition with a cognate tRNA. The aminoglycoside paromomycin increases the error rate for the UAA stop codon only and streptomycin induces higher error rates for both UAA and AUA codons indicating a differential misreading propensity for when the codon-anticodon mismatch is in the first or the second position (29). When we look at the importance of the first vs the second position degeneracy for all the codons, the second position

maximal accuracy was highest out of the three positions for tRNA selection (32), and the second position base pairing stability was found to be the determining factor affecting the degree third position degeneracy (33). This suggests a potential explanation for the base flipping observed in our structure with the more impactful second position U•U mismatch but not for the structure with the mismatch in the first position. This observation also correlates with previous results where the rates of peptidyl-tRNA drop-off and peptide hydrolysis are the highest when the U•U mismatch is in the second position as compared to when the mismatch is in the first or third positions (17).

Protein synthesis normally terminates when the ribosome reaches a stop codon (UAA, UAG, or UGA) in the reading frame of the mRNA. In bacteria, release factors RF1 (specific for UAG and UAA) and RF2 (specific for UGA and UAA) recognize the stop codons in the A site and hydrolyses the nascent chain on the P-site peptidyl tRNA to release the newly synthesized protein from the ribosome. The two release factors have different recognition mechanisms for their target stop codons but in *E. coli* only RF2 has noncanonical termination activity, participating in both post PT QC activity (17) and in ribosome rescue on truncated mRNAs (34). Time-resolved cryo-EM structures show that RFs are compact in solution then extend upon binding to the stop codon to reach the peptidyl transferase center for hydrolysis activity (35, 36). Unlike the Watson-Crick interactions of the aminoacyl-tRNA with the A-site mRNA codon, release factors interact with both the Watson-Crick and the Hoogsteen edges on the mRNA nucleotides (37). Release factors form extensive interactions with the first two nucleotides of the stop codon while the third mRNA codon nucleotide flips out of the three-nucleotide codon stack. Both RF1 and RF2 bind to the first uracil in the stop codon with a hydrogen bonding network

mimicking the A-U Watson Crick interaction, accounting for the necessity of the first position uracil in all three stop codons, but RF1 and RF2 have different second position recognition mechanisms that allows RF2 to have distinct activities from RF1. Our structures also have a first position U in a phenylalanine UUC codon, but it is not clear how release factors interact with sense codons. There is a ribosome structure of RF1 binding to a UAU (Tyr) codon and RF1 is able to recognize the tyrosine UAU codon similarly to the UAG stop codon because of their structural similarity (38). Our structures also have an ASL occupying the A site, so the mRNA nucleotides are primed to adopt the three-nucleotide stacking interaction with the tRNA anticodon. The A-site mRNA mispositioning in response to the second position U•U mismatch in the P-site destabilizes tRNA binding and contributes to the loss of fidelity of the A site, but we don't know how RF2 can interact with this UUC codon. The second position uracil in this phenylalanine sense codon should be very unfavorable for release factor binding, but RF2 clearly recognizes this context for increased post PT QC (17). In the absence of an RF2 bound to the A-site sense codon and a P site containing a second position U•U mismatch ribosome structure, it is difficult to postulate the downstream effects that connect the loss of the A-site fidelity and premature translation termination.

### 4.5. Materials and Methods

*Thermus thermophilus* 70S ribosomes were purified from cell pellets by fractionation over a sucrose gradient as previously described (21). Ribosome complexes were formed sequentially by adding 70S ribosomes, messenger RNA (IDT), CC-puromycin, P-site $tRNA^{Lys}$ (Chemical Block), and $ASL^{Phe}$ (IDT) in buffer (5 mM HEPES-KOH pH 7.5, 50 mM KCl, 10 mM $NH_4Cl$, 10 mM $Mg(CH_3COO)_2$, 6 mM β-mercaptoethanol (β-Me)) at 55° C.

CC-puromycin is an RNA-antibiotic conjugate that mimics the CCA-end of the ribosome. The initial addition of CC-puromycin ensures $tRNA^{Lys}$ positioning in the P site. Ribosome crystals were grown in 0.1 M Tris-HOAc pH 7.0, 0.2 M KSCN, 4-4.5% (v/v) PEG 20K, 4.5-5.5% (v/v) PEG 550MME, 10 mM $Mg(OAc)_2$, 2.8 μM Deoxy BigCHAP. X-ray diffraction data sets were collected at the 24ID-C beamline at the Advanced Photon Source at Argonne National Laboratory. Datasets were processed in XDS and PHENIX, with molecular replacement performed using PDB code 4Y4O. The coordinated were finalized through iterative rounds of refinements in PHENIX (39) and model building in Coot (40). Figures were created in PyMol (41).

**Data deposition:** X-ray crystallography, atomic coordinates, and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB codes 7N8Z, 7N90)

## 4.6. Figures and Tables

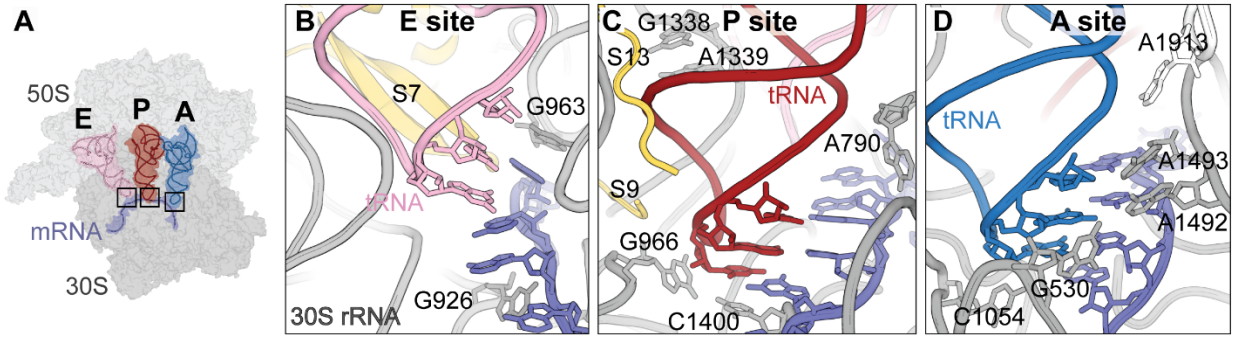


**Figure 4.1. The ribosome extensively probes the tRNA-mRNA interaction at the A site but does not monitor P- and E-site tRNA-mRNA base pairing.**

(**A**)**.** An overview of the tRNA binding sites on the 70S ribosome. (**B**) **E site**: as the tRNA (pink) is about to exit the ribosome, it makes minimal contacts with the mRNA (purple) and the small subunit of the ribosome (gray). The anticodon stem loop (ASL) of the tRNA contacts the β-hairpin at the C-terminus of 30S ribosomal protein S7 (yellow). 30S rRNA bases G963 and G926 interacts with the mRNA but have not been shown to participate in proofreading. (**C**) **P site**: the tRNA (red) is gripped by two 30S rRNA loops: G1338-A1339 and A790, and the C-termini tails of proteins S9 and S13. The G1338, A1339, and A790 rRNA bases are critical for P-site tRNA binding, controls the translocation of tRNA from the P to the E site, and is essential for translation *in vivo*. Additionally, G966 packs against the backbone of the P-site ASL, and C1400 base stacks with tRNA nucleotide 34 and play critical roles in P-site binding. All mentioned P-site ribosomal bases are universally conserved in bacteria. (**D**) **A site**: tRNAs enter the ribosome first at the A site, and the ribosome monitors the tRNA-mRNA base pairing extensively. 16S rRNA bases A1492, A1493, C1054 and G530 probe the anticodon-

codon interaction and incorrect tRNAs are rejected. Figure was generated using a crystal structure of Thermus thermophilus ribosome with tRNAs bound (PDB ID 6OF6).

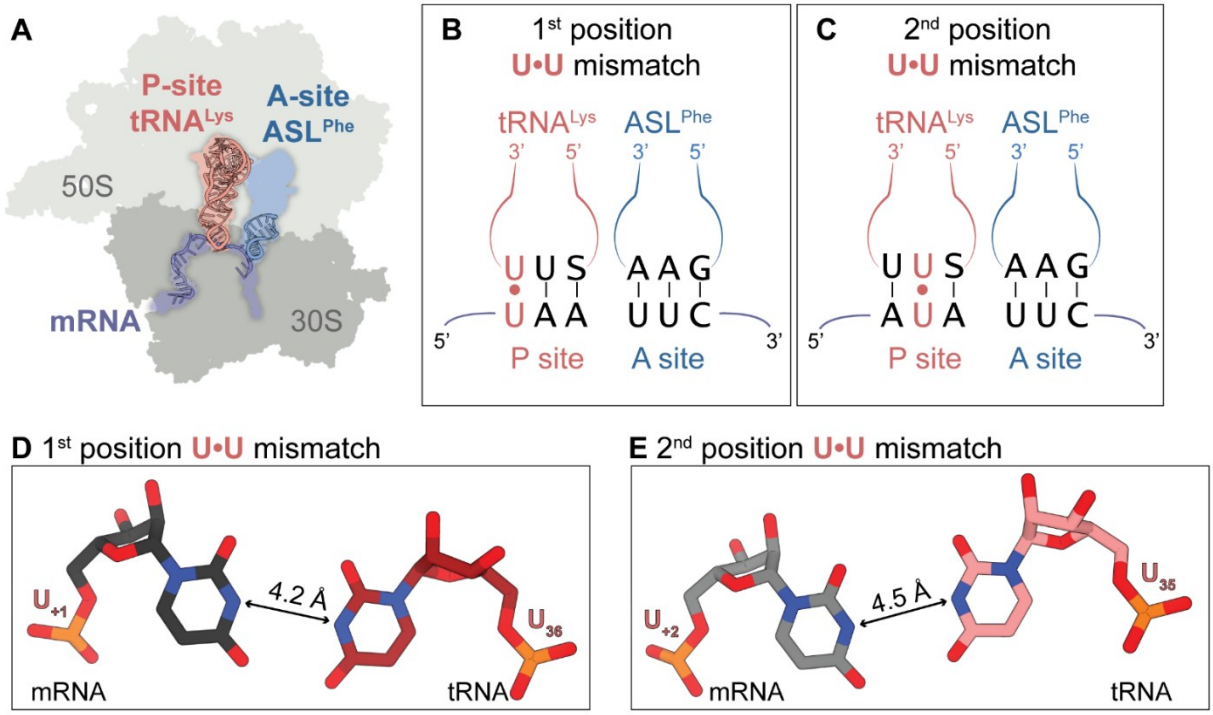


**Figure 4.2. Structures of the 70S ribosome containing codon-anticodon mismatches in the peptidyl (P) site.**

(**A**) An overview of the 70S ribosome complexes in this study showing the mRNA in purple along with the $tRNA^{Lys}$ (red) in the P site and the anticodon stem loop (ASL, blue) of $tRNA^{Phe}$ in the aminoacyl (A) site. (**B**) $tRNA^{Lys}$ binds to the UAA codon resulting in a 1st position U·U mismatch in the P site. (**C**) $tRNA^{Lys}$ binds to the AUA codon resulting in a 2nd position U·U mismatch in the P site. (**D**) The 1st position U·U mismatch is Watson-Crick-like in their orientation but is too wide for hydrogen bonding. (**E**) Similar to the 1st position mismatch, the 2nd position U·U mismatch also adopts a wider Watson-Crick-like base pairing. S = 5-methylaminomethyl-2-thiouridine ($mnm^5s^2U$).

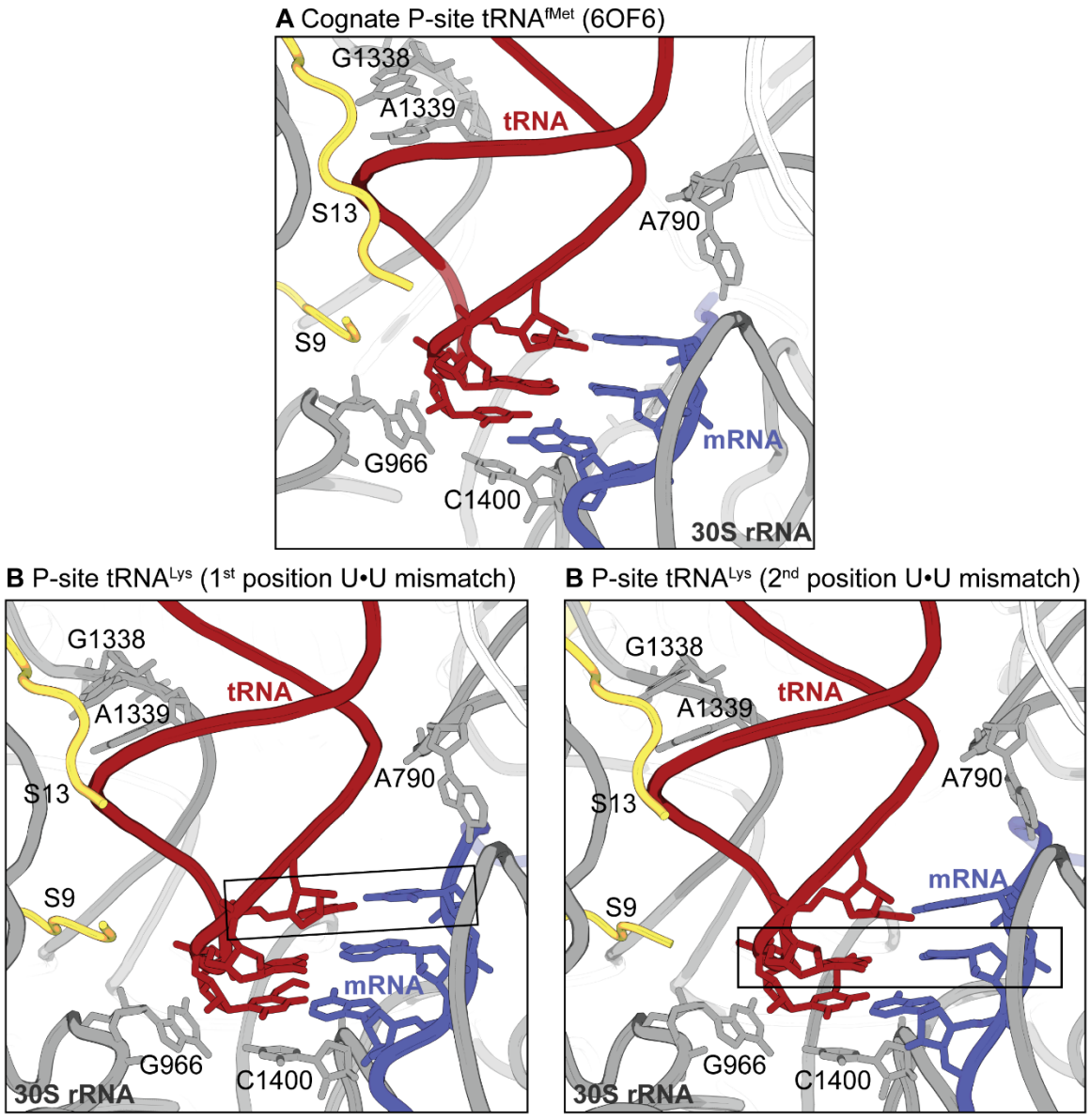


**Figure 4.3. The ribosomal P site does not recognize the mismatches in the anticodon-codon interaction.**

The ribosomal environment of a cognate P-site $tRNA^{fMet}$ bound to the start codon in the P site (PDB ID 6OF6) (A) is similar to the two structures of the mismatches in the P-site (B and C). The three base pairs of the anticodon-codon interaction is not extensively probed in the P site, and the P site is not known to have proofreading capability like that

of the A site. In all three structures shown here, the P site rRNA bases and proteins critical for tRNA binding and translocation do not have any structural response to the mismatch in the first (B) and second position (C) of the codon-anticodon interaction: RNA bases G966 and C1400 pack against the third anticodon-codon interaction similarly in all three structures, A790 is in the same conformation, G1338 and A1339 grips the stem of the anticodon loop, and r-proteins S9 and S13 tails form tRNA binding site in the same way as when a correct tRNA is bound (A).

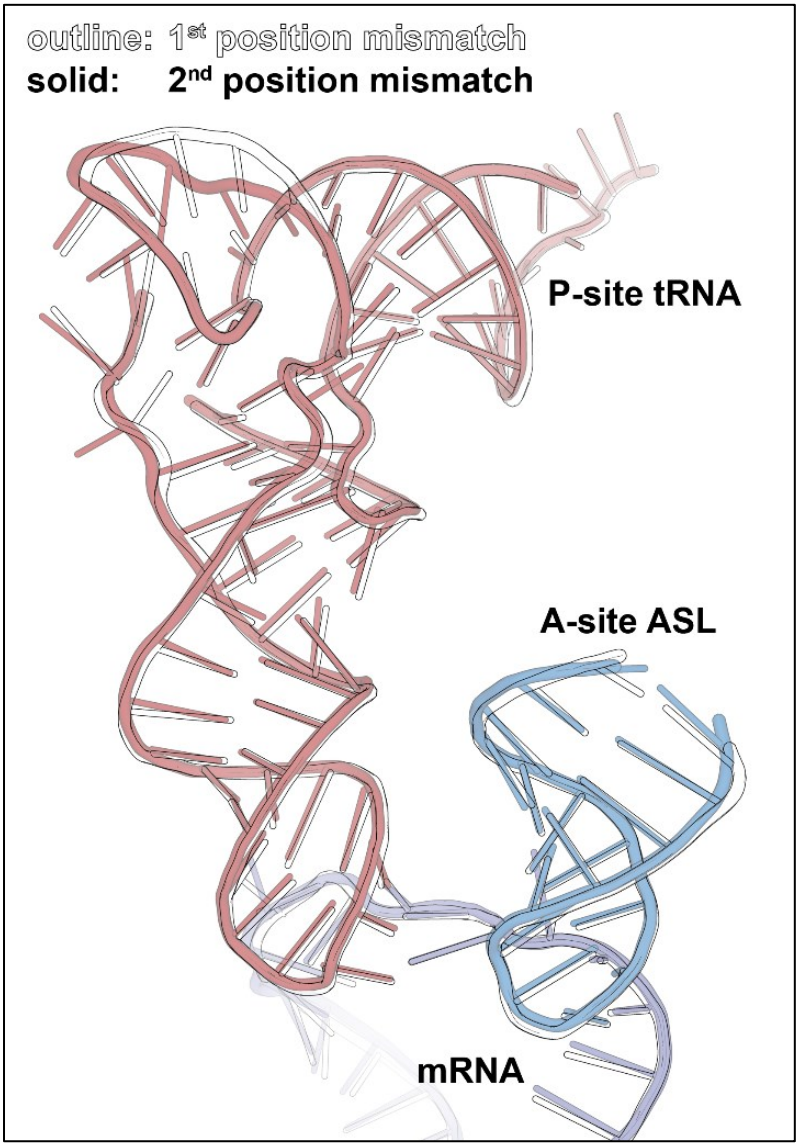


**Figure 4.4. Overlay of the P-site tRNA of these two structures show similar tRNA conformations.**

The $tRNA^{Lys}$ bound to the 1st position mismatch (in black outline) look the same as when it is bound to the 2nd position mismatch (in solid pink). The A-site ASL and the mRNA are also every similar, with the overall Cα RMSD of the P-site tRNA, A-site ASL, and the mRNA between the structures being 0.357Å as measured using the 'align' tool in PyMOL.

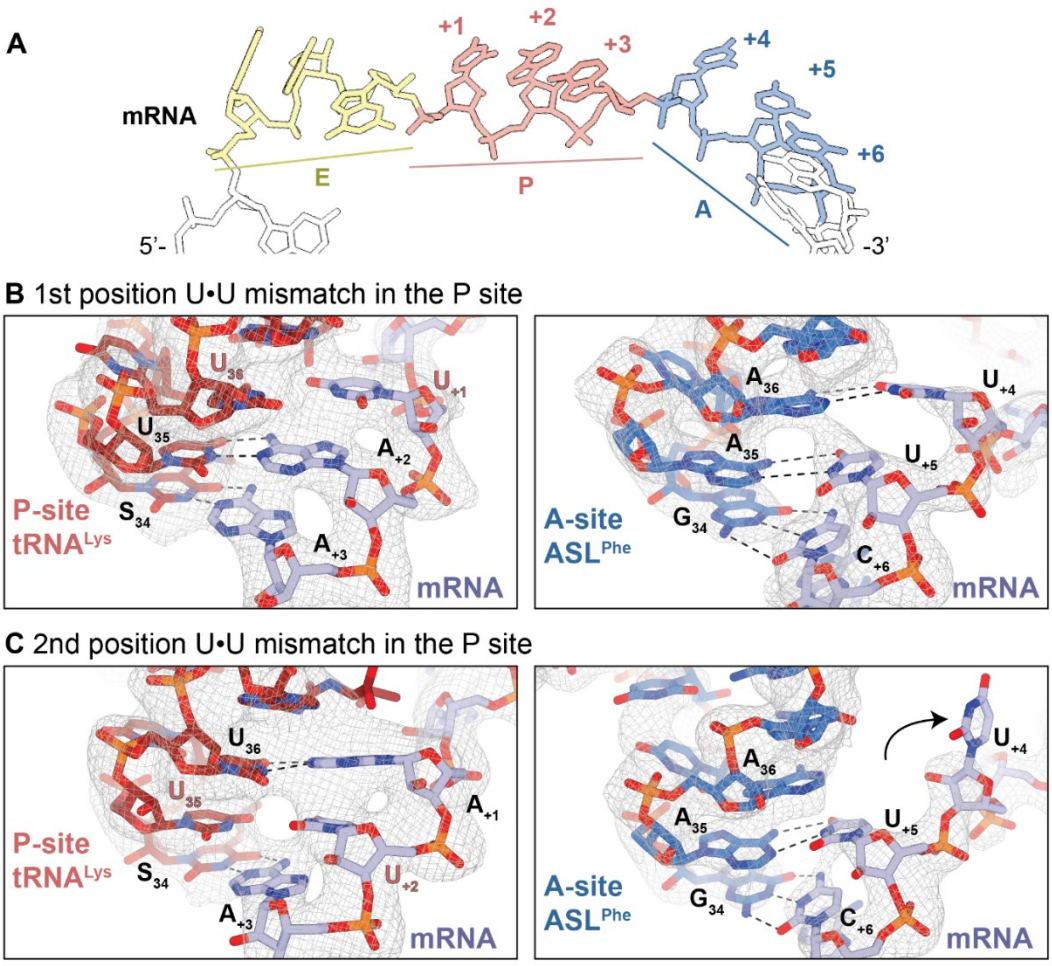


**Figure 4.5. The context of the P-site mismatch results in mRNA mispositioning.**

(**A**) The typical path of the mRNA through the ribosome showing the codon positioning in the ribosomal tRNA binding sites: exit (E) site in yellow, peptidyl (P) site in red, aminoacyl (A) site in blue. The nucleotides are numbered with the first nucleotide in the P site denoted +1. (**B**) When the U·U mismatch is at the first nucleotide of the P-site anticodon-codon pairing, the mRNA (purple) is positioned correctly in both the P and A sites. Left: the P-site codon nucleotides ($U_{+1}$, $A_{+2}$, $A_{+3}$) is paired to the $tRNA^{Lys}$ (red) anticodon. Right: the A-site codon nucleotides form three Watson-Crick base pairs to the A-site $ASL^{Phe}$ (blue) anticodon ($U_{+4}$-$A_{34}$, $U_{+5}$-$A_{35}$, $C_{+6}$-$G_{36}$). (**C**) When the U·U mismatch is located in at the second position of the P-site mRNA-tRNA codon-anticodon interaction,

the mRNA is mispositioned in the A site. Left: the P-site codon ($U_{+1}$, $A_{+2}$, $A_{+3}$) is correctly paired to the P-site $tRNA^{Lys}$. Right: the first nucleotide of the A-site codon (U+4) is flipped away from the A-site $ASL^{Phe}$ causing the A-site anticodon-codon interaction is disrupted. Only two Watson-Crick base pairing interaction is maintained: $U_{+5}$-$A_{35}$, $C_{+6}$-$G_{34}$. S = 5-methylaminomethyl-2-thiouridine ($mnm^5s^2U$).

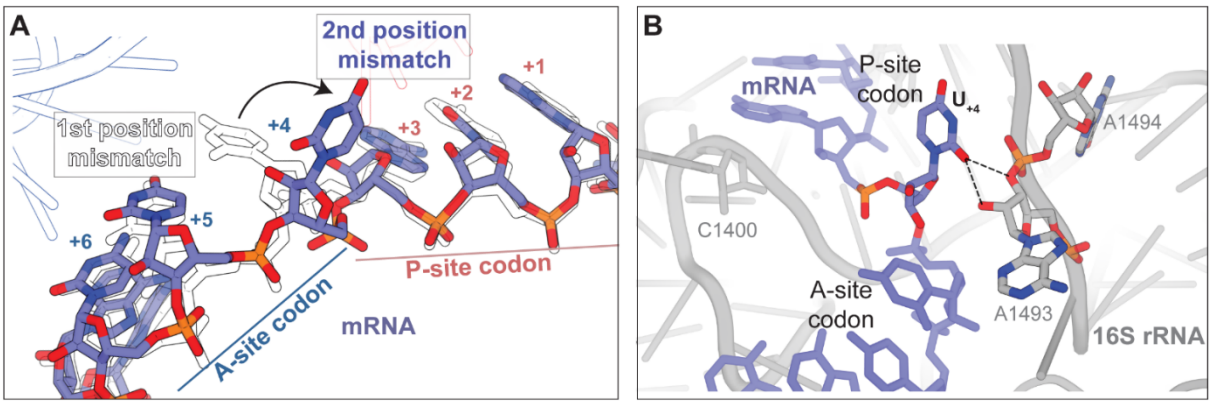


**Figure 4.6. The mRNA mispositioning disrupts the A-site codon-anticodon interactions.**

(A) Overlay of the mRNA paths from both the structures indicate the backbone paths are very similar**.** The first nucleotide of the A site is flipped away from the normal position when the mismatch is at the second position of the P-site anticodon-codon interaction by ~90°. The nucleotides from the 1$^{st}$ position mismatch are in black outlines, and the solid nucleotides represent the 2$^{nd}$ position mismatch structure. (**B**) When there is 2$^{nd}$ position U·U mismatch in the P site, the first nucleotide of the A site (U+4) flips out and interacts with the 16S rRNA A1493 nucleotide instead of pairing with the tRNA anticodon.

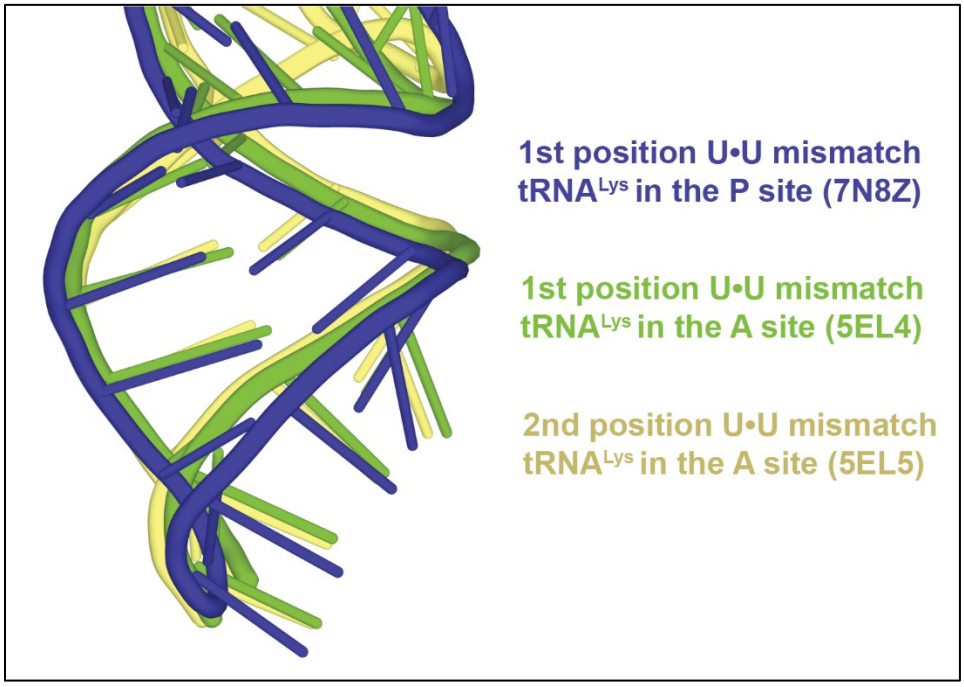


**Figure 4.7. The P-site mismatches do no alter overall ASL conformation.**

The anticodon stem loop (ASL) conformation of the P-site tRNA$^{Lys}$ 1$^{st}$ position mismatch is similar to that of the A-site tRNA$^{Lys}$ with a 1$^{st}$ (green) and 2$^{nd}$ (yellow) position U•U mismatches (24).

**Table 4.1. Data collection and refinement statistics**

| | **$tRNA^{Lys}$ UAA** | **$tRNA^{Lys}$ AUA** |
|---|---|---|
| PDB ID | 7N8Z | 7N90 |
| **Data collection** | | |
| Wavelength (Å) | 0.97890 | 0.97890 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell dimensions | | |
| *a, b, c (Å)* | 209.7, 449.5, 617.3 | 210.6. 447.2, 618.1 |
| *α, β, γ (°)* | 90, 90, 90 | 90, 90, 90 |
| Resolution (Å) | 137.3-3.58 (3.708-3.58) | 173.5-3.64 (3.77-3.64) |
| $R_{merge}$ | 0.15 (0.86) | 0.19 (1.07) |
| $R_{pim}$ | 0.094 (0.52) | 0.096 (0.54) |
| $I/\sigma I$ | 8.66 (1.70) | 7.28 (1.55) |
| $CC_{1/2}$ | 0.997 (0.448) | 0.997 (0.438) |
| Completeness (%) | 98.3 (99.0) | 95.7 (97.5) |
| Redundancy | 3.4 (3.5) | 4.4 (4.4) |
| **Refinement** | | |
| Total reflections | 2,270,192 | 2,696,566 |
| Reflections used in refinement | 667,719 | 618,599 |
| $R_{work}/R_{free}$ (%) | 20.6/25.4 | 23.1/27.9 |
| No. atoms | | |
| Macromolecules | 291,264 | 291,330 |
| Ligands | 922 | 1,267 |
| B-factors | | |
| Macromolecules | 119.8 | 124.3 |
| Ligands | 61.1 | 50.6 |
| Clashscore | 12.7 | 5.5 |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.005 | 0.004 |
| Bond angles (°) | 0.93 | 0.87 |
| Ramachandran plot (%) | | |
| Favored regions (%) | 91.2 | 92.6 |
| Allowed regions (%) | 7.60 | 6.87 |
| Outliers (%) | 1.18 | 0.51 |

## 4.7. Acknowledgments

Support for this work was provided by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program (CEF), NIH Training Grant T32 GM8367 (CEF), and NIH (R01GM093278 (CMD). CMD is a Burroughs Wellcome Fund Pathogenesis of Infectious Diseases Fellow. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## 4.8. References

1. H. S. Zaher, R. Green, Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**, 746-762 (2009).
2. V. Ramakrishnan, Ribosome Structure and the Mechanism of Translation. *Cell* **108**, 557-572 (2002).
3. R. M. Voorhees, V. Ramakrishnan, Structural basis of the translational elongation cycle. *Annu Rev Biochem* **82**, 203-236 (2013).
4. D. A. Drummond, C. O. Wilke, The evolutionary consequences of erroneous protein synthesis. *Nat Rev Genet* **10**, 715-724 (2009).
5. J. M. Ogle, V. Ramakrishnan, Structural insights into translational fidelity. *Annu Rev Biochem* **74**, 129-177 (2005).
6. M. V. Rodnina, K. B. Gromadski, U. Kothe, H. J. Wieden, Recognition and selection of tRNA in translation. *FEBS letters* **579**, 938-942 (2005).
7. K. B. Gromadski, T. Daviter, M. V. Rodnina, A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. *Mol Cell* **21**, 369-377 (2006).

8. T. Daviter, K. B. Gromadski, M. V. Rodnina, The ribosome's response to codon-anticodon mismatches. *Biochimie* **88**, 1001-1011 (2006).

9. M. V. Rodnina, N. Fischer, C. Maracci, H. Stark, Ribosome dynamics during decoding. *Philos Trans R Soc Lond B Biol Sci* **372** (2017).

10. J. M. Ogle, F. V. Murphy, M. J. Tarry, V. Ramakrishnan, Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721-732 (2002).

11. A. B. Loveland, G. Demo, N. Grigorieff, A. A. Korostelev, Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546**, 113-117 (2017).

12. M. V. Rodnina, Translation in Prokaryotes. *Cold Spring Harbor perspectives in biology* **10** (2018).

13. M. V. Rodnina, T. Pape, R. Fricke, L. Kuhn, W. Wintermeyer, Initial binding of the elongation factor Tu.GTP.aminoacyl-tRNA complex preceding codon recognition on the ribosome. *J Biol Chem* **271**, 646-652 (1996).

14. T. Pape, W. Wintermeyer, M. Rodnina, Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. *Embo J* **18**, 3800-3807 (1999).

15. K. B. Gromadski, M. V. Rodnina, Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* **13**, 191-200 (2004).

16. F. Bouadloun, D. Donner, C. G. Kurland, Codon-specific missense errors in vivo. *Embo J* **2**, 1351-1356 (1983).

17. H. S. Zaher, R. Green, Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161-166 (2009).

18. J. Precup, J. Parker, Missense misreading of asparagine codons as a function of codon identity and context. *J Biol Chem* **262**, 11351-11355 (1987).

19. H. S. Zaher, R. Green, Kinetic basis for global loss of fidelity arising from mismatches in the P-site codon:anticodon helix. *RNA* **16**, 1980-1989 (2010).

20. H. S. Zaher, R. Green, A primary role for release factor 3 in quality control during translation elongation in Escherichia coli. *Cell* **147**, 396-408 (2011).

21. M. Selmer *et al.*, Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942 (2006).

22. L. Lancaster, H. F. Noller, Involvement of 16S rRNA nucleotides G1338 and A1339 in discrimination of initiator tRNA. *Mol Cell* **20**, 623-632 (2005).

23. K. G. McGarry, S. E. Walker, H. Wang, K. Fredrick, Destabilization of the P site codon-anticodon helix results from movement of tRNA into the P/E hybrid state within the ribosome. *Mol Cell* **20**, 613-622 (2005).

24. A. Rozov *et al.*, Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat Commun* **7**, 10457 (2016).

25. A. B. Loveland, G. Demo, A. A. Korostelev, Cryo-EM of elongating ribosome with EF-Tu*GTP elucidates tRNA proofreading. *Nature* **584**, 640-645 (2020).

26. G. Z. Yusupova, M. M. Yusupov, J. H. Cate, H. F. Noller, The path of messenger RNA through the ribosome. *Cell* **106**, 233-241 (2001).

27. S. J. Nasvall, P. Chen, G. R. Bjork, The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons in vivo. *RNA* **10**, 1662-1673 (2004).

28. P. F. Agris, F. A. Vendeix, W. D. Graham, tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol* **366**, 1-13 (2007).

29. E. B. Kramer, P. J. Farabaugh, The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. *RNA* **13**, 87-96 (2007).

30. N. Manickam, N. Nag, A. Abbasi, K. Patel, P. J. Farabaugh, Studies of translational misreading in vivo show that the ribosome very efficiently discriminates against most potential errors. *RNA* **20**, 9-15 (2014).

31. H. J. Grosjean, S. de Henau, D. M. Crothers, On the physical basis for ambiguity in genetic coding interactions. *Proc Natl Acad Sci U S A* **75**, 610-614 (1978).

32. M. Johansson, J. Zhang, M. Ehrenberg, Genetic code translation displays a linear trade-off between efficiency and accuracy of tRNA selection. *Proc Natl Acad Sci U S A* **109**, 131-136 (2012).

33. J. Lehmann, A. Libchaber, Degeneracy of the genetic code and stability of the base pair at the second position of the anticodon. *RNA* **14**, 1264-1269 (2008).

34. D. Kurita, Y. Chadani, A. Muto, T. Abo, H. Himeno, ArfA recognizes the lack of mRNA in the mRNA channel after RF2 binding for ribosome rescue. *Nucleic Acids Res* **42**, 13339-13352 (2014).

35. Z. Fu *et al.*, The structural basis for release-factor activation during translation termination revealed by time-resolved cryogenic electron microscopy. *Nature communications* **10**, 2579 (2019).

36. E. Svidritskiy, G. Demo, A. B. Loveland, C. Xu, A. A. Korostelev, Extensive ribosome and RF2 rearrangements during translation termination. *eLife* **8** (2019).

37. A. A. Korostelev, Structural aspects of translation termination on the ribosome. *RNA* **17**, 1409-1421 (2011).

38. E. Svidritskiy, G. Demo, A. A. Korostelev, Mechanism of premature translation termination on a sense codon. *J Biol Chem* **293**, 12472-12479 (2018).

39. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).

40. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta crystallographica* **66**, 486-501 (2010).

41. Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

# Chapter 5.

## Conclusions and Future Directions

Translation is a process where mRNA is decoded three nucleotides at a time to synthesize all the cellular proteins. We are still discovering new mechanisms of regulation of translation speed and fidelity (**Figure 5.1**) (1) and ways to exploit these processes for the recoding and expansion of the genetic code.

tRNAs were initially postulated to be 'generic' adaptors in translation (2) but it is now clear they exert significant control over the regulation of translation speed and accuracy. In *E. coli*, the 45 tRNAs decode 61 sense codons and these tRNAs have remarkable diversity in sequence and post-transcription modifications despite being structurally very similar and performing the same function in protein synthesis (**Figures 1.4, 1.5**) (3). It is also known that codons are not translated with the same speed and accuracy due to factors such as the differences in the strengths of the codon-anticodon or different chemical properties of amino acids and tRNA modifications (4-6). Our lab, along with many others, are interested in understanding the molecular mechanisms with which certain tRNA elements can influence the recognition of the correct vs. incorrect mRNA-tRNA pairs by the bacterial ribosome.

tRNAs are the most highly modified nucleic acids, both in frequency and diversity in modification chemistry (7). While the anticodon is the main determinant of the tRNA identity, other parts of the tRNA have also evolved to maintain the tRNA's roles in translation through the steps of aminoacylation, EF-Tu binding and peptide bond formation on the ribosome (8-11). Specifically, the role of the $m^1G37$ modification in mRNA frame maintenance and how $tRNA^{Pro}$ mutants can decode a four-nucleotide codon

has been a topic of special interest to our lab (12-15). Initially, it was thought that these tRNAs decoded a four-nucleotide codon through quadruplet base pairing where four nucleotides of the anticodon paired with four nucleotides in the codon (13). However, the additional methyl group on G37 should preclude the possibility of a CGGm$^1$G (positions 34 to 37) forming a four-nucleotide anticodon to base pair with a four-nucleotide codon for quadruplet base pairing. U33 also forms a highly conserved U-turn in the ASL away from the mRNA when it is bound to the ribosome (16, 17) so it would not be able to participate in the codon-anticodon interaction. Since both the 5' (position 37) and the 3' (position 33) nucleotides are not able to act as an expanded anticodon, it was unlikely that quadruplet decoding was the mechanism for the frameshifting. Structures of frameshifting complexes have also confirmed that the codon-anticodon pairings only form 3 base pairs in the ribosomal P and A sites. The ribosome environment of the P and A sites do not have the expanded space available for four base pairs to form (18-23). After decades of biophysical, biochemical, and structural data, the current model for modification deficient tRNA$^{Pro}$ +1 frameshift is that in the absence of the m$^1$G37 modification, the tRNA is destabilized on the ribosome such that the ribosome is unable to maintain all of the interactions controlling the tRNA movement through the ribosome presenting opportunities for the tRNAs to slip during translocations between the ribosomal tRNA binding sites (18, 20, 22, 23).

My research presented here in Chapter 2 (21) and another studies from our lab (18, 20, 23) have worked with the major proline isoacceptor in *E. coli*, $tRNA^{Pro}_{CGG}$ also known as ProK, but there two other proline isoacceptors: $tRNA^{Pro}_{GGG}$ (ProL) and $tRNA^{Pro}_{UGG}$ (ProM) (**Figure 1.7**). We also now have structures of the ProL tRNA lacking a modification in the

P site and single-molecule FRET data tracking the movement of the ribosome in response to this tRNA (unpublished), and comparison of $m^1G37$-deficient proK and proL tRNAs indicates there are differences in the ribosomal response to the two isoacceptors (23). There are also now cryo-EM structures and further biochemical investigations into the +1 frameshifting mechanism of $tRNA^{SufB2}$, the suppressor derivative of $tRNA^{Pro}_{GGG}$ (24). Similar to our studies with $tRNA^{Pro}_{CGG}$ and its suppressor variant $tRNA^{SufA6}$ (23), $tRNA^{SufB2}$ also causes +1 frameshifting due to the perturbations in ribosome movement during the late stages of translocation on the slippery mRNA codon . They also showed that this tRNA can be used in recoding experiments to incorporate synthetic amino acid analogs such as *cis*-hydropro and azetidine on the CCC-C codon (24). The final proline isoacceptor, $tRNA^{Pro}_{UGG}$, is also found to slip on the mRNA CCC-A codon during translocation through an elongation factor G (EF-G) mediated intermediate (22). Cryo-EM structures of this complex undergoing translocation showed that the decoding center is disrupted in the pre-translocation complex such that rRNA base G530 is shifted away from A1492 and the bound $tRNA^{Pro}_{UGG}$, leading to destabilized ribosome interactions with the tRNA-mRNA pair. Interestingly, in one of the translocation intermediates, they found a previously unobserved state of the mRNA where it has bulged out of the canonical mRNA path so the first C nucleotide in the CCC-A codon flips out into the space between the P and E sites to base stack with the 16S rRNA G926, resulting in the +1 frame in the P site with the tRNA base pairing to the CCA codon. This is distinct from the frameshifted structure of $tRNA^{Pro}_{CGG}$ where the tRNA alone is able to facilitate the +1 frameshift, independent of EF-G catalyzed translocation (23). Furthermore, the P-site $tRNA^{Pro}_{CGG}$ moves towards the E site on the slippery CCC-U codon, adopting an e*/E state (e*: transitional state between

the 30S P and E sites, E: canonical E-site tRNA binding on the 50S) coupled with spontaneous head domain swiveling and tilting. These conformational rearrangements indicate the interactions of the ribosome with the $tRNA^{Pro}_{CGG}$ bound to the P-site slippery codon is impacted. These data indicate there are subtle differences in the +1 frameshifting mechanisms of the three $tRNA^{Pro}$ isoacceptors. Further there are differences in how $m^1G37$ contributes to maintaining the correct frame as well as how the suppressor tRNAs, $tRNA^{SufA6}$ and $tRNA^{SufB2}$, function to decode the +1 frame on these slippery codons. Bacteria has a special elongation factor P (EF-P) to facilitate the movement of stalled ribosomes on proline codons and suppresses the +1 frameshift by $tRNA^{Pro}$ (25-27). Our lab has collected structural data of EF-P bound to the ribosome containing a modification-deficient P-site $tRNA^{Pro}$ and we hope this structure will elucidate the mechanism of EF-P-mediated correction of the frameshift induced by the lack of $m^1G37$ on $tRNA^{Pro}$. It would also be fascinating to investigate whether the EF-P mechanism is similar across the three $tRNA^{Pro}$ isoacceptors and the suppressor variants.

$tRNA^{Pro}_{UGG}$ is naturally frameshift-prone, and it is the only $tRNA^{Pro}$ with a modification in its anticodon (28). The uridine at the wobble position in $tRNA^{Pro}_{UGG}$ is post-transcriptionally modified as 5-oxyacetyl uridine ($cmo^5U$). This modification can base pair with any of the RNA bases so this tRNA decodes all four proline codons (CCA, CCU, CCG, CCC) (29) (**Figure 1.7**). $tRNA^{Pro}_{UGG}$ also contains the $m^1G37$ modification and interestingly, there seems to be a hierarchal relationship between the modification on the G37 and the terminal methylation of $cmo^5U$ into $mcmo^5U$ (30). Out of the proline isoacceptors, there is currently no known suppressor variant of $tRNA^{Pro}_{UGG}$. Considering that $tRNA^{SufA6}$ and $tRNA^{SufB2}$ contain an identical insertion of G37.5, it would be logical to

see if engineering an insertion at this position will confer additional decoding capabilities and result in a new suppressor tRNA. $tRNA^{Pro}_{UGG}$ and $tRNA^{Pro}_{CGG}$ both have a U•A pair at the 32-38 pairing, but $tRNA^{Pro}_{GGG}$ has the A•U pair, similar to that of $tRNA^{Ala}_{GGC}$ studied in Chapter 3. Some unnatural amino acid incorporation optimizations resulted in changing of the 32-38 pairing (31) so we should explore whether swapping these 32-38 pairings will also influence miscoding. Most tRNA engineering efforts have focused on a 'top-down' directed evolution based on the premise of natural mutations giving rise to new capabilities (32), but these avenues of rational tRNA design would overcome the low incorporation efficiencies these other methods encounter *in vivo*.

Expanding on the importance of the 32-38 pairing, in Chapter 3 I looked at how disrupting the 32-38 pairing in $tRNA^{Ala}_{GGC}$ can impair decoding accuracy and render the ribosome unable to distinguish between correct vs. incorrect tRNAs. The 32-38 pairing is evolutionarily tuned corresponding to the anticodon nucleotide sequence to modulate the binding of the tRNAs so they have similar binding affinities to the ribosome (33). $tRNA^{Ala}_{GGC}$ is an ideal model system to study the impact of changes to the 32-38 pairing because it does not contain any modifications in the anticodon loop, and alanyl-tRNA synthetase (AlaRS) only recognized the G3:U70 base pair in the acceptor stem (9, 34, 35) so the effects observed are due to the changes in the 32 and 38 pairing only. $tRNA^{Ala}_{GGC}$ has an A32-U38 pairing that is relatively rare in tRNAs (33, 36, 37). Disruptions to the 32-38 pairing are so catastrophic to the cell that over expression the minor isoacceptor $tRNA^{Ala}_{GGC}$ with different 32-38 pairings C•A, U•A, or U•U (• denotes a non-Watson-Crick base pair) overwhelms the canonical translation machinery causing cell death (38). Impressively, the U32•A38 variant is the same base pair just swapped, but it is sufficient

to induce high levels of miscoding. While there was extensive biochemical characterization showing the 32-38 pairing prevented miscoding in $tRNA^{Ala}_{GGC}$, it was unclear how this base pair in the anticodon loop could actually prevent miscoding. We solved four ribosome crystal structures with $tRNA^{Ala}_{GGC}$ and its U32-A38 variant bound the cognate GCC codon or the near-cognate GCA codon to see how the swapping the 32-38 base pair abolish correct tRNA recognition. Typically, the 32-38 base pair in the anticodon loop forms a single hydrogen bond (Chapter 3, **Fig. S3**) (39). Surprisingly, the A32-U38 pair in $tRNA^{Ala}_{GGC}$ forms a widened conformation of the Watson-Crick base pairing with two hydrogen bonds (**Fig. 3.2**). This stable Watson-Crick base pair is destabilized when the tRNA is bound to a near-cognate codon containing a G•A mismatch, and this causes a conformation change in one of the monitoring bases of the decoding center, the 23S rRNA A1913. A1913 disengages away from the tRNA, and this loss of interaction would weaken the interaction of the decoding center with the tRNA priming this tRNA to be rejected by the ribosome. This rejection of the incorrect tRNA corresponds with the kinetic proofreading mechanisms of decoding where the incorrect tRNA dissociates rapidly from the ribosome, even after GTP hydrolysis has already happened (40, 41). However, when the A32-U38 pairing is reversed to U32-A38, the recognition of the mismatch in the codon-anticodon is lost. The 32-38 pairing in $tRNA^{Ala}_{GGC}$ no longer responds to the mismatch in the tRNA-mRNA interaction, and the ribosome accepts the mismatch as if it is correct. In other words, changing the 32-38 pairing allowed the near-cognate decoding to bypass proofreading mechanisms. This is the first time 23S rRNA A1913 has been observed to adopt this alternate conformation in response to a mismatch. Previously, cryo-EM

structures showed A1913 becoming dynamic during near-cognate decoding (42), but in our structure, A1913 stably adopts the position away from the tRNA (43).

In Chapter 4, I identified a possible mechanism with which post peptidyl transfer quality control (post PT QC) can be triggered when a mismatch is present in the P site. As discussed in Chapter 2, the mRNA-tRNA pairing is significantly probed in the A site with a multitude of kinetic proofreading steps to ensure the correct tRNA is accepted (**Fig. 4.1**). The ribosome 'grips' the tRNA in the P site (44), and controls its movement from the P site to the E site (45, 46) where it exits the ribosome, but it does not seem to actively monitor the tRNA identity past the A site. However, work in the Green lab showed that there are post peptidyl transfer quality control (post PT QC) mechanisms once a miscoding event happens and the mismatched mRNA-tRNA pair is translocated to the P site (47, 48). The P-site mismatches induce a decrease in the fidelity of the A site resulting in successive misincorporation that triggers premature termination. We solved two ribosome structures with $tRNA^{Lys}$ bound to a U•U mismatch in the first and second positions of the codon-anticodon interaction in the P site. The first-position U•U mismatch appear to cause a misposition of the mRNA in the A site. We hypothesize this mRNA mispositioning is a post PT QC mechanism that destabilizes tRNA binding to the A site such that cognate and near-cognate tRNAs can no longer be distinguished. Furthermore, it can also cause stalling at the A site such that rescue mechanisms can also cause premature termination to release the erroneous polypeptide for degradation. It was unclear how the ribosome could signal to the A site for P-site mismatches, and this provides the first structural mechanism for this phenomenon. However, it is still not known how release factors can then be recruited on sense codons to trigger premature

termination in post PT QC (49). Furthermore, it is perplexing that the first-position mismatch is not recognized, but the second-position mismatch was. It would be interesting to investigate the impact of mismatches aside from the U·U mismatch and whether there is a difference in response to miscoding errors on sense vs. nonsense codons.

In summary of this dissertation, I have explored three additional ways tRNAs can impact translation fidelity. First, in Chapter 2, I studied how a modification ($m^1G37$) and insertion (G37.5) in $tRNA^{Pro}_{CGG}$ can affect tRNA binding, an indicator of whether the ribosome will accept or reject the tRNA, and how changing the insertion can restore an important base pair in the anticodon stem loop of tRNAs to stabilize the binding and structure of $tRNA^{Pro}_{CGG}$ (50). Then in Chapter 3, I examined how the 32-38 pairing in $tRNA^{Ala}_{GGC}$ can act as a signal for a mismatch in the codon-anticodon interaction and how a ribosomal RNA nucleotide A1913 responds to this mismatch; disruption to this 32-38 pairing ablates this recognition mechanism (43). Finally, in Chapter 4, I propose that a mismatch in the P site causes mRNA mispositioning to disrupt A-site binding triggering post-peptidyl transfer quality control. All these contributions (**Fig. 5.2**) further our understanding of bacterial translation dysregulation and could be investigated as novel therapeutic targets or expanding our biotechnological toolkit using unnatural amino acid incorporation.

Antibiotic resistance is a public health crisis with enormous costs. There were over 2.8 million antibiotic-resistant infections resulting in 35,000 deaths in the US alone in 2019 (51, 52). The estimated cost to treat of antibiotic-resistant infections in 2017 was $4.6 billion in the US (53). With the lack of new antibiotics being developed, untreatable

infections are becoming a reality (54), therefore there is a critical need for novel antibiotic targets. The enzyme catalyzing the transfer of the methyl group from *S*-adenosyl-*L*-methionine (SAM) to G37 in tRNAs is called TrmD (55, 56). It is essential in bacteria but is structurally dissimilar to its human homologue Trm5 (57) making it a prime candidate for a novel antibiotic target. Previously, a fragment-based drug screen targeting the SAM binding pocket of TrmD identified potent TrmD inhibitors but they lacked antibiotic activity (58). More recently, a high-throughput screening assay using a bioluminescence readout found a number of inhibitors with antimicrobial activity against Gram-positive, Gram-negative, and even mycobacteria (59, 60). Understanding how disrupting the $m^1G37$ modification affects the cell would be beneficial to understand the mechanisms of the drug, potential downstream effects, and even to predict possible resistance pathways.

The incorporation of unnatural amino acids is a fascinating bioengineering feat to produce special proteins to expand our repertoire of techniques to investigate and manipulate biological processes (61). Proteins with an unnatural amino acid incorporated can have new reactive sites for bioconjugation (62, 63), generate native proteins with uniform post-translational modifications (64), and even one day to produce a completely unnatural protein with new functions. Most commonly, the stop codons are targeted to recode for a sense codon (65) using an orthogonal engineered tRNA/synthetase system (66). Pyrrolysine is a lysine derivative originally discovered in archaeal *Methanosarcina barkeri* that is encoded by the UAG ‘amber’ stop codon (67) that has now become the most well-known unnatural amino acid to be incorporated into proteins. The orthogonal PylRS/$tRNA^{CUA}$ pair is compatible in *E. coli*, yeast, mammalian cells, *Caenorhabditis elegans* and *Drosophila melanogaster* making it a ubiquitous synthetic biology tool

especially since the release factor recognizing the UAG stop codon (RF1) can also be engineered (68). Even ribosomes can be engineered to have specialized ribosomes synthesizing proteins outside of normal cellular translation, so there two sets of distinct translation machineries in the cell (69, 70). Recently, the Chin lab was successful in developing a synthetic *E. coli* strain with a recoded genome (71). They were able to remove the need for three codons (two sense and one stop) so this strain only uses 59 codons to encode the 20 amino acids, making a canonically essential tRNA ($tRNA^{Ser}_{GCU}$) redundant. This discovery opens exciting possibilities for unnatural amino acid incorporation by providing a method to minimize natural genomes to release additional codons for recoding efforts (72).

## Figures

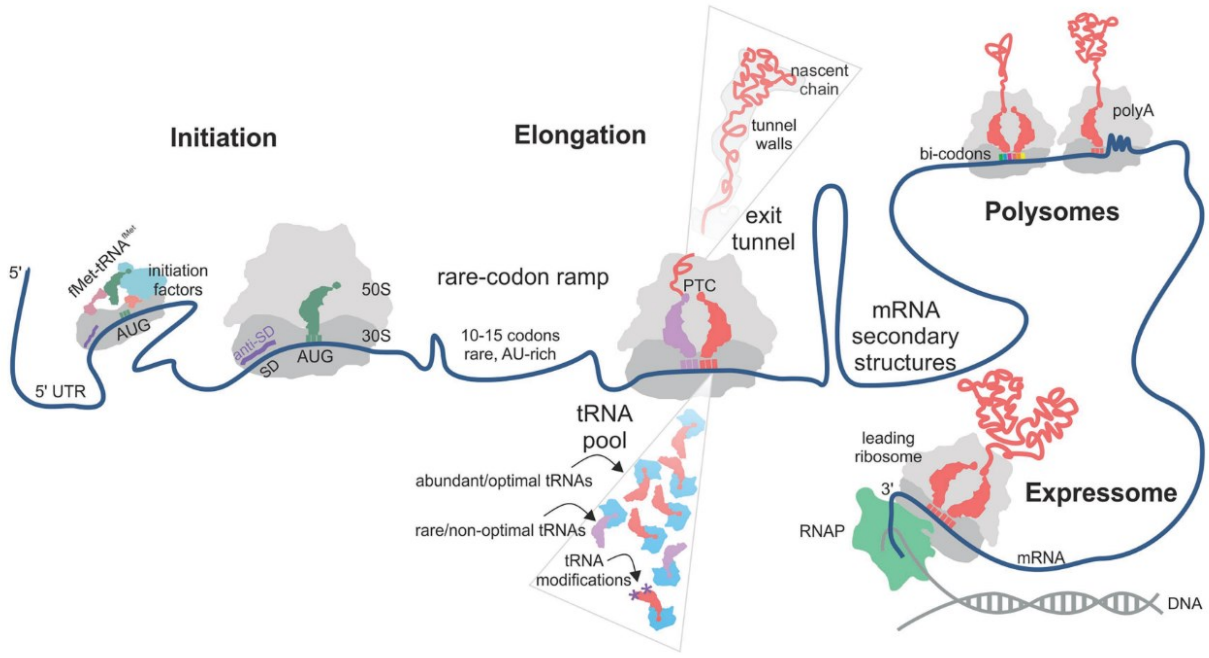


**Figure 5.1. Regulation mechanisms of translation speed and fidelity in bacteria.** There are many factors contributing to the regulating the accuracy and speed of protein synthesis. Initiation is controlled by having a specific initiation-only tRNA and initiation factors that ensure the correct positioning of the start codon. The mRNA contains a number of regulatory elements such as the ribosome binding site positioning the start codon, rare codons slowing the rate of translation to allow for co-translation protein folding, mRNA secondary structures that can serve as sequestration for certain sequences. tRNAs can regulate translation through variations in their levels and modifications. Additionally, the nascent chain can also interact with the exit tunnel influencing its own translation. Figure taken from reference (1).

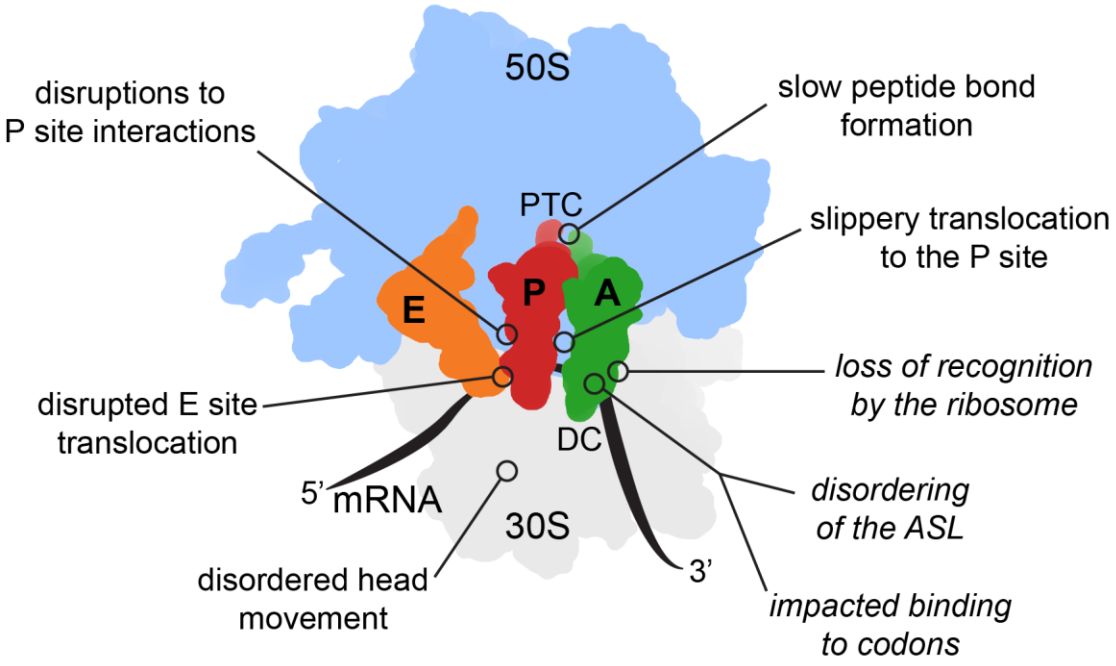


**Figure 5.2. tRNA-mediated mechanisms of translation fidelity regulation.**

The mechanisms of dysregulation of frame maintenance mediated by tRNAs as discussion in Chapter 1. The contributions of this dissertation are italicized.

**References**


1. E. Samatova, J. Daberger, M. Liutkute, M. V. Rodnina, Translational Control by Ribosome Pausing in Bacteria: How a Non-uniform Pace of Translation Affects Protein Production and Folding. *Frontiers in Microbiology* **11** (2021).
2. F. H. Crick, The origin of the genetic code. *J Mol Biol* **38**, 367-379 (1968).
3. V. Y. Vare, E. R. Eruysal, A. Narendran, K. L. Sarachan, P. F. Agris, Chemical and Conformational Diversity of Modified Nucleosides Affects tRNA Structure and Function. *Biomolecules* **7** (2017).
4. D. D. Nedialkova, S. A. Leidel, Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell* **161**, 1606-1618 (2015).
5. J. A. Dunkle, C. M. Dunham, Mechanisms of mRNA frame maintenance and its subversion during translation of the genetic code. *Biochimie* **114**, 90-96 (2015).
6. S. Rudorf, R. Lipowsky, Protein Synthesis in E. coli: Dependence of Codon-Specific Elongation on tRNA Concentration and Codon Usage. *PLOS ONE* **10**, e0134994 (2015).
7. H. Grosjean, M. Sprinzl, S. Steinberg, Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. *Biochimie* **77**, 139-141 (1995).
8. A. Louie, N. S. Ribeiro, B. R. Reid, F. Jurnak, Relative affinities of all Escherichia coli aminoacyl-tRNAs for elongation factor Tu-GTP. *J Biol Chem* **259**, 5010-5016 (1984).
9. K. Gabriel, J. Schneider, W. H. McClain, Functional evidence for indirect recognition of G.U in tRNA(Ala) by alanyl-tRNA synthetase. *Science* **271**, 195-197 (1996).
10. S. S. Phelps, O. Jerinic, S. Joseph, Universally Conserved Interactions between the Ribosome and the Anticodon Stem-Loop of A Site tRNA Important for Translocation. *Molecular Cell* **10**, 799-807 (2002).
11. N. Manickam, K. Joshi, M. J. Bhatt, P. J. Farabaugh, Effects of tRNA modification on translational accuracy depend on intrinsic codon-anticodon strength. *Nucleic Acids Res* **44**, 1871-1881 (2016).
12. G. R. Bjork, P. M. Wikstrom, A. S. Bystrom, Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science* **244**, 986-989 (1989).
13. G. E. Sroga, F. Nemoto, Y. Kuchino, G. R. Bjork, Insertion (sufB) in the anticodon loop or base substitution (sufC) in the anticodon stem of tRNA(Pro)2 from Salmonella typhimurium induces suppression of frameshift mutations. *Nucleic Acids Res* **20**, 3463-3469 (1992).

14. T. G. Hagervall, T. M. Tuohy, J. F. Atkins, G. R. Bjork, Deficiency of 1-methylguanosine in tRNA from Salmonella typhimurium induces frameshifting by quadruplet translocation. *J Mol Biol* **232**, 756-765 (1993).

15. Q. Qian, G. R. Bjork, Structural alterations far from the anticodon of the tRNAProGGG of Salmonella typhimurium induce +1 frameshifting at the peptidyl-site. *J Mol Biol* **273**, 978-992 (1997).

16. S. S. Ashraf *et al.*, The uridine in "U-turn": contributions to tRNA-ribosomal binding. *RNA* **5**, 503-511 (1999).

17. M. Sundaram, P. C. Durant, D. R. Davis, Hypermodified nucleosides in the anticodon of tRNALys stabilize a canonical U-turn structure. *Biochemistry* **39**, 12575-12584 (2000).

18. T. Maehigashi, J. A. Dunkle, S. J. Miles, C. M. Dunham, Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. *Proc Natl Acad Sci U S A* **111**, 12740-12745 (2014).

19. C. E. Fagan, T. Maehigashi, J. A. Dunkle, S. J. Miles, C. M. Dunham, Structural insights into translational recoding by frameshift suppressor tRNASufJ. *RNA* **20**, 1944-1954 (2014).

20. S. Hong *et al.*, Mechanism of tRNA-mediated +1 ribosomal frameshifting. *Proc Natl Acad Sci U S A* 10.1073/pnas.1809319115 (2018).

21. H. A. Nguyen, E. D. Hoffer, C. M. Dunham, Importance of tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in tRNA(Pro) CGG for decoding. *J Biol Chem* 10.1074/jbc.RA119.007410 (2019).

22. G. Demo *et al.*, Structural basis for +1 ribosomal frameshifting during EF-G-catalyzed translocation. *Nature Communications* **12** (2021).

23. E. D. Hoffer *et al.*, Structural insights into mRNA reading frame regulation by tRNA modification and slippery codon–anticodon pairing. *eLife* **9** (2020).

24. H. Gamper *et al.*, Insights into genome recoding from the mechanism of a classic +1-frameshifting tRNA. *Nature Communications* **12** (2021).

25. H. B. Gamper, I. Masuda, M. Frenkel-Morgenstern, Y. M. Hou, Maintenance of protein synthesis reading frame by EF-P and m(1)G37-tRNA. *Nat Commun* **6**, 7226 (2015).

26. T. Katoh, I. Wohlgemuth, M. Nagano, M. V. Rodnina, H. Suga, Essential structural elements in tRNAPro for EF-P-mediated alleviation of translation stalling. *Nature Communications* **7**, 11657 (2016).

27. P. Huter *et al.*, Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. *Mol Cell* **68**, 515-527 e516 (2017).

28. H. B. Gamper, I. Masuda, M. Frenkel-Morgenstern, Y. M. Hou, The UGG Isoacceptor of tRNAPro Is Naturally Prone to Frameshifts. *Int J Mol Sci* **16**, 14866-14883 (2015).

29. S. J. Nasvall, P. Chen, G. R. Bjork, The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons in vivo. *RNA* **10**, 1662-1673 (2004).

30. I. Masuda *et al.*, Selective terminal methylation of a tRNA wobble base. *Nucleic Acids Research* 10.1093/nar/gky013, gky013-gky013 (2018).

31. T. Katoh, Y. Iwane, H. Suga, tRNA engineering for manipulating genetic code. *RNA Biol* **15**, 453-460 (2018).

32. H. Xiao, P. G. Schultz, At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code. *Cold Spring Harb Perspect Biol* **8** (2016).

33. M. Olejniczak, T. Dale, R. P. Fahlman, O. C. Uhlenbeck, Idiosyncratic tuning of tRNAs to achieve uniform ribosome binding. *Nat Struct Mol Biol* **12**, 788-793 (2005).

34. Y. M. Hou, P. Schimmel, A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* **333**, 140-145 (1988).

35. Y. E. Chong *et al.*, Distinct ways of G:U recognition by conserved tRNA binding motifs. *Proc Natl Acad Sci U S A* **115**, 7527-7532 (2018).

36. S. Ledoux, M. Olejniczak, O. C. Uhlenbeck, A sequence element that tunes Escherichia coli tRNA(Ala)(GGC) to ensure accurate decoding. *Nat Struct Mol Biol* **16**, 359-364 (2009).

37. M. Sprinzl, T. Hartmann, J. Weber, J. Blank, R. Zeidler, Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* **17 Suppl**, r1-172 (1989).

38. H. Murakami, A. Ohta, H. Suga, Bases in the anticodon loop of tRNA(Ala)(GGC) prevent misreading. *Nat Struct Mol Biol* **16**, 353-358 (2009).

39. P. Auffinger, E. Westhof, Singly and bifurcated hydrogen-bonded base-pairs in tRNA anticodon hairpins and ribozymes. *J Mol Biol* **292**, 467-483 (1999).

40. K. B. Gromadski, M. V. Rodnina, Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* **13**, 191-200 (2004).

41. K. B. Gromadski, T. Daviter, M. V. Rodnina, A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. *Mol Cell* **21**, 369-377 (2006).

42. A. B. Loveland, G. Demo, N. Grigorieff, A. A. Korostelev, Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546**, 113-117 (2017).

43. H. A. Nguyen, S. Sunita, C. M. Dunham, Disruption of evolutionarily correlated tRNA elements impairs accurate decoding. *Proceedings of the National Academy of Sciences* **117**, 16333-16338 (2020).

44. S. J. Nasvall, K. Nilsson, G. R. Bjork, The ribosomal grip of the peptidyl-tRNA is critical for reading frame maintenance. *J Mol Biol* **385**, 350-367 (2009).

45. M. Levi, K. Walak, A. Wang, U. Mohanty, P. C. Whitford, A steric gate controls P/E hybrid-state formation of tRNA on the ribosome. *Nature Communications* **11** (2020).

46. L. Lancaster, H. F. Noller, Involvement of 16S rRNA nucleotides G1338 and A1339 in discrimination of initiator tRNA. *Mol Cell* **20**, 623-632 (2005).

47. H. S. Zaher, R. Green, Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161-166 (2009).

48. H. S. Zaher, R. Green, Kinetic basis for global loss of fidelity arising from mismatches in the P-site codon:anticodon helix. *RNA* **16**, 1980-1989 (2010).

49. A. D. Petropoulos, M. E. Mcdonald, R. Green, H. S. Zaher, Distinct Roles for Release Factor 1 and Release Factor 2 in Translational Quality Control. *Journal of Biological Chemistry* **289**, 17589-17596 (2014).

50. H. A. Nguyen, E. D. Hoffer, C. M. Dunham, Importance of a tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in tRNACGGPro for decoding. *Journal of Biological Chemistry* **294**, 5281-5291 (2019).

51. CDC (2019) Antibiotic Resistance Threats in the United States. (U.S. Department of Health and Human Services, Atlanta, GA).

52. J. A. Jernigan *et al.*, Multidrug-Resistant Bacterial Infections in U.S. Hospitalized Patients, 2012–2017. *New England Journal of Medicine* **382**, 1309-1319 (2020).

53. R. E. Nelson *et al.*, National Estimates of Healthcare Costs Associated With Multidrug-Resistant Bacterial Infections Among Hospitalized Patients in the United States. *Clinical Infectious Diseases* **72**, S17-S26 (2021).

54. L. Chen, R. Todd, J. Kiehlbauch, M. Walters, A. Kallen, Notes from the Field: Pan-Resistant New Delhi Metallo-Beta-Lactamase-Producing Klebsiella pneumoniae - Washoe County, Nevada, 2016. *MMWR Morb Mortal Wkly Rep* **66**, 33 (2017).

55. H. J. Ahn *et al.*, Crystal structure of tRNA(m1G37)methyltransferase: insights into tRNA recognition. *EMBO J* **22**, 2593-2603 (2003).

56. Y. M. Hou, R. Matsubara, R. Takase, I. Masuda, J. I. Sulkowska, TrmD: A Methyl Transferase for tRNA Methylation With m(1)G37. *Enzymes* **41**, 89-115 (2017).

57. T. Christian, H. Gamper, Y. M. Hou, Conservation of structure and mechanism by Trm5 enzymes. *RNA* **19**, 1192-1199 (2013).

58. P. J. Hill *et al.*, Selective Inhibitors of Bacterial t-RNA-(N1G37) Methyltransferase (TrmD) That Demonstrate Novel Ordering of the Lid Domain. *Journal of Medicinal Chemistry* **56**, 7278-7288 (2013).

59. W. Zhong *et al.*, Thienopyrimidinone Derivatives That Inhibit Bacterial tRNA (Guanine37-N(1))-Methyltransferase (TrmD) by Restructuring the Active Site with a Tyrosine-Flipping Mechanism. *J Med Chem* 10.1021/acs.jmedchem.9b00582 (2019).

60. W. Zhong *et al.*, Targeting the Bacterial Epitranscriptome for Antibiotic Development: Discovery of Novel tRNA-(N(1)G37) Methyltransferase (TrmD) Inhibitors. *ACS Infect Dis* **5**, 326-335 (2019).

61. J. W. Chin, Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem* **83**, 379-408 (2014).

62. J. S. Italia *et al.*, Mutually Orthogonal Nonsense-Suppression Systems and Conjugation Chemistries for Precise Protein Labeling at up to Three Distinct Sites. *Journal of the American Chemical Society* **141**, 6204-6212 (2019).

63. M. Koh *et al.*, Site-Specific Incorporation of a Dithiolane Containing Amino Acid into Proteins. *Bioconjug Chem* **30**, 2102-2105 (2019).

64. S. Venkat *et al.*, Genetically Incorporating Two Distinct Post-translational Modifications into One Protein Simultaneously. *ACS Synthetic Biology* **7**, 689-695 (2018).

65. C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, P. G. Schultz, A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182-188 (1989).

66. D. R. Liu, T. J. Magliery, M. Pastrnak, P. G. Schultz, Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. *Proc Natl Acad Sci U S A* **94**, 10092-10097 (1997).

67. G. Srinivasan, C. M. James, J. A. Krzycki, Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**, 1459-1462 (2002).

68. W. H. Schmied, S. J. Elsasser, C. Uttamapinant, J. W. Chin, Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J Am Chem Soc* **136**, 15577-15583 (2014).

69. N. A. Aleksashin *et al.*, A fully orthogonal system for protein synthesis in bacterial cells. *Nat Commun* **11**, 1858 (2020).

70. A. E. d'Aquino, D. S. Kim, M. C. Jewett, Engineered Ribosomes for Basic Science and Synthetic Biology. *Annu Rev Chem Biomol Eng* 10.1146/annurev-chembioeng-060817-084129 (2018).

71. J. Fredens *et al.*, Total synthesis of Escherichia coli with a recoded genome. *Nature* **569**, 514-518 (2019).

72. W. E. Robertson *et al.*, Creating custom synthetic genomes in Escherichia coli with REXER and GENESIS. *Nature Protocols* **16**, 2345-2380 (2021).